

David Luengo, Sergey Subbotin,
Peter Arras, Yevgeniy Bodyanskiy,
Karsten Henke, Ivan Izonin,
Vitaly Levashenko, Volodymyr Lytvynenko,
Anzhelika Parkhomenko, Andreas Pester,
Natalya Shakhovska, Alexei Sharpanskykh,
Galyna Tabunshchuk, Carsten Wolff,
Heinz-Dietrich Wuttke, Elena Zaitseva

Computer Modeling and Intelligent Systems

Proceedings of the Second International Workshop
CMIS-2019

**Zaporizhzhia, Ukraine
April 15-19
2019**

Luengo D., Subbotin S., Arras P., Bodyanskiy Ye., Henke K., Izonin I., Levashenko V., Lytvynenko V., Parkhomenko A., Pester A., Shakhovska N., Sharpanskykh A., Tabunshchik G., Wolff C., Wuttke H.-D., Zaitseva E., (Eds.): The Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019). Zaporizhzhia, Ukraine, April 15-19, 2019, CEUR-WS.org, online

This volume represents the proceedings of the Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019), held in Zaporizhzhia, Ukraine, in April, 2019. It comprises 81 contributed papers that were carefully peer-reviewed and selected from 130 submissions.

Copyright © 2019 for the individual papers by the papers' authors.
Copying permitted only for private and academic purposes. This
volume is published and copyrighted by its editors.

Preface

It is our pleasure to present you the proceedings of the Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019), held in Zaporizhzhia, Ukraine on 15-19 of April 2019.

The main purpose of the CMIS-2019 workshop is a discussion of the recent research results in all areas of Computer Sciences and Information Technologies.

The workshop is soliciting literature review, survey and research papers including, whilst not limited to, the following areas of interest:

- Computer modeling;
- Computational intelligence;
- Signal and image processing;
- Information Systems and Technologies;
- CAD / CAM / CAE;
- IoT / IoE;
- ICT for Engineering, Medicine and Education.

The workshop joined scientists from Austria, Belarus, Belgium, China, Czech Republic, Germany, India, Netherlands, Poland, Romania, Slovakia, Spain, Swaziland, and Ukraine.

The language of CMIS-2019 Workshop is English. The workshop took the form of oral and poster presentations of peer-reviewed regular papers.

The Organizing Committee have received 130 submissions, out of which 81 were accepted for presentation as a regular papers. These papers were published in this volume of CMIS-2019 proceedings.

The workshop would not have been possible without the support of many people. First of all, we would like to thank our partners Ligth IT and Computools LLC. We would like to thank the Workshop Chairs and Program Committee Members, organization staff, and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the workshop and wish them every success in their research. It is the humble wish of the workshop organizers that the professional dialogue among the researchers, scientists, and engineers continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

April, 2019

David Luengo, Sergey Subbotin,
Peter Arras, Yevgeniy Bodyanskiy,
Karsten Henke, Ivan Izonin,
Vitaly Levashenko, Volodymyr Lytvynenko,
Anzhelika Parkhomenko, Andreas Pester,
Natalya Shakhovska, Alexei Sharpanskykh,
Galyna Tabunshchuk, Carsten Wolff,
Heinz-Dietrich Wuttke, Elena Zaitseva

Programme Committee

Co-chairs

Prof. **David Luengo**, Universidad Politecnica de Madrid, Spain

Prof. **Sergey Subbotin**, Zaporizhzhia National Technical University, Ukraine

Members

Dr. **Peter Arras**, Katholieke Universiteit Leuven, Belgium

Prof. **Yevgeniy Bodyanskiy**, Kharkiv National University of Radio Electronics, Ukraine

Dr. **Karsten Henke**, Technical University of Ilmenay, Germany

Dr. **Ivan Izonin**, Lviv Polytechnic National University, Ukraine

Prof. **Vitaly Levashenko**, University of Zilina, Slovak Republic

Prof. **Volodymyr Lytvynenko**, Kherson National Technical University, Ukraine

Dr. **Anzhelika Parkhomenko**, Zaporizhzhia National Technical University, Ukraine

Prof. **Andreas Pester**, Carinthia University of Applied Sciences, Austria

Prof. **Natalya Shakhovska**, Lviv Polytechnic National University, Ukraine

Dr. **Alexei Sharpanskykh**, Delft University of Technology, Netherlands

Dr. **Galyna Tabunshchyk**, Zaporizhzhia National Technical University, Ukraine

Prof. **Carsten Wolff**, Dortmund University of Applied Sciences and Arts, Germany

Dr. **Heinz-Dietrich Wuttke**, Technical University of Ilmenay, Germany

Prof. **Elena Zaitseva**, University of Zilina, Slovak Republic

Local Organizational Committee

Prof. **Sergey Subbotin**, Zaporizhzhia National Technical University, Ukraine

Dr. **Anzhelika Parkhomenko**, Zaporizhzhia National Technical University, Ukraine

Dr. **Galyna Tabunshchyk**, Zaporizhzhia National Technical University, Ukraine

Organizers



Zaporizhzhia National Technical University,
Zaporizhzhia, Ukraine
<http://www.zntu.edu.ua>



Department of Software Tools
of Zaporizhzhia National Technical University,
Zaporizhzhia, Ukraine
<http://pz.zntu.net>



Universidad Politécnica de Madrid,
Madrid, Spain
<http://www.upm.es/>



Katholieke Universiteit Leuven,
Leuven, Belgium
<https://www.kuleuven.be>



Delft University of Technology,
Delft, Netherlands
<https://www.tudelft.nl>



Ilmenau University of Technology,
Ilmenau, Germany
<https://www.tu-ilmenau.de>



Dortmund University of Applied Sciences and Arts,
Dortmund, Germany
<http://www.fh-dortmund.de/>



Carinthia University of Applied Sciences,
Carinthia, Austria
<http://www.fh-kaernten.at/>



Tadeusz Kościuszko University of Technology,
Cracow, Poland
<https://pk.edu.pl/>



University of Žilina,
Žilina, Slovak Republic
<https://www.uniza.sk>



Lviv Polytechnic National University,
Lviv, Ukraine
<http://www.lp.edu.ua>



Kharkiv National University of Radioelectronics,
Kharkiv, Ukraine
<http://nure.ua/>



Kherson National Technical University,
Kherson, Ukraine
<http://kntu.net.ua/>



State institution "M. Sytenko Institute of Spine and Joint Pathology of Academy of Medical Sciences of Ukraine»,
Kharkiv, Ukraine
<http://sytenko.org.ua/>



Israel Inter University Computation Center,
Israel
<https://www.iucc.ac.il/>



Sami Shamoon Academic College of Engineering,
Israel
<https://sce.ac.il/>



Holon Institute of Technology.
Holon, Israel
<https://www.hit.ac.il>

Partners



International project "Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees" (BIOART) of Erasmus+ Programme co-funded by the European Union (Ref. No. 586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP)



Ligth IT,
Zaporizhzhia, Ukraine
<https://light-it.net/>



Computools LLC,
Zaporizhzhia, Ukraine
<https://computools.com>

Parallel genetic method for the synthesis of recurrent neural networks for using in medicine

Serhii Leoshchenko¹[0000-0001-5099-5518], Andrii Oliinyk²[0000-0002-6740-6078],
Stepan Skrupsky³[0000-0002-9437-9095], Sergey Subbotin⁴[0000-0001-5814-8268] and
Viktor Lytvyn⁵[0000-0003-4061-4755]

^{1,2,4,5} Dept. of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia
69063, Ukraine

³Dept. of Computer Systems and networks, Dept. of Software Tools, Zaporizhzhia National
Technical University, Zaporizhzhia 69063, Ukraine

sergleo.zntu@gmail.com, olejnikaa@gmail.com,
sskrupsky@gmail.com, subbotin@zntu.edu.ua,
lytvynviktor.a@gmail.com

Abstract. In modern medicine, information technologies are widely used in the diagnosis and treatment of various diseases. The main task of creating such systems is to improve the quality of diagnosis and treatment. Therefore, the work aimed at finding new solutions in the creation of such systems are relevant. Despite all the advantages of neural networks, there are many difficulties in their implementation in medicine. In this paper are presented methods for solving the problem of recurrent neural network synthesis, which can be used as models in medical diagnostics.

Keywords: medical diagnosis, neural networks, synthesis, parallel, genetic method.

1 Introduction

Intensive development of medical science, expansion of possibilities of deepening in etiology, pathogenesis of disease, increase in data on markers of various pathological conditions dictates the necessity of searching new approaches to the processing of the results. Today, it is important to quickly analyze a large number of data and make the right decision, which can affect the prognosis, course and outcomes of the disease. In this case, more and more attention is paid to information technologies (IT), and in the scope of medicine it can be talked about electronic medicine [1]. IT is implemented in the form of special medical systems for various purposes and individual automated diagnostic and treatment devices.

The using IT allows to solve various tasks, which include prediction the risks of diseases, complications and treatment effectiveness, early diagnosis, treatment planning, monitoring the patient's health, automated analysis and statistical processing of clinical material. Medical systems significantly simplify the work in such situations when it is impossible to present the problem in numerical form, there is no certainty

or accuracy in the studied parameters or there is no one-digit algorithm for solving problems [2]. These characteristics are suitable for solving medical problems, which are a large amount of multidimensional, complex and sometimes contradictory clinical data obtained in the process of censored observations.

Currently, the use of statistical methods of data processing prevails in medical research. The most common descriptive methods used in traditional statistical studies are survival analysis and multivariate complex analysis classified as discriminant, cluster, factor, and correlation.

The fact that artificial neural network (ANN) are very successfully used in various fields, where it is necessary to solve the problems of forecasting, classification and management, explains the undying interest to the ANN methods, which has been observed recently. ANN have the ability to nonlinear modeling in combination with a relatively simple implementation and this makes them indispensable in solving complex multidimensional problems, including medical [3].

Today, there are many models of using ANN's architectures, which differ in their computational complexity, the degree of similarity with the living neurons of the brain, as well as having an exclusive and unique in its creation. Therefore, ANN are not subject to any classification standards in comparison with traditional statistical methods.

Existing ANN are able to work with both numerical data lying in a certain limited range and non-numerical parameters, for example, graphic images of various configurations. However, the non-standard scale of quantitative characteristics, the presence of missing values, the variability of nominal variables, the conversion of qualitative parameters into a numerical function or declaring them insignificant, create additional problems in the operation of the ANN and distort the output result.

From both scientific and practical point of view, one of the main advantages of using ANN is its ability to learn with data analysis, the establishment of complex and hidden connections and the subsequent presentation of independent results [4], [5]. In the process of training, when a large number of errors appear, it is possible to revise both the configuration of the network itself and to change the parameters included in its training [1].

Thus, the advantages of using ANN are:

- ability to learn from multiple qualitative and quantitative examples with unknown patterns between input and output data without fragmentation of the data sample. A more accurate description of the parameters, the ability to display the dynamics of the statistical properties of various indicators;
- effective data compression due to the construction of nonlinear mappings and the ability to visualize in the space of a smaller number of nonlinear principal component neural networks built;
- ability to make decisions based on absolute resistance to noise of input data and adaptation to environmental changes;
- modelling real situations solving tasks is done by analyzing the knowledge from their own experience of the ANN for an independent but. Minimal or complete absence of subjective factor influence on the final result. The ability to manually edit

- the values of individual parameters and their properties of an ANN, as well as other ways to include expert knowledge in the network;
- potential fault tolerance in hardware implementation of ANN;
- the possibility of using in situations that require the immediate adoption of the solution.

However, the use of ANN technologies for solving practical problems is associated with many difficulties. One of the dominant problems in the application of ANN's models is the unknown architecture of the projected neural network and its degree of complexity, which will be sufficient for the reliability of the result.

2 Review of the literature

In a number of works [6–16] was presented different algorithms to perform the ANNs training stage. The most common the Backpropagation method (BP), which allows you to adjust the weight of multi-layer complex ANNs using training sets. On the recommendation of E. Baum and D. Hassler [7, 8], the volume of the training set is directly proportional to the number of all ANN weights and inversely proportional to the proportion of erroneous decisions in the operation of the trained network [9, 10].

It should be noted that the BP method was one of the first methods for ANNs training. Most of all brings trouble indefinitely long learning process. In complex tasks, it can take days or even weeks to train a network, and it may not train at all. The cause may be one of the following [6, 11, 12].

1. Network paralysis. During network training, the weights can become very large as a result of the correction. This can cause all or most neurons to function at very high OUT values, in an area where the derivative of the compression function is very small. Since the error sent back in the learning process is proportional to this derivative, the learning process can practically freeze.
2. Local minimum. The network can hit a local when there are much deeper lows nearby. At the point of the local minimum, all directions lead up, and the network is unable to get out of it. Statistical training techniques can help avoid this trap, but they are slow.
3. Step size. The step size should be taken as final. If the step size is fixed and very small, the convergence is too slow, if it is fixed and too large, paralysis or constant instability may occur.

It should also be noted the possibility of retraining the network, which is rather the result of erroneous design of its topology. With too many neurons, the property of the network to generalize information is lost. The training set will be examined by the network, but any other sets, even very similar ones, may be misclassified.

The Backpropagation through time (BPTT) method has become a continuation, which is why it is faster. Moreover, it solves some of the problems of its predecessor. However, the BPTT experiences difficulties with local optima. In recurrent neural networks (RNN), the local optimum is a much more significant problem than in feed-

forward neural networks. Recurrent connections in such ANNs tends to create chaotic reactions in the error surface, resulting in local optima appearing frequently. Also in the blocks of RNN, when the error value propagates back from the output, the error is trapped in the part of the block. This is referred to as the “error carousel”, which constantly feeds the error back to each of the valves until they become trained to cut off this value. Thus, regular back propagation is effective when training an RNN unit to memorize values for very long durations [13, 14].

The main difference between genetic programming and genetic algorithms is that each individual in the population now encodes not the numerical characteristics that provide the optimality of the problem, but some solution to the problem. The term solution here refers to the configuration of the neural network.

At the moment, there are several reasons to criticize the genetic algorithm and genetic programming using. Below is a list of the main drawbacks of this approach.

- Re-evaluation of fitness function for complex problems is often a factor limiting the use of artificial evolution algorithms. Finding the optimal solution for a complex high-dimensional problem often requires a very costly evaluation of the fitness function.
- Genetic algorithms are poorly scalable to the complexity of the problem to be solved.
- Despite attempts to formalize genetic algorithms and the neuroevolutionary approach in particular [17], the theoretical basis remains scant.
- The solution is only more suitable than other solutions. As a result, the stopping criterion of the algorithm is unclear for each problem.
- In many problems, genetic algorithms tend to converge to a local optimum or even to controversial points, instead of a global optimum for the given problem [18].

However, many of these shortcomings can be corrected. For example, to prevent premature convergence, it is necessary to correctly select such parameters of the genetic algorithm as the population size and the percentage of individuals subjected to mutation. In addition, new variants of genetic operators are constantly being developed. The additional cost of recalculating the fitness function value can be avoided by using flags for cases where the fitness function value does not change over time.

However, in addition to the above drawbacks, genetic algorithms have quite a significant list of advantages.

- Scalability. Genetic algorithms can be easily adapted for parallel and multicores programming, so that due to the peculiarities of this approach, the corresponding overhead costs are significantly reduced.
- Universality. Genetic algorithms do not require any information about the response surface, they work with almost any tasks.
- Genetic algorithms may be used for tasks in which the value of the fitness function changes over time or depends on various changing factors.
- Even in cases where existing techniques work well, interesting results can be achieved by combining them with genetic algorithms, using them as a complement to proven methods.

- Gaps existing on the response surface have little effect on the full efficiency of optimization, which also allows to further expand their use.

Thus, taking into account all the advantages and disadvantages of genetic algorithms, it is possible to obtain a sufficiently universal system for solving the necessary problems [19] and, in particular, for optimization of the neural network.

3 Sequential modified genetic method of recurrent neural networks synthesis

In the method, which is proposed to find a solution using a population of neural networks: $P = \{NN_1, NN_2, \dots, NN_n\}$, that is, each individual is a separate ANN $Ind_i \rightarrow NN_i$ [18–20]. During initialization population divided into two halves, the genes $g_{Ind_i} = \{g_1, g_2, \dots, g_n\}$ of the first half of the individuals is randomly assigned $g_{Ind_i} = \{g_1 = \text{Rand}, g_2 = \text{Rand}, \dots, g_n = \text{Rand}\}$. Genes of the second half of the population are defined as the inversion of genes of the first half $g_{Ind_i} = \{\overline{g_1 = \text{Rand}}, \overline{g_2 = \text{Rand}}, \dots, \overline{g_n = \text{Rand}}\}$. This allows for a uniform distribution of single and zero bits in the population to minimize the probability of early convergence of the method ($p \rightarrow \min$).

After initialization, all individuals have coded networks in their genes without hidden neurons (N_h), and all input neurons (N_i) are connected to each output neuron (N_o). That is, at first, all the presented ANNs differ only in the weights of the interneuron connection w_i . In the process of evaluation, based on the genetic information of the individual under consideration, a neural network is first built, and then its performance is checked, which determines the fitness function ($f_{fitness}$) of the individual. After evaluation, all individuals are sorted in order of reduced fitness, and a more successful half of the sorted population is allowed to cross, with the best individual immediately moving to the next generation. In the process of reproduction, each individual is crossed with a randomly selected individual from among those selected for crossing. The resulting two descendants are added to the new generation $G = P' = \{Ind_1, Ind_2, \dots, Ind_n\}$. Once a new generation is formed the mutation operator starts working. However, it is important to note that the selection of the truncation significantly reduces the diversity within the population, leading to an early convergence of the algorithm, so the probability of mutation is chosen to be rather large ($p_{mut} = 15\text{-}25\%$) [20].

If the best individual in the population does not change for a certain number of generations (by default, it is proposed to set this number at eight), this individual is forcibly removed, and a new best individual is randomly selected from the queue. This makes it possible to realize the exit from the areas of local minima due to the relief of the objective function, as well as a large degree of convergence of individuals in one generation. The general scheme of the method demonstrated at Fig.1.

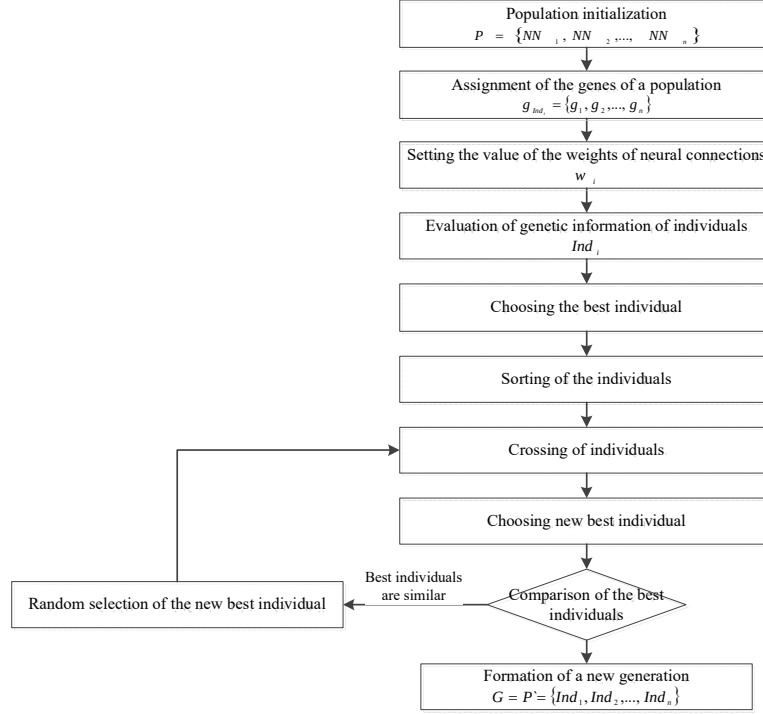


Fig. 1. The general scheme of the method

3.1 The calculation of the output layer of ANN

On condition using the support vector machine, the optimality criterion for calculating the output weights may not be specified. If the value of the mean square error is replaced by the criterion of the maximum separation of the support vectors, then the optimal linear weights of the output can be estimated using, for example, quadratic programming, as in the traditional method of support vectors, for this it is advisable to use the Evoke operator [21], by the formula:

$$y(t) = w_0 + \sum_{i=1}^k \sum_{j=0}^{l_i} w_{i,j} K(\phi(t), \phi^i(j)) \quad (1)$$

where $\phi(t) \in R^n$ is the output of a recurrent neural network $f(\cdot)$ at a time t ; $K(\cdot, \cdot)$ is a predefined kernel function; $w_{i,j}$ is weights corresponding to k training sequences ϕ^i , each length l_i , and are calculated using the support vector machine.

The value of the mean square error is replaced by the criterion of maximum separation of support vectors. In this case, the optimal linear weights can be estimated using quadratic programming, as in the traditional support vector machine.

One of the problems of neuroevolutionary method realization is the algorithm of ANN output calculation with arbitrary topology.

ANN can be represented as a directed planar graph. Based on the fact that the network structure can be any, loops and cycles containing any nodes are allowed in the graph, except for the nodes of the corresponding input neurons. Let denote the set of nodes of the graph by $V = \{v_i | i \in [0; N_v - 1]\}$, and a set of arcs through $E = \{e_j | j \in [0; N_e - 1]\}$, where N_v and N_e are accordingly, the number of nodes and arcs in the graph, and $N_v \equiv N_s$, and $N_e \equiv N_c$. The arc, which goes from node k to node l denote by an ordered pair $c_{k,l} = (v_k, v_l)$, the weight of the corresponding link will be denoted by $w_{k,l}$.

Give the index to the nodes of the graph as neurons, that is, the nodes that are the input neurons, called input. have an index out of range $[0; N_l - 1]$. By analogy, the indexes of outgoing nodes belong to the interval $[N_l; N_l + N_o - 1]$, and indexes for hidden nodes will be set in the interval $[N_l + N_o; N_v - 1]$.

Let introduce an additional characteristic for all nodes of the graph equal to the minimum length of the chain to any of the input nodes and denote it ch_i . Let's call ch_i the layer to which the i^{th} node belongs. Thus, all input nodes belong to the 0^{th} layer, not all input nodes that have input arcs from the input belong to the 1^{st} layer, all other nodes with input arcs from nodes of the 1^{st} layer will belong to the layer with index 2, etc. in this case, there may be situations when the node does not have input arcs, we will call it a hanging node with the layer number $ch_i = -1$.

For arcs, we also introduce an additional characteristic $b_{k,l}$ for the arc $c_{k,l}$, which is necessary to determine whether the arc corresponds to forward or reverse. It will be calculated as follows:

$$b_{k,l} = \begin{cases} 1, & ch_l - ch_k > 0 \\ -1, & ch_l - ch_k \leq 0 \end{cases} \quad (2)$$

That is, if the index of the layer of the end node of the arc is greater than the index of the layer of the beginning node, then we will consider such an arc as a straight line, otherwise we will consider the arc as an inverse.

Since each node of the graph represents a neuron, we denote by sum_i the value of the weighted sum of inputs, and through o_i is the value of the output (the value of the activation function of the i^{th} neuron-node). Then, $o_i = f_{fitness}(sum_i)$ where $f_{fitness}$ is the function of neuron activation.

Let's divide the whole process of signal propagation from the input nodes into stages, and during one such stage the signals manage to pass only one arc. The number of the stage is denoted by s . For the very first stage $s=1$. For short assumed that all arcs have the same length, and the signals are sewn on them instantly. We denote the

feature that the output of node i was updated at this stage through a_i , that is, if $a_i \equiv 1$, then the output of the node at stage s is calculated, otherwise, if $a_i \equiv 1$ is not.

Let's introduce one more designation $X = \{x_i | i \in [0; N_l - 1]\}$ it is vector of input signals. Then the algorithm for calculating the ANN output is as follows:

1. $o_i = x_i$, $a_i = 1$, for all $i \in [0; N_l - 1]$;
2. $o_i = 0$, for all $i \in [N_l; N_s - 1]$;
3. $s=1$;
4. $sum_i = 0$, $a_i = 1$, for all $i \in [N_l; N_s - 1]$;
5. if $s \equiv 1$, than go to the step number 7;
6. calculation of the feedback network. For all input feedbacks $c_{j,k}$ node v_k , where $k \in [N_l; N_s - 1]$: $sum_k = sum_k + o_j$, if $ch_j < s$;
7. if $a_i \equiv 0$, than $fn(i)$ for all $i \in [N_l; N_s - 1]$;
8. if the stop criterion is not met, than $s=s+1$ and go to the step number 4.

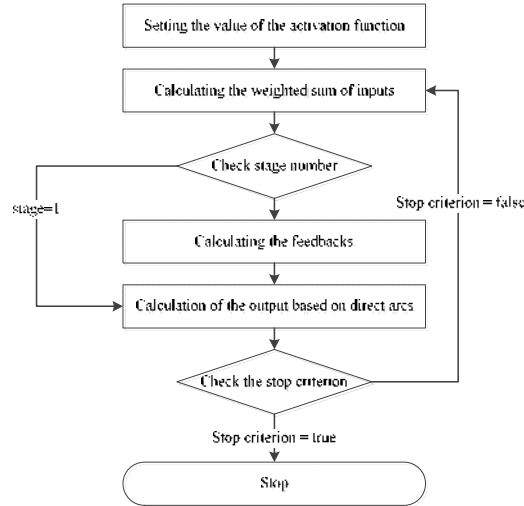


Fig. 2. The general scheme of the calculation of the output layer of ANN

Here $fn(i)$ is a recursive function that calculates the output of the 1st node taking into account all straight arcs. Works on the following algorithm:

1. if $ch_i < 0$, than go to the step number 3;
2. for all input arcs $c_{k,l}$ node v_i : if $a_k = 1$, than $sum_i = sum_i + o_k$, else $fn(k)$;
3. $o_i = f(sum_i)$;
4. exit.

The stopping criterion of the ANN output calculation algorithm can be one of the following:

- stabilization of values at the output of ANN;
- s exceeds the set value.

It is more reliable to calculate the output until the values at the output of ANN do not change, but for the case when the network contains cycles and/or loops, its output may never become stable. Therefore, the required additional stopping criteria limiting the maximum number of stages of calculation of network output.

4 Parallel genetic modified method for the synthesis of recurrent neural networks

Considering the features of the proposed modified genetic method for RNN synthesis, its parallel form can be represented as in Fig. 3. All stages of the method can be divided into 3 stages, separated by points of barrier synchronization. At the first stage, the main core initializes the population P , and adjusts the initial parameters of the method, namely: the stopping criterion, the population size, the criterion for adaptive selection of mutations. Next, the distribution of equal parts of the population (sub-populations) and initial parameters to the cores of the computer system is performed. Initialization of the initial population cannot be carried out in parallel on the cores of the system, because the generated independent populations intersect thus increasing the search for solutions. The second stage of the proposed method is performed in parallel by the cores of the system. All cores perform the same sequence of operations on their initial population. After the barrier synchronization, the main core receives the best solutions from the other cores and checks the stopping criterion. If it is, then the next generation (G) is formed. Otherwise, after changing the initial parameters, allowing the cores of the system getting the other solutions, return to the distribution of the initial parameters to the cores on the system is performed. And then the cores perform parallel calculations according to the second stage of the method.

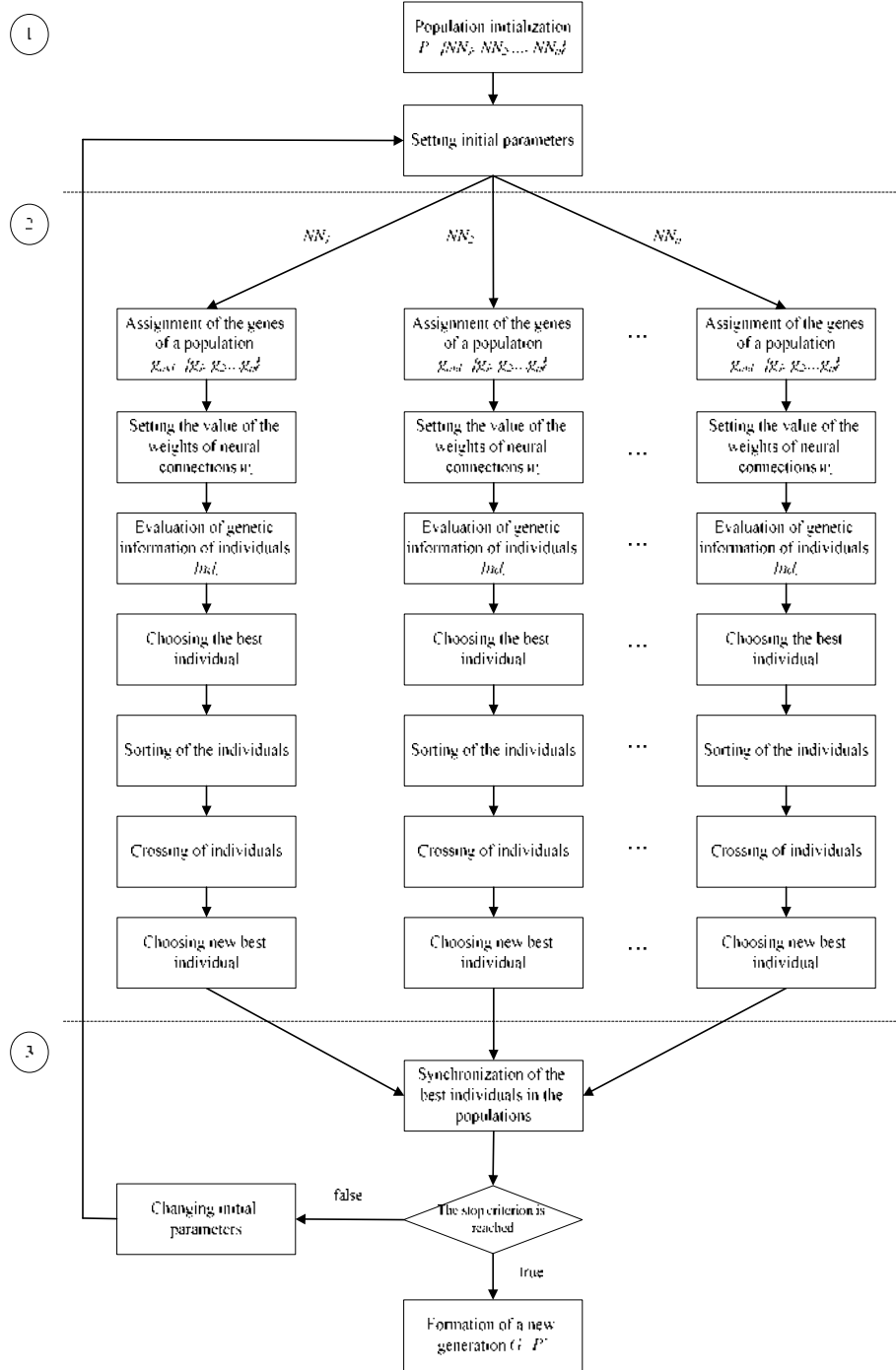


Fig. 3. Parallel genetic method for RNN synthesis

The proposed parallel method for RNN synthesis can be applied both on MIMD-systems [22] (clusters and supercomputers) and on SIMD (for example, graphics processors programmed with CUDA technology).

5 Experiments

The following hardware and software have been used for experimental verification of the proposed parallel genetic method for RNN synthesis [23]:

1. cluster of Pukhov Institute for Modeling in Energy Engineering National Academy of Sciences of Ukraine (IPME), Kyiv: processors Intel Xeon 5405, RAM – 4×2 GB DDR-2 for each node, communication environment InfiniBand 20Gb/s, middle-ware Torque and OMPI. MPI and Java threads programming models;
2. the computing system of the Department of software tools of Zaporizhzhya national technical university (ZNTU), Zaporizhzhya: Xeon processor E5-2660 v4 (14 cores), RAM 4x16 GB DDR4, the programming model of Java threads.
3. Nvidia GTX 960 graphics processor (GPU) with 1024 cores, which are programmed using CUDA technology.

During testing, the main task is to track the speed of the proposed method, quality and stability. Since synthesized RNN can be further used as diagnostic models for medical diagnosis, testing should be carried out on the relevant test data.

Data for testing were taken from the open repository – UC Irvine Machine Learning Repository. Data sample was used: Parkinson's Disease Classification Data Set [24]. The data used in this study were gathered from 188 patients with PD (107 men and 81 women) with ages ranging from 33 to 87 (65.1 ± 10.9). The data used in this study were gathered from 188 patients with PD (107 men and 81 women) with ages ranging from 33 to 87 (65.1 ± 10.9) at the Department of Neurology in Cerrahpasa Faculty of Medicine, Istanbul University. The control group consists of 64 healthy individuals (23 men and 41 women) with ages varying between 41 and 82 (61.1 ± 8.9). During the data collection process, the microphone is set to 44.1 KHz and following the physicians examination, the sustained phonation of the vowel was collected from each subject with three repetitions. Table 1 shows the main characteristics of the data sample.

Table 1. Main characteristics of the Parkinson's Disease Classification Data Set

Criterion	Characteristic	Criterion	Characteristic
Data Set Characteristics	Multivariate	Number of Instances	756
Attribute Characteristics	Integer, Real	Number of Attributes	754

6 The results analysis

In the Fig. 4 and 5 are graphs of the execution time (in minutes) of the proposed method on computer systems, which depends on the number of cores involved. It can be seen from the graphs that the proposed method has an acceptable degree of parallelism and is effectively performed on both MIMD and SIMD systems. This way, the IPME cluster was able to reduce the method execution time from 1565 minutes (on one core) to an acceptable 147 minutes on 16 cores. On the ZNTU the computing system, the method execution time was reduced from 1268 minutes on a single core to 110 minutes on 16 cores. The differences in the performance of the systems are due to their architectural features: in the cluster cores are connected by means of the Infini-Band communicator, and in the multi-core computer they are located on a single chip, which explains the smaller impact of overhead (transfers and synchronizations). In addition, the processor in multi-core computer supports Turbo Boost technology [25], making the time of the method execution on the single core much less than the execution time on the core of the cluster that does not support this technology. On a GPU with 960 cores involved, the execution time was 326.4 minutes, which can be adequately compared with the four cores of an IPME cluster or a ZNTU computing system.

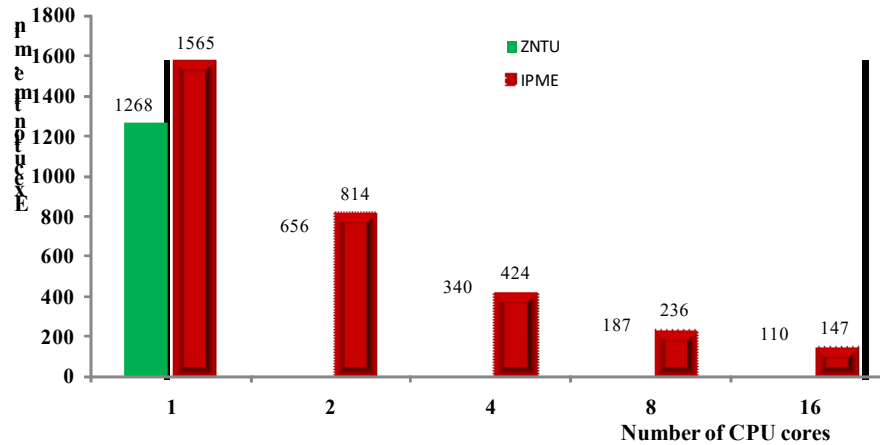


Fig. 4. Dependence the execution time of the proposed method to the number of involved cores of IPME cluster and ZNTU the computing system

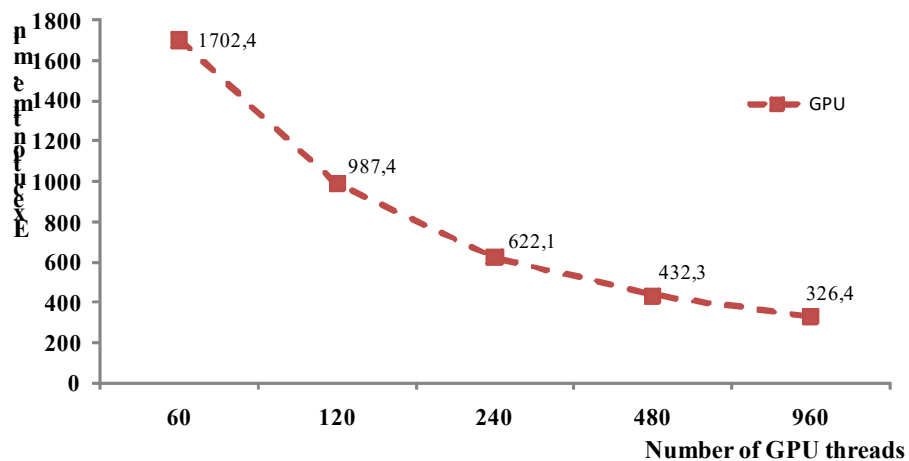


Fig. 5. Dependence the execution time of the proposed method to the number of GPU cores involved

The speedup graphics of calculations on a cluster IPME, ZNTU computing system and the GPU are shown in Fig. 6 and 7.

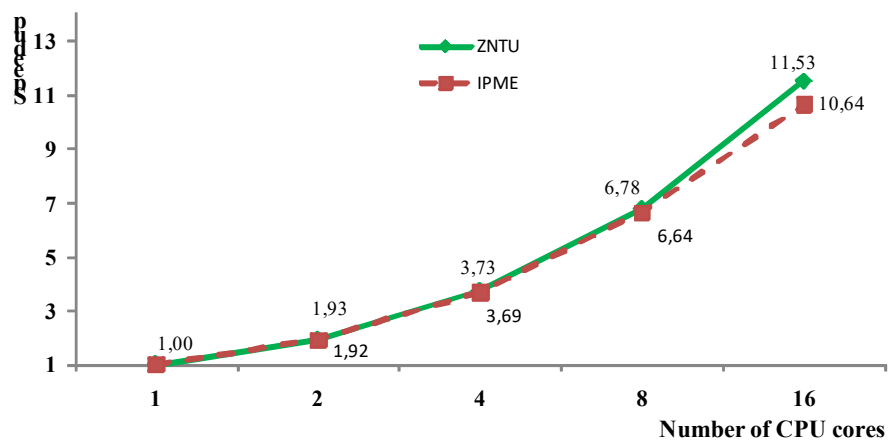


Fig. 6. The speedup graphics of calculations on a cluster IPME and ZNTU computing system

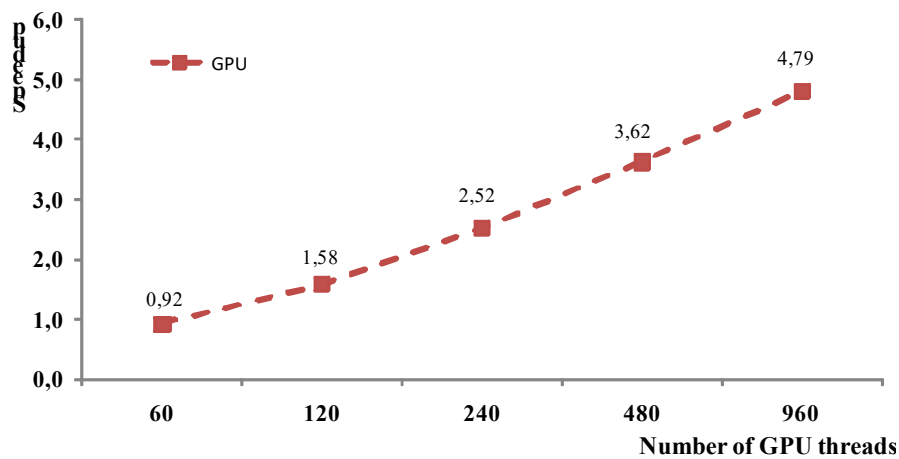


Fig. 7. The speedup graphics of calculations on a GPU

From the figures it is noticeable that the acceleration, though not linear, but approaches to linear. This is explained by the fact that communication overhead of the proposed method execution on computer systems is relatively small (Fig. 8, 9), and the number of parallel operations significantly exceeds the number of serial operations and synchronizations. In communication overhead, is understood the ratio of the time spent by the system for transfers and synchronization among cores to the time of target calculations on a given number of cores.

The graph of efficiency of computer systems IPME and ZNTU is presented in Fig. 10. It shows that the using of even 16 cores of computer systems for the implementation of the proposed method retains the efficiency at a relatively acceptable level and indicates the potential, if necessary and possibly, to use even more cores.

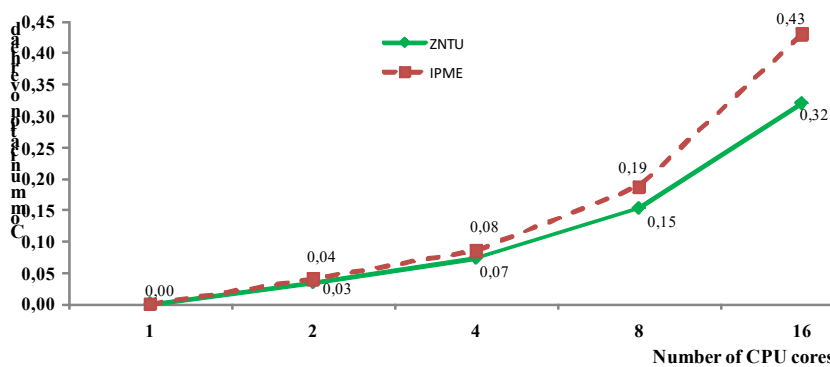


Fig. 8. Communication overhead performing the proposed method to the number of cores involved of IPME cluster and ZNTU the computing system

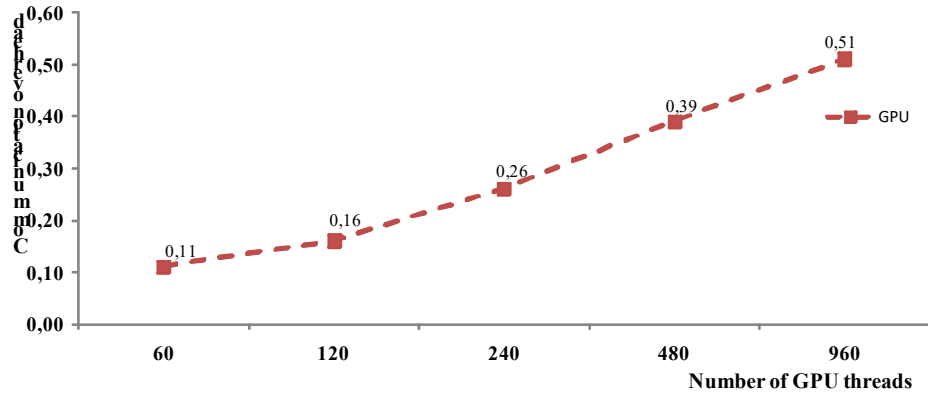


Fig. 9. Communication overhead performing the proposed method to the number of GPU cores involved

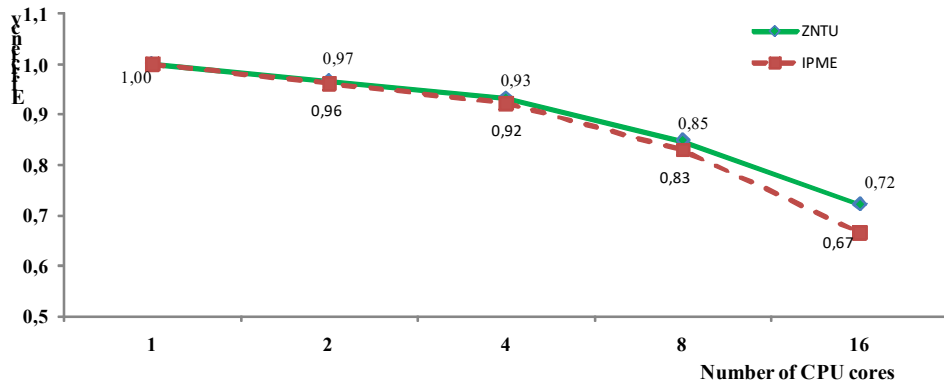


Fig. 10. The efficiency graph of IPME and ZNTU computing systems when executing the proposed method

Thus, the proposed method is well parallelized on modern computer architectures, which can significantly reduce the task: generate the models for future medical diagnosis execution time.

7 Conclusion

The problem of finding the optimal method of synthesis of ANN requires a comprehensive approach. Existing methods of ANNs training are well tested, but they have a number of nuances and disadvantages. The paper proposes a mechanism for the use of a modified genetic algorithm for its subsequent application in the synthesis of ANNs.

A model of parallel genetic method of RNS synthesis is proposed, which in comparison with the sequential implementation significantly speeds up the synthesis process.

ess. In the developed model is proposed to parallelize the most resource-intensive operations: the generation of RNS populations, the calculation of genetic information about individuals, which can significantly accelerate the process of finding the best solution in the synthesis of networks.

Based on the analysis of the experimental results, it can be argued about the good work of the proposed method. However, to reduce iterativity and improve accuracy, it should be continued to work towards parallelization of calculations.

Acknowledgment

The work was performed as part of the project “Methods and means of decision-making for data processing in intellectual recognition systems” (number of state registration 0117U003920) of Zaporizhzhia National Technical University.

References

1. Volchek, Y.A., Shyshko, V.M., Spiridonova, O.S., Mokhort, T.V.: Position of the model of the artificial neural network in medical expert systems. *Juvenis scientia* (9), 4–9. Scientia, Saint Petersburg (2017).
2. Shkaruplyo, V., Skrupsky, S., Oliinyk, A., Kolpakova T.: Development of stratified approach to software defined networks simulation. *EasternEuropean Journal of Enterprise Technologies*, vol. 89, issue 5/9, pp. 67–73 (2017). doi: 10.15587/1729-4061.2017.110142
3. Leoshchenko, S., Oliinyk, A., Subbotin, S., Gorobii, N., Zaiko, T.: Synthesis of artificial neural networks using a modified genetic algorithm. *Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018)*, pp. 1-13 (2018). dblp key: conf/iddm/PerovaBSKR18
4. Graves, A., Fernández, S., Gomez, F., Schmidhuber, J.: Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. In *Proceedings of the 23rd international conference on Machine learning (ICML '06)*. ACM, New York, pp. 369–376 (2006). doi: <https://doi.org/10.1145/1143844.1143891>.
5. Kotsur, M., Yarymbash, D., Kotsur, I., Bezverkhnia, Yu.: Speed Synchronization Methods of the Energy-Efficient Electric Drive System for Induction Motors. *IEEE: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) 2018*, pp. 304-307, Lviv-Slavske, Ukraine (2018). doi:10.1109/TCSET.2018.8336208
6. Van Tuc, N.: Approximation contexts in addressing graph data structures. *University of Wollongong Thesis Collection*, 30–55 (2015).
7. Barkoulas, J. T., Baum, Ch. F.: Long Term Dependence in Stock Returns. *Economics Letters*, vol. 53, no. 3, 253–259 pp. (1996).
8. Barkoulas, J. T., Baum, Ch. F., Travlos, N.: Long Memory in the Greek StockMarket. *Applied Financial Economics*, vol. 10, no. 2, 177–184 pp. (2000).
9. Kolpakova, T., Oliinyk, A., Lovkin, V.: Improved method of group decision making in expert systems based on competitive agents selection. *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Institute of Electrical and Electronics Engineers, pp. 939–943, Kyiv (2017). doi: 10.1109/UKRCON.2017.8100388

10. Stepanenko, O., Oliinyk, A., Deineha, L., Zaiko, T.: Development of the method for decomposition of superpositions of unknown pulsed signals using the second-order adaptive spectral analysis. *Eastern European Journal of Enterprise Technologies*, vol. 2, no 9, pp. 48–54 (2018). doi: 10.15587/1729-4061.2018.126578.
11. Handa, A., Patraucean, V.: Backpropagation in convolutional LSTMS. University of Cambridge Cambridge, pp. 1–5 (2015).
12. Boden M.: A guide to recurrent neural networks and backpropagation. Halmstad University, pp. 1–10 (2001).
13. Guo, J.: BackPropagation Through Time, pp. 1–6 (2013).
14. Yue, B., Fu, J., Liang, J.: Residual Recurrent Neural Networks for Learning Sequential Representations. *Information*, 9, 56 (2018).
15. Erofeeva, V.A.: Review of the theory of data mining based on neural networks [Obzor teorii intellektualnogo analiza dannykh na baze neyronnykh setey. Stokhasticheskaya optimizatsiya v informatike], 11 (3), pp. 3–17 (2015).
16. Yarymbash, D., Kotsur, M., Subbotin, S., Oliinyk, A.: A New Simulation Approach of the Electromagnetic Fields in Electrical Machines. *IEEE: The International Conference on Information and Digital Technologies*, July 5th - 7th, Zilina, Slovakia, 2017, Catalog Number CFP17CDT-USB, pp. 452–457, (2017). DOI: 10.1109/DT.2017.8024332
17. Balakrishnan K., Honavar V.: Properties of Genetic Representation of Neural Architectures. Iowa State University (1995).
18. Whitley D. Genetic Algorithms and Neural Networks. *Genetic Algorithms in Engineering and Computer Science*, pp. 203–216, (1995).
19. Mitchell M. An introduction to Genetic Algorithm. MIT Press (1996).
20. Hochreiter, S., Schmidhuber, J.: Long Short-Term Memory. *Neural Computation*, vol. 9, issue 8, pp. 1735–1780 (1997).
21. Schmidhuber, J., Wierstra, D., Gagliolo, M., Gomez, F.: Training Recurrent Networks by Evolino. *Neural computation*. vol. 19(3), 757–779 pp. (2007). doi: 10.1162/neco.2007.19.3.757.
22. Skillicorn, D.: Taxonomy for computer architectures. *Computer* (21), pp. 46–57 (1988). doi: 10.1109/2.86786.
23. Alsayaydeh, J.A., Shkaruplyo, V., Hamid, M.S., Skrupsky, S., Oliinyk, A.: Stratified Model of the Internet of Things Infrastructure, *Journal of Engineering and Applied Science*, vol. 13, issue 20, pp. 8634–8638, (2018). doi: 10.3923/jeasci.2018.8634.8638.
24. Parkinson's Disease Classification Data Set, <https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification>
25. Intel Turbo Boost Technology 2.0, <https://www.intel.com/content/www/us/en/architecture-and-technology/turbo-boost/turbo-boost-technology.html>

Online Neuro Fuzzy Clustering of Data with Omissions and Outliers based on Completion Strategy

Yevgeniy Bodyanskiy¹[0000-0001-5418-2143], Alina Shafronenko¹[0000-0002-8040-0279], and Diana Rudenko¹[0000-0002-1792-5080]

¹Kharkiv National University of Radio Electronics, Nauky Ave., 14, Kharkiv, Ukraine
yevgeniy.bodyanskiy@nure.ua,
alina.shafronenko@nure.ua,
diana.rudenko@nure.ua

Abstract. In the paper new recurrent adaptive algorithms for fuzzy clustering of data with missing values are proposed. This algorithm is based on fuzzy clustering procedures and self-learning Kohonen's rule using principle "Winner-Takes-More" with Cauchy neighborhood function. Using proposed approach it's possible to solve clustering task in on-line mode in situation when the amount of missing values in data is too big.

Keywords: fuzzy clustering, Kohonen self-organizing network, learning rule, incomplete data with missing values.

1 Introduction

The problem of data sets described by vector-images clustering often occurs in many applications associated with Data Mining, but recently the focus on fuzzy clustering [1-3], when processed vector-image with different levels of probabilities, possibilities or memberships, can belong to more than one class.

However, there are situations when the data sets contain omissions and outliers, the information that is lost. In this situation more effective is to use mathematical apparatus of Computational Intelligence [4] and, first of all artificial neural networks [5], that solve the task of restoring the lost observations, and modifications of the popular method of fuzzy c-means [6], which solve the problem of clustering without recovery of data.

Existing approaches [7,8] for data processing with omissions and outliers, are efficient in cases when the massive of the original observations is given in batch form and does not change during the processing. At the same time, there is a wide class of problems in which the data that arrive to the processing, have the form of sequence that is feed in real time as it occurs in the training of Kohonen self-organizing maps

[9] or their modifications [10]. In this regard we introduced [11] the adaptive neuro-fuzzy Kohonen network to solve the problem of clustering data with gaps based on the strategy of partial distances (PDS FCM). However, in situations where the number of such omissions and outliers is so much, the strategy of partial distances [12] may be not effective, and therefore it may be necessary, along with the solution of fuzzy clustering simultaneously estimate the missing observations. In this situation, a more efficient is approach that is based on the optimal expansion strategy (OCS FCM) [6]. This work is devoted to the task of on-line data clustering using the optimal expansion strategy, adapted to the case when information is processed in a sequential mode, and its volume is not determined in advance.

2 Probabilistic adaptive fuzzy clustering with missing data based on the optimal completion strategy

Baseline information for solving the task of clustering in a batch mode is the sample of observations, formed from N n -dimensional feature vectors $X = \{x_1, x_2, \dots, x_N\} \subset R^n$, $x_k \in X$, $k = 1, 2, \dots, N$. The result of clustering is the partition of original data set into m classes ($1 < m < N$) with some level of membership $U_q(k)$ of k -th feature vector to the q -th cluster ($1 \leq q \leq m$). Incoming data previously are centered and standardized by all features, so that all observations belong to the hypercube $[-1, 1]^n$. Therefore, the data for clustering form array $\tilde{X} = \{\tilde{x}_1, \dots, \tilde{x}_k, \dots, \tilde{x}_N\} \subset R^n$, $\tilde{x}_k = (\tilde{x}_{k1}, \dots, \tilde{x}_{ki}, \dots, \tilde{x}_{kn})^T$, $-1 \leq \tilde{x}_{ki} \leq 1$, $1 < m < N$, $1 \leq q \leq m$, $1 \leq i \leq n$, $1 \leq k \leq N$ that is, all observations $\tilde{x}_{ki}$ are available for processing.

Introducing the objective function of clustering [1]

$$E(U_q(k), w_q) = \sum_{k=1}^N \sum_{q=1}^m U_q^\beta(k) D^2(\tilde{x}_k, w_q)$$

with constraints $\sum_{q=1}^m U_q(k) = 1$, $0 < \sum_{k=1}^N U_q(k) < N$ and solving standard nonlinear programming problem, we get the probabilistic fuzzy clustering algorithm [2, 3]

$$\left\{ \begin{array}{l} U_q^{(\tau+1)}(k) = \frac{(D^2(\tilde{x}_k, w_q^{(\tau)}))^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (D^2(\tilde{x}_k, w_l^{(\tau)}))^{\frac{1}{1-\beta}}}, \\ w_q^{(\tau+1)} = \frac{\sum_{k=1}^N (U_q^{(\tau+1)}(k))^{\beta} \tilde{x}_k}{\sum_{k=1}^N (U_q^{(\tau+1)}(k))^{\beta}}, \end{array} \right. \quad (1)$$

where w_q - prototype (centroid) of q -th cluster, $\beta > 1$ - parameter that is called fuzzyfier and defines "vagueness" of boundaries between classes, $D^2(\tilde{x}_k, w_q)$ - the distance between $\tilde{x}_k$ and w_q in adopted metric, $\tau = 0, 1, 2, \dots$ - index of epoch of information processing which is organized as a sequence of $w_q^{(0)} \rightarrow U_q^{(1)} \rightarrow w_q^{(1)} \rightarrow U_q^{(2)} \rightarrow \dots$. The calculation process continues until satisfy the condition

$$\|w_q^{(\tau+1)} - w_q^{(\tau)}\| \leq \varepsilon \quad \forall 1 \leq q \leq m,$$

(here ε - defines threshold of accuracy) or until the specified maximum number of epochs Q ($\tau = 0, 1, 2, \dots, Q$).

Note also that when $\beta = 2$ and

$$D^2(\tilde{x}_k, w_q) = \|\tilde{x}_k - w_q\|^2,$$

we get a popular algorithm of Bezdek's fuzzy c-means (FCM) [1].

The process of fuzzy clustering can be organized in on-line mode as sequentially processing. At this situation batch algorithm (1) can be rewritten in recurrent form [12]

$$\left\{ \begin{array}{l} U_q(k+1) = \frac{(D^2(\tilde{x}_{k+1}, w_q(k)))^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (D^2(\tilde{x}_{k+1}, w_l(k)))^{\frac{1}{1-\beta}}}, \\ w_q(k+1) = w_q(k) + \eta(k+1) U_q^{\beta}(k+1) (\tilde{x}_{k+1} - w_q(k)), \end{array} \right. \quad (2)$$

where $\eta(k+1)$ - learning rate parameter, $U_q^{\beta}(k+1)$ - bell-shaped neighborhood function of neuro-fuzzy Kohonen network (Cauchy function), designed to solve the problems of fuzzy clustering [10, 11], based on the principle "Winner Takes More» (WTM) [9].

In the presence of an unknown number of missing values in vector images $\tilde{x}_k$, that form array $\tilde{X}$, following [6], we introduce the sub-arrays:

$$X_F = \{\tilde{x}_k \in \tilde{X} \mid \tilde{x}_k - \text{vector containing all components}\};$$

$$X_P = \{\tilde{x}_{ki}, 1 \leq i \leq n, 1 \leq k \leq N \mid \text{values } \tilde{x}_k, \text{ available in } \tilde{X}\};$$

$$X_G = \{\tilde{x}_{ki} = ?, 1 \leq i \leq n, 1 \leq k \leq N \mid \text{values } \tilde{x}_k, \text{ absent in } \tilde{X}\}.$$

The optimal completion strategy consists in the fact that the elements of sub-array X_G are considered as additional variables, which are estimated by minimization of objective function E . Thus, in parallel with clustering (optimization E by $U_q(k)$ and w_q) estimation of missing observations is made (optimization E by $\tilde{x}_{ki} \in X_G$). In this case, the algorithm of fuzzy c-means based on the optimal expansion strategy can be written as the following sequence of steps [6]:

1. Setting the initial conditions for the algorithm: $\beta > 0$; $1 < m < N$; $\varepsilon > 0$;

$$w_q^{(0)}; 1 \leq q \leq m;$$

$$\tau = 0, 1, 2, \dots, Q; X_G^{(0)} = \{-1 \leq \hat{x}_{ki}^{(0)} \leq 1\}, \text{ where } X_G^{(0)} - N_G (1 \leq N_G \leq (n-1)N)$$

arbitrary initial estimates $\hat{x}_{ki}^{(0)}$ of missing values $\tilde{x}_{ki} \in X_G$;

2. Calculation of membership levels by solving the optimization problem:

$$U_q^{(\tau+1)}(k) = \arg \min_{U_q(k)} E(U_q(k), w_q^{(\tau)}, X_G^{(\tau)}) = \frac{(D^2(\hat{x}_k^{(\tau)}, w_q^{(\tau)}))^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (D^2(\hat{x}_k^{(\tau)}, w_l^{(\tau)}))^{\frac{1}{1-\beta}}}$$

$$= \frac{(\|\hat{x}_k^{(\tau)} - w_q^{(\tau)}\|^2)^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (\|\hat{x}_k^{(\tau)} - w_l^{(\tau)}\|^2)^{\frac{1}{1-\beta}}}$$

(here vector $\hat{x}_k^{(\tau)}$ differs from $\tilde{x}_k$ by replacing missing values $\tilde{x}_{ki} \in X_G$ by estimates $\hat{x}_{ki}^{(\tau)}$ that are calculated for the τ -th epoch of data processing);

3. Calculation the centroids of clusters:

$$w_q^{(\tau+1)} = \arg \min_{w_q} E(U_q^{(\tau+1)}(k), w_q, X_G^{(\tau)}) = \frac{\sum_{k=1}^N (U_q^{(\tau+1)}(k))^{\beta} \hat{x}_k^{(\tau)}}{\sum_{k=1}^N (U_q^{(\tau+1)}(k))^{\beta}};$$

4. Checking the stop conditions:

if $\|w_q^{(\tau+1)} - w_q^{(\tau)}\| < \varepsilon \ \forall \ 1 \leq q \leq m$ or $\tau = Q$, then the algorithm terminates, otherwise go to step 5;

5. Estimation of missing observations by solving the optimization problem:

$$X_G^{(\tau+1)} = \arg \min_{X_G} E(U_q^{(\tau+1)}(k), w_q^{(\tau+1)}, X_G)$$

or, equivalently

$$\frac{\partial E(U_q^{(\tau+1)}(k), w_q^{(\tau+1)}, X_G)}{\partial \hat{x}_{ki}} = 0,$$

that leads to

$$\hat{x}_{ki}^{(\tau+1)} = \sum_{q=1}^m (U_q^{(\tau+1)}(k))^\beta w_{qi}^{(\tau+1)} / \sum_{q=1}^m (U_q^{(\tau+1)}(k))^\beta.$$

Information processing with this algorithm is organized as a sequence

$$\begin{aligned} w_q^{(0)} \rightarrow U_q^{(1)} \rightarrow \hat{x}_{ki}^{(1)} \rightarrow w_q^{(1)} \rightarrow U_q^{(2)} \rightarrow \dots \rightarrow w_q^{(\tau)} \rightarrow U_q^{(\tau+1)} \rightarrow \\ \rightarrow \hat{x}_{ki}^{(\tau+1)} \rightarrow w_q^{(\tau+1)} \rightarrow \dots \rightarrow w_q^{(Q)}. \end{aligned}$$

3 Possibilistic adaptive fuzzy clustering with missing data based on the optimal completion strategy

The main disadvantage of conventional probabilistic algorithms is connected with the constraints on membership levels which sum has to be equal unity. This reason has led to the creation of possibilistic fuzzy clustering algorithms [14].

In possibilistic clustering algorithms the objective function has the form

$$E(U_q(k), w_q, \mu_q) = \sum_{k=1}^N \sum_{q=1}^m U_q^\beta(k) D^2(\tilde{x}_k, w_q) + \sum_{q=1}^m \mu_q \sum_{k=1}^N (1 - U_q(k))^\beta$$

(where the scalar parameter $\mu \geq 0$ determines the distance at which level of membership equals to 0.5, i.e. if $D^2(\tilde{x}_k, w_q) = \mu_q$, then $w_q(k) = 0.5$) and its minimization by $U_q(k)$, w_q , μ_q leads to the result:

$$\left\{ \begin{aligned} U_q^{(\tau+1)}(k) &= 1 / (1 + (D^2(\tilde{x}_k, w_q^{(\tau)}) / \mu_q^{(\tau)})^{\frac{1}{\beta-1}}), \\ w_q^{(\tau+1)} &= \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta \tilde{x}_k / \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta, \\ \mu_q^{(\tau+1)} &= \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta D^2(\tilde{x}_k, w_q^{(\tau+1)}) / \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta. \end{aligned} \right. \quad (3)$$

In the on-line processing recurrent form of the algorithm (3) can be written as [10,13]:

$$\begin{cases} U_q(k+1) = 1 / \left(1 + \left(\frac{D^2(\tilde{x}_{k+1}, w_q(k))}{\mu_q(k)} \right)^{\frac{1}{\beta-1}} \right), \\ w_q(k+1) = w_q(k) + \eta(k+1) U_q^\beta(k+1) (\tilde{x}_{k+1} - w_q(k)), \\ \mu_q(k+1) = \sum_{p=1}^{k+1} U_q^\beta(p) D^2(\tilde{x}_p, w_q(k+1)) / \sum_{p=1}^{k+1} U_q^\beta(p), \end{cases} \quad (4)$$

and the second relation in (4) differs only in the form of the neighborhood function $U_q^\beta(k+1)$. Thus, the recurrent possibilistic fuzzy clustering is based on Kohonen competitive self learning.

Considering the situation with missing observations and using the optimal completion strategy, as the objective function of possibilistic fuzzy clustering use the expression:

$$E(U_q(k), w_q, \mu_q, X_G) = \sum_{k=1}^N \sum_{q=1}^m U_q^\beta(k) D^2(\hat{x}_k, w_q) + \sum_{q=1}^m \mu_q \sum_{k=1}^N (1 - U_q(k))^\beta$$

and its minimization leads to the system of obvious relations:

$$\begin{cases} \partial E(U_q(k), w_q, \mu_q, X_G) / \partial U_q(k) = 0, \\ \nabla_{w_q} E(U_q(k), w_q, \mu_q, X_G) = \vec{0}, \\ \partial E(U_q(k), w_q, \mu_q, X_G) / \partial \hat{x}_{ki} = 0, \\ \partial E(U_q(k), w_q, \mu_q, X_G) / \partial \mu_q = 0, \end{cases}$$

or

$$\begin{cases} U_q^{(\tau+1)}(k) = 1 / \left(1 + \left(\frac{D^2(\hat{x}_k^{(\tau)}, w_q^{(\tau)})}{\mu_q^{(\tau)}} \right)^{\frac{1}{\beta-1}} \right), \\ w_q^{(\tau+1)} = \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta \hat{x}_k^{(\tau)} / \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta, \\ \hat{x}_{ki}^{(\tau+1)} = \sum_{q=1}^m (U_q^{(\tau+1)}(k))^\beta w_{qi}^{(\tau+1)} / \sum_{q=1}^m (U_q^{(\tau+1)}(k))^\beta, \\ \mu_q^{(\tau+1)} = \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta D^2(\hat{x}_k^{(\tau+1)}, w_q^{(\tau+1)}) / \sum_{k=1}^N (U_q^{(\tau+1)}(k))^\beta. \end{cases} \quad (5)$$

Similarly to probabilistic fuzzy clustering based on optimal completion strategy we can organize the possibilistic clustering process with missing observations

$$\left\{ \begin{array}{l} U_q^{(\tau+1)}(k+1) = \frac{1}{1 + \left(\frac{\|\hat{x}_{k+1}^{(\tau)} - w_q(k)\|^2}{\mu_q^{(\tau)}} \right)^{\frac{1}{\beta-1}}}, \\ \hat{x}_{k+1,i}^{(\tau+1)} = \frac{\sum_{q=1}^m (U_q^{(\tau+1)}(k+1))^\beta w_{qi}(k)}{\sum_{q=1}^m (U_q^{(\tau+1)}(k+1))^\beta}, \\ \mu_q^{(\tau+1)}(k+1) = \frac{\sum_{p=1}^{k+1} (U_q^{(\tau+1)}(p))^\beta \|\hat{x}_p^{(\tau+1)} - w_q(k)\|^2}{\sum_{p=1}^{k+1} (U_q^{(\tau+1)}(p))^\beta}, \\ w_q(k+1) = w_q(k) + \eta(k+1)(U_q^{(Q)}(k+1))^\beta (\hat{x}_{k+1}^{(Q)} - w_q(k)), \end{array} \right. \quad (6)$$

or in accelerate time

$$\left\{ \begin{array}{l} U_q^{(\tau+1)}(k+1) = \frac{1}{1 + \left(\frac{\|\hat{x}_{k+1}^{(\tau)} - w_q^{(\tau)}(k+1)\|^2}{\mu_q^{(\tau)}} \right)^{\frac{1}{\beta-1}}}, \\ w_q^{(0)}(k+1) = w_q^{(Q)}(k), \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1)(U_q^{(\tau+1)}(k+1))^\beta * \\ \quad * (\hat{x}_{k+1}^{(\tau)} - w_q^{(\tau)}(k+1)), \\ \hat{x}_{k+1,i}^{(\tau+1)} = \frac{\sum_{q=1}^m (U_q^{(\tau+1)}(k+1))^\beta w_{qi}^{(\tau+1)}(k+1)}{\sum_{q=1}^m (U_q^{(\tau+1)}(k+1))^\beta}, \\ \mu_q^{(\tau+1)}(k+1) = \frac{\sum_{p=1}^{k+1} (U_q^{(\tau+1)}(p))^\beta \|\hat{x}_p^{(\tau+1)} - w_q^{(\tau+1)}(k+1)\|^2}{\sum_{p=1}^{k+1} (U_q^{(\tau+1)}(p))^\beta}. \end{array} \right. \quad (7)$$

It's understandable that the algorithm (7) from computational point of view is the most unwieldy; however, its advantage is that it can be used in on-line mode to detect the emergence of new clusters. In the procedures (6), (7) is unnecessary accumulate the processed sample, which is important in problems (for example, Web Mining), where large amounts of information have to be processed.

4 Experiments

Experimental research conducted on two standard samples of data such as Wine and Iris UCI repository.

To estimate the quality of the algorithm we used quality criteria partitioning into clusters such as: Partition Coefficient (PC), Classification Entropy (CE), Partition Index (SC), Separation Index (S), Xie and Beni's Index (XB), Dunn's Index (DI).

We also compared the results of our proposed algorithm with other more well-known such as Fuzzy C-means (FCM) clustering algorithm and Gustafson-Kessel clustering algorithm.

As seen from the experimental results (Table 1, Table 2 and Table 3), the proposed algorithm shown better results than the FCM and Gustafson-Kessel clustering algorithm.

Table 1. Results of experiments with 10 missing values

Algorithms	Gaps	Iris UCI repository							Wine UCI repository					
		PC	CE	SC	S	XB	DI	PC	CE	SC	S	XB	DI	PC
Adaptive fuzzy possibilistic clustering data with missing values	10	9,1249e-07	-2,3308e-04	0,3733	48,7067	24,4028	0,2005	0,582e-13	-2,28375e-04	3,6936e+05	2,7115e+08	1,359e+08	0,0109	4,56245e-07
FCM		0,3808	0,21415	0,00715	0,7473e-04	1,92845	0,01375	0,3954	0,1903	3,6674e-04	3,42085e-06	2,8555	0,00585	0,38085
Gustafson-Kessel		0,4731	0,05725	0,2395	0,0032	1,7309	0,1699	0,27535	0,31965	4,29665	0,02415	0,5375	0,05075	0,4731

Table 2. Results of experiments with 50 missing values

Algorithms	Gaps	Iris UCI repository							Wine UCI repository					
		PC	CE	SC	S	XB	DI	PC	CE	SC	S	XB	DI	PC
Adaptive fuzzy possibilistic clustering data with missing values	50	9,1249e-07	-4,6617e-04	0,3775	48,7067	48,8301	0,3365	7,5181e-12	-5,4992e-04	1,1834e+04	4,1981e+06	4,2024e+06	0,0240	9,1249e-07
FCM		0,7399	0,4632	0,0174	1,8345e-04	4,4887	0,0355	0,7892	0,3838	7,6110e-04	7,1760e-06	8,8618	0,0237	0,7399
Gustafson-Kessel		0,9422	0,1177	0,5219	0,0035	3,4413	0,3341	0,5824	0,6010	4,7678	0,0268	1,1703	0,1030	0,9422

Table 3. Results of experiments with 100 missing values

Algorithms	Gaps	Iris UCI repository							Wine UCI repository					
		PC	CE	SC	S	XB	DI	PC	CE	SC	S	XB	DI	PC
Adaptive fuzzy possibilistic clustering data with missing values	100	1,7429e-06	-4,6617e-04	0,1209	25,5	22,7008	0,3497	2,1059e-11	-5,7912e-04	4,3497e+03	1,4587e+06	1,5006e+06	0,0236	1,7329e-06
FCM		0,8120	0,3555	0,0181	9,2846e-05	3,9633	0,0579	0,7798	0,3731	7,6342e-04	7,3466e-06	8,2453	0,3590	0,8131
Gustafson-Kessel		0,8335	0,2850	0,9994	0,0060	2,9995	0,1236	0,7974	0,4854	7,1553	0,0144	1,3627	0,0974	0,8254

5 Conclusions

The problem of probabilistic and possibilistic fuzzy adaptive clustering, containing a priori unknown number of gaps, based on the optimal completion of data strategy is considered. The proposed algorithms are based on the recurrent optimization of a special type of goal functions. Missing observations are replaced by their estimates also obtained in the solution of optimization problem. Centroids of recovered clusters are tuned using a procedure close to the T.Kohonen WTM-rule with the function of the neighborhood (membership), having the Cauchian form.

References

1. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981).
2. Hoepfner, F., Klawonn, F., Kruse, R., Runkel, T.: Fuzzy Clustering Analysis: Methods for Classification, Data Analysis and Image Recognition. Chichester, John Wiley & Sons (1999)
3. Xu, R., Wunsch, D.C.: Clustering. Hoboken, N.J. John Wiley & Sons, Inc. (2009)
4. Rutkowski, L.: Computational Intelligence. Methods and Techniques. Berlin-Heidelberg: Springer-Verlag (2008)
5. Marwala, T.: Computational Intelligence for Missing Data Imputation, Estimation, and Management: Knowledge Optimization Techniques. Hershey-New York, Information Science Reference (2009)
6. Hathaway, R.J., Bezdek, J.C.: Fuzzy c-means clustering of incomplete data. IEEE Trans. on Systems, Man, and Cybernetics, №5, 31, pp. 735-744 (2001)
7. Zagoruyko, N.G.: Empirical predictions. Novosibirsk, Nauka (1979)
8. Zagoruyko, N.G.: Applied Data Analysis and Knowledge. Novosibirsk (1999)
9. Kohonen, T.: Self-Organizing Maps. Berlin: Springer-Verlag (1995)
10. Gorshkov, Ye., Kolodyazhnyi, V., Bodyanskiy, Ye.: New recursive learning algorithms for fuzzy Kohonen clustering network. Proc. 17th Int. Workshop on Nonlinear Dynamics of Electronic Systems. (Rapperswil, Switzerland, June 21-24, 2009) Rapperswil, Switzerland, pp. 58-61 (2009)
11. Bodyanskiy, Ye., Shafronenko, A., Volkova, V.: Adaptive clustering of incomplete data using neuro-fuzzy Kohonen network. In "Artificial Intelligence Methods and Techniques for Business and Engineering Applications" – Rzeszow-Sofia: ITHEA, pp. 287-296 (2012)
12. Shafronenko, A., Dolotov, A., Bodyanskiy, Y., Setlak, G.: Fuzzy Clustering of Distorted Observations Based on Optimal Expansion Using Partial Distances. In "2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)", pp. 327 – 330 (2018)
13. Bodyanskiy, Ye.: Computational intelligence techniques for data analysis. Lecture Notes in Informatics. Bonn: GI (2005)
14. Krishnapuram, R., Keller, J.M.: A possibilistic approach to clustering. Fuzzy Systems, №2, pp.98-110 (1993)

Optimization on Combinatorial Configurations Using Genetic Algorithms

Sergiy Yakovlev¹[0000-0003-1707-843X], Oleksii Kartashov²[0000-0002-6282-553X],

Oksana Pichugina²[0000-0002-7099-8967]

¹ National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
svsyak7@gmail.com

² National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
alexeykartashov@gmail.com

³ National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
oksanapichugina1@gmail.com

Abstract. An optimization problem on a set of Euclidean combinatorial configurations is formulated. Peculiarities of applying genetic algorithms to solving this class of problems are explored. Principles of formation of an initial population and selection mechanisms are described. A choice of crossover and mutation operators is justified. Examples of the construction of crossover operators for sets of Euclidean configurations are given. A genetic algorithm of optimization on permutation configurations sets is presented. Based on the algorithm, a random search approach is offered for optimization on spherically-located and well-described sets. The algorithm was tested on a problems of balancing masses of rotating objects.

Keywords: combinatorial optimization, genetic algorithm, crossover, mutation, permutation, balancing of masses.

1 Introduction

Traditionally, combinatorial optimization problems are considered as difficult [1]-[3], which leads to the necessity of developing effective approximate methods for their solution. Nowadays, development of theory and methods of computational intelligence regarding problems of combinatorial optimization is of interest of researchers. Of particular importance is a class of evolutionary methods [4]-[6], to which genetic algorithms belong [7]-[9]. Modern publications in this direction [10]-[14] prove the effectiveness of applying genetic and other evolutionary algorithms in solving combinatorial optimization problems.

In the development of combinatorial optimization theory, an important place is occupied by an area of formalization of concepts such as “combinatorial set”, “combinatorial object”, “combinatorial configuration”. Not of less importance is investigating properties of functions given on these sets. Combinatorial configuration is one of the fundamental concepts. Depending on classes of combinatorial configuration sets,

various optimization problems arise. Respectively, methods of their solving are highly determined by the properties of these configurations' sets.

In this paper, we consider a class of so-called Euclidean combinatorial configurations as the basis of developing new approaches to solving combinatorial optimization problems by genetic algorithms.

2 The Euclidean Combinatorial Configurations

By a configuration [15] we mean a mapping

$$\psi : U \rightarrow V \quad (1)$$

of some finite initial set U of elements of arbitrary nature into an abstract set V with a certain structure if a given set Ω of constraints holds. When both U and V are finite, the configuration (1) is called combinatorial.

We represent the combinatorial configuration of a tuple $\langle \psi, U, V, \Omega \rangle$, where $U = \{u_1, \dots, u_n\}$ – is initial set, $V = \{v_1, \dots, v_k\}$ – is resulting set, ψ is a mapping of type (1), Ω – is a given system of constraints on the mapping ψ .

The papers [16-18] are devoted to the study of combinatorial configurations. Further development of the concept of a combinatorial configuration was made by weakening the conditions on the finiteness of V [11, 18]. Thus, it is assumed that the resulting set V can be countable, and a combinatorial configuration, in this case, is called a combinatorial object. In [11, 18] another direction for a generalization of the concept is proposed. Namely, combinatorial objects of order k are defined that allows to expand significantly the range of real-world problems that can be formalized.

We focus on considering a class of combinatorial configurations where elements of the resulting set V are numerical vectors. The selection of such a class is justified by a wide range of real problems, in which the elements of the initial set U are characterized by a certain set of numerical parameters (for example, physical and metric characteristics). First of all, it concerns the placement problems for geometric objects and other problems of Geometric Design. The synthesis of spatial configurations is based on the concept of configuration spaces of geometric objects proposed in [19-22].

Let us $\mathbf{B}$ be the set of vectors of a space R^m of the same dimension m , i.e., $\mathbf{b}_l = (b_{1l}, \dots, b_{ml})^T \in R^m$, $l \in J_k$. Let $\mathbf{B}$ be the resulting set V . Then the configuration π will be an ordered sequence of vectors of $\mathbf{b}_{j_1}, \mathbf{b}_{j_2}, \dots, \mathbf{b}_{j_n} \in \mathbf{B}$. Each configuration $\pi = \{\mathbf{b}_{j_1}, \mathbf{b}_{j_2}, \dots, \mathbf{b}_{j_n}\}$ is put in a one-to-one correspondence to a vector $\mathbf{x} = (x_1, \dots, x_N) \in R^N$, $N = nm$, i.e., there exists a bijection φ such that:

$$\mathbf{x} = \varphi(\pi), \quad \pi = \varphi^{-1}(\mathbf{x}). \quad (2)$$

For example, such correspondence can be specified as follows:

$$\mathbf{x} = (x_1, \dots, x_N) = (b_{1j_1}, \dots, b_{1j_n}, b_{2j_1}, \dots, b_{2j_n}, \dots, b_{mj_1}, \dots, b_{mj_n}).$$

Definition. A Euclidean combinatorial configuration (an e-configuration) is a mapping $\varphi: \langle \chi, U, \mathbf{B}, \Lambda \rangle \rightarrow R^N$, where $\mathbf{B}$ is the resulting set; χ is a mapping of the form $\chi: U \rightarrow \mathbf{B}$; Λ are constraints on the mappings χ, φ .

We represent the Euclidean configuration by a tuple $\langle \chi, U, \mathbf{B}, \Lambda \rangle$. Let Π be a set whose elements are all possible combinatorial configurations for the given $U, \mathbf{B}$, which satisfy a system of constraints Ω . Then Π will be a Euclidean combinatorial set [23], and its image $E_\varphi = \varphi(\Pi)$ of the set Π in R^N will be a set of all Euclidean combinatorial configurations that satisfy (2). The choice of the class of sets of e-configurations (C-sets) is justified by some specific properties that these combinatorial sets possess if they are mapped into R^N .

3 Genetic Algorithms for Optimization on C-Sets

Let us consider an optimization problem on a C-set E as follows:

$$f(x) \rightarrow \min, \quad x \in E. \quad (3)$$

Methods of solving optimization problems on C-sets, as a rule, are based on applying the theory of convex extensions and extremal properties of functions defined on vertex-located sets, e.g., sets coinciding with its convex hull vertex sets [23-26]. Among vertex-located sets, there are many those inscribed into a hypersphere. Such sets are called polyhedral-spherical, properties of which allow developing specific optimization methods [27] - [32].

Let us consider features of an implementation of genetic algorithms for optimization on C-sets. Genetic algorithms operate with a variety of solutions (populations) formed by a sample of individuals. Chromosomes form individuals with parameters of the problem coded in them. Chromosomes are ordered sequences of genes. A gene is an atomic element of a genotype, in particular, of a chromosome. A set of chromosomes of each individual determines its genotype (a structure). Thus, the individuals of the population can be either genotypes or single chromosomes. The totality of external and internal signs corresponding to a given genotype determines a phenotype of the individual, i.e., decoded structure or a set of the considered problem parameters.

For the class of combinatorial optimization problems under consideration, we assume that the genotype and phenotype of the population individuals coincide, the chromosome is a feasible e-configuration $x = (x_1, \dots, x_N)$, and the genes are the values of its components in the sequence. Solutions are positioned in the population by their position on the surface of the function being examined. In this case, new solu-

tions are generated successively as different combinations of parts of the existing individuals of the populations.

Typically, the generation of the initial population involves a random selection of individuals. For our class of problems, we are talking about a random choice of feasible e-configurations. This generation problem is of independent interest and is solved depending on the class of configurations under consideration.

The next step is to select the parent pairs. As a rule, the elite selection is applied in this case, i.e., it is selected k individuals with the best found so far values of the objective function $f(x)$ and parent pairs are composed of them. If one selects all possible combinations of parental pairs, there will be $k(k-1)/2$ pairs in total. A specifics of the class of problems under consideration make it possible to offer the following approach to the choice of parental pairs. Evaluating the Euclidean distances between the best k individuals, one can cluster the searching domain. Let there are choose $k_0 \leq k$ clusters, in each of which nearby individuals are grouped. Parent pairs are selected from only one cluster. Naturally, in this case, the number of descendants will be less than with a full search of pairwise combinations. However, different rules of crossing, as well as mutations allow getting the required amount of offspring.

In connection with this, we describe the methods for the formation of the crossover operator based on the properties of various classes of sets of e-configurations. Suppose two individuals – e-configurations $x = (x_1, \dots, x_N)$ and $y = (y_1, \dots, y_N)$ – are chosen for crossing. The most common methods of crossbreeding are single-point, two-point and, in general, k -point crossovers. In this case, the parents are divided into the points $j_1, j_2, \dots, j_k$, where $j_1 < j_2 < \dots < j_k$, and their parts alternate in the offspring. Also, a uniform crossover is well-known for which the value of the component is taken from the first parent with probability p and the second parent with probability $(1-p)$.

A generalized crossover is of interest in which a special bit mask vector determines which a child gene inherits from the parent.

A. Quasy-Crossover Operator

The complex combinatorial structure of the set E leads to the fact that the resulting descendants $\tilde{x}$ and $\tilde{y}$, as a rule, do not satisfy the constraint system Λ . Therefore, the crossover operators demonstrated above will be called *quasi-crossover* ones. The result z of the quasi-crossover under Euclidean combinatorial configurations $x = (x_1, \dots, x_N)$ and $y = (y_1, \dots, y_N)$ is represented in the form $z = H(x, y)$.

To form feasible e-configurations by the quasi-crossover of parental individuals, special transformations of z can be required. In this regard, we propose the following approach to choose a Euclidean combinatorial configuration $\tilde{z} = (\tilde{z}_1, \dots, \tilde{z}_N)$ that satisfies the constraint system Λ and is closest to the individual $z = (z_1, \dots, z_N)$ obtained

as a result of quasi-crossover. Thus, we have the auxiliary problem of projecting a point $\mathbf{z}$ onto the C -set E , which solution is $\tilde{\mathbf{z}} = Pr_E \mathbf{z}$.

Consequently, the crossover operator for the pair $\mathbf{x} = (x_1, \dots, x_N)$ and $\mathbf{y} = (y_1, \dots, y_N)$ of e-configurations is representable as $\tilde{\mathbf{z}} = Pr_E H(\mathbf{x}, \mathbf{y})$.

A search of $\tilde{\mathbf{z}}$ implies solving the optimization problem

$$\|\tilde{\mathbf{z}} - \mathbf{z}\| \rightarrow \min, \mathbf{z} \in E. \quad (4)$$

A specifics of different C -sets allows including many of them in a class of well-described ones [33], i.e., sets on which linear problems are polynomially solvable. If, in addition, E is inscribed into a hypersphere, the problem (4) is polynomially solvable as well. To show this, let us introduce the following class of sets.

A set $E \subset R^n$ is said to be spherically-located if there exist such $\boldsymbol{\tau} \in R^n$ and a number $r > 0$ such that for any $\mathbf{z} \in E$

$$\|\mathbf{z} - \boldsymbol{\tau}\| = r. \quad (5)$$

Let E be a spherically located C -set, then, by (5), for any $\mathbf{z} = (z_1, \dots, z_N) \in E$ and $\mathbf{z}^0 = (z_1^0, \dots, z_N^0)$ there is

$$\|\mathbf{z} - \mathbf{z}^0\|^2 = \|\mathbf{z} - \boldsymbol{\tau}\|^2 + \|\mathbf{z}^0 - \boldsymbol{\tau}\|^2 - 2(\mathbf{z} - \boldsymbol{\tau}, \mathbf{z}^0 - \boldsymbol{\tau}) = \sum_{i=1}^N c_i z_i + b,$$

where

$$c_i = -2(z_i^0 - \tau_i), b = r^2 + \sum_{i=1}^N (z_i^0 - \tau_i)^2 + 2 \sum_{i=1}^N \tau_i (z_i^0 - \tau_i).$$

Thus, the solution of problem (5) reduces to finding the minimum of the linear function $f(\mathbf{z}) = \sum_{i=1}^N c_i z_i$ on the set E , equivalently, on the polyhedron $\text{conv } E$.

Developing this approach, we propose the following ways of obtaining descendants for individuals $\mathbf{x} = (x_1, \dots, x_N)$ and $\mathbf{y} = (y_1, \dots, y_N)$. Considering that individuals - Euclidean combinatorial configurations - are elements of Euclidean space, we use the property of linearity of this space. To search for offspring, we will choose a linear combination of individuals $\mathbf{x}$ and $\mathbf{y}$, and then perform projecting onto E .

B. Crossover Operator

Let us consider a general approach for the formation of a crossover operator that takes into account different schemes for constructing linear combinations $\mathbf{x}$ and $\mathbf{y}$:

- a simple linear combination of parental pairs: $\tilde{\mathbf{z}} = Pr_E(\mathbf{x} + \mathbf{y})$;
- a weighted linear combination of parental pairs in accordance with the values

of the function $f(x)$ at these points

$$\tilde{z} = Pr_E (x f(x) + y f(y)). \quad (6)$$

Moreover, in the maximization problem, a larger coefficient corresponds to a larger value of the function;

- a randomized weighted linear combination:

$$\tilde{z} = Pr_E (p x f(x) + (1 - p) y f(y)),$$

where p is a random variable uniformly distributed on a segment $[0, 1]$.

In general, it makes sense to assume that the descendant retains the genes of its parents as well as of other ancestors. Then the crossover operator with a weighted linear combination of k individuals will take the form:

$$\tilde{z} = Pr_E \left(\sum_{i=1}^k x_i f(x_i) \right)$$

or a randomized weighted linear combination:

$$\tilde{z} = Pr_E \left(\sum_{i=1}^k p_i x_i f(x_i) \right),$$

where $\sum_{i=1}^k p_i = 1$.

4 A Genetic Algorithm of Optimization on the Permutation Set

Let us consider the application of the described approach in solving the optimization problem on the combinatorial set of permutations (without repetitions). In this case, the Euclidean combinatorial configuration $\langle \varphi, \chi, U, \mathbf{B}, \Lambda \rangle$ is given by a bijective mapping χ and a set of constraints $\Lambda = \emptyset$. Let the set $\mathbf{B}$ be such that $m = 1$, $k = N = n$, i.e., $\mathbf{B} = \{b_1, b_2, \dots, b_n\}$ is the set real numbers ordered as follow $b_1 < b_2 < \dots < b_n$. Then the Euclidean combinatorial configuration $z = (z_1, \dots, z_n) \in R^n$ is an ordered set of numbers from $\mathbf{B}$. In particular, we can choose $\mathbf{B} = J_n$.

The set of all Euclidean combinatorial configurations satisfying the above property is called the basic permutation (without repetitions) C-set, which we denote by $E(B)$. It is known [34, 35] that the set $E(B)$ is polyhedral-spherical and coincides with the set of solutions of the system of linear and quadratic constraints:

$$\sum_{i=1}^n z_i = \sum_{i=1}^n b_i,$$

$$\sum_{i \in W} z_i \geq \frac{|W|}{\sum_{i=1}^n b_i}, \quad \forall W \subset J_n,$$

$$\sum_{i=1}^n (z_i - \tau)^2 = \sum_{i=1}^n (b_i - \tau)^2,$$

$$\tau = \frac{1}{n} \sum_{i=1}^n b_i$$

where $|W| = \text{card } W$.

Also, note that the linear function $f(\mathbf{z}) = \sum_{i=1}^n c_i z_i$ attains its minimum on the set

$E(B)$ at the point $\tilde{z} = (\tilde{z}_1, \dots, \tilde{z}_n)$, where $\tilde{z}_{\pi_i} = b_i$, $i \in J_n$, and the sequence

$\{\pi_1, \dots, \pi_n\}$, $\pi_i \in J_n$, $\pi_i \neq \pi_j \quad \forall i, j \in J_n, i \neq j$ is such that $c_{\pi_1} \geq \dots \geq c_{\pi_n}$. Thus,

$E(B)$ is well-described set.

Since the set $E(B)$ is polyhedral-spherical and well-described, then for any point

$z^0 = (z_1^0, \dots, z_n^0) \in R^n$ it is possible to find the nearest point of $E(B)$ in the closed

form. Namely, it is representable in the form $\tilde{z} = (\tilde{z}_1, \dots, \tilde{z}_n)$, where $\tilde{z}_{\pi_i} = b_{n-i+1}$,

$i \in J_n$ and the sequence $\{\pi_1, \dots, \pi_n\}$, $\pi_i \in J_n$, $\pi_i \neq \pi_j \quad \forall i, j \in J_n, i \neq j$ is such

that $z_{\pi_1}^0 \geq \dots \geq z_{\pi_n}^0$.

The results are directly generalized to the case if $k < n$, but E is still consists of permutation configurations induced by the same multiset. In this case, E is the basic permutation with repetitions C-set. This two classes of permutation C-sets are united in a class of basic generalized permutation C-sets to which the above results are applicable. Also, the Boolean C-set B_n along with its special subclasses such as the Boolean permutation C-set, the Boolean half-cube C-set, permutation matrices set are spherically-located and well-described. Respectively, the above results are extendable to the classes.

5 A Hybrid Approach to Optimization on C-Sets

Consider the optimization problem (3) on a C-set $E \subset R^n$ such as

- E is spherically-located, namely,

$$E \subset S_r(a). \quad (7)$$

Note that together with the finiteness of E , the condition (4) means that E is vertex-located, i.e.,

$$E = \text{vert } P, \quad (8)$$

where $P = \text{conv } E$ - is a polyhedron;

- E is well-described (9)
- and respectively, by (4), projecting onto the set is conducted effectively;
- for P , the vertex adjacency criterion is known. (10)

Among the combinatorial sets with properties (7)-(10), there are the mentioned above generalized set of permutations, the set of partial permutations and combinations induced by two numbers. The same holds for the direct product and direct sum, as well as for certain subsets and particular cases of the listed sets, such as Boolean and binary sets, sets of polypemutations and permutation matrices, sets of even and odd permutations, a set of vertices of a demicube, and so on.

Thus, (3), (7)-(10) covers a wide class of combinatorial problems that are, typically, NP-hard, starting with Boolean problems [2] and ending with optimization problems on composite images [17] of the above combinatorial sets. On the other hand, these problems have plenty of practical applications [1-3].

So, a search for new approaches to the solution of the problem (3), (7)-(10) is of interest both from the theoretical and practical point of view.

An interesting feature of optimization over sets of type (8) is the possibility of reducing the problem (3) with an arbitrary function to optimization of a convex function, called a convex extension of $f(x)$ from E [22, 23]. This feature is very important when new methods are developed because it allows getting estimates of the accuracy of the solutions found and, accordingly, constructing approximation algorithms based on heuristics similar to the one proposed in this paper.

We offer the following hybrid method of a random search for solving (3), (7)-(10) using some ideas of the above genetic algorithm.

Step 0. Put parameter $m \in \mathbb{Z}_+$;

Step 1. Initial iteration: $i = 0$. Generate M -element sampling (an initial population) from $E : X^i = \{x_j^i\}_{j \in J_M}$, where $M = C_m^2$.

The initial record is $f^{\min} = \min_{j \in J_M} f(x_j^i)$, $x^{\min} = \arg \min_{j \in J_M} f(x_j^i)$.

Step 2. Perform a descent from the individuals along the adjacent vertices to local minimizers of $f(x)$: $Y^i = \{y_j^i\}_{j \in J_M}$, where y_j^i - is a local minimizer of $f(x)$ obtained from x_j^i .

Step 3. Form a basis $S(Y^i)$ of the multiset Y^i , i.e., a set of its various elements. From $Y^i = \{y_j^i\}_{j \in J_M}$, choose m the best ones by values of $f(x)$ and form a set

$Z^i = \{z_j^i\}_{j \in J_m}$ from them.

Step 4. Try to improve the record:

$$\text{if } f^{\min} > \min_{j \in J_m} f(z_j^i), \text{ then } f^{\min} = \min_{j \in J_m} f(z_j^i), x^{\min} = \min_{x \in Z^i} f(x).$$

Step 5. Make the crossing within Z^i , namely: $\forall j < j'$ find a center of the segment

$$\left[z_j^i, z_{j'}^i \right] - z_{jj'}^i = \frac{z_j^i + z_{j'}^i}{2}$$

and project $z_{jj'}^i$ onto E creating in such a way a new M -element sampling (a new population) from E :

$$Z'^i = \{z_{jj'}^i\}_{1 \leq j < j' \leq m}, |Z'^i| = M, \forall j, j' \quad z_{jj'}^i = Pr_E z_{jj'}^i.$$

Step 6. Set $i = i + 1$. Check the termination condition. If it does not hold, set $X^i = Z'^{i-1}$ and go to **Step 2**.

As a termination criterion, an achievement of the maximum number of iterations, non-improvement of the current record for a prescribed number of iterations, and so on can be applied.

The advantage of the offered method is that it uses the structural specifics of each particular combinatorial set. On the one hand, it allows, when crossing distant elements, obtaining feasible elements, differ significantly from the elements, then performing a search in a vicinity of the new elements, thus decreasing the probability of omitting an exact solution. On the other hand, if the points are subjected to the crossing, which are already sufficiently close to each other, then the new points will "inherit" general properties of both "parents". For instance, if two points belong to the same hyperface of a polytope, then as a result of their crossing is also a point on the hyperface.

6 Simulation and Numerical Results

Consider the use of the proposed genetic algorithm in solving the problem of balancing solids. There are a set of points $A_j = (x_j, y_j, z_j)$, $j \in J_N = \{1, 2, \dots, N\}$ in the space

$\mathbf{R}^3$ and a set of geometric objects S_i , $i \in J_N$ with masses m_i , $i \in J_N$, whose centers of mass are at points $T_i = (x_i, y_i, z_i)$, $i \in J_N$ respectively. It is necessary to place the center of gravity of each of the objects S_i , $i \in J_N$ in one of the points A_i , $i \in J_N$ so, that the deviation of the center of gravity of the system relative to the point $A_0 = (x_0, y_0, z_0)$ was minimal. If the deviation of the center of gravity of the placed

objects from a point A_0 is considered in the Euclidean metric, then the objective function of the problem can be written as follows:

$$F(\mathbf{m}) = \sqrt{\alpha^2 + \beta^2 + \gamma^2} ,$$

where

$$\alpha = x_0 - \frac{\sum_{i=1}^N m_i x_i}{m_\Sigma}, \quad \beta = y_0 - \frac{\sum_{i=1}^N m_i y_i}{m_\Sigma}, \quad \gamma = z_0 - \frac{\sum_{i=1}^N m_i z_i}{m_\Sigma},$$

$$m_\Sigma = \sum_{i=1}^N m_i, \quad \mathbf{m} = (m_1, \dots, m_N) .$$

Since each point $T_i, i \in J_N$, must be placed in one and only one of the points $A_j, j \in J_N$, the given task belongs to the class of assignment problems. Therefore, it can be formulated as an optimization problem on the set of N -permutations induced by a masses' set $\{m_1, \dots, m_N\}$.

This class includes a problem of balancing masses of rotating parts, occurred in a turbine construction, power plant engineering, etc.

Problem statement: on a perfectly balanced disk, it is necessary to place the blades with the specified angular pitch so that the total unbalance of the system is minimal. The objective function of the problem is determined by the static moments of the blades with respect to a pair of mutually perpendicular axes. Let $m_i, i \in J_N$ be the static moments of the blades about axes of their coordinate systems. If the blade is at an angle φ_k to the axis Ox , then, about this axis, its moment is equal $m_i \cos \varphi_k$ and about the axis Ox - $m_i \sin \varphi_k$ in a coordinate system associated with the disk. Then the total imbalance of the system of the blades $S_i, i \in J_N$ is as follows:

$$f(\mathbf{m}) = \left[\left(\sum_{i=1}^N m_{\pi_i} \cos \varphi_i \right)^2 + \left(\sum_{i=1}^N m_{\pi_i} \sin \varphi_i \right)^2 \right]^{1/2}, \quad (11)$$

where $\varphi_i, i \in J_N$ are the given angles corresponding to spots of the blades on the disk. Thus, a permutation of the masses $\{m_1, \dots, m_N\}$ uniquely determines the value of the problem objective function.

We tested the proposed genetic algorithms for solving the balancing problem (11) when placing from 50 to 300 masses uniformly distributed within an interval (0, 100). Coordinates of points $A_i, i \in J_N$ were also generated randomly within an interval (-50, 50) and $A_0 = (0, 0, 0)$. To perform the calculations PC with characteristics i3/8G/SSD 256G was used. The average runtime for solving the balancing problem

for 100 masses was 9 seconds. In the series of test samples, an unbalance does not exceed 0.1. The results were compared with a random search for a series of samples solved for the same running time as the genetic algorithm required. The best unbalance results obtained using a random search belong to the interval (5.10), which is significantly worse than using the genetic algorithm.

Also, it was solved a test problem of balancing rotating masses offered in [36]. The unbalance was calculated using the formula (2), where 96 vanes were placed on a disk with an equal angular steps $\varphi_i = 2\pi i / N$, $i \in J_N$. A crossover operator of type (6) was used. There were chosen 1000 permutations in population. The best 50 of them were selected for crossing. The optimal unbalance of 0,054 was achieved in six generations and 0,83 seconds. The corresponding permutation of static moments of the vanes is as follows {42; 7; -63; 7; -9; 3; -10; 7; -14; 17; -11; 22; 17; -30; 5; -28; 77; 19; -6; -46; 0; 25; -31; -6; 11; 3; 19; 8; 22; -26; 20; -4; -38; 2; -26; -14; 49; -27; 12; -4; 16; -7; -18; -55; 5; 9; -24; -33; -18; 2; -9; 11; 37; -25; -14; -2; -7; -16; 3; -53; 48; 14; 30; 29; 48; 0; 17; -36; -69; -2; 13; -5; -26; -4; 13; 5; 12; 42; -9; -3; -10; 0; 6; 7; -9; -40; 11; -30}. The total number of the objective function evaluation is 9552. The best value of the function attained is 5.491.

7 Conclusions

The report offers a new approach to implementing genetic algorithms in Combinatorial Optimization. A notion of a Euclidean combinatorial configuration is introduced as a mapping of a finite abstract set of an arbitrary nature into Euclidean space. As a result of this mapping, an optimization problem over a combinatorial configuration set is equivalently formulated as a discrete optimization problem on a finite point configuration which elements are Euclidean combinatorial configurations. Consideration is given to the specifics of implementing genetic algorithms to solving this class of problems: methods for the formation of the initial population and selection mechanisms are proposed, and the choice of crossover and mutation operators is formalized and justified. As an example, it is considered optimization on Euclidean permutational configurations. Based on the genetic algorithm, a random search method is offered for optimization over spherically-located and well-described sets that cover a wide class of problems in theoretical and practical domains.

References

1. Korte, B., Vygen, J.: Combinatorial Optimization: Theory and Algorithms, 6th ed. 2018 edition. New York, NY: Springer (2018).
2. Pardalos, P.M. . Du, D.-Z., Graham, R.L.(Eds.): Handbook of combinatorial optimization. New York: Springer, (2013).
3. Sergienko, I.V., Shilo, V.P.: Problems of discrete optimization: problems, methods of solution, research. Kiev: Naukova Dumka (2003).

4. Spears, W.M.: Evolutionary Algorithms: The Role of Mutation and Recombination, Berlin, New York: Springer (2000).
5. Simon, D.: Evolutionary Optimization Algorithms, Hoboken, New Jersey: Wiley (2013).
6. Petrowski, A., Ben-Hamida, S.: Evolutionary Algorithms, Hoboken, NJ: Wiley-ISTE, (2017).
7. Mitchell, M.: An Introduction to Genetic Algorithms, Cambridge, Mass.: MIT Press (1998).
8. Back, T.: Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms, 1 edition. New York: Oxford University Press (1996).
9. Kramer, O.: Genetic Algorithm Essentials, 1st ed. 2017 edition. New York, NY: Springer (2017).
10. Pintea, C.-M.: Advances in Bio-inspired Computing for Combinatorial Optimization Problems. Heidelberg: Springer (2014).
11. Gulianitsky, L.F., Sergienko, I.V.: Meta-evolutionary method of deformed polyhedron in combinatorial optimization. Cybernetics and system analysis, vol. 44, no. 6, pp. 70-79 (2007).
12. Hussain, M.N., et al.: Metaheuristic research: a comprehensive survey. Artificial Intelligence Review, pp. 1-43 (2018).
13. Kritbodin Phiwhorm, Kanda Runapongsa Saikaew: A hybrid genetic algorithm with multi-parent crossover in fuzzy rule-based. International Journal of Machine Learning and Computing, vol. 7, no. 5, pp. 114-117 (2017).
14. Ding, Y., Fu, X.: Kernel-based fuzzy c-means clustering algorithm based on genetic algorithm,” Neurocomputing, vol. 188, no. Supplement C, pp. 233-238 (2016).
15. Berge, C.: Principes de combinatoire. Paris: Dunod (1968).
16. Sachkov, V.N.: Combinatorial methods of discrete mathematics. M: Nauka (1975).
17. Grebennik, I.V., Lytvynenko O.S.: Generating combinatorial sets with given properties Cybernetics and system analysis, vol. 48, no. 6, pp. 890–898 (2012) doi:1060-0396/12/4806-0890.
18. Hulianytskyi, L., Riasna, I.: Formalization and classification of combinatorial optimization problems. Springer Optimization and its Applications, vol. 130, pp. 239-250 (2017) doi: 10.1007/978-3-319-68640-0_11.
19. Yakovlev, S.V.: The method of artificial space dilation in problems of optimal packing of geometric objects. Cybernetics and Systems Analysis, vol. 53 no. 5, pp. 725-732 (2017) doi:10.1007/s10559-017-9974-y.
20. Stoyan, Y.G., Yakovlev, S.V.: Configuration space of geometric objects. Cybernetics and Systems Analysis, vol. 54, no. 5, 716-726 (2018) doi: 10.1007/s10559-018-0073-5.
21. Yakovlev, S.V.: On some classes of spatial configurations of geometric objects and their formalization. J. of Autom. and Inform. Sciences, vol. 50, no. 9, 38-50 (2018) doi:10.1615/JAutomatInfScien.v50.i9.30.
22. Yakovlev, S.V.: Formalization of spatial configuration optimization problems with a special function class. Cybernetics and Systems Analysis 55(4), 512-523 (2019).
23. Yakovlev, S.: Convex extensions in combinatorial optimization and their applications. Springer Optimization and its Applications, vol. 130, pp. 567-584 (2017) doi:10.1007/978-3-319-68640-0_27.

24. Yakovlev, S.V.: The theory of convex continuations of functions on vertices of convex polygons. *Comp. Math. and Math. Physics* 34, 959-965 (1994)
<https://dl.acm.org/citation.cfm?id=196926>.
25. Yakovlev, S.V.: Bounds on the minimum of convex functions on Euclidean combinatorial sets. *Cybernetics*, vol. 25, no. 3, pp. 385-391 (1989) doi:10.1007/BF01069996.
26. Yakovlev, S.V., Valuiskaya, O.A.: Optimization of linear functions at the vertices of a permutation polyhedron with additional linear constraints. *Ukrainian Mathematical Journal*, vol. 53, no.9, pp. 1535-1545 (2001) doi:10.1023/A:1014374926840.
27. Stoyan, Y.G., et al.: Quadratic optimization on combinatorial sets in R^n . *Cybernetics and Systems Analysis*, vol. 27, no. 4, pp. 561–567 (1991) doi:10.1007/BF01130367.
28. Stoyan, Y.G., et al.: Construction of convex continuations for functions defined on a hypersphere. *Cybernetics and Systems Analysis*, vol. 34, no. 2, pp. 27–36 (1998). doi:10.1007/BF02742066.
29. Yakovlev, S.V., Grebennik, I.V.: Localization of solutions of some problems of nonlinear integer optimization. *Cybernetics and Systems Analysis*, vol. 29, no. 5, pp. 727-734 (1993) doi:10.1007/BF01125802.
30. Yakovlev, S.V., Pichugina, O.S.: Properties of combinatorial optimization problems over polyhedral-spherical sets. *Cybernetics and Systems Analysis*, vol. 54. no. 1, pp. 99-109 (2018) doi:10.1007/s10559-018-0011-6.
31. Pichugina, O., Yakovlev, S.: Optimization on polyhedral-spherical sets: theory and applications. *Proceedings of the IEEE First Ukraine Conference on Electrical and Computer Engineering*, pp. 1167-1175 (2017) doi:10.1109/UKRCON.2017.8100436
32. Yakovlev, S.V., Pichugina, O.S., Yarovaya, O.V.: Polyhedral spherical configuration in discrete optimization. *Journal of Automation and Information Sciences*, vol. 51, no. 1, pp. 38-50 (2019).
33. Berstein, Y., et al.: Parametric nonlinear discrete optimization over well-described sets and matroid intersections. *Math. Program.*, vol. 124, no. 1–2, pp. 233–253 (2010).
34. Yemelichev, V.A., et al.: *Polytopes, graphs and optimisation*. Cambridge University Press, Cambridge (1984).
35. Pichugina, O.S., Yakovlev, S.V.: Functional and analytic representations of the general permutations. *Eastern-European Journal of Enterprise Technologies*, vol. 1, no. 4, pp. 27-38 (2016) doi:10.15587/1729-4061.2016.58550.
36. Stoyan, Y.G., et al.: Method of balancing rotating discretely distributed masses. *Energomashinostroenie*, no. 2, pp. 4–5 (1982).

Method for Parametric Identification of Gaussian Mixture Model Based on Clonal Selection Algorithm

Eugene Fedorov¹[0000-0003-3841-7373], Valentyna Lukashenko¹[0000-0002-6749-9040],
Tetyana Utkina¹[0000-0002-6614-4133], Andriy Lukashenko²[0000-0002-6016-1899],
Kostiantyn Rudakov¹[0000-0003-0000-6077]

¹ Cherkasy State Technological University, Cherkasy, Shevchenko blvd., 460, 18006, Ukraine
{ckc, k.rudakov, t.utkina}@chdtu.edu.ua, fedorovee75@ukr.net

² E. O. Paton Electric Welding Institute, Kyiv, Bozhenko str., 11, 03680, Ukraine
ineks-kiev@ukr.net

Abstract. The problem of increasing the efficiency of parametric identification of Gaussian mixture model (GMM) is considered. A method for identifying GMM parameters based on clonal selection algorithm with preliminary formation of a learning set taking into account the structure of vocal sounds, which increases the likelihood of speaker recognition, is proposed. The parameter identification method is intended for software implementation on the GPU using the CUDA technology, which speeds up the process of parametric identification. This method has been studied on the TIMIT database and serves for intelligent systems of biometric personal identification.

Keywords: quasi-periodic signal, speaker recognition, gaussian mixture model, learning set formation, parametric identification, clonal selection algorithm.

1 Introduction

Automated biometric identification of a person means making decisions based on acoustic and visual information, which improves the quality of recognition of the person under study. Unlike the traditional approach [1], computer biometric identification speeds up and improves the accuracy of the recognition process, which is especially critical in the conditions of limited time.

A special class of biometric identification of a person is formed by methods based on the analysis of acoustic information [2].

The well-known voice identification methods, such as dynamic programming [3-4], analyze the entire signal, which increases the likelihood of recognition, but require entire storage of learning signals and long comparison of the analyzed signal with all learning signals. Other well-known methods, such as vector quantization [5-6], neural network method [7-8] and decision tree [9], store only generalized characteristics of learning signals and carry out a quick classification of signal components, but analyze these components without interconnecting them, which reduces the recognition probability. Network methods, including those that use Gaussian mixture models (GMM)

[10-13], which quickly analyze the entire signal and store only generalized characteristics of learning signals, are a compromise between the above groups of methods. At the same time, GMM-based methods use the identification of its parameters based on local search, which reduces the probability of recognition.

The aim of the work is to increase the efficiency of parametric identification method of Gaussian mixture model (GMM) due to a metaheuristic clonal selection algorithm, with preliminary formation of a learning set.

To achieve this goal, it is necessary to solve the following tasks:

1. to develop a method for learning set formation;
2. to create a method of GMM parametric identification;
3. to develop an algorithm of GMM parametric identification;
4. to conduct a numerical study of the proposed method of parametric identification.

2 Formal problem statement

Let a learning set be defined for a particular speaker.

Then the problem of increasing the probability of speaker recognition by Gaussian mixture model (GMM) is presented as the problem of finding such a vector of parameters for this model that satisfies the criterion.

3 Literature review

Biometric identification methods based on dynamic programming [3-4] compare the signal representing the recognizable speaker with all signals representing the well-known speakers. The advantage of these methods lies in their ability to analyze the entire signal and compare signals of different lengths, which increases the probability of recognition. The disadvantage of these methods consists in the entire storage of signals representing the well-known speakers and the duration of the procedure for comparison of the recognizable signal with all available signals.

Biometric identification methods based on vector quantization [5-6] form the codebook vectors as averaged components of the signals representing the well-known speakers and compare components of the signal representing the recognizable speaker with these vectors. The advantage of these methods lies in the storage of only the codebook vectors and quick comparison of the recognizable signal components with these vectors. The disadvantage of these methods consists in the ability to analyze only single components of the signal without interconnecting them, which reduces the likelihood of recognition.

Biometric identification methods based on artificial neural networks [7-8] identify the neural network parameters, which store information about signal components representing the well-known speakers and recognize signal components representing the recognizable speaker through the neural network. The advantage of these methods lies in the storage of only the neural network parameters and quick classification of the recognizable signal components. The disadvantage of these methods consists in

the identification of neural network parameters based on local search and the ability to analyze only single components of the signal without interconnecting them, which reduces the likelihood of recognition.

Biometric identification methods based on decision trees [9] automatically identify threshold values (for example, formants) at decision tree nodes representing speech characteristics of the well-known speakers and compare threshold values of the signal component representing the recognizable speaker with threshold values of the decision tree. The advantage of these methods lies in the storage of only threshold values and quick classification of the recognizable signal components. The disadvantage of these methods consists in the complexity and subjectivity of threshold values formation and the ability to analyze only single components of the signal without interconnecting them, which reduces the likelihood of recognition.

Biometric identification methods based on Gaussian mixture models (GMMs) [10-13] identify GMM parameters, which store information about signal components representing the well-known speakers, and recognize the signal representing the recognizable speaker by means of these GMMs. The advantage of these methods lies in the ability to quickly recognize the entire signal and store only GMM parameters. The disadvantage of these methods consists in the identification of GMM parameters based on local search, which reduces the likelihood of recognition.

A common feature of all methods listed above consists in the following: they form a learning set of signal patterns and analyze a signal without taking into account the structure of vocal sounds, which leads to a decrease in the probability of recognition.

Therefore, the increase of the efficiency of the method of GMM parametric identification with preliminary formation of a learning set, taking into account the structure of vocal sounds, is an urgent task.

4 Method of learning set formation

The method of learning set formation includes the following steps:

- partitioning of a quasi-periodic signal into discrete patterns;
- shifting of discrete patterns in time and amplitude;
- interpolation of discrete patterns;
- shifting and scaling of continuous patterns in time;
- shifting and scaling of continuous patterns in amplitude;
- sampling of continuous patterns.

1. Partitioning of a quasi-periodic signal into discrete patterns

Let's define a finite set of discrete patterns of a quasi-periodic signal, described by a family of integer bounded finite discrete functions $X = \{x_i \mid i \in \{1, \dots, I\}\}$, in the form:

$$x_i(n) = \begin{cases} f(n), & n \in \{N_i^{\min}, \dots, N_i^{\max}\} \\ 0, & n \notin \{N_i^{\min}, \dots, N_i^{\max}\} \end{cases}, \quad i \in \{1, \dots, I\},$$

$$A_i^{\min} = \min_n f(n), \quad n \in \{N_i^{\min}, \dots, N_i^{\max}\}, \quad i \in \{1, \dots, I\},$$

$$A_i^{\max} = \max_n f(n), \quad n \in \{N_i^{\min}, \dots, N_i^{\max}\}, \quad i \in \{1, \dots, I\},$$

where $A_i^{\min}, A_i^{\max}$ – the minimum and maximum values of the function x_i on the compact $\{N_i^{\min}, \dots, N_i^{\max}\}$.

2. Shifting of discrete patterns in time and amplitude

Let's define a finite set of discrete patterns shifted in time and amplitude, which are described by a finite family of integer bounded finite discrete functions

$X^S = \{x_i^S \mid i \in \{1, \dots, I\}\}$, in the form:

$$x_i^S(n) = \begin{cases} x_i(n + N_i^{\min}) - A_i^{\min}, & n \in \{0, \dots, N_i\}, \\ 0, & n \notin \{0, \dots, N_i\} \end{cases}, \quad i \in \{1, \dots, I\},$$

$$N_i = N_i^{\max} - N_i^{\min}, \quad A_i = A_i^{\max} - A_i^{\min}.$$

3. Interpolation of discrete patterns

A linear interpolation, which requires the least computational complexity, has been chosen in the article. Let's define a finite set of continuous patterns, obtained as a result of linear interpolation and described by a finite family of real-valued bounded finite continuous functions $\Psi = \{\psi_i \mid i \in \{1, \dots, I\}\}$, in the form:

$$\begin{aligned} \forall t \in [-\Delta t, T_i + \Delta t] \quad \psi_i(t) = \\ = \sum_{n=-1}^{N_i} \chi_{(t_n, t_{n+1})}(t) \left(x_i^S(n) + \frac{x_i^S(n+1) - x_i^S(n)}{\Delta t} (t - t_n) \right) + \sum_{n=-1}^{N_i+1} \chi_{\{t_n\}}(t) x_i^S(n), \\ i \in \{1, \dots, I\}, \quad \forall t \notin [-\Delta t, T_i + \Delta t] \quad \psi_i(t) = 0, \quad T_i = N_i \Delta t, \quad t_n = n \Delta t, \end{aligned}$$

$$\chi_B(t) = \begin{cases} 1, & t \in B \\ 0, & t \notin B \end{cases} - \text{indicator function},$$

where Δt – the quantization step in time.

4. Shifting and scaling of continuous patterns in time

Let's define a finite set of shifted and scaled in time continuous patterns, described by a finite family of real-valued bounded finite continuous functions

$\Psi^S = \{\psi_i^S \mid i \in \{1, \dots, I\}\}$, in the form:

$$\forall t \in [\tilde{T}^{\min}, \tilde{T}^{\max}] \quad \psi_i^s(t) = \psi_i \left(T_i \frac{t - \tilde{T}^{\min}}{\tilde{T}^{\max} - \tilde{T}^{\min}} \right),$$

$$\forall t \notin [\tilde{T}^{\min}, \tilde{T}^{\max}] \quad \psi_i^s(t) = 0,$$

where $[\tilde{T}^{\min}, \tilde{T}^{\max}]$ – the compact, set and single for all patterns.

The article proposes to define $\tilde{T}^{\min}, \tilde{T}^{\max}$ as follows:

$$\tilde{T}^{\min} = \Delta t,$$

$$\tilde{T}^{\max} = \text{round} \left(\frac{f_d}{f_{\min}} \right) \Delta t,$$

where f_d – the sampling frequency in Hz (for vocal sounds 8 kHz is enough), $f_{\min}$ – the minimum frequency of frequency range of speech sound (for vocal sounds 50 Hz is enough), $\text{round}()$ – the function, rounding the number to the nearest integer.

5. Shifting and scaling of continuous patterns in amplitude

Let's define a finite set of continuous patterns shifted and scaled in amplitude, which are described by a finite family of real-valued bounded finite continuous functions $\Psi^{ss} = \{\psi_i^{ss} \mid i \in \{1, \dots, I\}\}$, in the form:

$$\forall t \in [\tilde{T}^{\min}, \tilde{T}^{\max}] \quad \psi_i^{ss}(t) = \tilde{A}^{\min} + \frac{\tilde{A}^{\max} - \tilde{A}^{\min}}{\tilde{A}_i^{\max} - \tilde{A}_i^{\min}} \psi_i^s(t),$$

$$\forall t \notin [\tilde{T}^{\min}, \tilde{T}^{\max}] \quad \psi_i^{ss}(t) = 0,$$

$$\tilde{A}_i^{\max} = \max_t \psi_i^s(t),$$

$$\tilde{A}_i^{\min} = \min_t \psi_i^s(t),$$

$$t \in [\tilde{T}^{\min}, \tilde{T}^{\max}],$$

where $\tilde{A}_i^{\min}, \tilde{A}_i^{\max}$ – the minimum and maximum values of the function ψ_i^s on the compact $[\tilde{T}^{\min}, \tilde{T}^{\max}]$; $\tilde{A}^{\min}, \tilde{A}^{\max}$ – the minimum and maximum values, set and single for all patterns, on a given compact $[\tilde{T}^{\min}, \tilde{T}^{\max}]$.

The article proposes to define $\tilde{A}^{\min}, \tilde{A}^{\max}$ as follows: $\tilde{A}^{\min} = 0$, $\tilde{A}^{\max} = 2^b - 1$, where b – the number of level quantization bits.

6. Sampling of continuous patterns

Let's define a finite set of discrete patterns, obtained from continuous ones by sampling in time and described by a finite family of integral bounded finite discrete functions $S = \{s_i \mid i \in \{1, \dots, I\}\}$, in the form:

$$s_i(n) = \text{round}(\psi_i^{ss}(n\Delta t)), \quad n \in \{\tilde{N}^{\min}, \dots, \tilde{N}^{\max}\},$$

$$\tilde{N}^{\min} = \tilde{T}^{\min} / \Delta t, \quad \tilde{N}^{\max} = \tilde{T}^{\max} / \Delta t, \quad N = \tilde{N}^{\max} - \tilde{N}^{\min}.$$

Each received discrete sample is considered as a feature vector and is located in a single amplitude-time window.

5 Method for parametric identification of Gaussian mixture model

Each GMM associated with a particular speaker is defined on the basis of the total probability formula as follows:

$$P(s) = \sum_{k=1}^K P(k) p(s \mid k),$$

$$p(s \mid k) = \frac{1}{\sqrt{(2\pi)^N \det C_k}} \exp\left(-\frac{1}{2}(s - m_k)^T C_k^{-1} (s - m_k)\right),$$

where $P(k)$ – a priori unconditional probabilities (weights of mixture components), $p(s \mid k)$ – densities of multi-dimensional Gaussian distribution, $m_k = (m_{k1}, \dots, m_{kN})$ – vector of mathematical expectations, C_k – covariance matrix, K – the number of mixture components.

In the work, for GMM parametric identification the target function, which means the choice of such values of parameter vector Θ that deliver likelihood functions to the maximum logarithm, is chosen:

$$F = \ln P(S \mid \Theta) = \sum_{i=1}^I \ln \sum_{k=1}^K P(k) p(s_i \mid k) \rightarrow \max_{\Theta},$$

where $\Theta = ((P(1), \dots, P(K)), (m_1, \dots, m_K), (C_1, \dots, C_K))$ – GMM parameters vector.

Since the traditional EM algorithm used for GMM parametric identification implements only local search, so currently for GMM parametric identification are actively applied evolutionary, swarm and immune metaheuristics, which use the parameter vector Θ as an individual of the population [14-16]. However, due to the large dimensionality of the covariance matrix C , only a diagonal matrix C is used in these metaheuristics, which reduces the likelihood of identification. Therefore, in this work, the joint probability vector $(P(s_1, 1), \dots, P(s_i, k), \dots, P(s_I, K))$, where

$P(s_i, k) = P(k)p(s_i | k)$, is used as an individual of the population, which allows to work with the off-diagonal matrix C .

GMM parametric identification is based on clonal selection algorithm proposed in [17-23] and is presented in the following form:

1. The number of mixture components K , the mutation parameter β , the maximum number of iterations $\tau^{\max}$ are set.
2. The intermediate population $U = \{u\}$ of the power $\mu + \lambda + \gamma$ is created, each antibody of the population being represented as $u = (u_{11}, \dots, u_{IK})$.

Each element u_{ik} is defined as $u_{ik} = \text{rand}()$, where $\text{rand}()$ – a function that returns a uniformly distributed random number in the range $[0, 1]$.

A finite set of discrete patterns S is given.

3. The iteration number $\tau = 1$ is set, the maximum value of the target function $b^{\text{old}} = 0$.
4. A posteriori conditional probabilities for each antibody u are calculated:

$$P(k | s_i) = \frac{u_{ik}}{\sum_{z=1}^K u_{iz}}, \quad k \in \{1, \dots, K\}, \quad i \in \{1, \dots, I\}.$$

5. A priori unconditional probabilities (weights of the mixture components) for each antibody u are calculated $P(k) = \frac{1}{I} \sum_{i=1}^I P(k | s_i)$, $k \in \{1, \dots, K\}$.
6. Expectation vectors for each antibody u are calculated:

$$m_k = \frac{\sum_{i=1}^I P(k | s_i) s_i}{\sum_{i=1}^I P(k | s_i)}, \quad k \in \{1, \dots, K\}.$$

7. Covariance matrices for each antibody u are calculated:

$$C_k = \frac{\sum_{i=1}^I P(k | s_i) (s_i - m_k)^T (s_i - m_k)}{\sum_{i=1}^I P(k | s_i)}, \quad k \in \{1, \dots, K\}.$$

8. The densities of multidimensional Gaussian distribution for each antibody u are calculated:

$$p(s_i | k) = \frac{1}{\sqrt{(2\pi)^N \det C_k}} \exp\left(-\frac{1}{2} (s_i - m_k)^T C_k^{-1} (s_i - m_k)\right), \quad k \in \{1, \dots, K\}, \quad i \in \{1, \dots, I\}.$$

9. The total probabilities for each antibody u are calculated $P(s_i) = \sum_{k=1}^K P(k)p(s_i | k)$.

10. The value of the target function for each antibody u is calculated:

$$F(u) = \sum_{i=1}^I \ln P(s_i).$$

11. The intermediate population $U = \{u\}$ is ordered in descending of the target function value.

12. The current population $H = \{h\}$ of the power μ is created by means of reproduction and replacement operators, with each antibody of this population being represented as $h = (h_1, \dots, h_{IK})$.

As antibodies of the current population $H = \{h\}$, the first μ (the best in the target function) antibodies of the intermediate population $U = \{u\}$ are taken. This corresponds to the reduction operator with a selection scheme that provides the search directionality (the best antibodies are preserved). Replacement (randomly generated) antibodies can be among the best antibodies. This corresponds to the replacement operator.

13. The value of the target function of the first antibody of the current population $H = \{h\}$ is set by the minimum value of the target function a .

14. The value of the target function of the last antibody of the current population $H = \{h\}$ is set by the maximum value of the target function b .

15. If $|b - b^{old}| > \varepsilon$ and $\tau < \tau^{\max}$, then $\tau = \tau + 1$, $b^{old} = b$, going to 16, otherwise parametric identification is complete.

16. The affinity for each antibody h is calculated.

The affinity determines the proximity of the current antibody to the best one and is calculated based on the utility function in the form:

$$\Phi(h) = \frac{F(h) - a}{b - a}.$$

If $\Phi(h) = 1$, then the i -th antibody is the best one.

If $\Phi(h) = 0$, then the i -th antibody is the worst one.

17. The mutation probability for each antibody h is calculated $p(h) = \exp(-\beta\Phi(h))$.

The larger β , the less likely is the mutation probability.

18. A set of clones $C = \{c\}$ of the power λ is created by means of a cloning operator, each clone of this set being represented as $c = (c_{11}, \dots, c_{IK})$.

The cloning operator plays a role similar to the reproduction operator of genetic algorithm and is applied to the current population $H = \{h\}$. The number of clones for each antibody h is defined as λ/μ .

19. A set of mutated clones $\hat{C} = \{\hat{c}\}$ of the power λ is formed by means of the proposed mutation operator, each mutated clone of this set being represented in the form $\hat{c} = (\hat{c}_{11}, \dots, \hat{c}_{IK})$.

The mutation operator allows to obtain new antibodies with sharply different properties from antibody clones.

Each element $\hat{c}_{ik}$ is defined as:

$$r_{ik} = rand(), \hat{c}_{ik} = \begin{cases} c_{ik}, & r_{ik} \geq p(h) \\ round(), & r_{ik} < p(h) \end{cases}, k \in \{1, \dots, K\}, i \in \{1, \dots, I\}.$$

The features of the proposed variant of the mutation operator are the following: it does not require the use of binary potential solutions, i.e. it doesn't need to convert probabilistic potential solutions into binary ones before the mutation and to convert binary potential solutions into probabilistic ones after the mutation, which reduces the computational complexity of the mutation operator and allows to quickly find a solution.

20. A set of replacement antibodies $D = \{d\}$ of the power γ is created, each antibody of the set is represented as $d = (d_{11}, \dots, d_{IK})$.

Each element d_{ik} is defined as $d_{ik} = rand()$.

21. The intermediate population $U = \{u\}$ of the power $\mu + \lambda + \gamma$ is formed.

The antibodies of the current population $H = \{h\}$ and a set of mutated clones $\hat{C} = \{\hat{c}\}$ are taken as the first $\mu + \lambda$ antibodies of the intermediate population $U = \{u\}$. The antibodies of a set of the replacement antibodies $D = \{d\}$ are taken as the last γ antibodies of the intermediate population $U = \{u\}$. Going to 4.

As a result of the method, GMM parameters will be defined.

6 Creation of an algorithm for parametric identification of Gaussian mixture model based on clonal selection algorithm

The algorithm for GMM parametric identification based on clonal selection algorithm, which is intended for implementation on the GPU using the CUDA technology, is presented in Fig. 1. This block diagram functions as follows.

Step 1 – Set the number of mixture components K , the mutation parameter β , the maximum number of iterations $\tau^{\max}$.

Step 2 – Set the intermediate population U and a finite set of discrete patterns S .

Step 3 – Set the iteration number $\tau = 1$, the maximum value of the target function $b^{old} = 0$.

Step 4 – Calculate the total probabilities $P(s_i)$ for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot (\mu + \lambda + \gamma)$ one-dimensional

blocks. In each block, based on the reduction, a sum from K elements of the form u_{ik} is calculated.

Step 5 – Calculate a posteriori conditional probabilities for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks, where N^S is the number of threads in the block. Each thread computes $P(k | s_i) = u_{ik} / P(s_i)$.

Step 6 – Calculate the sum of a posteriori conditional probabilities q_k for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma)$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $P(k | s_i)$ is calculated.

Step 7 – Calculate a priori unconditional probabilities (weights of the mixture components) $P(k)$ for each antibody u using $K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks. Each thread computes $P(k) = q_k / I$.

Step 8 – Calculate expectation vectors m_k for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma)$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $P(k | s_i) s_i / q_k$ is calculated.

Step 9 – Calculate the covariance matrices C_k for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma)$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $P(k | s_i) (s_i - m_k)^T (s_i - m_k) / q_k$ is calculated.

Step 10 – Calculate the densities of Gaussian multidimensional distribution $p(s_i | k)$ for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks. Each thread computes:

$$p(s_i | k) = \frac{1}{\sqrt{(2\pi)^N \det C_k}} \exp \left(-\frac{1}{2} (s_i - m_k)^T C_k^{-1} (s_i - m_k) \right).$$

Step 11 – Calculate the total probabilities $P(s_i)$ for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot (\mu + \lambda + \gamma)$ one-dimensional blocks. In each block, based on the reduction, a sum from K elements of the form $P(k) p(s_i | k)$ is calculated.

Step 12 – Calculate the target function values $F(u)$ for each antibody u using $I \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $\mu + \lambda + \gamma$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $\ln P(s_i)$ is calculated.

Step 13 – Order the target functions by the value based on parallel sorting by merging into the intermediate population U using $\mu + \lambda + \gamma$ threads that are grouped into one one-dimensional block.

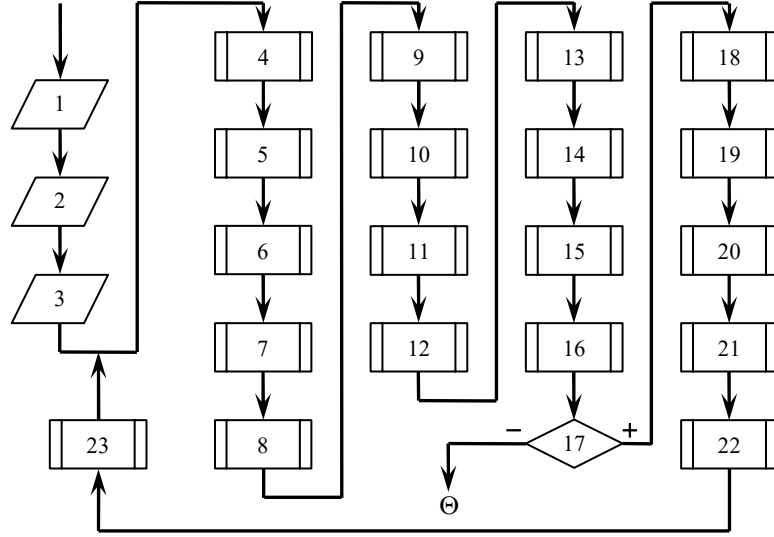


Fig. 1. Block diagram of the algorithm for parametric identification of Gaussian mixture model based on clonal selection algorithm.

Step 14 – Create the current population $H = \{h\}$ of the power μ by means of element-by-element μ copying of the first antibodies of the intermediate population $U = \{u\}$ using $I \cdot K \cdot \mu$ threads that are grouped into $I \cdot K \cdot \mu / N^s$ one-dimensional blocks.

Step 15 – Set the value of the target function of the first antibody of the current population $H = \{h\}$ by the minimum value of the target function a .

Step 16 – Set the value of the target function of the last antibody of the current population $H = \{h\}$ by the maximum value of the target function b .

Step 17 – If $|b - b^{old}| > \varepsilon$ and $\tau < \tau^{\max}$, then $\tau = \tau + 1$, $b^{old} = b$, going to 18, otherwise parametric identification is completed.

Step 18 – Calculate the affinity for each antibody h using $\mu + \lambda + \gamma$ threads that are grouped into one one-dimensional block. Each thread computes:

$$\Phi(h) = \frac{F(h) - a}{b - a}.$$

Step 19 – Calculate the mutation probability for each antibody h using λ threads that are grouped into one one-dimensional block $p(h) = \exp(-\beta\Phi(h))$.

Step 20 – Create a set of clones $C = \{c\}$ of the power λ by means of element-by-element copying of the antibodies of the current population $H = \{h\}$ using $I \cdot K \cdot \lambda$ threads that are grouped into $I \cdot K \cdot \lambda / N^s$ one-dimensional blocks.

Step 21 – Form a set of mutated clones $\hat{C} = \{\hat{c}\}$ of the power λ by means of the proposed mutation operator, using $I \cdot K \cdot \lambda$ threads that are grouped into $I \cdot K \cdot \lambda / N^s$ one-dimensional blocks. Each thread computes:

$$r_{ik} = rand(), \quad \hat{c}_{ik} = \begin{cases} c_{ik}, & r_{ik} \geq p(h) \\ round(), & r_{ik} < p(h) \end{cases}.$$

Step 22 – Create a set of random antibodies $D = \{d\}$ of the power γ using $I \cdot K \cdot \gamma$ threads that are grouped into $I \cdot K \cdot \gamma / N^s$ one-dimensional blocks. Each thread computes $d_{ik} = rand()$.

Step 23 – Generate the intermediate population $U = \{u\}$ of the power $\mu + \lambda + \gamma$ by means of element-by-element copying of the antibodies of the current population $H = \{h\}$, a set of mutated clones $\hat{C} = \{\hat{c}\}$, a set of replacement antibodies $D = \{d\}$, using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot K \cdot (\mu + \lambda + \gamma) / N^s$ one-dimensional blocks. Going to 4.

7 Experiments and results

Numerical experiments were carried out using the CUDA technology of parallel processing of information on the GeForce 920M video card with the number of threads in the block $N^s = 1024$. Parametric identification was carried out for each of the 100 GMMs, corresponding to 100 speakers. In the work it has been accepted that the power of a set of patterns of vocal speech sounds $I = 1024$, the power of the current population $\mu = 20$, the power of a set of clones $\lambda = 1000$, the power of a set of random antibodies $\gamma = 0.2$, $\mu = 4$, the mutation parameter $\beta = 2.3$, the maximum number of iterations $\tau^{\max} = 1000$, the parameter $\varepsilon = 10^{-6}$.

The dependence of the probability of incorrect recognition of the speaker on the number of components of the mixture is shown in Fig.2. This dependence shows that the probability of incorrect recognition of the speaker decreases with increasing number of components of the mixture.

Table 1 presents the speaker's recognition probabilities obtained from the TIMIT database using a neural network based on radial basis functions (RBFNN), trained on the basis of error correction with the MFCC feature system, Gaussian mixture model (GMM), trained on the basis of EM-algorithm with the MFCC feature system, Gaussian mixture model, trained on the basis of clonal selection algorithm, with features obtained by means of the proposed method of learning set formation.

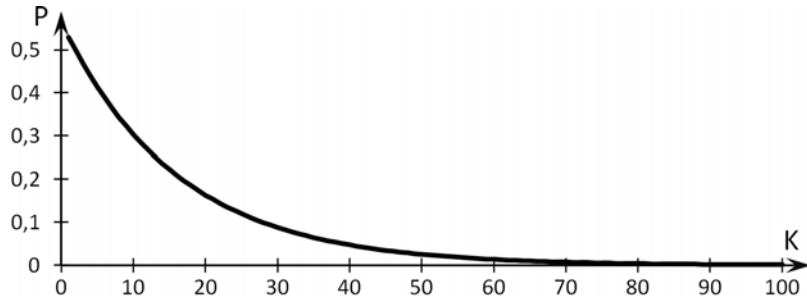


Fig. 2. The dependence of the probability of incorrect recognition of the speaker on the number of components of the mixture.

According to Table 1, the best results are given by GMM, trained on the basis of clonal selection algorithm with features obtained by means of the proposed method of learning set formation.

Table 1. Speaker recognition probability.

Model + learning algorithm + feature system	Probability of recognition
RBFNN + error correction + MFCC	0.81
GMM + EM-algorithm + MFCC	0.92
GMM + proposed method + proposed feature system	0.98

This is due to the fact that error correction and the EM algorithm only perform a local search, which increases the likelihood of falling into a local extremum, and learning set formation based on the MFCC is done without taking into account the structure of vocal sound.

8 Conclusion

The article deals with the problem of increasing the efficiency of parametric identification of Gaussian mixture model (GMM). The method of learning set formation, which uses shifting, scaling, interpolation and sampling of signal patterns that correspond to quasi-periods to locate them in a single amplitude-time window, which allows to take into account the quasi-periodic signal structure and increase the probability of speaker recognition, has been improved. The method of GMM parametric identification, which is based on clonal selection algorithm that allows to reduce the likelihood of falling into a local extremum, has reached further development. The proposed method allows probabilistic individuals of the population (potential solutions) in the mutation operator, which accelerates parametric identification (it is not necessary to perform the operations of converting real individuals into binary ones and vice versa), and uses not the parameter vector, but the vector of joint probabilities as an individual of the population, which allows to work with a non-diagonal covariance

matrix and increases the likelihood of speaker recognition. The algorithm of GMM parametric identification, which is intended for software implementation on the GPU using the CUDA technology that speeds up the GMM learning process, has been created. The software that implements the proposed algorithm has been developed and investigated on the TIMIT database. The experiments have confirmed the operability of the developed software and allow to recommend it for use in practice. Prospects for further research are to test the proposed methods on a wider set of test databases.

References

1. Singh, N., Khan, R. A., Shree, R.: Applications of Speaker Recognition. In: Rajesh, R., Ganesh, K. (eds.). International Conference on Modelling Optimization and Computing, vol. 38, pp. 3122–3126. Elsevier Procedia Engineering (2012) doi: 10.1016/j.proeng.2012.06.363
2. Campbell, J. P.: Speaker Recognition: a tutorial. Proceedings of the IEEE 85(9), 1437–1462 (1997) doi: 10.1109/5.628714
3. Togneri, R., Pullela, D.: An Overview of Speaker Identification: Accuracy and Robustness Issues. IEEE Circuits And Systems Magazine 11(2), 23–61 (2011) doi: 10.1109/MCAS.2011.941079
4. Beigi, H.: Fundamentals of Speaker Recognition, Springer, New York (2011) doi: 10.1007/978-0-387-77592-0
5. Reynolds, D. A.: An Overview of Automatic Speaker Recognition Technology. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 4072–4075. IEEE, Orlando, FL, USA (2002) doi: 10.1109/ICASSP.2002.5745552
6. Kinnunen, T., Li, H.: An Overview of Text-Independent Speaker Recognition: From Features to Supervectors. Speech Communication 52(1), 12–40 (2010) doi: 10.1016/j.specom.2009.08.009
7. Reynolds, D. A., Rose, R. C.: Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models. IEEE Transactions on Speech and Audio Processing 3(1), 72–83 (1995) doi: 10.1109/89.365379
8. Zeng, F.-Z., Zhou, H.: Speaker Recognition based on a Novel Hybrid Algorithm. In: 25th International Conference on Parallel Computational Fluid Dynamics, vol. 61, pp. 220–226. Elsevier Procedia Engineering (2013) doi: 10.1016/j.proeng.2013.08.007
9. Jeyalakshmi, C., Krishnamurthi, V., Revathi, A.: Speech Recognition of Deaf and Hard of Hearing People Using Hybrid Neural Network. In: 2010 2nd International Conference on Mechanical and Electronics Engineering, pp. 83–87. IEEE, Kyoto, Japan (2010) doi: 10.1109/ICMEE.2010.5558589
10. Rabiner, L. R., Jang, B. H.: Fundamentals of speech recognition, Prentice Hall, NJ, USA (1993)
11. Nayana, P. K., Mathew, D., Thomas, A.: Comparison of Text Independent Speaker Identification Systems Using GMM and i-Vector Methods. In: 7th International Conference on Advances in Computing & Communications, pp. 47–54. Elsevier

Procedia Engineering, Cochin, India (2017)
doi: doi.org/10.1016/j.procs.2017.09.075

12. Chauhan, V., Dwivedi, Sh., Karale, P., Potdar, S. M.: Speech to text converter using Gaussian Mixture Model (GMM), International Research Journal of Engineering and Technology (IRJET) 3(5), 160–164 (2016)
13. Reynolds, D. A.: Automatic Speaker Recognition Using Gaussian Mixture Speaker Models. The Lincoln Laboratory Journal 8(2), 173–195 (1995)
14. Lin, L., Wang, Sh.: A New Genetically Optimized GMM for Speaker . In: Published in 6th World Congress on Intelligent Control and Automation, pp. 10235–10239 (2006) doi: 10.1109/WCICA.2006.1714005
15. Zablotskiy, S., Pitakrat, T., Zablotskaya, K., Minker, W.: GMM Parameter Estimation by Means of EM and Genetic Algorithms. In: Human-Computer Interaction. Design and Development Approaches: 14th international conference on Human-computer interaction: design and development approaches, pp. 527–536. Proceeding, Orlando, FL (2011) doi: 10.1007/978-3-642-21602-2_57
16. Saeidi, R., Mohammadi, H. R. S., Ganchev, T., Rodman, R. D.: Particle Swarm Optimization for Sorted Adapted Gaussian Mixture Models. IEEE Transactions on Audio, Speech, and Language Processing 17(2), 344–353 (2009) doi: 10.1109/TASL.2008.2010278
17. de Castro, L. N., von Zuben, F. J.: The Clonal Selection Algorithm with Engineering Applications. In: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '00), Workshop on Artificial Immune Systems and Their Applications, pp. 36–39. Las Vegas, Nv (2000)
18. de Castro, L. N., von Zuben, F. J.: Learning and Optimization Using Clonal Selection Principle. IEEE Transactions on Evolutionary Computation 6(3), 239–251 (2002) doi: 10.1109/TEVC.2002.1011539
19. Babayigit, B. A., Guney, K., Akdagli, A.: Clonal Selection Algorithm for Array Pattern Nulling by Controlling the Positions of Selected Elements. Progress in Electromagnetics Research B 6, 257–266 (2008) doi: 10.2528/PIERB08031218
20. White, J. A., Garrett, S. M.: Improved Pattern Recognition with Artificial Clonal Selection? In: Timmis J., Bentley P. J., Hart E. (eds). 2nd International Conference on Artificial Immune Sysyems, vol. 2787, pp. 181–193. Springer, Berlin (2003) doi: 10.1007/978-3-540-45192-1_18
21. Alba, E., Nakib, A., Siarry, P.: Metaheuristics for Dynamic Optimization, Springer-Verlag, Berlin (2013)
22. Du, K.-L., Swamy, M. N. S.: Search and Optimization by Metaheuristics. Techniques and Algorithms Inspired by Nature, Springer, Birkhäuser Basel (2016) doi: 10.1007/978-3-319-41192-7
23. Brownlee, J.: Clever Algorithms: Nature-Inspired Programming Recipes, Melbourne, Brownlee (2012)

Operative Recognition of Standard Signals in the Presence of Interference with Unknown Characteristics

Viktor Avramenko¹[0000-0002-6317-6711], Alona Moskalenko²[0000-0003-3443-3990]

^{1,2}Sumy State University, Rimsky-Korsakov st., 2, Sumy, 40007, Ukraine

¹avramenko1938@gmail.com, ²a.moskalenko@cs.sumdu.edu.ua

Abstract. The purpose of this work is to develop new methods for the operative recognition of standard signals in the presence of interference with unknown characteristics. To solve this problem, the disproportion functions are used.

With the exception of the method based on the use of current spectra, the proposed methods do not require observation of the analyzed signal over a certain period of time. For a given set of standard signals, recognition is performed by the current values of the analyzed signal and its derivatives. The problem is solved for several cases.

The cases are considered when additive or multiplicative interference appears and disappears at random times. The solution of the problem is given when the standard signals can be pulses of different types. An option is also considered when the interference with an unknown spectral characteristic is superimposed on the standard signal at some obviously unknown frequencies. In addition, the case is considered when additive interference takes constant values at random times. In all cases, the standard signal is a part of the analyzed signal with some constant, previously unknown scale factor. The proposed methods make it possible to recognize a fragment of the standard signal if it is in the analyzed signal at the current time.

Keywords: signal recognition, disproportion functions, integral disproportion function, additive interference, multiplicative interference, standard function, pulses, signal fragment, spectrum, harmonic.

1 Introduction

There is a wide class of tasks for which solution it is necessary to recognize the standard signals. Such problems arise in technical diagnostics and in solving other technical and technological problems. For example, in flaw detection, it is necessary to recognize the oscillograms characteristic of defects of a certain type [1].

There is a problem of recognition of acoustic pulse signals against the background of technogenic noise is described [2].

When radio signals are retransmitted in automatic mode, they are scanned in order to identify one of the pre-defined sound images [3]. If they are detected, commands are generated to change the mode of operation of the repeater.

The waveform recognition is also used in asynchronous address communication systems (AACC). Each channel (subscriber) is assigned a certain waveform, which is the hallmark of this subscriber [4].

If an analysed signal is a sum of standard signal and of the interference, it is usually unknown with what coefficient the standard signal is included in the analyzed

one. It's a serious obstacle to recognize the standard signal. In addition, often the decision-making system imposes restrictions on the time required for recognition. Therefore, the development of recognition methods that would be operational and invariant with respect to the scale factors is an important task.

This paper describes methods for recognizing standard signals in the presence of random interference with unknown characteristics. The cases when additive or multiplicative interference appears and disappears at random time are considered.

A variant is considered when the interference with an unknown spectral characteristic is superimposed on the standard signal at some obviously unknown frequencies.

It is also considered the case when the standard signals can be different types of pulses.

In addition, the problem is solved for the conditions when the additive interference at random times takes constant values, with the result that its first and second derivatives are equal to zero. All the standard functions are continuous, smooth. In all the cases considered, the standard signal is included in the analyzed one with some constant, previously unknown scale factor.

2 Formal Problem Statement

Below is the mathematical formulation of the problem. There are the finite set of standard signals that are described by the functions $f_i(t)$, where $t \in [0; T_i]$, $i = 1, 2, \dots, M$.

These functions are smooth, continuous, having first derivatives.

In the presence of additive interference, the analyzed signal is described by the expression:

$$y(t) = kf_i(t + \tau_i) + \eta(t), \quad (1)$$

where $f_i(t)$ is the i -th standard function; $\tau_i \in [0; T_i]$ is the time shift between the signal and the i -th standard; $\eta(t)$ is an additive interference, which is known only that it can disappear and appear at random points in time; k is a coefficient whose value is unknown.

The expression for the signal with multiplicative interference is

$$y(t) = kf_i(t + \tau_i)\eta(t). \quad (2)$$

It is necessary to determine by the current values of the signal and its first derivative, which of the standard functions is present at a given time in the analyzed signal.

3 Literature Review

There are many different ways to recognize signals. In [1] a neuron is proposed that can find a signal in the presence of interference. A deterministic neural network based on this neuron is proposed.

However, this network needs a certain number of samples of the processed signal. In addition, for each neuron, it is necessary to experimentally select the so-called coefficient of generalization.

In [2], the problem of recognizing acoustic pulsed signals against a background of technogenic noise is solved using samples of seismoacoustic pulses and pulses of industrial noise, as well as from sound recordings during coal mining.

A non-classical approach for solving the signal recognition task, which occurs during automated radio monitoring, is described in [5]. It is proposed to use the decision rule for selection and recognition of specified signals in the presence of unknown signals, which is based on the signal description, operating in the frequency channel, by a probability model in the form of orthogonal expansions.

The operative recognition of fragments and complexes of signals and the allocation of video data objects by means of object systems of wireless networks is developed in [6]. In this case, so-called essential counts are used. Among them, with the help of information parameters, the significant weighty counts are selected. They are a basis for the operative processing of data.

A problem of acceleration of objects detection process on the images is solved in [7]. A multiscale scanning is used. For the solution of this task, it is offered to use preliminary processing of candidates with using integrated characteristics. This processing is realized as the first stage of the classifiers cascade of the mixed type.

The artificial neural networks are used widely to solve the recognition problem [8,9].

The Wavelet analysis is also used to recognize signals [10]. However, its use also requires observations of the analyzed signal over a certain period of time.

In practice, the decision-making system often requires operative detection of signs that a fragment of one of a given set of standard signals is present in the signal being processed. To do this, it's necessary to filter the interference. However, this can be done only if certain information about the interference is known, for example, the spectral characteristic. In practice, obtaining such information can be an independent task.

The widely used correlation methods can show a close correlation for similar in shape, but different standard signals, which prevents their recognition.

There are several methods for solving a problem that satisfies the conditions set. All of them are based on using the disproportion functions, proposed in [11].

For the case when the analyzed signal is described by expression (1) or (2), the first order derivative disproportion function is used for numerical functions that are parametrically defined [12]. The first-order derivative disproportion function of $y(t)$ with respect to standard function $f_j(t + \tau_j)$ is defined as follows:

$$\begin{aligned} @d_{f_j(t+\tau_j)}^{(1)} y(t) &= \frac{y(t)}{f_j(t+\tau_j)} - \frac{y'(t)}{f_j'(t+\tau_j)} = \frac{kf_i(t+\tau_i) + \eta(t)}{f_j(t+\tau_j)} - \frac{kf_i'(t+\tau_i) + \eta'(t)}{f_j'(t+\tau_j)}. \\ &= k @d_{f_j(t+\tau_j)}^{(1)} f_i(t+\tau_i) + @d_{f_j(t+\tau_j)}^{(1)} \eta(t) \end{aligned} \quad (3)$$

The symbol "@" is selected to denote the operation of computing the disproportion. The left-hand side of (3) is read "at d one $y(t)$ with respect to $f_j(t + \tau_j)$ ".

4 Method of recognition of a continuous standard signal when impulse interference appears and disappears at random times

In this case, it is necessary to calculate the disproportion (3) for each of a given set of standards $f_j(t + \tau_j)$ while gradually increasing the shift τ_j from zero to T_j .

For the case, when $j = i$, the disproportion (3) has the form:

$$@d_{f_i(t+\tau_i)}^{(1)} y(t) = \frac{\eta(t)}{f_i(t + \tau_i)} - \frac{\eta'}{f_i'} = @d_{f_i(t+\tau_i)}^{(1)} \eta(t). \quad (4)$$

Obviously, with the disappearance of interference when $\eta(t) = 0$, $\eta'(t) = 0$, as well as with a properly selected time shift $\tau_j \in [0; T_i]$, the disproportion (4) is zero. Thus, the zero of disproportion (3) indicates that at the moment of time t the interference disappeared, and the standard function $f_i(t)$ presents in the analyzed signal with shift by τ_i . For other standard functions, the disproportion (3) will not be zero for any shifts in time. An exception may be the case when several standards have matching fragments.

When the analyzed signal is described by expression (2), the disproportion (3) $y(t)$ with respect to $f_i(t + \tau_i)$ is:

$$@d_{f_i(t+\tau_i)}^{(1)} y(t) = -k\eta'(t) \frac{f_i(t + \tau_i)}{f_i'(t + \tau_i)}. \quad (5)$$

At the moment when the derivative of interference $\eta'(t) = 0$, the disproportion (5) becomes equal to zero. That is, in this case, operative recognition of the standard signal occurs even in the presence of multiplicative interference.

5 The method of recognition of pulsed standard signals with the appearance and disappearance of impulse interference at random times

In [12], the case is considered when pulsed standard signals are recognized: rectangular, trapezoidal, etc., for which the first derivative does not always exist or a significant time interval it's equal to zero. As a result, the method discussed above cannot be applied since the disproportion (3) cannot always be calculated. For these conditions, it is proposed to use the first-order integral disproportion for the functions defined

parametrically [14]. In this case, the derivatives are not used. This disproportion of the analyzed signal $y(t)$ with respect to the standard $f_i(t + \tau_i)$ is:

$$@I^{(1)} f_i(t + \tau_i) y(t) = \frac{\int_{t-h}^t y(t) dt}{\int_{t-h}^t f_i(t + \tau_i) dt} - \frac{y(t)}{f_i(t + \tau_i)} \quad (6)$$

where h is the preset time interval. In the discrete representation of signals, this is a time quantization step.

At the time when the interference disappears, a fragment of the standard signal is automatically recognized.

If in general, we denote any standard function by $x(t)$, then it is represented by an array of samples $x_0, x_1, \dots, x_q, \dots, x_N$. The corresponding array $y_0, y_1, \dots, y_q, \dots, y_N$ is the analyzed signal.

We assume that they are obtained with the same quantization step. The defined integrals in (6) are approximately calculated by the trapezium formula. Then the integral disproportion (6) of the function $y(t)$ with respect to $x(t)$ is:

$$@I_x^1 y = \frac{y_{q-1} + y_q}{x_{q-1} + x_q} - \frac{y_q}{x_q} \quad (7)$$

To illustrate the operation of the proposed algorithm, a signal was simulated, which is the sum of a rectangular pulse multiplied by a scale factor and of some random interference. The waveform of the signal is shown in fig. 1. On intervals from 18 to 38 and from 218 to 238, the fragments of rectangle pulse are visible because the interference disappears in this time. But the automatic detection of these fragments is possible only due to the calculation of the disproportion (7). The disproportion function (7), that is appropriate for the analyzed signal is shown in the fig.2.

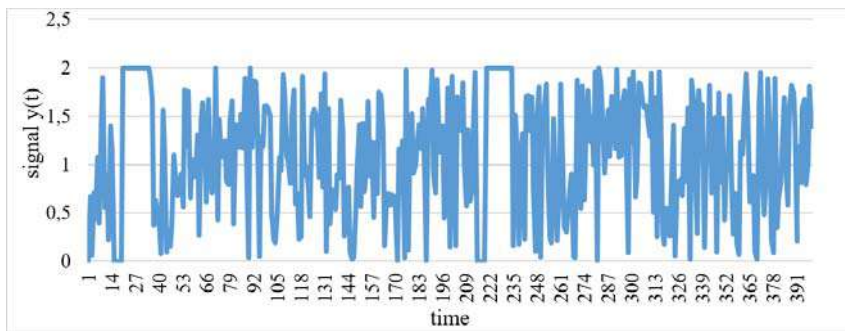


Fig. 1. Rectangular impulse with additive interference

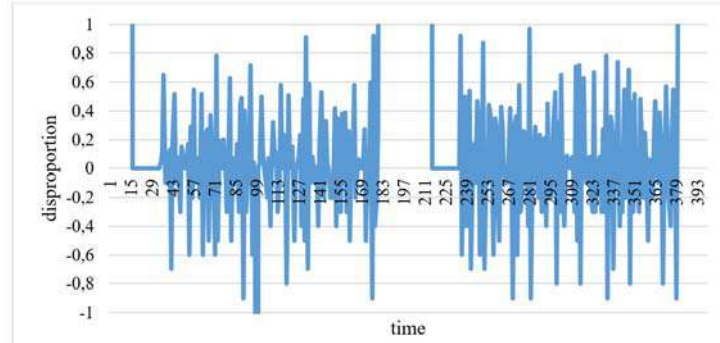


Fig. 2. The disproportion of a rectangular pulse with presence addition interference with respect to a rectangular standard one.

As can be seen from the fig.2, for points where the interference is not superimposed on the standard signal, the disproportion (7) is equal to zero, despite the presence of an unknown scale factor. Therefore, we can conclude that the analyzed signal contains a rectangular pulse.

6 Signal recognition with partial overlapping of its spectrum by the interference spectrum

The case is considered when the interference with an unknown spectral characteristic is superimposed on a standard signal at unknown frequencies. Let the set of spectra of the standard periodic signals be given. The analyzed signal is the sum of an unknown standard signal from a given set with an unknown coefficient and of random interference. The interference is an unknown periodic signal with an unknown spectrum. It is only known that the interference spectrum is partially intersected with the spectra of the standard signals at unknown frequencies. It is necessary to develop a method of recognition of the standard signal in the analyzed one by the current values of the spectrum of the analyzed signal at the current time.

In this case, the mathematical formulation of the problem can also be represented by the expression (1). However, as the standard functions, the spectra $s_i(\omega)$ of these signals are used, where ω is the frequency, $i = 1, 2, \dots, n$. Each of the functions by the condition is periodic with a period T_i , $i = 1, 2, \dots, n$.

In fact, the ordinates of the spectrum for discrete frequencies are known.

The spectrum $s_y(\omega, t)$ of the analyzed signal $y(t)$ (1) is calculated on the current interval $[t - T_i]$.

It is necessary to execute the operational recognition of a fragment spectrum of the standard signal, that is included in the analyzed one.

Let there are M harmonics of analyzed signal, that belong to standard signal only. The interference spectrum is equal to zero on these frequencies. In general, if the standard signal is a part of the analyzed one, it can be expected that there will be har-

monics m such that $m \in M$. Then the obtained amplitudes of the harmonics will differ from the amplitude of the harmonics of the standard function by the scale factor c , the value of which is unknown.

$$A_{yi}(\omega, t) = cA_i(\omega), \quad (8)$$

here, $A_i(\omega)$ is the amplitude of the i -th standard signal for the harmonic m , $A_{yi}(\omega, t)$ is the corresponding amplitude of the analyzed signal at the time instant t .

According to the condition, the frequencies $m = 1, 2, \dots, q$, for which the interference spectrum does not overlap with the spectrum of the i -th standard function are unknown. Therefore, the value of the coefficient c from equation (8) cannot be found. Also, the presence of interference does not make it possible to determine this coefficient, with the help of the normalization of the analyzed signal spectrum with respect to the spectrum of the standard one, since it is not known which fragment of the standard spectrum is included into the spectrum of the analyzed signal.

Thus, for recognition, it is necessary to detect the presence of a proportional relationship between fragments of the amplitude spectra of the standard and analyzed signals although the coefficient of proportionality c is unknown.

Therefore, to detect a proportional relationship between the amplitudes of the current spectra of the signal $y(t)$ and the standard $f_j(t)$, the first order derivative disproportion function is used for the functions specified parametrically [11].

At the current time, the disproportion (3) is calculated for the amplitude spectra of the signals $y(t)$ and $f_j(t)$. These characteristics are functions of frequency ω :

$A_{yi}(\omega, t)$ and $A_i(\omega)$, that is, the parameter in both cases is frequency.

The spectrum of the analyzed signal includes the harmonics of both the standard and the interference. At the same time, the spectrum of the analyzed signal may be wider than the spectrum of the standard one. Therefore, in order to carry out recognition for each harmonic of the analyzed signal in this case, it is preferable to calculate the disproportion of the amplitude spectrum of the standard signal with respect to the spectrum of the analyzed signal.

Accordingly, the disproportion for them is

$$@d_{A_{yi}(\omega, t)}^{(1)} A_i(\omega) = \frac{A_i(\omega)}{A_{yi}(\omega, t)} - \frac{dA_i / d\omega}{dA_{yi} / d\omega} \quad (9)$$

If the recognition of i -th standard occurs, its decomposition into a Fourier series occurs on the current interval of length T_i . The frequency changes discretely. Therefore, disproportion (9) is calculated for discrete values of the frequency $k\omega_i$, where k is the harmonic number, $k > 0$. Since discrete spectra are used, it is proposed to calculate the integral disproportion (7).

For the problem under consideration, this disproportion has the form:

$$@ I_{A_{yi}(\omega,t)}^{(1)} A_i = \frac{A_i[(k-1)\omega] + A_i[k\omega]}{A_{yi}[(k-1)\omega_i t] + A_i[k\omega_i t]} - \frac{A_i[d\omega_i]}{A_{yi}[k\omega_i t]} \quad (10)$$

The disproportion (10) equals to zero for proportional harmonics of both spectra.

Example.

Let a set of standard functions consists of two functions $f_1(t)$ and $f_2(t)$.

$$f_1(t) = \cos(t) - 0.5 \cos(2t) + 1.2 \sin(3t) + 2.2 \cos(4t) + 2 \cos(5t + 1) + 4.3 \cos(6t) - 3 \sin(7t) + 1.1 \cos(8t) - 1.5 \sin(9t) \quad (11)$$

$$f_2(t) = -1.6 \sin(t) + \cos(2t) - 2.7 \cos(3t) + 3 \sin(4t) - 1.7 \cos(5t) - 0.3 \sin(6t) + 0.8 \cos(8t) + 1.6 \cos(10t) \quad (12)$$

The interference is

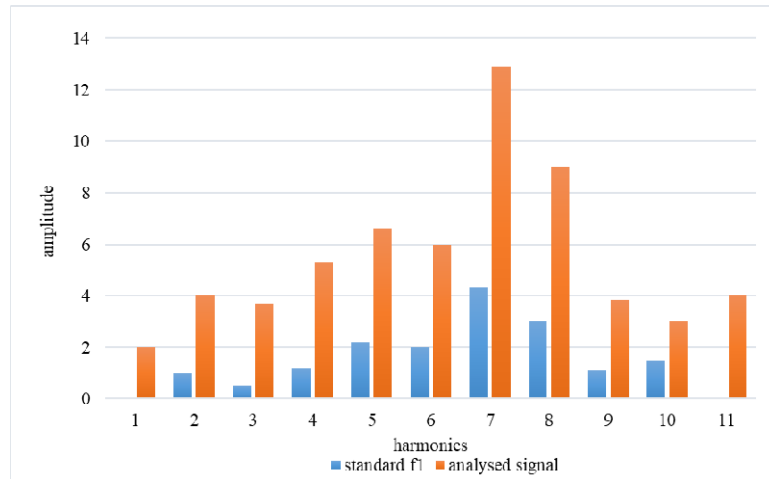
$$\eta(t) = 1 + \cos(t) + 3.2 \sin(2t - 3) - 1.7 \sin(3t + 3) - \cos(8t - 2) + 2.8 \cos(9t + 4) - 2 \cos(10t) \quad (13)$$

The choice of $f_1(t)$ and $f_2(t)$ was due to the need to model such standard signals so that the interference partially crosses both standards at some frequencies.

The analyzed signal (14) presents the first standard with a scale factor of $k = 3.12$.

$$y(t) = 3.12 f_1(t) + \eta(t) \quad (14)$$

The fig. 3 shows the graphs of the amplitude spectra of both standards against the background of the spectrum of the analyzed signal. It is obvious, that it's difficult to identify the relationship between the spectra of standard and analyzed signals visually.



a)

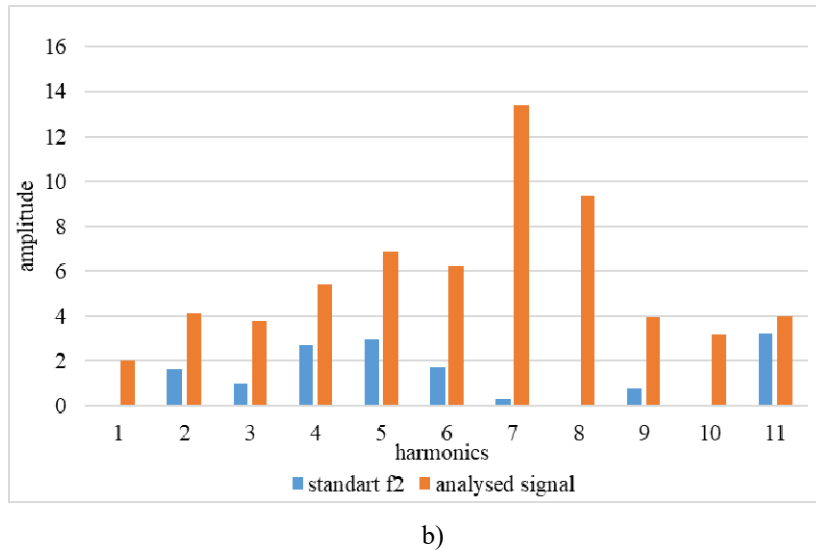


Fig. 3. Graphs of the spectra for the first time reference of the time of the first (a) and second (b) standards against the background of the signal under study

For the solving task, the values of disproportion (10) are calculated for the amplitude spectra of the standards $f_1(t)$, $f_2(t)$ with respect to the spectrum of the analyzed signal $y(t)$. The results are shown in fig. 4 and fig. 5.

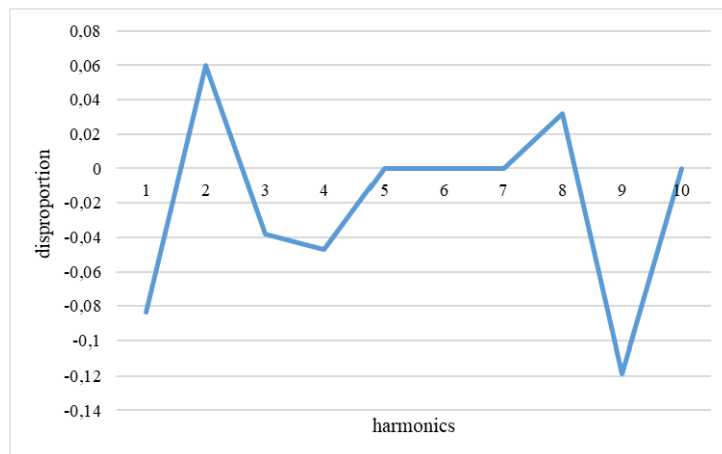


Fig. 4. Graph of the disproportion of the first standard $f_1(t)$ with respect to the signal under study

An analysis of fig. 4 and fig. 5 shows that for the first standard, the disproportion is zero at harmonics 5, 6 and 7. At the same time, for the second standard, it turned out

to be non-zero for all harmonics. This suggests that it is the first standard is a part of the signal under study.

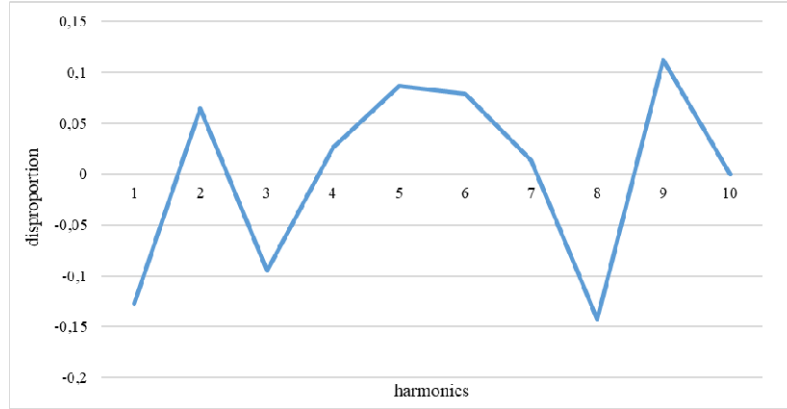


Fig. 5. Graph of the disproportion of the second standard $f_2(t)$ i -th respect to the signal under study

It is advisable to compare the proposed method of recognition of the standard signal with the correlation one. To do this, it is necessary to determine the presence or absence of a correlation between the harmonics of the analyzed and standard signals. To improve the accuracy, 21 harmonics of every standard spectrum were used to calculate the pair correlation coefficient. The critical (tabular) value is $r = 0.423$ at a significance level of $p = 0.05$ [16, 17].

For this example, the following pair correlation coefficients were obtained: between the first standard and the signal under study $r_1 = 0.837$; between the second standard and the signal under study $r_2 = -0.447$. The absolute values of both coefficients exceed the critical value. It means that there is a correlation between the spectra of the signal under study and both standards. Although the coefficient of pair correlation for the second standard signal is less than for the first one, however, this does not give grounds for asserting that the namely first standard is included in the analyzed signal.

7 Statement of the problem for the case when the standard functions are smooth, but the interference takes constant values

In the mentioned works, the case is not considered when the standard signal is continuous and is described by a smooth function, but the interference is of a pulsed nature. Specifically, the interference may be a sequence of rectangular or trapezoidal pulses with arbitrary parameters. There may be other impulse interferences, for example, the sinusoid with cut vertices due to passing through non-linear devices with

amplitude limits. In all these cases, the interference can take constant values at random times.

Thus, a constant value of the interference and a time-varying one of the standard signals arrive at the input of the discriminator at random times. The analyzed signal is described by expression (1), but now the first and second derivatives of the interference $\eta(t)$ become zero at random times.

In this case, it is proposed to investigate the first derivative of the analyzed signal (1)

$$y'(t) = kf_i'(t + \tau_i) + \eta'(t); \quad (15)$$

Instead of (3), it is necessary to calculate the current values of the first-order derivative disproportion of the function $y'(t)$ with respect to $f_j(t + \tau_j)$

For the case when $j = i$, this disproportion is:

$$\begin{aligned} \text{zyf}(t) &= @d_{f_i'(t+\tau_i)} y'(t) = \left(\frac{kf_i'(t + \tau_i) + \eta'(t)}{f_i'(t + \tau_i)} + \frac{kf_i''(t + \tau_i) + \eta''(t)}{f_i''(t + \tau_i)} \right) = \\ &= \frac{\eta'(t)}{f_i'(t + \tau_i)} - \frac{\eta''(t)}{f_i''(t + \tau_i)} = @d_{f_i''(t+\tau_i)}^1 \eta'(t) \end{aligned} \quad (16)$$

Obviously, if the interference disappears or its value is constant, $\eta'(t) = \eta''(t) = 0$. So, if the time shift is selected properly in this case, the disproportion (16) is equal to zero. In practice, the disproportion (16) also may be close to zero, if a denominator modulo tends to infinity. Therefore, it is also required to calculate the coefficient

$$k_i(t) = \frac{y'(t)}{f_i'(t + \tau_i)} \quad (17)$$

If coefficient (17) is not equal to zero and it doesn't tend to infinity, and disproportion (16) is equal to zero, this indicates that the interference has a constant (including zero) value at time t , and a fragment of the standard function $f_i(t)$ is included to the analyzed signal. It may be if the time shift τ_i is selected properly.

For the other standard functions, these two conditions will not be fulfilled at the same time for any shifts of time. As in the previous cases, the exception may be the case when several standards have matching fragments.

Now consider an example, when the analyzed signal $y(t)$ is defined as follow:

$$y(t) = kf(t) + \eta(t), \quad (18)$$

where $k = 3$;

$$f(t) = \exp(-0.1t) \cos(t). \quad (19)$$

Here, for simplicity, the shift $\tau = 0$ is assumed.

The interference $\eta(t)$ is a sine wave with cut vertices.

The disproportion $zyf(t)$ (16) for $y(t)$ (18) with respect to $f(t)$ (19) at $\tau = 0$ is showed in fig.6.

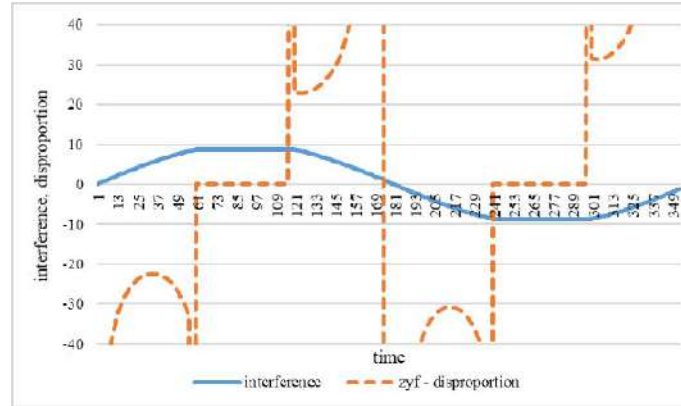


Fig. 6. Dependence of the disproportion $zyf(t)$ on the interference $\eta(t)$

From fig. 6 it can be seen that the disproportion $zyf(t)$ (16) equal to zero in the intervals, where the interference has constant values. The coefficient (17) is equal to three. Thus, fragments of the standard signal $f(t)$ are recognized at these intervals.

Let's consider another example when the standard function $f_3(t)$ is defined as

$$f_3(t) = \exp(\sin(t)) + \sin(t + 0.1). \quad (20)$$

This standard signal isn't a part of the analyzed one.

The disproportion $zyf_3(t)$ (16) for $y(t)$ (18) with respect to $f_3(t)$ (20) at $\tau = 0$ is shown in fig.7.

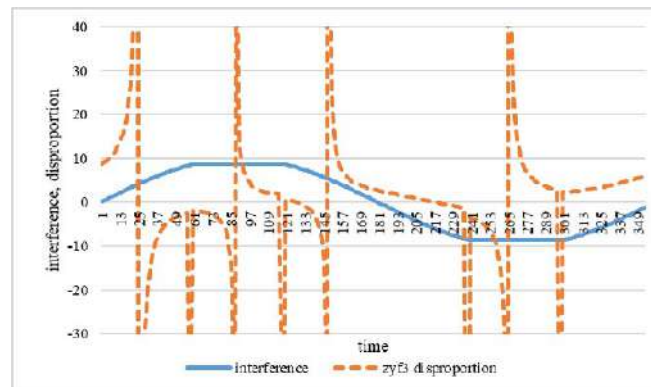


Fig. 7. The dependence of the disproportion $zyf_3(t)$ on the interference $\eta(t)$

In this example, fig.7 shows that the disproportion (16) $zyf_3(t)$ differs from zero on intervals where the interference is constant. In some cases, its values approach zero as a derivative of the standard function tends to infinity. In this case, the coefficient (17) tends to zero. All these circumstances indicate the absence of a fragment of the standard signal $f_3(t)$ in analyzed signal $y(t)$ for the current value of t .

Figures 8, 9 show the graphs of disproportions $zyf(t)$ and $zyf_3(t)$ for another interference. In this case, the interference is a sequence of rectangular pulses with pauses between them. In fact, in this example, the case when the interference disappears is also considered.

The standard and analyzed signals are previous.

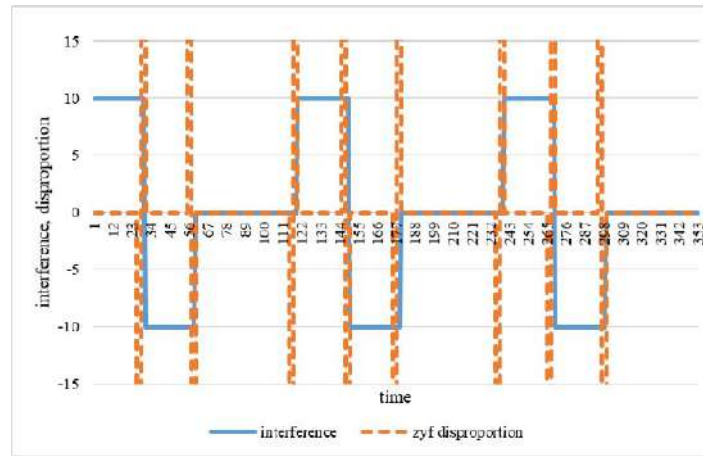


Fig. 8. The dependence of the disproportion $zyf(t)$ on the interference $\eta(t)$

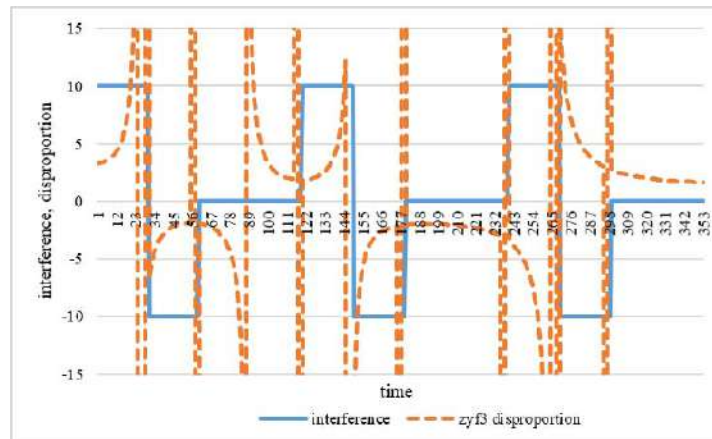


Fig. 9. The dependence of the disproportion $zyf_3(t)$ on the interference $\eta(t)$

As can be seen in fig. 8, with the exception of transients, the disproportion $zyf(t)$ is zero for almost the entire interval of change of t . In this case, the coefficient (17) is equal to three and constant. This indicates recognition of a fragment of the signal $f(t)$ at time t .

Unlike the previous case, the fig. 9 shows that the disproportion $zyf_3(t)$ differs from zero. The standard signal $f_3(t)$ does not present in the analyzed signal $y(t)$ at the time t .

8 Analysis of the methods considered

1. The advantage of the proposed methods is the recognition of a fragment of one of a given set of standard signals by the instantaneous values of the analyzed signal and its derivatives. The methods are invariant with respect to the coefficient before the standard signal.

In the case of using an analog computing device, recognition is performed instantly at the current time. For digital devices, several measurements are required to obtain derivatives by numerical methods. In particular, when computer modeling in this work, the derivatives were calculated using three points.

2. The proposed methods require large computational resources in the case of a large number of standard signals, as well as if it is necessary to select a phase shift. However, it should be noted that the methods are easily implemented using the parallelization of calculations. Indeed, it is possible to solve the problem in parallel for each of the entire specified set of standard signals and, in turn, for each of them to do this for several shift values.

3. The proposed methods allow detecting the standard signal both at individual points in time and on segments of a certain length. Therefore, the final decision on which signal is present at the moment should be made by the decision-making system, for which these methods allow to obtain initial information. In some cases, you can use information about how often the disproportion function is zero for a particular signal, and in others, how long it takes.

4. In the presence of interference, the proposed methods work only under certain conditions. In some cases, recognition is performed only in cases where interference appears and disappears. In another case, it is required that the interference take constant values.

9 Conclusions

The proposed methods allow you to automatically recognize a fragment of one of the standard signals that is a part of the analyzed at the current time. Only current data are used. All methods have invariance with respect to amplitudes of the standard and analyzed signals.

The integral disproportions make it possible to work with signals for which the calculation of the first derivative is impossible.

Separately, it should be noted the possibility of fast recognition of standard signals in case of interference in the form of a sequence of rectangular pulses that appear and disappear at random times. In this case, recognition is possible almost constantly, regardless of the amplitude of the interference and the level of the standard signal. The effectiveness of the proposed methods is verified as a result of computer simulation.

References

1. Barhatov, V.A.: Obnaruzhenie signalov i ih klassifikacija s pomow'ju raspoznavanija obrazov. Defectoscopya, vol.4, pp.14–27 (2006)
2. Deglina, JU.B.: Nejrosetevoj algoritm raspoznavanija signalov akusticheskoy jemissii. SHtuchnij intelekt, vol.4, pp. 731–734 (2006)
3. Spravochnaja sistema po moduljam Digispot II. URL: <http://redmine.digispot.ru/.../digispot.ru/projects/digispot/wiki/WikiStart>
4. Venediktov, M. D., Markov, V. V., Eyduš, G. S.: Asinhronnye adresnye sistemy svyazi. Svyaz, M.(1968)
5. Bezruk V.M., Ivanenko S.A.: Selection and recognition of the specified radio signals in the SW band. J. Information and Telecommunication Sciences, vol. 9, number 2, pp.21–25 (2018)
6. Shevchuk B. M., Zadiraka V.K., Fraier S.V., Luts V.K.: Operativne rozpoznavannia fragmentiv i kompleksiv syhnaliv ta vydilennia ob'ektiv videodanykh zasobamy ob'ektnykh system bezprovidnykh merezh.J. Iskusstvennyy intellekt, vol.3. pp.275–283 (2013)
7. Murugin K.V.: Obnaruzhenie avtomobilnykh nomernykh znakov s ispolzovaniem predvaritel'noy obrabotki kandidatov. J. Iskusstvennyy intellekt, vol.3. pp.193-199 (2013)
8. Artificial Neural Networks: Concepts and Theory, IEEE Computer Society Press, 1992
9. Osovskij, S.: Nejronnye seti dlja obrabotki informacii. Finansy i statistika, M.(2004)
10. Lazorenko, O.V., Lazorenko, S.V., CHernogor, L.F.: Primenenie vejvlet-analiza k zadache obnaruzhenija sverhshirokopolosnykh signalov na fone pomeh. Radiofizika i radioastronomija, vol.1, t. 7, pp. 46–63 (2002)
11. Avramenko, V.V.: Harakteristiki neproporcional'nostey i ih primeneniya pri reshenii zadach diagnostiki. Vestnik SumGU, vol. 16, pp.12–20 (2000)
12. Avramenko, V.V., Slepshko, N.JU.: Raspoznavanie jetalonnykh signalov pri nepolnoj informacii o harakteristikah pomeh. Visnik SumDu, vol.4, pp. 13–18 (2009)
13. Avramenko, V.V., Konopljanchenko, A.E., Prohnenko, JU.I.: Raspoznavanie jetalonnykh signalov pri pojavlenii i ischeznovenii pomehi v sluchajnye momenty vremeni. ScienceRise, vol.3/2(20), pp. 38–42 (2016)
14. Karpenko, A.P.: Integral'nye harakteristiki neproporcional'nosti chislovykh funkcij i ih primeneniye v diagnostike. Vestnik SumGU, vol. 16, pp. 20–25(2000)
15. Avramenko,V.V., Prohnenko, Yu.I.: Raspoznavanie signala pri chastichnom perekrytii ego spektra pomehoy. Visnik SumDu, vol.3, pp. 135–140 (2011)
16. Pustyl'nik, E.I.: Statisticheskie metody analiza i obrabotki nablyudeniy. Nauka, M.(1966)
17. Stanton, A Glantz: Primer of Biostatistics, McGraw-Hill Education, New York(2011)

Method for Recognizing and Processing Complex Signals

Sergey Rassomakhin ¹[0000-0003-1394-3588], Vadim Serbin ²[0000-0002-7748-1688],

Tetiana Kuznetsova ¹[0000-0001-6154-7139]

¹ V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, 61022, Ukraine
rassomakhin@karazin.ua, kuznetsova.tatiana17@gmail.com

² Yuzhnoye State Design Office, 3, Krivorozhskaya St., Dnipro, 49008, Ukraine
buba75@i.ua

Abstract. The methods of processing digital samples of complex structure signals with unknown parameters are considered. With the use of algebraic methods, the following tasks are sequentially solved: clock synchronization, determining the range of carrier frequencies, the multiplicity of phase modulation and obtaining a stream of information bits. The methods for improving the quality of processing digital samples of signals based on solving special overdetermined systems of linear equations are proposed. The estimation of efficiency of the offered method is carried out by an imitation statistical modeling. The advantages of the proposed methods of signal processing for the telecommunications and radio monitoring systems are shown.

Keywords. Orthogonal Frequency Division Multiplexing, digital sampling, linear algebraic equations, correlation convolutions.

1 Introduction

The construction of effective information transmission systems is inextricably linked with the problem of intensifying the usage of the time and frequency-energy resource of communication channels. One of the ways to solve this problem is using the complex signals with combined types of modulation in combination with the methods of spectrum narrowing and noise-resistant coding [1-5]. In this connection, the structure of the signals used to transmit information is becoming complicated, and, consequently, the algorithms of their processing are becoming complicated as well [6-11].

The most promising type of signal code constructions in wireless networks is OFDM (Orthogonal Frequency Division with Multiplexing) [2-6, 12]. The basic idea of building such signals is arranging a set of mutually orthogonal frequency subchannels so that, on the one hand, one subchannel does not interfere with the other, and on the other hand, the spectra of the subchannels overlap. Due to the orthogonality of the linear subchannels, each of them can be considered independently of the others. Errors caused by the interference in one of the subchannels do not lead to errors in the other. As a result, only a small part of the transmitted information is distorted. Error-correcting coding being used the errors can be corrected [13-18]. The structure of

signals with multiple simultaneously operating subcarrier frequencies has well established itself in conditions of heterogeneity of the propagation medium. In recent years, the capabilities of systems with OFDM signals have evolved significantly. Such signals began to be used in a wide variety of telecommunication systems operating in different radio frequency bands.

The complex structure of such signals, the a priori uncertainty of the channel properties cause significant difficulties in solving the problems of radio control and radio monitoring. A distinctive feature of such tasks is the absence of data on the structure and informative parameters of the measured signals. This information should be obtained from the results of the study, with high accuracy and as soon as possible. Therefore, the tasks of developing mathematical methods for analyzing complex signals based on digital measurement sequences are highly relevant.

2 Mathematical model of OFDM signals

For correct choice of the methods for digital analysis of the primary parameters of OFDM signals, a brief description of their basic properties is necessary. Arbitrary OFDM signal $S_j(t)$ on j -th modulation interval T_p is formed by algebraic summation of the several harmonic oscillations of the same amplitude. Each of the oscillations has m options of modulation phase shift. The value m determines the multiplicity of the used phase (PM) modulation and corresponds to the base of the numerical source code. Commonly, $m = 2^k$ where k is the number of binary symbols (bits) represented by the elementary signal on one modulation interval. When using relative phase coding and a unit value of the amplitude of the oscillation subcarriers, the mathematical model of the signal can be represented as the following sequence:

$$S_j(t) = \sum_{i=0}^{n_f-1} \sin \left\{ 2\pi \left(f_0 + \frac{i}{T} \right) \cdot \left(t - T_p \left\lfloor \frac{t}{T_p} \right\rfloor \right) + \varphi_{j,i} \right\}, \quad (1)$$

where t – current time; f_0 – the lowest subcarrier frequency in the signal spectrum; $T = 1/\Delta f$ – inverse of the minimum subcarrier spacing Δf ; n_f – the number of frequencies used; $\varphi_{j,i}$ – the value of the manipulation angle of i -th fluctuation on j -th modulation interval. This angle can take one of m values depending on the manipulation code used. The informative features in the signal described by model (1) are relative phase jumps in carrier frequencies. These jumps are measured for each of the frequency subcarriers separately: $\varphi_{j,i} - \varphi_{j-1,i}$, $i = 0, \dots, n_f - 1$. The time parameters of the modulation interval used in model (1) are tied by the relation:

$$T_p = T + \Delta T = \frac{1}{\Delta f} + \Delta T, \quad (2)$$

where ΔT - the duration of the prefix part of the signal. The prefix part (hereinafter – the prefix) is a repeating (to the exact sign) initial part of the signal added at the end of the modulation interval T_p . Prefix structure is used to facilitate synchronization in the presence of channel irregularities. For OFDM synchronization violations the signal can be correctly received and processed for any segment with the duration of T within the full modulation interval $\dots T_p \dots$. Commonly, the choice of the prefix duration on the modulation interval corresponds to the ratio $\Delta T = (0,1 \div 0,5)T$. The sign of the prefix depends on the value of the following parameter:

$$P = f_0 \bmod(\Delta f). \quad (3)$$

For existing OFDM standards, parameter (3) can take two values that determine the sign of the cyclic prefix: when $P = 0$ the prefix is positive and when $P = \Delta f/2$ – inverse. Fig. 1 gives a qualitative idea of the form of the signal envelope constructed in accordance with model (1) on two adjacent modulation intervals with $n_f = 16$ and $T_p = 1,47 \cdot T$. At the end of each of the intervals T_p the inverse cyclic continuation of the signal with duration ΔT is located, which repeats, up to a sign, the shape of the initial segment of the signal on the modulation interval. For the example in question $P = \Delta f/2$ therefore the prefix part is the inverse of the initial part of the signal.

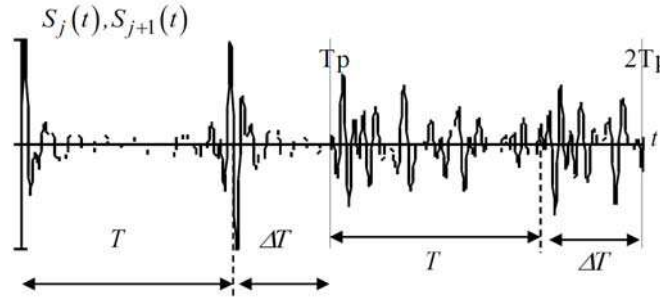


Fig. 1. Example of OFDM signal

3 The main stages of the structural analysis of OFDM signals

A comprehensive analysis of the properties of complex signals is advisable to implement on the basis of phased processing. At each stage, only a part of the signal parameters is determined. Given the fact that OFDM signals contain a prefix, it is advisable to use the correlation method for determining structural time parameters T_p and T at the first stage. This technique is based on the principle of "sliding" time window. This makes it possible to determine the following parameters of an OFDM signal: the value of the orthogonality interval, the duration of the modulation interval, and the value of the frequency spacing between the channels.

At the second stage of the analysis, the tasks of determining the number and the values of service and information channel frequencies, as well as, the signal phase demodulation.

The two-stage processing results in the possibility to extract an information flow from signals of an a priori unknown structure without using traditional fast Fourier transform algorithms (FFT – Fast Fourier Transformation).

3.1 Correlation method for determining the time parameters of OFDM signals

We propose a correlation method for determining structural time parameters T_p and T . The basis is the "sliding time window" principle. The most probably value of the time interval between the most correlated segments (with the "+" or "-" sign) of the segments from the digital sample of signal measurements is determined. The assessment of T_p – the most likely period of the emergence of "bursts" of correlation in the process of moving the viewing window on the samples of the array of measurements $Q = \{q_0, q_1, \dots\}$ is determined as well. The scheme of the calculation procedure is presented in Fig. 2. For the correlation analysis the two vectors, each containing K elements of array Q in two non-overlapping time observation windows of the signal are formed,

$$\begin{aligned} Y0 &= \{q_j, q_{j+1}, \dots, q_{j+K-1}\}, \\ Y1 &= \{q_{j+i+K}, q_{j+i+K+1}, \dots, q_{j+i+2K-1}\}, \end{aligned} \quad (4)$$

removed from each other by i , $i = 0, \dots, M$. The position of the second time window corresponding to the vector $Y1$ is determined by the successive change in the offset index $i = 0, \dots, M$ that ensures its "slip" along the signal sample Q at each of the values $j = 0 \dots L$.

For a wide range of analyzed signals, for example, for the 0.3-3.4 KHz frequency band with minimum quality ADC, the most universal limits of the values of these parameters, resulting in a quick and accurate assessment, are $K = 10 \div 30$, $M = 200 \div 300$ and $L = 1000$.

At each value of index j (moving the window slip area) the $M+1$ dimensional vector is being formed

$$V_j = \{v_0^j, v_1^j, \dots, v_M^j\}, \quad (5)$$

the elements of which are the coefficients of mutual correlation of vectors $Y0$ and $Y1$. The calculations (according to the Fig. 2) are performed after centering and normalizing the vectors by the formulas:

$$\begin{aligned}
Y0_N &= \left[Y0_j - \frac{1}{K} \sum_{i=0}^{K-1} Y0_i \right] \cdot \left[\frac{1}{K} \sum_{i=0}^{K-1} Y0_i \right]^{-1} ; \\
Y1_N &= \left[Y1_j - \frac{1}{K} \sum_{i=0}^{K-1} Y1_i \right] \cdot \left[\frac{1}{K} \sum_{i=0}^{K-1} Y1_i \right]^{-1} .
\end{aligned} \tag{6}$$

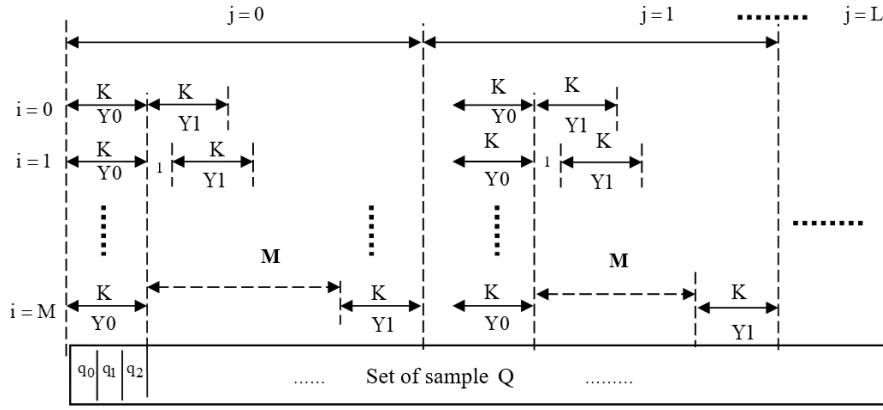


Fig. 2. Calculation scheme

The resulting vectors $Y0_N$ and $Y1_N$ in normalized space have the same length, equal to $\sqrt{K}$, and the cosine of the angle between these vectors is equal to the cross-correlation coefficient:

$$\cos(Y0_N, Y1_N) = \sum_{i=0}^{K-1} Y0_{ni} \cdot Y1_{ni} = r(Y0_N, Y1_N) . \tag{7}$$

Since the prefix is a repetitive (up to sign) part of the OFDM signal, ideally the correlation coefficient between these parts is ± 1 .

Fig. 3 shows the distribution of the values of the elements of the vector V_j calculated according to a specific implementation OFDM signal (16 carrier frequencies, a modulation rate – 75 bauds) at $M = 200$. The presence of pronounced extreme values which are close in magnitude to unity is obvious. According to the results of calculations when $j = 0 \dots L$ the two new vector $V1 = \{v1_0, v1_1, \dots, v1_L\}$ and $V2 = \{v2_0, v2_1, \dots, v2_L\}$ are formed. Their elements are calculated according to the rules:

$$\begin{aligned}
v1_j &= \text{match} \left[\min(V_j), V_j \right]_0 + K ; \\
v2_j &= \text{match} \left[\max(V_j), V_j \right]_0 + K ;
\end{aligned} \quad j = 0 \dots L . \tag{8}$$

Here the function $match[x, \mathbf{X}]$ calculates the indices of the elements of the vector $\mathbf{X}$ equal to x , where the index 0 in the function (8) indicates a selection of the element with a minimum sequence number, if there are several such elements in the vector. The elements of the vector $\mathbf{V1}$ represent the number of sampling intervals that fit between the initial elements $\mathbf{Y0}$ and $\mathbf{Y1}$ with minimal (negative) correlation on j -th step of moving the observation window. Accordingly, the elements of the vector $\mathbf{V2}$ are calculated for the maximum (positive) correlation of the vectors $\mathbf{Y0}$ and $\mathbf{Y1}$. Simultaneous determination of the maximum and minimum is necessary to reveal the value of function (3). It is obvious that the elements of the vectors $\mathbf{V1}$ and $\mathbf{V2}$ defined by expression (8), can take values only in the range $K \dots K + M$. To study the statistical distribution of the values of the elements the histograms for the elements of the vectors $\mathbf{V1}$ and $\mathbf{V2}$ are formed:

$$\begin{aligned} \mathbf{H1} &= \{h1_0, h1_1, \dots, h1_{K+M}\}, \\ \mathbf{H2} &= \{h2_0, h2_1, \dots, h2_{K+M}\}. \end{aligned} \quad (9)$$

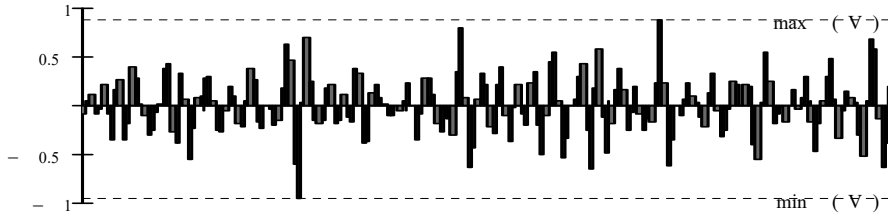


Fig. 3. Cross-correlation coefficients

Obtaining distributions (9) gives an opportunity to estimate the most likely value Num – the number of sampling intervals between the initial measurements of the segments of the digital sample $\mathbf{Q}$ with maximal (positive or negative) correlation:

$$Num = \begin{cases} match[\max(\mathbf{H1}), \mathbf{H1}]_0; \\ \langle \max(\mathbf{H1}) \geq \max(\mathbf{H2}) \rangle \\ match[\max(\mathbf{H2}), \mathbf{H2}]_0. \\ \langle \max(\mathbf{H1}) < \max(\mathbf{H2}) \rangle \end{cases} \quad (10)$$

The value Num determines the number of sampling intervals that fit on the orthogonality interval of the signal T . It gives an opportunity to find two interrelated OFDM signal parameters: the orthogonality interval and the minimum carrier frequency spacing:

$$T = \frac{Num}{f_o} \text{ and } \Delta f = T^{-1}. \quad (11)$$

Besides, choosing the corresponding method of detecting the Num value according to the condition specified in (10) automatically determines the value of the function (3), and, consequently, the ratio between the frequency parameters f_0 and Δf . If the value Num is determined by the first line of expression (10), then $f_0 = \left(k + \frac{1}{2}\right) \cdot \Delta f$. Otherwise the minimal carrier frequency is multiple to the spacing of the carrier frequencies: $f_0 = k \cdot \Delta f$ where k – any positive integer.

As a result of processing histograms (9) according to (10) based on the values of function (3), only one of the vectors $\mathbf{V1}$ or $\mathbf{V2}$ is left for the further analysis, hereinafter denoted $\mathbf{V}^*$. This is possible because the prefix repetition sign is defined. Based on the elements of the vector $\mathbf{V}^*$ another vector $\mathbf{V3}$ is formed for the analysis:

$$\mathbf{V3} = match[Num, \mathbf{V}^*]. \quad (12)$$

The elements of this vector are equal to the numbers of the elements of the vector $\mathbf{V}^*$ in which the numbers Num are located. The feature of the vector $\mathbf{V}^*$ provided that the analyzed signal belongs to the OFDM class, is that it contains a sequence of periodic series of numbers which are close or equal to Num . Therefore, the values of the elements of $\mathbf{V3}$ in order of increasing their indices will be the segments (series) of an ordinary positive integer sequence with some gaps in the sequence. Small gaps can be observed inside the series too. A possible approximation of the sequence of the elements of $\mathbf{V3}$ is illustrated by the following expression:

$$\mathbf{V3} = \left\{ \underbrace{11, 12, 14, 15, 16}_{\text{series \#1}}, \underbrace{105, 107, 108, 110}_{\text{series \#2}}, \dots, \underbrace{620, 621, 622, 623, 624}_{\text{series \# } n_c} \right\}.$$

The length of series of consecutive numbers (position numbers) may differ due to measurement errors, features of the signal envelope and rounding during calculations. However, when the vector length is sufficient, averaging results converges to the true estimation in accordance with the law of large numbers. To exclude "fragmentation" of the series, small gaps between adjacent numbers of the series must be ignored. It has been empirically found that in most cases the number N_i should be considered to belong to the current series of numbers if $N_i - N_{i-1} \leq \varepsilon$ where $\varepsilon = 3 \div 5$. In general, the structure of the vector $\mathbf{V3}$ can be depicted as shown in Fig. 4.

$$\left\{ \underbrace{v3_0, \dots, v3_{k_1}}_{\text{series \#1}}, \underbrace{v3_{n_2}, \dots, v3_{k_2}}_{\text{series \#2}}, \dots, \underbrace{v3_{n_{(n_c-1)}}, \dots, v3_{k_{(n_c-1)}}}_{\text{series \# } (n_c - 1)}, \underbrace{v3_{n_{(n_c)}}, \dots, v3_{k_{(n_c)}}}_{\text{series \# } n_c} \right\}$$

Fig. 4. Example of the $\mathbf{V3}$ series structure

In Fig. 4: $v3_{n_j}$, $v3_{k_j}$ – initial and final elements of the j -th series of the consecutive numbers in the vector $\mathbf{V3}$; n_c – the total number of the identified series. Using the presented structure and values of the elements of the vector $\mathbf{V3}$, it is possible to determine the number of sampling intervals that fit between adjacent pairs of mutually correlated segments of signal measurements, i.e. a period of "bursts" of correlation:

$$Num1 = \frac{\left(v3_{k_{(n_c-1)}} - v3_{k_2}\right) - \left(v3_{n_{(n_c-1)}} - v3_{n_2}\right)}{2(n_c - 2)}. \quad (13)$$

Using the obtained value $Num1$, we can determine the average value of the modulation interval T_p and, therefore, the average modulation rate W :

$$T_p = \frac{Num1}{f_o}, \quad W = T_p^{-1}. \quad (14)$$

For the final determination of the time-frequency structure of the signal, we must find the number of carrier frequencies n_f and the vector of their nominal values $\mathbf{F} = \{f_0, \dots, f_{n_f-1}\}$. This can be done on the basis of previously obtained values $T_p, T, \Delta f$ when the position of the element of the array $\mathbf{Q}$ corresponding to the beginning of the first full modulation interval is determined correctly. The beginning of a reliably identified clock interval could be most correctly associated with the beginning of the second series of maximal responses of correlators in the vector $\mathbf{V3}$, since, due to the randomness of the beginning of the observation, the first series may be incomplete. It should be taken into account that the beginning of a series of maximal responses of correlation of the segments from K samples must appear before the next modulation interval actually begins. Therefore, to fall within the interval with the duration T_p (taking into account that demodulation can be performed on any segment T within T_p) it is necessary to add the number $K/2$ to the starting sample, at least. Then the beginning of the modulation interval can be assumed to coincide with the next element number in the sequence

$$n_T^0 = \text{round} \left\{ \left[\left[v3_{n_2} \bmod \left(\frac{f_o}{W} \right) \right] + \frac{K}{2} \right] \right\}. \quad (15)$$

Here $\text{round}(x)$ – the rounding function to the nearest integer. The lowest possible frequency f_0 in the group of carrier frequencies is determined by the value of function (3) and the fulfillment of the corresponding condition in (10):

$$f_0 = \begin{cases} \frac{1}{2}\Delta f, & \max(\mathbf{H1}) \geq \max(\mathbf{H2}); \\ \Delta f, & \max(\mathbf{H1}) < \max(\mathbf{H2}). \end{cases} \quad (16)$$

The maximal number of subcarrier frequencies (or half the number of quadrature components) that can fit in the channel band F_{ef} is

$$n_{f \max} = \text{round}\left(\frac{F_{ef} - f_0}{\Delta f} + 1\right). \quad (17)$$

3.2 Determining the amount and nominal values of subcarrier frequencies of OFDM signals

The correlation method, considered above, allows making a reliable assessment of the main structural parameters of OFDM – T and T_p . The value of $\Delta f = T^{-1}$ uniquely defines the spacing of adjacent subcarrier frequencies. The minimal value of the subcarrier frequency and the maximal possible number of subcarriers placed within the signal bandwidth $n_{f \max}$ are determined from (16) and (17).

The number of samples N taken into account when analyzing a signal on one modulation interval, as well as, the harmonic quadrature $(2 \cdot n_{f \max})$ define the dimensions of the matrix of the linear algebraic equations system (SLAE) which can be compiled and solved to estimate the frequency range. Depending on the ratio of the vertical and horizontal dimensions of the matrix of coefficients, the system of equations can be overdetermined $(N > 2 \cdot n_{f \max})$, determined $(N = 2 \cdot n_{f \max})$ or underdetermined $(N < 2 \cdot n_{f \max})$. The simplest one is the $(N = 2 \cdot n_{f \max})$ case because then the SLAE is a joint one almost every time. The number of equations that matches the number of used elements of the digital sample $\mathbf{Q}$ equals to the number of unknowns $(2 \cdot n_{f \max})$ determining the amplitudes of quadrature components in the spectrum of carrier frequencies OFDM. For the correct solution of SLAU $(2 \cdot n_{f \max})$ uniformly spaced sample counts $\mathbf{Q}$ starting from the point of beginning of the observation of the first complete clock interval of signal n_T^0 should be selected on i -m modulation interval. For this the following rule is used:

$$n_T^i = n_T^0 + \text{round}(i \cdot \Delta) \text{ where } \Delta = T_p / t_\theta. \quad (18)$$

Square matrix of coefficients for unknown SLAE with size $(2 \cdot n_{f \max}) \times (2 \cdot n_{f \max})$ composed for quadrature components of subcarrier frequencies is formed according to the rule:

$$\begin{aligned}
\mathbf{A}_1 &= \|a_{i,j}\|, \quad i, j = 0, \dots, (2 \cdot n_{f \max} - 1); \\
a_{i,j} &= \sin[\omega_j \cdot t_i], \quad 0 \leq j \leq n_{f \max} - 1; \\
a_{i,j} &= \cos[\omega_j \cdot t_i], \quad n_{f \max} \leq j \leq 2 \cdot n_{f \max} - 1;
\end{aligned} \tag{19}$$

where

$$\omega_j = \omega_{(j+n_{f \max})} = 2\pi(f_0 + j \cdot \Delta f), \quad j = 0, \dots, n_{f \max}, \quad t_i = n_T^i + i \cdot t_\Delta.$$

The column-matrix of free members is formed as a vector of signal measurements on the duration of one orthogonality interval:

$$\mathbf{B}_1 = \{b_0, \dots, b_{(2 \cdot n_{f \max} - 1)}\}; \quad b_i = q_i, \quad i \cdot t_\Delta \in T_P^i. \tag{20}$$

Normal solution of normally defined SLAE

$$\mathbf{A}_1 \cdot \mathbf{X}_1 = \mathbf{B}_1 \Rightarrow \mathbf{X}_1 = \mathbf{A}_1^{-1} \cdot \mathbf{B}_1 \tag{21}$$

gives an estimation of the amplitude vector of quadrature components $\mathbf{X}_1 = \{x_0^1, \dots, x_{(2 \cdot n_{f \max} - 1)}^1\}$ which corresponds to the permissible values of carrier frequencies.

On the basis of this solution, it is possible to determine the power distribution vector of the signal between the harmonic oscillations of the carrier frequencies:

$$\mathbf{Y} = \{y_0, \dots, y_{(n_{f \max} - 1)}\}, \quad y_i = (x_i^1)^2 + (x_{(i+n_{f \max})}^1)^2, \quad i = 0, \dots, (n_{f \max} - 1). \tag{22}$$

The case of insufficiently defined SLAE ($N < 2 \cdot n_{f \max}$) is interesting for analyzing small samples of the signal. To solve such a SLAE, the pseudo inverse matrix method Moore-Penrose can be used. It is known that there is the normal solution of an underdetermined SLAE and is the only one. It is found by: $\mathbf{X}_1 = \mathbf{A}_1^+ \cdot \mathbf{B}_1$ where $\mathbf{A}_1^+$ – Moore-Penrose pseudo inverse matrix of size $2 \cdot n_{f \max} \times 2 \cdot n_{f \max}$ which is determined by the ratio: $\mathbf{A}_1 \cdot \mathbf{A}_1^+ \cdot \mathbf{A}_1 = \mathbf{A}_1$. In practice $\mathbf{A}_1^+$ can be found by the formula:

$$\mathbf{A}_1^+ = \mathbf{C}^+ \cdot \mathbf{D}^+ = \mathbf{C}^* \cdot (\mathbf{C} \cdot \mathbf{C}^*)^{-1} \cdot (\mathbf{D} \cdot \mathbf{D}^*)^{-1} \cdot \mathbf{D}^*.$$

The representation of the matrix $\mathbf{A}_1^+$ in the form of a product of two matrices with the size of $N \times r$ and $r \times N$ is used:

$$\mathbf{A}_1 = \mathbf{D} \cdot \mathbf{C} = \begin{pmatrix} d_{1,1} & \cdots & d_{1,r} \\ \vdots & \ddots & \vdots \\ d_{N,1} & \cdots & d_{N,r} \end{pmatrix} \cdot \begin{pmatrix} c_{1,1} & \cdots & c_{1,N} \\ \vdots & \ddots & \vdots \\ c_{r,1} & \cdots & c_{r,N} \end{pmatrix}.$$

With various skeletal decompositions of the matrix A the same solution for A^+ which can be written in the form: $\mathbf{X}_1 = \mathbf{A}_1^+ \cdot \mathbf{B}_1$ is derived. It is a pseudo solution giving a zero residual: $\|\mathbf{X}_1 - \mathbf{A}_1^+ \cdot \mathbf{B}_1\| = 0$.

The case of $(N > 2 \cdot n_{f \max})$ is the most advantageous for the maximal recording of signal information on the modulation interval. Due to using additional signal measurements from the sample $\mathbf{Q}$ the system which contains more equations with the same number of unknowns is formed. To form the matrix $\mathbf{A}_2$ and the vector $\mathbf{B}_2$ the maximal number of signal measurements determined by $Num \approx T_p/t_\theta$ on the duration T_p is used:

$$\begin{aligned} \mathbf{A}_2 &= \|a_{i,j}\|, \quad i = 0, \dots, (Num-1), \quad j = 0 \dots (2 \cdot n_{f \max} - 1); \\ a_{i,j} &= \text{Sin}[\omega_j \cdot t_i], \quad 0 \leq j \leq n_{f \max} - 1; \\ a_{i,j} &= \text{Cos}[\omega_j \cdot t_i], \quad n_{\max} \leq j \leq 2 \cdot n_{f \max} - 1; \end{aligned} \quad (23)$$

$$\mathbf{B}_2 = \{b_0, \dots, b_{(Num-1)}\}, \quad b_k = q_k, \quad k = 0, \dots, (Num-1). \quad (24)$$

SLAE has the form:

$$\mathbf{A}_2 \cdot \mathbf{X}_2 = \mathbf{B}_2, \quad (25)$$

and as a rule, has many solutions. To select the only one we need to use some criteria. In practice, the maximum likelihood criterion is used more often. In the case of a normal distribution of the vector $\mathbf{B}_2$ it is equivalent of the least square's criterion:

$$\mathbf{X}_2^* = (\mathbf{A}_2^T \cdot \mathbf{A}_2)^{-1} \mathbf{A}_2^T \cdot \mathbf{B}_2. \quad (26)$$

An approximate solution of system (26) gives a more accurate result than a strict solution of (25). The noise immunity of the solution is achieved by averaging the disturbing effect of interference when the number of signal measurements exceeds the required minimum. The obtained vector of amplitudes of the quadrature components $\mathbf{X}_2^*$, as well as $\mathbf{X}_1^*$ gives a possibility to calculate the power distribution signal in carrier frequencies using expression (22), wherein x_i^2 is used instead of x_i^1 .

For any type of SLAE determining the actual list of carrier frequencies in the OFDM spectrum is performed by comparing the elements of power distribution histograms with a threshold value. The obtained nominal values of frequencies determine

the last structural time-frequency parameter of the analyzed signal – the vector of working subcarrier frequencies $\mathbf{F}$.

Thus, the previously obtained signal parameters $T_p, T, \Delta f, W$ and the obtained in this subsection vector $\mathbf{F}$ identifies completely its structural properties and makes signal demodulation possible.

4 Conclusions

The considered statistical method for analyzing the structure and demodulation of OFDM signals under conditions of a priori uncertainty of solving radio monitoring tasks has been practically tested. It has demonstrated the high accuracy of parameter identification. The relatively low computational complexity of correlation and algebraic analysis makes it possible to identify the structure and the parameters of signals practically in seconds.

The noise immunity of the analysis is achieved by solving a SLAE with rectangular overdetermined matrixes of coefficients. To eliminate phase errors generated by asynchronous, with respect to the clock intervals of modulation, sampling the method for calculating phase corrections which takes into account the time-frequency parameters of the signal structure is proposed. The application of the phase correction method provides ideal conditions for identifying the modulation type of subcarrier oscillations. Mathematical formalization of solving the problem of determining the modulation multiplicity, based on generating the multimodal reference functions and sequential calculating the degree of mutual correlation, allows us completely automate the process of identifying the secondary parameters which are necessary for demodulating the signals of subcarrier frequencies.

The further researches can be focused on the generalization of the method for any structures of mono and poly frequency signals including those with a linear frequency modulation and also proposed author's method can be used in some different other areas [19-23].

References

1. Golomb, B.S., Gong, G.: Signal Design for Good Correlation for Wireless Communication, Cryptography and Radar. University of Southern California, University of Waterloo, Ontario (2005)
2. Shulze, H., Luders, C.: Theory and Applications of OFDM and CDMA, John Wiley & Sons, Ltd. (2005)
3. Mignone, V., Morello, A.: CD3-OFDM: A Novel Demodulation Scheme for Fixed and Personal Receivers. IEEE Transactions on Communications, 1996, vol. 44, pp. 1144–1150. (1996) doi:10.1109/26.536920
4. Li, Y., Cimini, L.J., Sollenberger, N.R.: Robust Channel Estimation for OFDM Systems with Rapid Dispersive Fading Channels. IEEE Transactions on Communications, 1998, vol. 46(7), pp. 902–915. (1998) doi:10.1109/26.701317

5. Zhao, Y., Huang, A.: A novel channel estimation method for OFDM mobile communication systems based on pilot signals end transform - domain processing. In: Proc. VTC vol 33 IEEE 47th, 1997, pp. 2089–2093. (1997) doi:10.1109/VETEC.1997.605966
6. Singh, I., Giridhar, K.: Enhancement in Spectral Efficiency using Transmit-side Channel Shortener for MISO-OFDM Systems. IEEE Signal Processing Letters, 2014, vol. 21(6), pp. 712–716. (2014) doi:10.1109/LSP.2014.2313885
7. Karpenko, O., Kuznetsov, A., Sai, V., Stasev, Y.: Discrete Signals with Multi-Level Correlation Function. Telecommunications and Radio Engineering, 2012, vol. 71(1), pp. 91–98. (2012) doi:10.1615/TelecomRadEng.v71.i1.100
8. Beaulieu, N.C., Young, D.J.: Statistical and intuitive nonlinear signal processing designs for multiuser interference mitigation in UWB systems. In: 2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), Montreal, QC, 2012, pp. 243–252. (2012) doi:10.1109/ISSPA.2012.6310554
9. Naumenko, N.I., Stasev, Y.V., Kuznetsov, A.A.: Methods of synthesis of signals with prescribed properties. Cybernetics and Systems Analysis, 2007, vol. 43(3), pp. 321–326. (2007) doi:10.1007/s10559-007-0052-8
10. Mallik, R.K., Singh, S., Murch, R.D., Mehra, S.: Signal Design for Multiple Antenna Systems with Spatial Multiplexing and Noncoherent Reception. IEEE Transactions on Communications, 2015, vol. 63(4), pp. 1245–1258. (2015) doi:10.1109/TCOMM.2015.2403831
11. Stasev, Y.V., Kuznetsov, A.A., Nosik, A.M.: Formation of pseudorandom sequences with improved autocorrelation properties. Cybernetics and Systems Analysis, 2007, vol. 43(1), pp. 1–11. (2007) doi:10.1007/s10559-007-0021-2
12. Jang, J., Choi, H.: Enhanced channel estimation for OFDM systems with multiple transmit antennas. In: 2009 15th Asia-Pacific Conference on Communications, Shanghai, 2009, pp. 402–405. (2009) doi:10.1109/APCC.2009.5375606
13. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North Holland Publishing Co. (1977)
14. Clark, G.C., Cain, J.B.: Error-Correction Coding for Digital Communications. Springer US (1981)
15. Brejza, M.F. et al.: Exponential Golomb and Rice Error Correction Codes for Generalized Near-Capacity Joint Source and Channel Coding. IEEE Access, 2016, vol. 4, pp. 7154–7175. (2016) doi:10.1109/ACCESS.2016.2584982
16. Kuznetsov, A., Serhiienko, R., Prokopovych-Tkachenko, D.: Construction of cascade codes in the frequency domain. In: 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2017, pp. 131–136. (2017) doi:10.1109/INFOCOMMST.2017.8246366
17. Stepanov, A.V., Matveev, S.A.: Methods of computer signal processing. Moscow: SOLON-Press (in Russian) (2003)
18. Tikhonov, A.N., Arsenin, V.A.: Methods for solving ill-posed problems, Moscow: Science (in Russian) (1979)
19. Kavun, S.: Conceptual fundamentals of a theory of mathematical interpretation. Int. J. Computing Science and Mathematics, 2015, vol. 6(2), pp. 107–121. (2015) doi:10.1504/IJCSM.2015.069459
20. Kavun, S.: Indicative-geometric method for estimation of any business entity. Int. J. Data Analysis Techniques and Strategies, 2016, vol. 8(2), pp. 87–107. (2016) doi:10.1504/IJDATS.2016.077486
21. Kovtun, V., Kavun, S., Zyma, O.: Co-Z Divisor Addition Formulae in Homogeneous Representation in Jacobian of Genus 2 Hyperelliptic Curves over Binary Fields. International

Journal Biomedical Soft Computing and Human Sciences, 2012, vol. 17(2), pp. 37–43.
(2012) doi:10.24466/ijbschs.17.2_37

22. Zamula, A., Kavun, S.: Complex systems modeling with intelligent control elements. Int. J. Model. Simul. Sci. Comput., 2017, vol. 08(01). (2017) doi:10.1142/S179396231750009X
23. Kavun, S., Zamula, A., Mikheev, I.: Calculation of expense for local computer networks. In: Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 2017 4th International, Kharkiv, Ukraine, 2017, pp. 146–151. (2017) doi:10.1109/INFOCOMMST.2017.8246369

SVM-based Technique for Mobile Malware Detection

Sergii Lysenko^[0000-0001-7243-8747], Kira Bobrovnikova^[0000-0002-1046-893X],
Andrii Nicheporuk^[0000-0002-7230-9475] and Roman Shchuka

Khmelnytsky National University, Khmelnytsky, Ukraine
{sirogyk@ukr.net, kirabobrovnikova@gmail.com,
andrey.nicheporuk@gmail.com, schuka.roman@gmail.com}

Abstract. A paper presents a new technique for the mobile malware detection based on the malware's network features analysis is proposed. It uses SVM for malicious programs detection. The novel approach provides the ability to detect malware in the mobile devices. As the inference engine for malware detection the support vector machine was used. The detection process is performed by taking into account the malware's features, captured in the mobile devices. Experimental research showed that the SVMs are able to produce the accurate classification results. Experiments demonstrated, that technique is able to detect different types of malware up to 98.21%.

Keywords: Malware, Mobile device, Cybersecurity, Behavior, Computer system, Network, Android, Cyberattack

1 Introduction

Last year the threats against mobile devices in particular backdoors, malicious cryptomining, banking trojans, especially during the second half of 2018, were growing [1]. The both new mobile device infection techniques (for example, DNS hijacking) and the use of tried-and-tested distribution schemes (for example, SMS spam) were applied. The one third of all mobile attacks - from building botnets, to stealing banking credentials, perpetrating click fraud, or threatening reputation - were performed by hidden apps. McAfee researchers identified 65 000 of fake mobile apps, up more than six-fold from the 10 000 detected only six months earlier. However Android remains by far the most popular mobile operating system, being used by four fifths of respondents. iOS is the only other mobile OS to reach double figures, at 15% [2].

2 Related works

Today variety of approaches to identify of android malware is widely present on the literature.

A malware detection technique called Personal Mobile Malware Guard – PMMG that classifies malwares based on the mobile user feedback is presented in [3]. PMMG

controls permissions of different applications and their behavior according to the user needs. These preferences are built incrementally on a personal basis according to the feedback of the user. Performance analysis showed that it is theoretically feasible to build PMMG tool and use it on mobile devices. However in case of sensitive mobile resources that require a large portion of permissions, the proposed technique may lead to increase a number of a false alarm.

Similar to above mentioned except permission, approach that use a combination of permissions and intents supplemented with multiple stages of classifiers for malware detection is proposed in [4]. Their experiments were performed on 1745 applications samples starting with a performance comparison between MLP, Decision Table, Decision Tree, Random Forest, Naive Bayes and Sequential Minimal Optimization classifiers. The Decision Table, MLP, and Decision Tree classifiers were then combined using three schemes: average of probabilities, product of probabilities and majority voting.

In work [5] authors have proposed an approach to detect Android malware using system call logs. The proposed technique mainly consists of three stages. At first using emulator Genymotion the system call log of applications are observed. To do this, each application was executed in emulator for approximately five minutes. To improve the quality of dataset on the second stage involves the using of filtering algorithm. The third step is the implementation of the dataset on machine-learning algorithms. The results of experiments have shown the high accuracy of detection, however authors didn't took into account capable of some applications identify a sandbox type environment.

In [6] a malware detection system that uses a deep convolutional neural network (CNN) is proposed. Malware classification is performed based on static analysis of the raw opcode sequence from a disassembled program. Features indicative of malware are automatically learned by the network from the raw opcode sequence thus removing the need for hand-engineered malware feature. Their experiments demonstrate that CNNs can effectively learn to detect sequences of opcodes that are indicative of malware.

Another static approach for android malware detection that uses multiple features to analyze is presented in [7]. A set of features such as hardware, permission, application components, filtered intents, opcodes and strings are extracted from the samples to form a vector space model. Feature selection methods such as Entropy based Category Coverage Difference (ECCD) and Weighted Mutual Information (WI) are used to choose the prominent features. The performance of the system is analyzed using classifiers, Random Forest, Rotation Forest and Support Vector Machine (SVM).

Authors in [8] propose a static-analysis based system that named MaMaDroid. The operation of a MaMaDroid goes through four stages. First, it extract the call graph from each app by using static analysis, then, it obtain the sequences of API calls using all unique nodes after which it abstract each call to class, package, or family. The third stage involves modeling of the behavior of each app by constructing Markov chains from the sequences of abstracted API calls for the app, with the transition

probabilities used as the feature vector to classify the app as either benign or malware using a machine learning classifier.

The framework TriFlow that use a triage mechanism to rank applications considering their potential risk is developed [9]. TriFlow combines a probabilistic model to predict the existence of information flows with a metric of how significant a flow is in benign and malicious apps. Based on this, TriFlow provides a score for each application that can be used to prioritize analysis. The experimental results show that it can predict the presence of information flows very accurately and that the overall triage mechanism enables significant resource saving.

The literature review has shown that the problem of android malware detection is extremely actual. However considered approaches have some common disadvantages which are manifested in ignoring packed malware and impossibility to protect the device from Zero Day attacks and malwares capable of modifying themselves.

In [10–13] an approaches for malware detection based on the network and hosting behavior analysis are presented. They use the self-adaptive idea for the malware detection and are based on usage of the semi-supervised fuzzy c-means clustering, where the objects of clustering were the feature vectors which elements may indicate the appearance of cyber threats networks and computer systems.

The mentioned above methods of the malware detection in the mobile devices have shown a high level of effectiveness, but also demonstrate high rate of false positives.

The common weakness of the aforementioned approaches is the requirement of large amounts of computing resources and the fact that they aren't able to respond adaptively to known and unknown attacks performed by malware against the mobile devices.

3 Support Vector Machine

The support vector machines (SVMs) are the high-potential approach for the object classification. SVMs are the supervised learning models with associated learning algorithms [14–15]. They are able to produce accurate and robust classification results, even when input data is non-monotone and non-linearly separable, and to evaluate more relevant information in a convenient way, providing high accuracy of classification with small training sets [16].

Basically, SVM performs classification by finding the hyperplane that maximizes the margin between two classes. Vectors (cases) that define the hyperplane are the support vectors. To define an optimal hyperplane we need to maximize the width of the margin:

$$w \cdot x + b = 0, \quad (1)$$

where x is a classification object lying on the hyperplane, w is normal to the hyperplane, b is the basis and $\frac{|b|}{\|w\|}$ is the perpendicular distance from the hyperplane to the origin, with $\|w\|$ the Euclidean norm of w [14, 17].

Let φ be a mapping function which projects the training data into a Hilbert high dimensional space H , $\varphi: R^q \rightarrow H$. The data point x is represented in space H as $\varphi(x)$.

In the context of SVM, the kernel function defines the hypothesis space, and is defined as:

$$K(x, x_i) = (\varphi(x) \cdot \varphi(x_i)), \quad (2)$$

where x_i is a training data.

It leads to decision functions of the following form:

$$f(x) = \text{sgn} \left(\sum_{i=1}^r \alpha_i \alpha_i K(x, x_i) + b \right), \quad (3)$$

where $\alpha_i, i=1, \dots, r$ are Lagrange multipliers, the maximal magnitude of which is governed by C .

In order to compute the separating hyperplane without explicitly carrying out the mapping into the feature space, different kernel functions can be used [18, 19]. In this approach, the classification process involved the kernels: linear, polynomial, Gaussian, exponential, and B-Spline.

The linear kernel is the simplest kernel function. It is given by the inner product (x, x_i) plus an optional constant c :

$$K(x, x_i) = x^T x_i + c, \quad c \in R. \quad (4)$$

The polynomial kernel is a non-stationary kernel, where the adjustable parameters are the slope α , the constant term c and the polynomial degree p :

$$K(x, x_i) = (\alpha x^T x_i + c)^p, \quad \alpha \in R, c \in R, p \in N. \quad (5)$$

The Gaussian kernel is an example of radial basis function kernel:

$$K(x, x_i) = e^{-\frac{1}{2\sigma^2} \|x - x_i\|^2}, \quad \sigma > 0, \quad (6)$$

where σ is the parameter that controls the width of the Gaussian kernel.

The exponential kernel is closely related to the Gaussian kernel, with only the square of the norm left out. It is also a radial basis function kernel:

$$K(x, x_i) = e^{-\frac{1}{2\sigma^2} \|x - x_i\|}. \quad (7)$$

The B-Spline is a radial basis function kernel, and is defined on the interval $[-1, 1]$. It is given by the recursive formula [18]:

$$K(x, x_i) = B_{2^{p+1}}(x - x_i), \text{ where } p \in N \text{ with } B_{i+1} := B_i \otimes B_0 \quad (8)$$

In order to perform the multi-class classification, the "one against all" and "one against one" SVM-based methods are used [16].

"One Against all" SVM Classifier. This method constructs k SVM models, where k is the number of classes. The m -th SVM is trained with all of the samples in the m -th class with positive labels, and all other samples with negative labels. The m -th SVM solves the task of training data mapping to a higher dimensional space for the given l training data $(x_1, a_1), \dots, (x_l, a_l)$, where $x_i \in R^n, i = 1, \dots, l, a_i \in \{1, \dots, k\}$ is the class of training data x_i .

In order to perform the classification for x_i , we use k decision functions, where k is the number of classes:

$$(w_i)^T \varphi(x_i) + b_i, \text{ where } i = 1, \dots, k, \quad (9)$$

where φ – is the mapping function, $\varphi: R^q \rightarrow H$.

The data x_i then belongs to class a , for which the above decision function has the largest value:

$$a \equiv \arg \max_{i=1, \dots, k} ((w_i)^T \varphi(x_i) + b_i). \quad (10)$$

"One Against One" SVM Classifier. The proposed technique uses the "one against one" SVM classification method as well. Here, $k(k-1)/2$ classifiers are to be constructed for each pair of classes and the max-win strategy is to be followed. Specifically, if $sqn(w_{jl})^T \varphi(x_i) + b_{jl}$ evaluates x_i to be in j -th class, then the vote for the j -th class is incremented by one, else that for the l -th class is increased by one. Finally, the training data vector x_i is predicted to belong to the class with maximum number of votes [16].

Taking into account the advantages of the SVM, it can be used as the inference engine for the making the decision about the presence of the malware in the mobile devices.

4 SVM-based Approach of the Mobile Malware Detection

We propose a new technique for the mobile malware detection. An approach presents the SVM-based system for malware detection with the ability to classify the malicious program and blocking them.

The proposed approach includes learning and detection stages. The learning stage consists of the following steps:

1. Knowledge formation based on the features that may indicate the mobile malware presence.

2. Presentation of the knowledge about the mobile malware behavior as a set of feature vectors.

3. Formation the set of mobile malware classes using SVM.

The monitoring stage consists of the following steps:

1. Gathering the features in the mobile device, which may indicate the mobile malware presence.

2. Construction of the feature vectors based on the obtained information.

3. The detecting stage involves the implementation of SVM classification of the obtained feature vectors in order to assign them to one of the mobile malware's class.

4. The blocking of the malicious program execution.

Let us take a closer look at each step of the method.

4.1 Knowledge Formation Based on the Features that May Indicate Mobile Malware Presence

Let us denote the set of malware's classes as $A = \{a_m\}_{m=1}^{N_A}$, where a_1 – the Trojans; a_2 – the backdoors; a_3 – the mobile botnets; a_4 – the spammers; a_5 – the spyware; a_6 – smartphone trackers; a_7 – the mobile proxy-servers; a_8 – SMS malware; a_9 – the exploits; a_{10} – the rootkits; a_{11} – the Adware; a_{12} – the DDoS attackers, where N_A – the number of attacks, performed by DDoS botnets [20].

Let us denote the set of features, that may indicate mobile malware's attacks against the device and are to be analyzed as $B = \{b_j\}_{j=1}^{N_B}$, where N_B – the number of features. The list of features is presented in Table 1. Let us denote the set of devices attacked by mobile malware as $H = \{h_i\}_{i=1}^{N_H}$, where N_H – the number of mobile devices. Thus, the function of the malware's attack identifying f can be presented as: $f : h_i \times b_j \rightarrow a_m$.

Table 1. The features, that take place in the malware's detection process [21].

Feature	Description
<i>The CPU resource consumption features</i>	
CPU_{PU}	the percentage of CPU time required for the user space by the operating system
CPU_{PK}	the percentage of CPU time required for the kernel space by the operating system
CPU_{PS}	the percentage of shared CPU assigned by the OS
CPU_{NP}	the total number of processes running for the application AUA
CPU_{NT}	the total number of threads running for the application AUA
<i>The memory resource consumption features</i>	
M_F	the global amount of idle memory
M_U	the global amount of used memory

M_A	the global amount of anonymous mapping maps, i.e., the area of the virtual memory (all processes) not backed by any file
M_S	the global amount of slab – a slab is the amount by which a cache can grow or shrink.
M_{TV}	the total amount of memory in swap and RAM
M_{PS}	the amount of application AUA memory shared with other processes, accounted in a way that the amount is divided evenly between the processes that share it
M_{RAM}	the size portion of memory occupied by the application AUA that is held in RAM
M_{UP}	the set of pages that are unique for the application AUA
<i>The storage resource consumption features</i>	
W_{kB}	the amount of information (expressed in kB) written on the storage by all the processes
R_{kB}	the amount of information (expressed in kB) read from storage by all the processes
R_B	the number of bytes which are read from storage by the application AUA
W_B	the number of bytes which are written into the storage by the application AUA
R_N	the value represents the number of read I/O operations
W_N	the value represents the number of write I/O operations
<i>The network resource consumption features</i>	
R_{AP}	the number of packets received by all the processes
T_{AP}	the number of packets transmitted by all the processes
R_{AB}	the number of bytes received by all the processes
T_{AB}	the number of bytes transmitted by all the processes
R_P	the number of packets received by the application AUA
T_P	the number of packets transmitted by the application AUA
R_B	the number of bytes received by the application AUA
T_B	the number of bytes transmitted by the application AUA
<i>DNS-based features</i>	
L_N	the length of the domain name
N_U	the number of unique characters in the domain name
E_N	entropy of the domain name
$T_{mod}, T_{med}, T_{aver}$	TTL-periods (mode, median, average value)
N_A	the number of A-records corresponding to domain name in the incoming DNS-message
N_{IP}	the number of IP-addresses concerned with the domain name
S_{IP}	the average distance between the IP-addresses concerned with domain name
S_A	the average distance between the IP-addresses in the set of A-records for domain name in the incoming DNS-message
N_{UA}	number of unique IP-addresses in sets of A-records corresponding to the domain name in the DNS-messages

S_{UA}	the average distance between unique IP-addresses in sets A-record corresponding to the domain name in the DNS-messages
N_D	number of domain names that share IP-address corresponding to the domain name
F_{UR}	the sign of the usage of uncommon types of the DNS-records, or DNS-records that are not commonly used by a typical client
E_R	entropy of the DNS-records, which are contained in the DNS-messages
L_P	maximum size of the DNS-messages about domain name
F_S	the sign of success of DNS-query

4.2 Presentation of the Knowledge About the Cyberattacks As the Set of the Feature Vectors

All the above-mentioned features are the base of the set of feature vectors $X = \{x_k\}_{k=1}^{N_x}$, where each of feature vector x_k describes the botnet' attack and the legitimate traffic, N_x – the number of the feature vectors.

4.3 Usage of the SVM-based inference engine for the malware detection

The main task of the SVM-based inference engine is to assign feature vector x_i to class a_t at, where $x_i \in X, a_t \in A, A = \{a_t\}_{t=1}^{N_A}$, N_A – is the number of classes, where each class corresponds to one specified type of attacks, performed by malware. The SVM-based inference engine makes a conclusion about the presence or absence of a malware in the mobile device and detects the possible type of the malware.

5 Experiments

In order to evaluate the effectiveness of the malware detection based on proposed method, experimental studies were conducted.

The experimental part of the study involved the solving of the following tasks:

- to investigate the applicability of SVM for making decisions about the presence of malware in the mobile device;
- to evaluate the detection efficiency of the malware detection in the mobile devices.

5.1 Evaluation Setting

In order to evaluate the technique's, a detection accuracy tests using malware's samples were conducted. For this purpose, a mobile malware datasets [22, 23] were employed. Used datasets were divided into training T and evaluation (test) E datasets. The training dataset includes 383 samples, 49.56% of which are malicious and the

reminder contains normal behaviors. The test dataset includes 403 samples, 55.77% of which represent malicious behaviors.

To retrieve information about the CPU consumption the top tool, which is able to provide an overview of the CPU activity in real time, was used.

To extract the features concerning the memory resource usage, system monitoring tool vmstat and procrank tool were used.

Since the tracking of the input-output activity information is disabled by default to retrieve it the modification of the specific configuration file (depending on the kernel and on the architecture of the device) was perfected and the kernel was recompiled. With purpose of the collection of input-output activity information the iostat command was used. To extract the features related to the use of the network, the means of open source tool Android Device Monitor was used.

Support vector machine was implemented using Matlab [24].

5.2 Performance Measures of the BotGRABBER System

The experimental results were estimated via standard sensitivity (SN), specificity (SP), and overall accuracy (Q) performance measures, taking into account the quantity measures of True Positives (TP), True Negatives (TN), False Positives (FP), False Negatives (FN):

$$SN = \frac{TP}{(TP + FN)}, SP = \frac{TN}{(TN + FP)}, Q = \frac{TP + TN}{TP + TN + FP + FN}. \quad (11)$$

In order to investigate the SVM-based inference engine efficiency we used different SVM kernel functions. Examples of classification results using linear, polynomial, Gaussian, exponential, and B-Spline kernels are presented in Fig. 1.

The process of classification is divided into several iterations.

In the first iteration, the classification objects are divided into two classes: malicious behavior and the benign one. Then classifiers divide objects into other two classes, for instance: malicious behavior and spam bot. The next iterations separate malicious behavior and other classes of malware and so on until all of them are totally divided.

Fig. 1 presents the placement of the classification objects on the 2-D plane and the objects' separation into two classes.

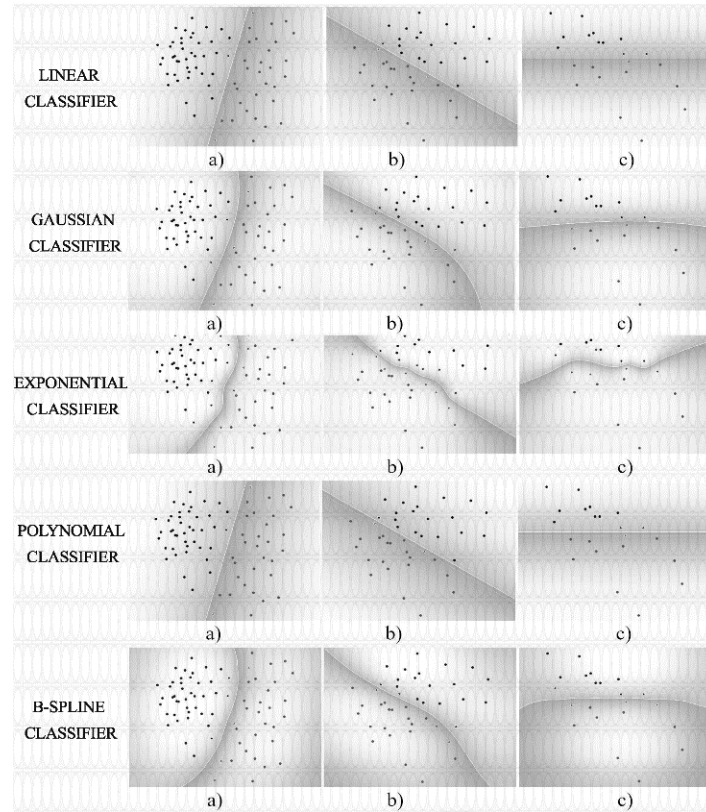


Fig. 3. Results of SVM classification using different kernel functions: a) malicious program/benign program; b) malicious program /spam bot; c) malicious program/trojn program

Experimental results of different SVM classifiers elucidated that the linear and polynomial classifiers had provided the worst results (Table 2). They were characterized by longer execution times, and higher rates of the overall classification accuracy.

Table 2. Experimental results for different SVM classifiers.

Classifier kernel	Classification accuracy		Execution time (sec)	Distance between hyperplanes
	one against	one against all SVM		
linear	92.53	96.35	0.5	0.18
Gaussian	96.18	97.39	0.4	0.25
B-spline	98.01	97.43	0.4	0.38
polynomial	93.56	96.79	0.5	0.37
exponential	97.02	97.64	0.5	0.33

Non-linear classifiers demonstrated better results, where B-Spline provided better results than others did.

For an experimental evaluation samples, the most effective classifier using the SVM was the B-spline since it provided the greatest distance between hyperplanes, the shortest time of evaluation, and the best accuracy of the classification. Given that, it was employed as a basic kernel function in the SVM-based inference engine of the proposed technique.

5.3 Results

This subsection presents overall results of technique's efficiency, taking into account sensitivity, specificity, and the overall accuracy. Employing the datasets [21, 22] this stage involved the evaluation of the SVM-based inference engine accuracy concerning each type of malware separately. The combined results are given in Table 3.

Table 3. Test results for different mobile malware's classes: number of training set samples T , number of evaluation samples E (including malicious and benign traffic samples), sensitivity (SN), specificity (SP), overall accuracy (Q) of the proposed approach, true positives (TP), true negatives (TN), false positives (FP), false negatives (FN), the rate of successful reconfigurations (SR).

Mobile malware's classes	T	E				Results		
		Malicious		Benign		SN , %	SP , %	Q , %
		TP	FN	TN	FP			
trojans	36	38	3	34	1	92,68	97,14	94,74
backdoors	31	33	2	29	2	94,29	93,55	93,94
mobile botnets	28	27	1	25	0	96,43	100,00	98,11
spammers	32	34	3	35	2	91,89	94,59	93,24
spyware	26	29	2	30	0	93,55	100,00	96,72
smartphone trackers	37	39	3	35	3	92,86	92,11	92,50
mobile proxy-servers	24	30	0	25	1	100,00	96,15	98,21
SMS malware	41	40	3	39	2	93,02	95,12	94,05
exploits	35	36	3	34	2	92,31	94,44	93,33
rootkits	30	35	6	30	1	85,37	96,77	90,28
adware	34	27	2	28	2	93,10	93,33	93,22
DDoS attackers	29	35	2	30	1	94,59	96,77	95,59

Table 3 shows that the overall accuracy of technique is in the range from 90.28% to 98.21%. Moreover, sensitivity and specificity are in the range of 85.37–100% and 92.11–100%, respectively. Therefore, this approach indicates the capability of SVM for the botnets classification.

6 Discussion

As the technique uses the SVM-based engine, there are several factors which may affect the prediction accuracy. One of them is the diversity of used training samples. Most conspicuously, that not all possible feature vectors, that describe different malware classes, are adequately represented in the training set. Thus, system may be fur-

ther improved by choosing more refined set of malicious samples for each malware classes.

In order to increase the classification accuracy the SVM prediction may be further improved by using different classification kernels, as well as the SVM optimization procedure and new mobile malware classes (and its feature vector selection) may also be improved.

Results of the experiments demonstrated, that the technique achieves the best results for detection of such mobile malware as DDoS, spyware, SMS malware, botnets, etc.

At the same time, the efficiency of the system concerning rootkits is rather lower. This is because the behavior of some malware is very similar to users' ones and some of malware's features weren't taken into account for the detection process.

7 Conclusions

A new technique for the mobile malware detection based on the malware's network features analysis is proposed. It uses SVM for malicious programs detection. The novel approach provides the ability to detect malware in the mobile devices.

As the inference engine for malware detection the support vector machine was used. The detection process is performed by taking into account the malware's features, captured in the mobile devices.

Experimental research showed that the SVMs are able to produce the accurate classification results. Implementation of the SVM-based inference engine into the mobile malware's detection process allowed to obtain its mean detection accuracy up to 98.01%. Experiments demonstrated, that technique is able to detect different types of malware in the range from 90.28 to 98.21%, while false positives is about 5%.

8 References

1. McAfee Mobile Threat Report Q1, 2019, <https://www.mcafee.com/enterprise/en-us/assets/reports/rp-mobile-threat-report-2019.pdf>
2. AV-Comparatives Security Survey, 2019, https://www.av-comparatives.org/wp-content/uploads/2019/02/Security_Survey_2019_en.pdf
3. Amro, B.: Personal Mobile Malware Guard PMMG: a mobile malware detection technique based on user's preferences. *IJCSNS International Journal of Computer Science and Network Security*, Vol. 18, No. 1, pp. 18–24 (2018)
4. Idrees, F., Rajarajan, M., Conti, M., Chen, T., Rahulamathavan, Y.: Pindroid: a novel android malware detection system using ensemble learning methods. *Computers & Security*, Vol. 68, pp. 36–46 (2017)
5. Chaba, S., Kumar, R., Pant, R., Dave, M.: Malware Detection Approach for Android systems Using System Call Logs, arXiv preprint arXiv:1709.0880 (2017)
6. McLaughlin, N., Martinez del Rincon, J., Kang, B, et al.: Deep android malware detection. In *Proc. of the Seventh ACM on Conference on Data and Application Security and Privacy*, pp. 301–308 (2017)

7. Varsha, M., Vinod, P., Dhanya, K.: Identification of malicious android app using manifest and opcode features. *Journal of Computer Virology and Hacking Techniques*, Vol. 13, Issue 2, pp. 125–138 (2016)
8. Mariconti, E., Onwuzurike, L., Andriotis, P., De Cristofaro, et al.: MaMaDroid: Detecting Android Malware by Building Markov Chains of Behavioral Model (Extended Version). *ACM Trans. Priv. Sec.*, Vol. 1, No. 1, pp. 1–33 (2019)
9. Mirzaei, O., Suarez-Tangil, G., Tapiador, J., M.de Fuentes, J.: Triflow: Triaging android applications using speculative information flows. In *Proc. of the 2017 ACM on Asia Conference on Computer and Communications Security*, pp. 640–651 (2017)
10. Lysenko, S., Pomorova, O., Savenko, O., Kryshchuk, A., Bobrovnikova, K.: DNS-based anti-evasion technique for botnets detection. In *Proc. of the 8th International Conference Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Vol. 1, pp. 453–458 (2015)
11. Pomorova, O., Savenko, O., Lysenko, S., Kryshchuk, A., Bobrovnikova, K.: Antievasion technique for the botnets detection based on the passive DNS monitoring and active DNS probing. In *Proc. of the 23rd International Conference, CN 2016, Computer Networks*, pp. 83–95 (2016)
12. Lysenko, S., Savenko, O., Bobrovnikova, K., Kryshchuk, A., Savenko, B.: Information Technology for Botnets Detection Based on Their Behaviour in the Corporate Area Network. In *Proc. of the 24th International Conference, CN 2017, Computer Networks*, pp. 166–181 (2017)
13. Lysenko, S., Savenko, O., Bobrovnikova, K., Kryshchuk, A.: Self-adaptive System for the Corporate Area Network Resilience in the Presence of Botnet Cyberattacks. In *Proc. of the 25th International Conference, CN 2018, Computer Networks*, pp. 385–401 (2018)
14. Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., Vapnik, V.: Feature selection for SVMs. In *Proc. of the 2000 Neural Information Processing Systems (NIPS) Conference*, pp. 668–674 (2001)
15. Chapelle, O., Vapnik, V., Bousquet, O., Mukherjee, S.: Choosing multiple parameters for support vector machines. *Machine learning*, Vol. 46, Issue 1–3, pp. 131–159 (2002)
16. Foody, G. M., Mathur, A.: A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on geoscience and remote sensing*, Vol. 42, Issue 6, pp. 1335–1343 (2004)
17. Deng, N., Tian, Y., Zhang, C.: *Support vector machines: optimization based theory, algorithms, and extensions*. Chapman and Hall/CRC, Data Mining and Knowledge Discovery Series, 363 p. (2012)
18. Hofmann, T., Schölkopf, B., Smola, A. J.: Kernel methods in machine learning. *The annals of statistics*, Vol. 36, No. 3(2008), pp. 1171–1220 (2008)
19. Larranaga, P., Atienza, D., Diaz-Rozo, J., et al.: *Industrial Applications of Machine Learning* / CRC Press, 336 p (2018)
20. Forensic Blog. Available: <https://forensics.spreitzenbarth.de/android-malware/>
21. Canfora, G., Medvet, E., Mercaldo, F., Visaggio, C. A.: Acquiring and analyzing app metrics for effective mobile malware detection. In *Proc. of the 2016 ACM on International Workshop on Security And Privacy Analytics. ACM.*, pp. 50–57 (2016)
22. The Drebin Dataset, <https://www.sec.cs.tu-bs.de/~danarp/drebin>
23. Android malware genome project, <http://www.malgenomeproject.org>
24. MathWorks, <https://www.mathworks.com/>

Application of global optimization methods to increase the accuracy of classification in the data mining tasks

Doroshenko A. V.^[0000-0002-7214-5108]

Lviv Polytechnic National University, Lviv, Ukraine
anastasia.doroshenko@gmail.com

Abstract. The article describes the solving of data mining task using neural-like structures of Successive Geometric Transformations Model (NLS SGTM). The main problems of this task are imbalanced dataset and different weigh of errors. Therefore, to take into account these features, the method of penalties and rewards was used, as well as a piecewise linear approach to classification. The supplement of the methods used by the final optimization procedure is proposed. The procedure of final optimization using simulated annealing.

Keywords: data mining, classification, imbalance problem, cost-sensitive learning, imbalanced data, principal components, neural-like structure of successive geometric transformations model, NLS SGTM, simulated annealing, analysis of the principal components, optimization methods.

1. Introduction

In the previous articles [1-2] the methods based on the combination of a Successive Geometric Transformations Model with the method of penalties and rewards was described. In addition, was developed the piecewise-linear approach to constructing separating surfaces in classification tasks.

The purpose of these methods is to solve the data mining tasks of the classification. The main features of these tasks are large-size datasets, imbalanced dataset and different weight of errors. The main goal of this research is increasing the accuracy of classification and minimize the number of penalty points.

In order to increase the accuracy of classification, we propose to supplement the developed methods with final optimization procedures, in particular by the method of random correction of decomposition elements and the method of simulated annealing.

2. Problem statement

Today the data mining tasks are widespread because most companies today have a huge amount of accumulated data with information about sales, customers, orders, and more. This information is a source of hidden knowledge. In turn, the possession of this knowledge allows this company to take a leading position in the market, to win

the competitive struggle. Among such tasks, one of the most popular is the task of classification. These tasks are formulated daily; in such spheres of life how as commerce, telecommunication, and chemical industry, target marketing, insurance, medicine, bioinformatics, and others. Researchers use different methods to solve classification problems [1, 2, 3, 4].

The main features of classification tasks in data mining are imbalanced data, different weight of error, huge amounts of data. These features require using some special additional methods to well-known methods of classification to provide high accuracy of classification.

Hereby as basic methods of classification, we used the neural-like structure of successive geometric transformations model (NLS SGTM) [5, 6]. As an additional method was used piecewise-linear approach [1] and cost-sensitive learning method [7, 8]. This allowed us to improve the classification accuracy and take into account the specifics of a specific task. This article proposes to apply global optimization methods to the neuro-like structure already trained as a result of previous experiments. This will allow us to find such parameters of the neural-like structure, in which the sum of points reaches the global max.

3. Increasing the accuracy of classification using random correction of decomposition elements.

3.1. Analysis of the principal components

The analysis of the principal components is the standard method used to reduce the dimensionality of data in statistical pattern recognition system and signal processing systems. However, it is also advisable to use the analysis of the principal components to solve the problems of data mining because of their high dimensionality [9, 10].

The main task of statistical recognition is the allocation of attributes - the process in which the data space is transformed into space of attributes, which theoretically has the same dimension as the input space. Conversions, however, are usually performed so that a reduced number of the most effective features can represent the data space. Consequently, only a substantial part of the information contained in the data remains, the dimension of the data is reduced. If this approach is applied to data mining task, we will reduce the size of the input data by extracting non-informative features without losing significant data. Consider a more detailed analysis of the principal components (in the theory of information is known as Karhunen-Loeve Transform) [11, 12].

Assume that there exists a vector X of dimension m , which we want to convey with the help of l numbers, where $l < m$. If we simply cut the vector x , this will cause the mean square error to be equal to the sum of the dispersions of the elements carved out of the vector x . It is necessary to find some linear transformation T , for which the value of the mean square error for the reduction of the vector X will be optimal. In this case, the transformation T must have the property of a small dispersion for its individual components. The analysis of the principal components maximizes the rate

of dispersion reduction and, accordingly, the probability of the correct choice [13, 14].

Let X be an m -dimensional random vector from the initial set of data. The mean value of this vector equal zero.

$$E(X)=0, \quad (1)$$

where E is the operator of statistical expectation. If X have a nonzero average, then you can calculate this value before the analysis begins. Let q be a unit vector with dimension m , the vector X is projected on vector q . This projection is defined as the product of the vectors X and q :

$$A=X^T q=q^T X \quad (2)$$

with restriction

$$\|q\|=(q^T q)^{0.5} = 1. \quad (3)$$

Projection A is a random variable that has an average value and variance that is related to random vector statistics X . Dispersion A equals

$$\sigma^2=E[A^2]=E[(q^T X)(X^T q)]=q^T E[(XX^T)]q = q^T R q \quad (4)$$

The matrix R of dimension $m \times m$ is the matrix of the correlation of the random vector X defined as the expectation of the product of the random vector X on itself:

$$R=E[XX^T] \quad (5)$$

The matrix R is symmetric, then:

$$R^T=R. \quad (6)$$

It follows from (6) that if a and b are arbitrary vectors with dimension $m \times 1$ then:

$$a^T R b = b^T R a \quad (7)$$

From equation (4) it follows that the projection A is a function of an odd vector q . So, we have:

$$\psi(q) = \sigma^2 = q^T R q, \quad (8)$$

$\psi(q)$ – dispersion probe.

Directly the principal components are defined as follows. Let the data vector x be a realization of the random vector X . Since there are m possible values of the single vector q , then we must consider m possible projections of the data vector x . By formula (2)

$$a_j = q_j^T x = x^T q_j, \quad j = 1, 2, \dots, m \quad (9)$$

where a_j – the projections of the vector x on the main guides, represented by single vectors q_j . These projections are called principal components, their number corresponds to the dimension of the data vector x . In this case formula (9) can be considered as the analysis procedure.

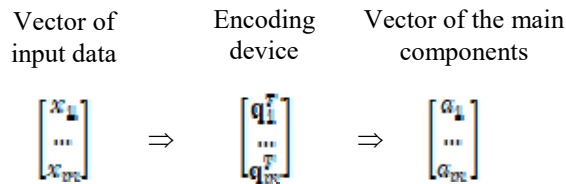


Fig. 1. Scheme of the procedure for analysis of the principal components. The encoding step

Allocation of the principal components (PC) is also proceeding on the outputs of the hidden layer of the neural-like structure of GTM. After selecting the PC, we create and teach an additional neural-like structure (Fig. 2), where the inputs are the vectors of the PC, and the outputs are the significance of the corresponding outputs of the initial training data [15, 16, 17].

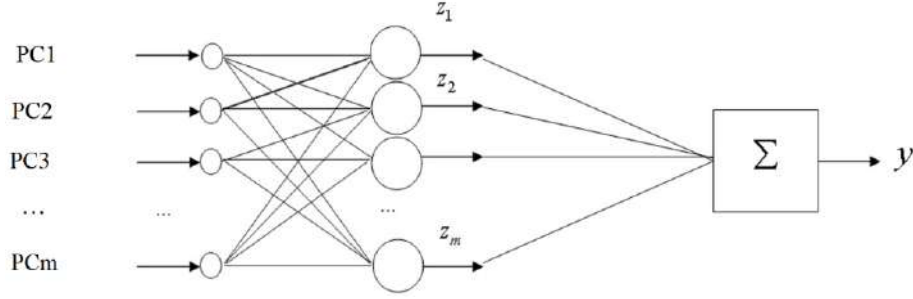


Fig. 2. The structure of neural network for obtaining decomposition elements

The feature of this neural network is the representation of the output of the synapse as $y = \sum_{i=1}^m z_i$ well as weak correlation z_i . In the mode of application on the inputs of the trained neural network, input vectors are:

$$\begin{aligned} 00\dots0 &\rightarrow y_0 \\ 10\dots0 &\rightarrow y_1 \\ 01\dots0 &\rightarrow y_2 \\ &\dots\dots\dots \\ 00\dots1 &\rightarrow y_m \end{aligned}$$

Define $\alpha_0 = y_0, \alpha_i = y_i - y_0, i = 1, 2, \dots, m$.

Get the elements of decomposition $z_0 = y_0, z_i = PC_i \alpha_i, i = 1, 2, \dots, m$.

Initially we have

$$y = \sum_{i=0}^m z_i \text{ or } y = \sum_{i=0}^m z_i k_i, \quad (10)$$

where z_i – the elements of decomposition, $k_i = 1, i = 1, 2, \dots, m$.

For classification tasks, appropriate indicators are optimized (percentage of properly classified specimens, number of penalty points, average arithmetic or mean square error) by random correction of elements of decomposition of coefficients k_i . Random correction of coefficients k_i can be performed simultaneously for all components or for each component independently. It should be noted that component-based random correction is appropriate for prediction since it has been experimentally confirmed that the components are practically independent [18, 19].

Consider a more detailed algorithm for random correction.

3.2. Increasing the accuracy of the classification tasks based on correction of decomposition elements by random correction

Let us consider the case of optimizing the accuracy of the classification problem solving for a case where the recognized data belong to one of the two classes.

Re-coding the input file: outputs of the elements belonging to class 1, assign 1, the outputs of the elements belonging to class 2, assign -1.

The problem of recognition is solved as a prediction problem by the formula (3), but after obtaining a value, we analyze it by the formula (11):

$$y \rightarrow \begin{cases} \text{class1}, y \geq 0, \\ \text{class2}, y < 0. \end{cases} \quad (11)$$

For the task of recognizing the criteria by which optimization can be carried out, there may be the number of penalty points, the percentage of incorrectly classified representatives of class 1, the percentage of incorrectly classified representatives of class 2.

Then the percentage of incorrectly classified representatives of classes is calculated as follows:

$$ErrC1 = 100\% NWC1/NC1, ErrC2 = 100\% NWC2/NC2 \quad (12)$$

where $NC1$ is the number of representatives in class 1 in the training sample, $NC2$ – number of representatives of class 2 in the training sample, $NWC1$ – number of incorrectly classified representatives of class 1, $NWC2$ – number of incorrectly classified representatives of class 2.

In order to calculate the number of penalty points, it is necessary to determine the fines that count for erroneous recognition. So if $P1$ is a fine charged if the element of class 1 is recognized as an element of class 2 and $P2$ is a fine charged if a class 2 element is recognized as an element of class 1.

Then the value of the penalty function - the total number of received penalty points (PP) is:

$$PP = ErrC1P1 + ErrC2P2. \quad (13)$$

Optimization method for classification tasks based on correction of decomposition elements by random correction:

1. Initial values $k_i=1$.
2. We calculate the value of the error, which is optimization (PP , $ErrC1$ or $ErrC2$).
3. Using a random number generator with a uniform distribution, we choose the value ΔD from the range $(-D, D)$.
4. Calculate new values $k_i = k_i + \Delta D$.
5. Calculate the value of outputs with new k_i .
6. Convert the resulting value to the designation of the class to which this element belongs.
7. Calculate the new value of the error, which is optimization (PP , $ErrC1$ or $ErrC2$).

8. Compare the value of the calculated error with the pre-calculated value. If the new value is less than the previous one, we will remember it as well as the current coefficients k_i . Go to step 3. Otherwise, go to step 3 without remembering.

9. Continue optimization until the predefined desired optimization value is reached, or until the time t has expired.

After performing the optimization method by randomly correcting the decomposition elements, the resulting coefficients are used for further classification.

4. Application of global optimization methods to increase the accuracy of classification in the tasks of the data mining

4.1. Method of simulated annealing

Annealing method is an algorithmic analogue of the controlled cooling process. It was proposed in 1953 by N. Metropolis and refined by numerous followers. Today it is considered one of the few methods by which one can practically find the global minimum of functions of several variables. Consider a more detailed method of simulated annealing [20, 21].

Algorithm of simulated annealing method:

1. Start the process from the starting point at a given initial temperature $T=T_{max}$.
2. As long as $T>0$, repeat L times the following actions:
 - choose a new solution w' from the vicinity w ;
 - calculate the change of target function; $\Delta = E(w') - E(w)$;
 - if $\Delta \leq 0$, take $w = w'$; otherwise, if $\Delta > 0$, take $w = w'$ with probability $\exp(-\Delta/T)$ by generating a random number R from the interval $(0,1)$, then comparing it with the value $\exp(-\Delta/T)$; if $\exp(-\Delta/T) > R$, take a new solution $w = w'$; in other case - ignore it.
3. Reduce the temperature ($T \leftarrow rT$) using the reduction coefficient r , selected from the interval $(0,1)$ and return to step 2.
4. After lowering the temperature to zero, apply one of the deterministic methods (Levenberg-Marquardt algorithm, error-return algorithm, fastest-speed algorithm, etc.) to achieve the minimum of the target function.

The concept of "temperature" in this algorithm is quite formal, since the presented optimization model is only a mathematical analogy of the annealing process.

The efficiency of the annealing algorithm has an extremely high impact with the choice of parameters such as the initial temperature T_{max} , the coefficient of reduction of temperature r and the number of cycles L , performed at each temperature level.

The main problem is to determine the threshold level optimal for each annealing simulation process. For some practical tasks, this level may have different meanings, but the overall range remains unchanged. As a rule, the initial temperature is selected so as to ensure the implementation of about 50% of the subsequent random changes in

the solution. Therefore, knowledge of the pre-distribution of such changes makes it possible to estimate the initial temperature approximately.

Numerous computer experiments [22] prove that in the case where the time limit is small, the best results give a single implementation. If simulation can be long lasting, then statistically better results can be achieved thanks to the multiple implementation of the annealing simulation if the value of the coefficient r is close to 1.

If we compare genetic algorithms with an annealing algorithm, then, in spite of the significant external difference between the algorithms, they are essentially similar in nature. An annealing algorithm according to [23] can be considered a genetic algorithm with a population consisting of one instance. Consequently, an algorithm for simulating annealing of a metal can be regarded as an algorithm that has only a mutation operation, but not cross-linking.

In addition, if we compare these two algorithms from the applied point of view, then it should be noted that, according to Kohonen's study [7], in the case when the initial solution is sufficiently close to optimal, the annealing algorithm of the metal has significant advantages over the genetic algorithms from a computational point of view.

Since in our study the initial data are pre-processed by methods of fines and incentives with sampling alignment and piecewise linear classification on the basis of the model of geometric transformations, then the initial solution of the problem is sufficiently close to the optimal. Accordingly, in this case, it is more appropriate to choose for optimization of the solution.

Improvement of accuracy of the decision of tasks of intellectual data analysis on the basis of correction of elements of decomposition by algorithm of simulated annealing

Let's consider a more detailed algorithm of simulated annealing of metal in combination with methods of fines and incentives and piecewise linear classification on the basis of a model of geometric transformations.

Fig. 3 depicts the structural scheme of the developed neural network based on the model of geometric transformations, where $x_1, x_2, \dots, x_n$ primary features of classification objects – input data, $PC_1, PC_2, \dots, PC_n$ – the principal components derived from input data, $w_1, w_2, \dots, w_n$ – weight coefficients, y – an output that indicate on belonging to certain classes.

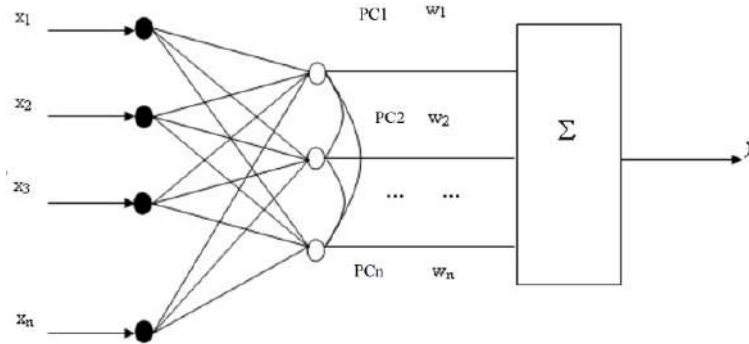


Fig. 3. The scheme of neural-like structure GTM

Functioning of such a neural network can be described by the formula 14.

$$y = \sum_{i=1}^n PCw_i \quad (14)$$

The method of simulating annealing of a metal is proposed to be used to optimize weight coefficients so that the resulting amount of penalty points is minimal, that is, the optimization parameter is the amount of penalty points [15, 17].

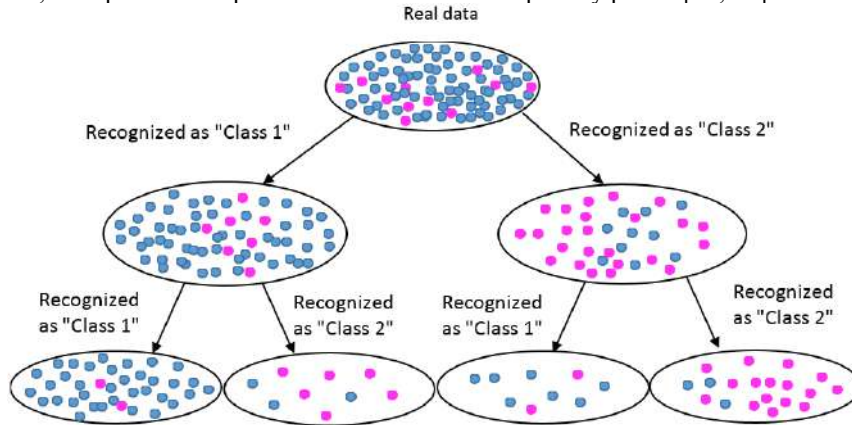


Fig. 4. The division tree into classes based on NLS SGTM

As can be seen from the flowchart describing the solution of the data mining problem by combining the method of fines and incentives, simulated annealing and piecewise linear approach, a modified annealing method for which the target function is the number of penalty points, will be applied separately for each cluster [22, 24].

Accordingly, if we have a two-step division into clusters for a sample of n classes, then we will have division into clusters, which is depicted in Fig. 4, and for each of the clusters a modified annealing algorithm will be implemented.

5. Experimental results

This article describes the solving of classification task, which was formulated in [1]. The training sample describes the transactions carried out by credit card holders within two days and consists of 284,807 lines and 31 columns. Also, the dataset contains one target feature 'Class', which shows the client's affiliation to one of two classes - frauds or ordinary clients. The main feature of the dataset is that the data set is highly unbalanced - only 492 transactions out of 284807 (0.172% of all transactions) have the value of the target field 1, that is, customers are fraudulent. The dataset has been collected and analyzed during a research collaboration of Worldline and the Machine Learning Group (<http://mlg.ulb.ac.be>) of ULB (Université Libre de Bruxelles) on big data mining and fraud detection [7].

According to the subject area, a matrix was formed. By analyzing this matrix, it can be seen that a properly classified vector that belongs to the "fraud" class has a much greater weight than a properly classified "ordinary client" vector. At the same time, the case where an ordinary customer is classified as fraud has the highest number of penalty points (Table 1).

Then we used the modified method of imitation of annealing of the metal with parameters: initial temperature $T=T_{max}=20895$, $L=100$, R is a random number from the interval $(0,1)$, $r=0,9$.

Table 1. Matrix of penalties and rewards solving task

Matrix of Rewards and Penalties	Values of Rewards and Penalties	
	The vector is recognized as class 1	The vector is recognized as class 2
The vector belongs to class 1	1	-3
The vector belongs to class 2	-2	5

After lowering the temperature to zero, apply one of the deterministic methods (Levenberg-Marquardt algorithm, error-return algorithm, fastest-speed algorithm, etc.) to achieve the minimum of the target function. The results of classification are described (Fig 5, Fig.6).

Classification accuracy using

method penalty and rewards (in vectors)

All Data

		Predicted		
		1	2	Σ
Actual	1	284278	37	284315
	2	113	379	492
		284391	416	284807

Cluster 1

		Predicted		
		1	2	Σ
Actual	1	284270	8	284278
	2	58	55	113
		284328	63	284391

Cluster 2

		Predicted		
		1	2	Σ
Actual	1	21	16	37
	2	3	376	379
		24	392	416

Classification accuracy using

method penalty and rewards with optimization (in vectors)

All Data

		Predicted		
		1	2	Σ
Actual	1	284295	20	284315
	2	105	387	492
		284400	407	284807

Cluster 1

		Predicted		
		1	2	Σ
Actual	1	284281	14	284295
	2	28	77	105
		284309	91	284400

Cluster 2

		Predicted		
		1	2	Σ
Actual	1	17	3	20
	2	9	378	387
		26	381	407

Classification accuracy in %

After optimization

ErrC1%	99,99297
ErrC2%	78,65854

ErrC1%	99,99508
ErrC2%	73,33333

ErrC1%	85
ErrC2%	97,67442

Fig. 5. Results of classification using NLS SGTM with the method of penalties and rewards and the method of simulated annealing (in vectors)

<u>All Data</u>					<u>All Data</u>				
Actual	1	Predicted		Σ	1	Predicted		Σ	
		1	2			1	2		
	2	284278	-111	284167	2	284295	-60	284235	
Actual	1	-226	1895	1669	Actual	1	-210	1935	
	2	284052	1784	285836		2	284085	1875	285960
<u>Cluster 1</u>					<u>Cluster 1</u>				
Actual	1	Predicted		Σ	1	Predicted		Σ	
		1	2			1	2		
	2	284270	-24	284246	2	284281	-42	284239	
Actual	1	-116	275	159	Actual	1	-56	305	
	2	284154	251	284405		2	284225	343	284568
<u>Cluster 2</u>					<u>Cluster 2</u>				
Actual	1	Predicted		Σ	1	Predicted		Σ	
		1	2			1	2		
	2	21	-48	-27	2	17	-9	8	
Actual	1	-6	1880	1874	Actual	1	-18	1890	
	2	15	1832	1847		2	-1	1881	1880
<u>Cluster 1</u> + <u>Cluster 2</u> = 286252					<u>Cluster 1</u> + <u>Cluster 2</u> = 286448				

Fig. 6. Results of classification using NLS SGTM with the method of penalties and rewards and the method of simulated annealing (in points)

6. Conclusion

The application of global optimization methods to classification in the data mining tasks allowed to increase the accuracy of classification, especially in combination with other methods of classification, such as neural-like structure of successive geometric transformations model. Also, the method of simulated annealing successfully combined with such methods as piecewise-linear approach and cost-sensitive learning method. The application of the method of simulated annealing made it possible to reach the point of the global maximum and minimize the amount of penalty points for this task.

References

1. Doroshenko, A.: Piecewise-Linear Approach to Classification Based on Geometrical Transformation Model for Imbalanced Dataset. In: Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, pp. 231-235 (2018)
2. Tkachenko, R., Doroshenko, A.: Classification of Imbalanced Classes using the Committee of Neural Networks. In: Proceedings of the XIIIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT), Lviv, 11-14 September 2018. (2018)
3. Tepla, T.L., Izonin, I.V., Duriagina, Z.A., Tkachenko, R.O., Trostianchyn, A.M., Lemishka, I.A., Kulyk, V.V., Kovbasyuk, T.M.: Alloys selection based on the supervised learning technique for design of biocompatible medical materials Archives of Materials Science and Engineering, 93/1, pp. 32-40 (2018)
4. Izonin, I., Trostianchyn, A., Duriagina, Z., Tkachenko, R., Tepla, T., Lotoshynska N.: The Combined Use of the Wiener Polynomial and SVM for Material Classification Task in Medical Implants Production. International Journal of Intelligent Systems and Applications (IJISA), Vol.10, No.9, pp.40-47 (2018).
5. Tkachenko, R., Yurchak, I., Polishchuk, U.: Neurolike networks on the basis of Geometrical Transformation Machine. In: Proceedings of the 2008 International Conference on Perspective Technologies and Methods in MEMS Design, Polyana, pp. 77-80 (2008). doi: 10.1109/MEMSTECH.2008.4558743
6. Polishchuk, U., Tkachenko, P., Tkachenko, R., Yurchak, I.: Features of the auto-associative neuro like structures of the geometrical transformation machine (GTM). In: Proceedings of the 2009 5th International Conference on Perspective Technologies and Methods in MEMS Design, Zakarpattya, pp. 66–67 (2009)
7. Pozzolo, A.D., Caelen, O., Johnson, R.A., Bontempi, G.: Calibrating probability with underdamping for unbalanced classification. In: Proceedings of the 2015 IEEE Symposium Series on Computational Intelligence, Cape Town, pp. 159–166 (2015) doi: 10.1109/SSCI.2015.33
8. Liu, X.: A Benefit-Cost Based Method for Cost-Sensitive Decision Trees. In: Proceedings of the WRI Global Congress on Intelligent Systems, Xiamen, pp. 463-467 (2009).
9. Ebied, H. M.: Feature extraction using PCA and Kernel-PCA for face recognition. In: Proceedings of the 2012 8th International Conference on Informatics and Systems (INFOS), Cairo, pp. MM-72-MM-77 (2012)
10. Subbotin, S.: Quasi-Relief Method of Informative Features Selection for Classification. In: Proceedings of the 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, pp. 318-321 (2018).
11. Amar, A., Leshem, A., Gastpar, M.: A greedy approach to the distributed Karhunen-Loève transform. In: Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, pp. 2970-2973 (2010)
12. Mattera, D., Palmieri, F., Haykin, S.: An explicit algorithm for training support vector machines. In: Proceedings of the IEEE Signal Processing Letters, vol. 6, no. 9, pp. 243-245 (1999).
13. Riznyk, O., Yurchak, I., Povshuk, O.: Synthesis of optimal recovery systems in distributed computing using ideal ring bundles. In: Proceedings of the 2016 XII Interna-

- tional Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), Lviv, pp. 220-222 (2016)
14. Li, Xin, Qian, Xu, Wang, Ziqiang: Classification rule mining using feature selection and genetic algorithm. In: Proceedings of the 2009 Asia-Pacific Conference on Computational Intelligence and Industrial Applications, Wuhan, pp. 107-110 (2009)
 15. Yue, S., Kaizhang, Liu, W., Wang, Y.: A Fuzzy Clustering Approach Using Reward and Penalty Functions. In: Proceedings of the 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery, Tianjin, pp. 18-21 (2009)
 16. Zhiwei, Ye, Juan, Yang, Xu, Zhang, Zhengbing, Hu: Remote Sensing Textual Image Classification based on Ensemble Learning. *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, Vol.8, No.12, pp.21-29 (2016).
 17. Tkachenko, R., Cutucu, H., Izonin, I., Doroshenko, A., Tsymbal, Y.. (2019) Non-iterative Neural-like Predictor for Solar Energy in Libya. In: Ermolayev V., Suárez-Figueroa M., Yakovyna V., Mayr H., Nikitchenko M., Spivakovsky A. (eds) In: Proceedings of the 14-th International Conference ICTERI 2018. Volume I: Main Conference. Kyiv, Ukraine, May 14-17, pp.35-45 (2018).
 18. Oleynik, A., Subbotin, S.: Parametrical synthesis of neural network models based on the evolutionary optimization. In: Proceedings of the 2009 10th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics, Lviv-Polyana, pp. 335-338 (2009).
 19. Feiden, D., Tetzlaff, R.: Iterative annealing: a new efficient optimization method for cellular neural networks. In: Proceedings of the 2001 International Conference on Image Processing (Cat. No.01CH37205), Thessaloniki, Greece, 2001, pp. 549-552 vol.1.
 20. Granville, V., Krivanek, M., Rasson, J.-P: Simulated annealing: A proof of convergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 16 (6): 652–656 (1994) doi:10.1109/34.295910.
 21. Huang, K., Hsieh, Y.: Very fast simulated annealing for pattern detection and seismic applications. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, pp. 499-502 (2011)
 22. Shafti, L. S., Pérez, E.: Data Reduction by Genetic Algorithms and Non-Algebraic Feature Construction: A Case Study. In: Proceedings of the 2008 Eighth International Conference on Hybrid Intelligent Systems, Barcelona, pp. 573-578 (2008)
 23. Zhengbing, Hu., Bodyanskiy, Y.V., Tyshchenko, O. K., Boiko, O. O.: An Ensemble of Adaptive Neuro-Fuzzy Kohonen Networks for Online Data Stream Fuzzy Clustering. *International Journal of Modern Education and Computer Science(IJMECS)*, Vol.8, No.5, pp.12-18 (2016). doi: 10.5815/ijmeecs.2016.05.02

The simulator and neuro-controller for small satellite attitude development

Nataliya Shakhovska^[0000-0002-6875-8534], Dmytro Kozii, Pavlo Mukalov

Lviv Polytechnic National University, Lviv 79013, Ukraine
nataliya.b.shakhovska@lpnu.ua, dmytruto@gmail.com,
pmykalov@gmail.com

Abstract. The paper describes the realization of simulator and neuro-controller for small satellite attitude. The main types of neuro-controllers are analyzed. A problem of proper neuroemulator choosing for neurocontroller training is analyzed. A new criterion on the basis of local control gradients analysis for input neuroemulator's neurons is proposed. Results of numerical simulations of neurocontroller training by a gradient descent method is given.

Keywords: satellite, neuro-controller, learning rate, attitude

1 Introduction

Neural control is a kind of adaptive control when artificial neural networks (NN) are used as building blocks of control systems. Neural networks have a number of unique properties that make them a powerful tool for building control systems: the ability to learn from examples and to summarize data, the ability to adapt to changing properties of the object of control and the environment, the suitability for the synthesis of nonlinear regulators. Over the past 20 years, a large number of neurological methods have been developed, the most popular among them are Model Reference Adaptive Neurocontrol and Adaptive Critics [2].

The method of neural control with a reference model, also known as a "circuit with neurotransmitter and neuro-controller" or "reciprocal distribution in time," was proposed in the early 1990s [1], [3 – 5]. This method does not require knowledge of the mathematical model of the control object. Instead, a separate neural network, a neuromuscular, studies the direct dynamics of the control object and then it is used to calculate derivatives when training a neuro-controller. At the same time, the trained neuro-emulators with the lowest mean square error of the simulation of the control object usually chooses from the set of trained neuro-emulators. However, is this criterion best if the neural network is used for further training another neural network, connected sequentially to the first, and not actually for modeling the control object?

The paper presents neurocontroller development for satellite rotation control.

2 State of arts

NN was proposed in 1943. McCullock and Pitts as the result of studying the structure and activity of biological neurons.

A typical structure of the automatic control system with the PID-regulator and the NN as an automatic adjustment unit is considered in the work [6]. NN acts as a functional transformation, which for each set of input signals the coefficients for the PID regulator are produced. The most complicated part of the design of an NN-based regulator is the training procedure, which reduces to the identification of unknown NN parameters, such as weighting factors and displacement of neurons. For NN training, the gradient search method uses a minimum criterion function, which depends on the parameters of the neurons. The search process is inertial, at each iteration, the search for all coefficients of the network occurs: first for the output layer, then for the previous and so on to the first.

The length of the learning process is a key issue when using NN methods for PID regulators [7]. In addition, when applying NN, there are difficulties due to the impossibility of predicting regulation errors for incoming actions that were not included in the set of training sequences by determining the structure of neurons in the network, the duration of training, the range and the number of training actions.

The main purpose of NN training is to choose the weighting factors of such a network to ensure consistency between input and output values. The neuron with the input $p = \{p_1, p_2, \dots, p_r\}$ is shown in Fig. 1. The initial value is equal to the scalar product of the vector W on the input vector p , the bias value b is added to the weighted sum of inputs [8].

Output signal is:

$$n = w_{11} \cdot p_1 + w_{12} \cdot p_2 + \dots + w_{1r} \cdot p_r + b \quad (1)$$

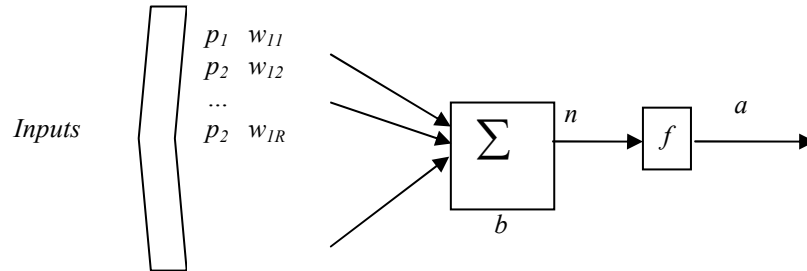


Fig. 1. Neuron structure

The choice of NN architecture is to determine the number of layers, the number of neurons in each of the layers, the form of the activation function of each layer, and information about the topological links of the neurons. Single-layer NNs are not suitable for solving complex problems [9], but combining several neurons into one or more layers has great potential. The two-layer NN, which in the first layer contains a

sigmoidal activation function, and in the second one linear, can be trained to approximate any function with finite number of breakpoints with arbitrary accuracy [9].

The purpose of identification is to determine the operator of the model, which converts the input action of the controlled object to the output value. Different identification methods are possible depending on the various forms of representation of mathematical models in the form of ordinary differential equations, difference equations, convolution equations [10], and others. However, none of the proposed methods is universal.

The paper [11] considers the use of NN as an alternative tool for the identification of dynamic objects. The use of NN is based on the fact that in practice modern electric drives are multi-mass systems with nonlinear links. Relevant linearized models built based on transfer functions, cannot always adequately reflect the state of the electric drive in all modes of its operation. The equivalence of a nonlinear system and its linear approximation will be equal in a limited time interval, and when transitioning the output system from one mode to another, it is expedient to use the linearization method and obtain a new linear system.

The paper [12] proposed the use of recurrent multi-layer N with external inputs NARX.

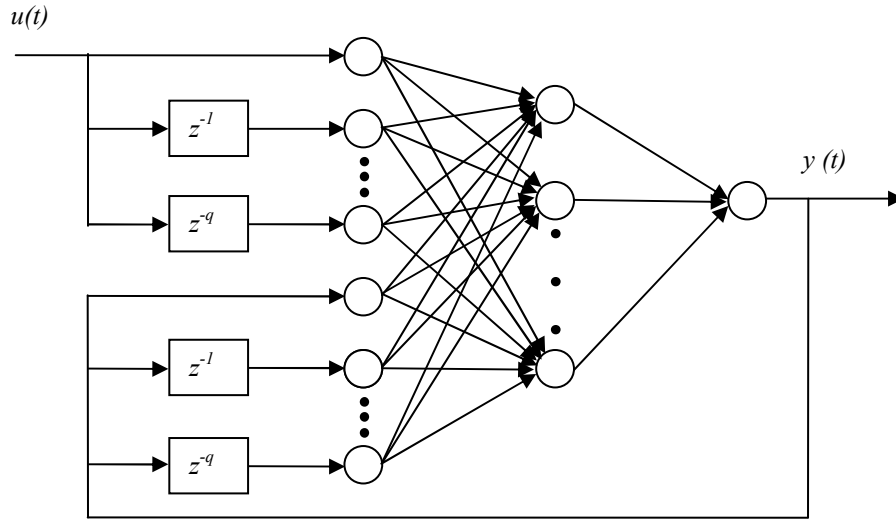


Fig. 2. The recurrent multilayer neural network NARX

The training model is given as

$$y(n+1) = f(y(n), \dots, y(n-q+1), u(n), \dots, u(n-q+1)), \quad (2)$$

where $y(n)$ is output vector, $u(n)$ is input vector, n is the discrete time moment, q is power of the system.

Such a NN, which has feedback with single delays, allows constructing on its basis a model of a dynamic object of arbitrary complexity. Using this method requires veri-

fication of trained NF for adequacy with the use of new data not included in the training sample. Such NN is associated with the possibility of re-training NN [12].

The Matlab [13] Neural Network Toolbox application suite contains the most popular neurocontrollers (NPCs) with

- Neural Predictive Control (NPC),
- Nonlinear Auto Regressive Moving Average (NARMA-L2) model,
- Model Reference Controller (MRC).

In [14], a mathematical description of predictive neuroization using MATLAB system tools is presented. In [15], the NARMA-L2 controller is used for automatic control of the vessel on a variable course. When solving the problem of guidance and stabilization of the armament of a light armored machine, the NARMA-L2 neuro regulator is used in the contour of speed. As the authors note, NARMA-L2 acts as a relay regulator, whose output is switched to opposite limits, resulting in significant fluctuations in speed (up to 40% of the maximum). However, these neuro-regulators are not connected with physical model of object.

The purpose of this work is to build model and neuro-controller to control small satellite with default amount of reaction wheels.

3 Materials and methods

The main tasks of the paper are to create:

1. Simple simulator of satellite rotations, controlled by 3 or 4 reaction wheels, placed in different configurations. The simulation model will be configurable and easy to read.
2. An Artificial Intelligence (AI) learning module which will trigger the simulator and learn autonomously from the behavior of the simulated satellite, how to control its rotations.
3. The AI module, after trained for different configurations of wheels, will get commands with desired 3D rotation speeds and control the wheels to achieve the desired rotation.

3.1 Satellite simulator design

Simulator is developed using C++ programming language.

The satellite simulator is created to solve the next tasks:

- To provide physical model of physical object;
- To provide physical model of satellite with reaction wheels for rotation control;
- To provide possibility to control satellite using reaction wheels during simulation.

Simulator is divided to the such layers of logical implementations:

- Core of simulation,

- Satellite simulation.

Core is a general simulation that grants us encapsulated logic for creating and moving of material object. It also allows us to configure simulation and to log information about all objects in simulation. Satellite simulation extends material object logic with reaction wheels and physical facts (friction, gravity, gyroscope effect etc.).

The class diagram is given in Fig. 3.

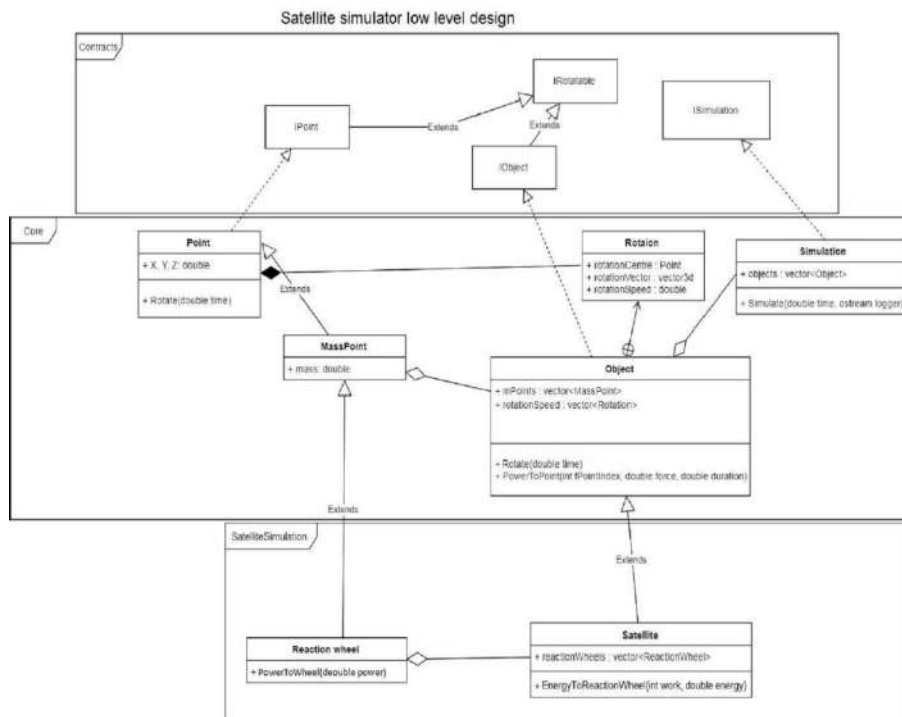


Fig. 3. Satellite simulator class diagram

Design layers:

- Contracts shows main entities of simulator and grants low coupling between their implementations. Contracts consists of abstractions;
- Core implements contracts. It contains primary physical model and Simulation entity.
- Satellite simulation extends Core with a dynamic of reaction wheel and satellite.

Entities:

- Point - provides an abstract point for further implementations;
- MassPoint - point which has mass and movement vector;
- Object - provides enumeration of points which interact with each other;

- ReactionWheel - inherited from MassPoint instances, is used for changing rotation speed of satellite by changing its angular momentum;
- Satellite - inherited from Object, provides simulated Satellite of arbitrary form, which moves and rotates using thrusters(ForcePoint) and reaction wheels;
- Simulator - provides enumeration of Object instances and configuration of scenario of their behavior.

The sphere in the Fig. 4 is a space, which limits the set of material points of the object. The center of mass is note center of the sphere, because its coordinates depends of coordinates and masses off other points.

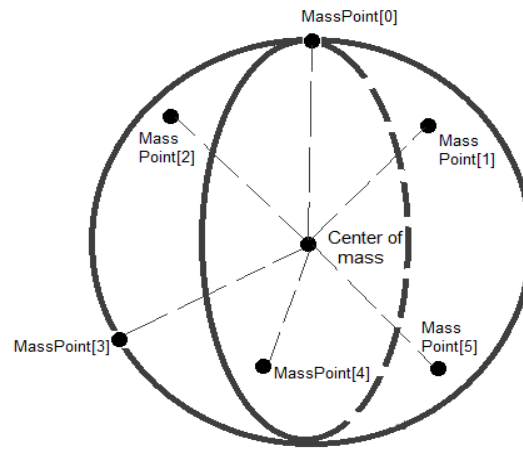


Fig. 4. The center of mass explanation

3.2 Neuro-controller design

The controller is created to solve the next tasks:

- Generate samples of satellite rotations.
- Train the Neural Network model, which can predict expected Energy on reaction wheel.
- Set the energy on reaction wheel based on training result.

During of training, the neural network must monitor and remember the dependence of the control signal $u(k-1)$ on the next value of the reaction of the control object that was before in the state $X(k-1)$. The values of the control signals and responses of the object are recorded and, on this basis, a training sample is formed.

$$U = \{P_i, T_i\}_{i=1}^M : P_i = [y(i)X(i-1)]^T, T_i = u(i) \quad (3)$$

We used and desired reaction.

In the training mode neural network must find and remember the dependence of control signal $u(k-1)$, in state before $S(k-1)$. When the object is controlled, the inverse neuro-emulator is connected as a controller and it is receiving the $rr(k)$ value from input $r(k+1)$:

$$rr(k) = [r(k+1)X(k)]^T. \quad (4)$$

The class diagram is given in Fig. 5.

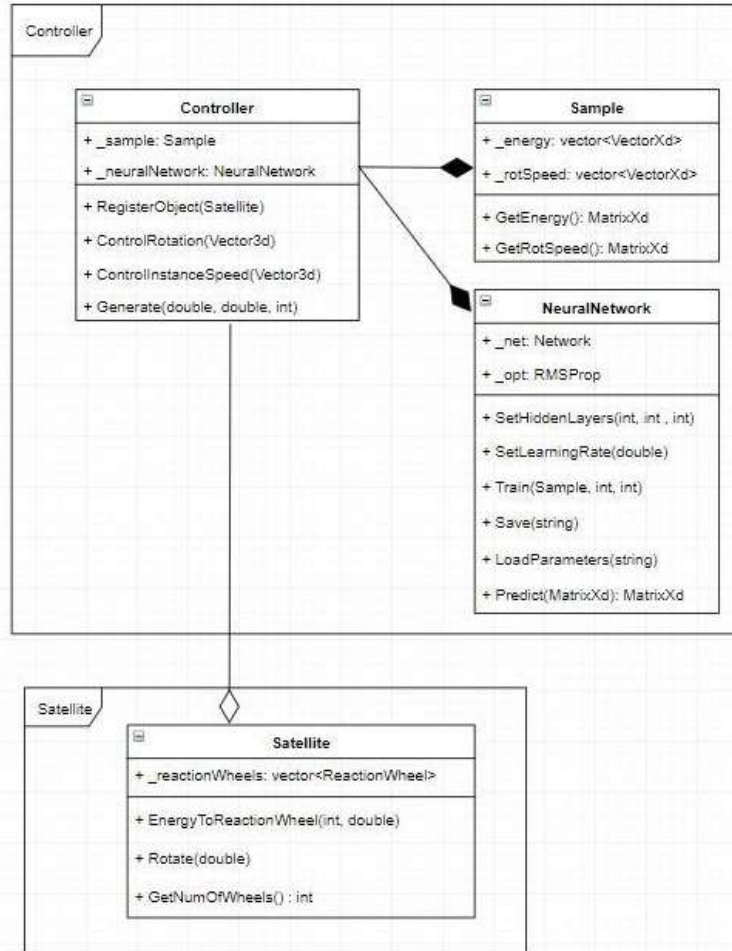


Fig. 5. Class controller diagram

Inputs in the control network is the satellite state (speed for each axes). The output is the control signal (torque) $u(t)$. This is energy level for each rotation wheel.

We used mini-batch gradient descent algorithm for neural network training.

The structure of neural network

The neural network structure for this task looks like this:

- Input layer – 3 neurons (for speed by x,y,z),
- Hidden layer – 15 full-connected neurons with sigmoid activation function,
- Output layer – n neuron with predicted energy level, where n is equal amount of rotation wheels,
- The bias is used too.

The architecture of neuro-controller is chosen experimentally and given in Fig. 6.

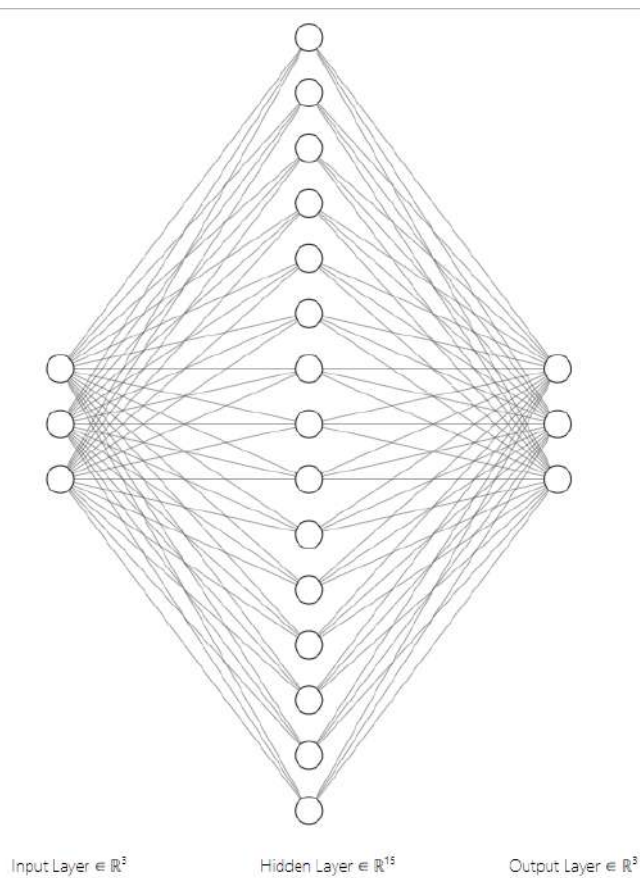


Fig. 6. Neural network architecture

We used mini-batch gradient descent in NN.

The goal of the algorithm is to find model parameters (e.g. coefficients or weights) that minimize the error of the model on the training dataset. It does this by making changes to the model that move it along a gradient or slope of errors down toward a minimum error value. This gives the algorithm its name of “gradient descent.”

Mini-batch gradient descent is a trade-off between stochastic gradient descent and batch gradient descent. In mini-batch gradient descent, the cost function (and therefore gradient) is averaged over a small number of samples, from around 10-500. This is opposed to the SGD batch size of 1 sample, and the BGD size of all the training samples.

Mini-batch gradient descent finally takes the best of both worlds and performs an update for every mini-batch of n training examples:

$$\theta = \theta - \eta \cdot \nabla \theta J(\theta; x(i:i+n); y(i:i+n)). \quad (5)$$

This allows us

- to reduce the variance of the parameter updates, which can lead to more stable convergence;
- can make use of highly optimized matrix optimizations common to state-of-the-art deep learning libraries that make computing the gradient w.r.t. a mini-batch very efficient. Common mini-batch sizes range between 50 and 256, but can vary for different applications.

4 Results

4.1 Stack of technologies

For neuro-controller realization we used

1. **Eigen** to provide vectors, matrixes, quaternions of different dimensions and working with them (it was mostly used in simulator) [16].
2. **MiniDnn** to provide neural network for creating controller of a satellite.

Parameters of NN is saved in NeuralConfig.h. These neural network parameters were chosen experimentally. We provided more than 500 training experiments with different neural network configuration. In the best attempts mean loss was be equal 0.013 and parameters there was:

NUMBEROFSAMPLES	1000
NUMBEROFHIDDENLAYERS	1
HIDDENLAYERSLENGTH	15
LEARNINGRATE	0.0007
BATCHSIZE	200
EPOCH	40000

Table 1. Training results

#	BatchSize	Epoch	LearningRate	Loss function
1	20	40000	0.0002	0.11
2	200	4000	0.0007	0.013
3	200	40000	0.002	0.008
4	200	40000	0.005	0.05
4	200	40000	0.001	0.07

Training time is appr 7 hours and 20 min. The computer configuration is given below:

Intel Core i3 (3,4 Ghz), 2 cores, NVidia GeForce, GT630, 2Gb

5 Conclusions

To sum up this article described how we could use neural networks for controlling satellites. Neural controllers is a very powerful method that allows us automate different processes and improve accuracy of its results.

An experimental study of the proposed criterion of 500 neuro-controllers was conducted, which showed its effectiveness compared to the traditional method (Loss function value is less than 0.05) of selecting neurotransmitters based on the least square root error method on the test data voter.

In the framework of further research, it is planned to test this criterion, along with other methods of neuro-control, which include the stage of preliminary neuro-identification of the control object: predictive model neuro management and hybrid neuro-PID control as well as using the Kalman cube filter.

6 References

1. Narendra, K. S., & Parthasarathy, K.: Identification and control of dynamical systems using neural networks. *IEEE Transactions on neural networks*, 1(1), 4-27 (1990)
2. Prokhorov, D. V., & Wunsch, D. C.: Adaptive critic designs. *IEEE transactions on Neural Networks*, 8(5), 997-1007 (1997)

3. Feldkamp, L. A., & Puskorius, G. V.: Training controllers for robustness: multi-stream DEKF. In *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)* 4, 2377-2382 (1994)
4. Prokhorov, D. V.: Toyota Prius HEV neurocontrol and diagnostics. *Neural Networks*, 21(2-3), 458-465 (2008)
5. Haykin, S. S., Haykin, S. S., Haykin, S. S., Elektroingenieur, K., & Haykin, S. S.: *Neural networks and learning machines* (Vol. 3). Upper Saddle River: Pearson. (2009)
6. Kawafuku, R., Sasaki, M., & Kato, S.: Self-tuning PID control of a flexible micro-actuator using neural networks. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)* Vol. 3, 3067-3072 (1998).
7. Burakov, M. V., & Kurbanov, V. G.: Neuro-PID control for nonlinear plants with variable parameters. *ARPJ Journal of Engineering and Applied Sciences*, 12(4), 1226-1229 (2017).
8. Anderson, D. F., Ermentrout, B., & Thomas, P. J.: Stochastic representations of ion channel kinetics and exact stochastic simulation of neuronal dynamics. *Journal of computational neuroscience*, 38(1), 67-82 (2015)
9. Zhernova, P. Y., Deineko, A. O., Bodyanskiy, Y. V., & Riepin, V. O.: Adaptive Kernel Data Streams Clustering Based on Neural Networks Ensembles in Conditions of Uncertainty About Amount and Shapes of Clusters. In *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)* 7-12 (2018)
10. Bodyanskiy, Y., Boiko, O., Zaychenko, Y., & Hamidov, G.: Evolving GMDH-neuro-fuzzy system with small number of tuning parameters. In *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* 1321-1326 (2017)
11. Ramachandran, R., Madasamy, B., Veerasamy, V., & Saravanan, L.: Load frequency control of a dynamic interconnected power system using generalised Hopfield neural network based self-adaptive PID controller. *IET Generation, Transmission & Distribution*, 12(21), 5713-5722 (2018)
12. Lee, C. C., Sheridan, S. C., Barnes, B. B., Hu, C., Pirhalla, D. E., Ransibrahmanakul, V., & Shein, K.: The development of a non-linear autoregressive model with exogenous input (NARX) to model climate-water clarity relationships: reconstructing a historical water clarity index for the coastal waters of the southeastern USA. *Theoretical and Applied Climatology*, 130(1-2), 557-569 (2017)
13. Medvedev V. S., Potjomkin V. G.: *Nejronnye seti. MATLAB 6*. Moscow, Dialog-MIFI, 496 p (2002)
14. Hwang, C. L., & Jan, C.: Recurrent-neural-network-based multivariable adaptive control for a class of nonlinear dynamic systems with time-varying delay. *IEEE transactions on neural networks and learning systems*, 27(2), 388-401 (2016)
15. Yang, Y., Xiang, C., Gao, S., & Lee, T. H.: Data-driven identification and control of nonlinear systems using multiple NARMA-L2 models. *International Journal of Robust and Nonlinear Control*, 28(12), 3806-3833 (2018)
16. Pukach, P., Il'kiv, V., Nytrebych, Z., Vovk, M., Shakhovska, N., & Pukach, P.: Galerkin Method and Qualitative Approach for the Investigation and Numerical Analysis of Some Dissipative Nonlinear Physical Systems. In *2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)* Vol. 1., 143-146 (2018)

Information-Extreme Learning of On-Board System for Recognition of Ground Vehicles

Igor Naumenko¹[0000-0003-2845-9246], Mykyta Myronenko²[0000-0001-5005-1672],
Vladyslav Piatachenko³[0000-0002-7464-3119]

¹ Scientific-research center of missile troops and artillery, Gerasim Kondratyev st. 165, Sumy, 40021, Ukraine, 790895@ukr.net

² Sumy State University, Rymskogo-Korsakova st. 2, Sumy, 40007 Ukraine, nikitam1996@ukr.net

³ Sumy State University, Rymskogo-Korsakova st. 2, Sumy, 40007 Ukraine, bronzeghoste@gmail.com

Abstract. The article deals with the task of information synthesis of on-board recognition system of ground vehicles. Machine training of the system is carried out within the framework of information-extreme intellectual technology of data analysis, which is based on maximizing the information ability of the recognition system. The modified Kulbak information measure is used as a criterion for optimizing the parameters of machine learning. The proposed algorithm of machine learning is realized on the example of recognition of monochrome coloring cars on approximately identical chassis.

Keywords. information-extreme machine learning, on-board recognition system, categorical model, training matrix, optimization, information criterion, vehicles.

1 Introduction

The widespread use of unmanned aerial vehicles to monitor the earth's surface makes it urgent to create autonomous on-board recognition systems (ORSs) for ground objects, including vehicles of various uses [1, 2]. The main way of solving this problem is to apply methods of machine learning and the theory of pattern recognition [3 – 5]. In this case, the functional efficiency of the ORS essentially depends on the method of processing images of the object being recognized and the method of machine learning. In addition, when creating the ORS it is necessary to ensure the invariance of the constructed in the process of machine learning decisive rules to the arbitrary position of the vehicle in the frame of the area of interest. One of the ways to solve this problem is to process images of recognition objects in a polar coordinate

system [6, 7]. But within the framework of this approach necessarily there is a need to solve the problem of determining the polar coordinate system on the ground vehicle. The overwhelming majority of the methods of the information synthesis of the ORS, which are taught, is based on the application of neural networks [8, 9]. In this case, there are complications of scientific and methodological nature, associated with arbitrary initial conditions for the formation of the input mathematical description, the intersection in the space of signs of recognition classes, a large measure of the space of signs and the complexity of retraining.

One of the promising ways of information synthesis of highly effective ORS is the application of ideas and methods of the so-called information-extreme intellectual technology (IEI-technology) of data analysis, which is based on maximizing the information ability of the recognition system in the process of its machine learning [10 – 12]. The main idea of the methods in the framework of IEI-technology as in neural networks is to adapt the input mathematical description of the recognition system to the maximum functional efficiency of machine learning. But in contrast to the neural networks built on the results of machine learning in the framework of the geometric approach, the decisive rules are practically invariant to many of the dimensionality of the recognition signspace.

The article discusses, in the framework of the IEI-technology, the formulation of the problem and the algorithms for the operation of the onboard vehicle recognition system in machine learning modes of the exam.

2 Formulation of the problem

Consider the formalized formulation of the task of information-extreme machine learning on-board recognition of a land vehicle. Let the alphabet $\{X_m^o \mid m = \overline{1, M}\}$ of the recognition classes characterizing the terrestrial vehicles and the input brightness $\|y_{m,i}^{(j)}\|$ of the pixels of the receptor field of the image objects of the recognized objects be given. In this case, the line $\{y_{m,i}^{(j)} \mid i = \overline{1, N}\}$ of the matrix, where N is the number of signs of recognition, is a structured vector-realization (hereinafter simply realization) of the image, and the matrix column is a random educational sample of $\{y_{m,i}^{(j)} \mid j = \overline{1, J}\}$ with a volume of J .

In the process of machine learning it is necessary:

in accordance with the concept of IEI-technology, convert the input training matrix into a binary working matrix $\|x_{m,i}^{(j)}\|$, which, through admissible transformations, can be adapted to the maximum full probability of making the correct classification decisions;

optimize according to the information criterion the parameters of the machine learning ORS, which for each class of recognition X_m^o are given by structured vector

$$g_m = \langle x_m, d_m, \delta \rangle, \quad (1)$$

where x_m is the averaged binary implementation of the recognition class X_m^o ; d_m is the radius of the hyperspherical container of the recognition class X_m^o , which is restored in the radial basis of the space of signs; δ – parameter, the value of which is equal to half of the symmetrical field of control tolerances on the signs of recognition; for determining the optimal geometric parameters of the classes of recognition of hyperspherical containers, determined by the machine learning process, to construct decisive rules.

At the same time, the following restrictions are set for the machine learning parameters:

- the range of values of the radius of the hyperspherical container of the recognition class d_m , is given by the inequality $d_m < d(x_m \oplus x_c)$, where $d(x_m \oplus x_c)$ is the inter-center coding distance between the reference implementation of class x_m of class X_m^o and the reference implementation of x_c neighboring class X_c^o nearest to it; $\oplus$ – symbol of the logical operation of adding by module 2;
- the region of values of parameter δ is given by inequality $\delta < \delta_H / 2$, where δ_H is the normalized tolerance field on the recognition signs.

In the functioning of the on-board system in the exam mode, it is necessary to confirm the high functional efficiency of the machine learning onboard recognition system.

3 Categorical models of machine learning

The categorical model of on-board computer training includes an input mathematical description of the on-board vehicle recognition system in terrain, which has the form

$$\Delta_B = \langle T, G, \Omega, Z, K, Y, X; \Phi_1, \Phi_2 \rangle,$$

where T is the set of moments of the time of obtaining information; G is the space of the functional states of the recognized object; Ω – the space of signs of recognition; Z – space of functional states of the recognition system; K – the set of frames of the electronic map of the area; Y is a sample set that forms the input training matrix; X – working binary training matrix; $\Phi_1 : G \times T \times \Omega \times Z \rightarrow Y$ – the operator of the formation of matrix Y ; $\Phi_2 : Y \rightarrow X$ – s an operator for transforming the input training matrix Y into a binary matrix X .

Fig. 1 shows a categorical model of information-extreme learning of recognition system with optimization of geometric parameters of containers of recognition classes and system of control tolerances on recognition signs.

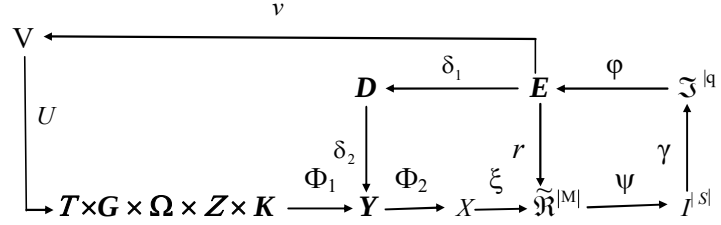


Fig. 1. Categorical model of machine learning

In Fig. 1 operator $\xi: X \rightarrow \tilde{\mathfrak{R}}^{|M|}$ displays the vectors of the implementation of recognition classes on the fuzzy partition of the $\tilde{\mathfrak{R}}^{|M|}$ binary traits of spaces, and the classification operator Ψ checks the basic statistical hypothesis of the validity of the class X_m^o mplementation and thus forms the set of hypotheses $I^{|l|}$, where l – is the number of statistical hypotheses. The operator γ , by evaluating the hypotheses received, forms the set of exact characteristics $\mathfrak{Z}^{|q|}$, where $q = l^2$, and operator ϕ calculates the set of values of the information criterion E , which is a function of the exact characteristics. The contour of the model, which is closed by the operator r , restores, at each step of the machine learning, the recognition class containers that are built in the radial basis of the feature space. In this case, the iterative process of optimizing the geometric parameters of the partition $\tilde{\mathfrak{R}}^{|M|}$ is carried out by finding the global maximum of the information criterion in the working (admissible) region of its function definition. In fig. 1 contour of optimization of control tolerances for recognition signs is closed through set D – the system of control tolerances for recognition signs and allows in the training process to change the value of the working binary training matrix X , adapting it to the maximum functional efficiency of the classifier. Shown in Fig. 1 categorical model implies, in accordance with the principle of deferred decisions, the transition to other types of radial-basic decision rules. For this purpose, its outer contour contains a set of V types of decisive rules, which are built using more complex radial-basic separation functions. The training process is regulated by the operator $U: V \rightarrow G \times T \times \Omega \times Z \times K$.

Fig. 2 shows a categorical model of the functioning of the on-board recognition system in the mode of examination, which tests the functional efficiency of machine-based learning.

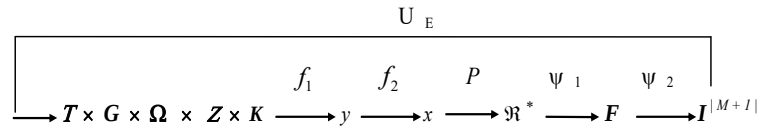


Fig. 2. Category model of functioning of on-board recognition system in the exam mode

In the categorical model (Fig. 2), the operator f_1 forms an incoming implementation of the recognizable object. Operator f_2 , obtained on the learning stage, obtains optimal control tolerances for recognition attributes, converts incoming readmissions into binary realization x , and operator ξ reflects the implementation of the recognizable object on the optimal breakdown of $\mathfrak{R}^*$ recognition classes built at the stage of machine learning. The operator ψ_1 for each vector-realization calculates the values of constructed at the learning stage of the deciding rules and forms the term-set F , and the operator ψ_2 by the maximum value of the deciding rule, assigns implementation to one of the classes of a given alphabet $\{X_m^o\}$. The purpose of the operator U_E is the regulation of the exam.

4 Machine learning

The idea of information-extreme machine learning of the pattern recognition system, as in neural networks in accordance with the work, is to adapt the input mathematical description to the maximum reliability of classification decisions. In this case, the transformation of the input a priori fuzzy distribution of image implementations in the clear is carried out in the process of optimization according to the information criterion of the learning parameters that affect the exact characteristics of the classification decisions. Based on the results obtained in the process of machine learning, the optimal geometric parameters of containers of recognition classes are based on decisive rules that allow the exam to take promptly reliable classification decisions.

Optimization of the parameters of the vector (1) is carried out by searching at each step the purposeful change of the radiuses of containers of the classes of recognition of the global maximum of the alphabet of the classes of recognition of the information criterion

$$\bar{E}^* = \frac{1}{M} \sum_{m=1}^M \max_{G_E \cap G_d} E_m(d), \quad (2)$$

where $E_m(d)$ – information criterion for optimizing the parameters of the training system to recognize the implementation of class X_m^o ; d – distance measure of radiuses of hyperspherical containers of recognition classes; G_E – working (admissible) area of determining the function of information criterion optimization of machine learning parameters; G_d is the permissible range for changing the radiuses of the recognition class containers.

As criteria of optimization of machine learning parameters in the methods of IEI-technology, modifications of the Kullback information measure and Shannon's entropy measure are mainly used. For example, the modified Kullback information measure for two alternatives to the a priori equivalence hypothesis has the form

$$E_m(d) = [1 - (\alpha_m(d) + \beta_m(d))] \log_2 \left(\frac{2 - (\alpha_m(d) + \beta_m(d))}{\alpha_m(d) + \beta_m(d)} \right), \quad (3)$$

where $\alpha_m(d)$ is a mistake of the first kind of acceptance of a classification decision at every step of machine learning; $\beta_m(d)$ is a mistake of the second kind.

On practice, when calculating the information criterion (3), when representing the amount of the training sample, it is necessary to use the estimates of the exact characteristics

$$\alpha_m^{(k)}(d) = \frac{K_{1,m}(d)}{n_{\min}}; \quad \beta_m^{(k)}(d) = \frac{K_{2,m}(d)}{n_{\min}}, \quad (4)$$

where $K_{1,m}(d)$ is the number of events that indicate the inappropriateness of "their" implementations of the recognition class X_m^o ; $K_{2,m}(d)$ – the number of events that indicate the belonging of "alien" implementations of class X_m^o , $n_{\min}$ – the minimum amount of representative sample of study.

The working modification of the Kullback criterion after the corresponding substitution of the estimates (4) in expression (3) takes the form

$$E_m^{(k)}(d) = \frac{1}{n} \{n - [K_1(d) + K_2(d)]\} \log_2 \left\{ \frac{2n + 10^{-r} - [K_1^{(k)}(d) + K_2^{(k)}(d)]}{[K_1^{(k)}(d) + K_2^{(k)}(d)] + 10^{-r}} \right\}, \quad (5)$$

where 10^{-r} is a sufficiently small number which is entered to avoid division into zero and in practice it is chosen in the interval $1 < r \leq 3$.

According to the categorical model (Fig. 1), the algorithm of ORS informational and emergency machine learning with optimization of the system of control tolerances can be represented as a two-cycle procedure for finding the global maximum of the information criterion (2)

$$\delta_K^* = \arg \{ \max_{G_\delta} \{ \max_{G_E \cap G_d} \bar{E}(d) \} \}, \quad (6)$$

Let's consider the main stages of implementing the algorithm (5) of information-extreme machine learning ORS. The input data is the array of the input learning matrix for a given alphabet of recognition classes and parameter:

1. zeroing the count of the recognition classes: $m := 0$;
2. increment of the count of recognition classes: $m := m + 1$;
3. zeroing the counter of the steps of changing the tolerance field parameter: $k := 0$;
4. $k := k + 1$;
5. zeroing the counter of steps to change the radius of the container of the recognition class: $d := 0$;
6. $d := d + 1$;
7. calculation of the lower $A_{KH,i}[k]$ and upper $A_{KB,i}[k]$ control tolerances for all signs, respectively, according to the formulas

$$A_{KH,i}[k] = y_{1,i} - \delta[k]; \quad A_{KB,i}[k] = y_{1,i} + \delta[k]; \quad (7)$$

8. formation of the binary training matrix $\|x_{m,i}^{(j)}\|$ by the rule

$$x_{m,i}^{(j)} = \begin{cases} 1, & \text{if } A_{KH,i} \leq y_{m,i}^{(j)} \leq A_{KB,i}; \\ 0, & \text{if else.} \end{cases}$$

9. calculation for class X_m^o binary averaging vector x_m according to the rule

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} > \rho_m; \\ 0, & \text{if else,} \end{cases}$$

where ρ_m is the level of selection of the coordinates of the averaged binary vector of the recognition class X_m^o , which by default is equal to $\rho_m = 0,5$;

10. pairwise decomposition of the set of averaged vectors of recognition classes by the method of closest neighbors;
11. formation for the breakdown of $\mathfrak{R}_m^{[2]} = \langle x_m, x_c \rangle$ educational matrix;
12. calculation of the information criterion for optimizing the parameters of machine learning GIS, for example, in the form of modification of the information measure of Kullback (2.5);
13. if $d < d(x_m \oplus x_c)$, then paragraph 6 is executed, otherwise, paragraph 14;
14. if $k < \delta_H / 2$, then paragraph 4, is executed, otherwise, paragraph 15;
15. if $m \leq M$, then paragraph 2 is executed, otherwise, paragraph 16;
16. calculation of averaged alphabet classes recognition information criterion (6) optimization of machine learning parameters;
17. determining the optimal value of parameter δ by the formula (5);
18. calculation of the formula (7) of the optimal lower and upper control tolerances for the diagnostic signs, respectively;
19. STOP.

Based on the results obtained in the process of machine learning, the optimal geometric parameters of the classes of recognition containers are based on decisive rules, which, when the recognition system operates directly in the operating mode, verifies the functional efficiency of machine learning. For hyperspherical containers of recognition classes, decisive rules have the form

$$(\forall X_m^o \in \mathfrak{R}^{[M]})(x^{(j)} \in \mathfrak{R}^{[M]})[\text{if } (\mu_m > 0) \& (\mu_m = \max\{\mu_m\}) \text{ then } x^{(j)} \in X_m^o], \quad (8)$$

where $x^{(j)}$ is an implementation that is recognized; the function of ownership of the container of the recognition class X_m^o .

In expression (8), the membership function for hyperspherical containers is determined by the formula [10]

$$\mu_m = 1 - \frac{d(x^{(j)} \oplus x_m)}{d_m^*}; \quad (9)$$

where d_m^* – obtained in the process of machine learning the optimal radius of the container of the recognition class X_m^o .

The algorithm of functioning of the ORS in the mode of examination has the following input data:

1. $\{x_m^* \mid m = \overline{1, M}\}$ is an array of reference binary vector-image implementations that determine the geometric centers of the corresponding optimal containers of recognition classes constructed at the stage of machine learning;
2. $\{d_m^*\}$ is an array of optimal radiuses constructed at the stage of training of the corresponding containers;
3. $\{x_s^{(j)} \mid s = \overline{1, SMAX}; j = \overline{1, n}\}$ is an array of binary vectors-implementations of identified frames, where $SMAX$ is the number of frames of a reconstructed terrain;
4. $\{\delta_{k,i}^* \mid i = \overline{1, N}\}$ is the optimal system of control tolerances for recognition signs, determined at the stage of training.

According to the categorical model (Fig. 2), the algorithm of the exam within the framework of the IEI-technology is based on the analysis of the values of the decisive rules formed at the stage of learning (8). If for all classes of recognition the maximum values of the function (9) are negative, then according to the deciding rules (8) the object is not identified;

The analysis of decisive rules (8) shows that they differ from other methods for recognizing low computational complexity, which allows BSC to take classification decisions in real time.

5 Results of physical modeling

Approbation of the proposed algorithm for information-extreme machine learning was carried out on the example of the recognition of three cars that were moving along the highway (Fig. 3). At the same time, in order to check the functional efficiency of machine learning ORS cars specially selected monotonous and with approximately the same contours.

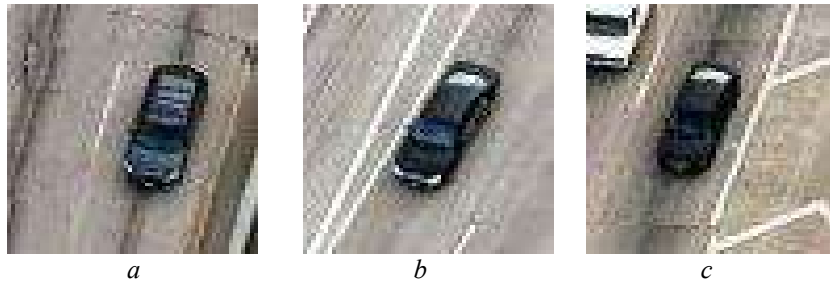


Fig. 3. Picture of cars of class X_1^o (a); class X_2^o (b); class X_3^o (c)

The formation of the input training matrix was carried out by processing images of cars in the polar coordinate system, which allowed to ensure the invariance of the decision of the rules to the arbitrary position of the object of recognition in the frame of the zone of interest. As a sign of recognition, the average value of the brightness of

the pixels of the reading range, built around the center of the polar coordinate system, was taken. At the same time, the definition of the center of the polar coordinate system on the car was carried out by the class SelectedObject, which, moreover, handles the image of the object in the polar coordinate system and forms the input training matrix ORS.

Fig. 4 shows a screenshot of the result of the Class SelectedObject program, which shows the center of the polar system on a class X_3^o (Fig. 3c) at different levels of quantization of the brightness of pixels of the area of interest frame.

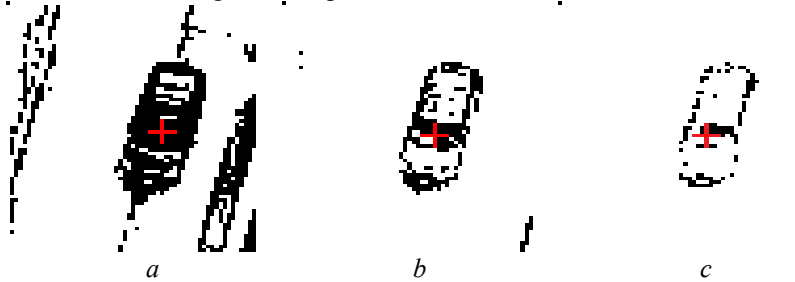


Fig. 4. Results of the centering of the car at the levels of quantization of brightness (in brightness gradations): $a - \eta = 25$; $b - \eta = 50$; $c - \eta = 75$

To form the implementation of the input training matrix, all pixels of the frame of the interest zone, which was accepted as the first quadrant of the Cartesian coordinate system, was numbered. This allowed us to determine polarization centers of cars as average arithmetic numbers of pixels whose brightness exceeded the corresponding quantization level. Then the center of the Cartesian coordinate system was transferred to the found center of the polar system, around which the area of the given radius was asked. In the given region, the coordinates of the pixels were converted to polar and formed arrays of pixels with the same radiuses. For each RGB-component image of the car for recognition, the average brightness value of the pixels of the corresponding array was taken.

Fig. 5 shows the graph of the dependence of the information criterion recognition (5) alpha of the information criterion recognition classes (5) on parameter δ , obtained during the machine learning process in accordance with the procedure (6) at the quantization level of brightness $\eta = 50$. In this case, parallel optimization was carried out, at which, at each step of machine learning, control tolerances changed simultaneously for all signs of recognition. When calculating the information criterion (5) parameters $n = 40$ and $r = 2$ were taken.

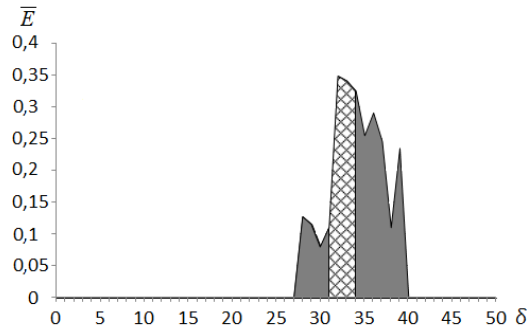


Fig. 5. The graph of the dependence of the information criterion (5) on the parameter of the field of control tolerances on the signs of recognition

Fig. 5 shows and then the double hatching indicates the working (admissible) area of definition of function (5), in which the value of errors of the first and second kind is less than the first and second reliability, respectively. Analysis of Fig. 5 shows that the maximum value of the informational criterion (5) in the work area for determining its function is equal to $\bar{E}^* = 0,35$. In this case, the optimal value of the parameter field of control tolerances on the recognition signs is equal to $\delta^* = 32$ brightness gradations.

To construct the decisive rules (7), it is necessary to know the optimal geometric parameters of the classes of recognition containers. Fig. 6 shows the graphs of the information criterion (5) dependence on the radiuses of the recognition class containers, obtained in the process of optimizing the control tolerances for recognition attributes.

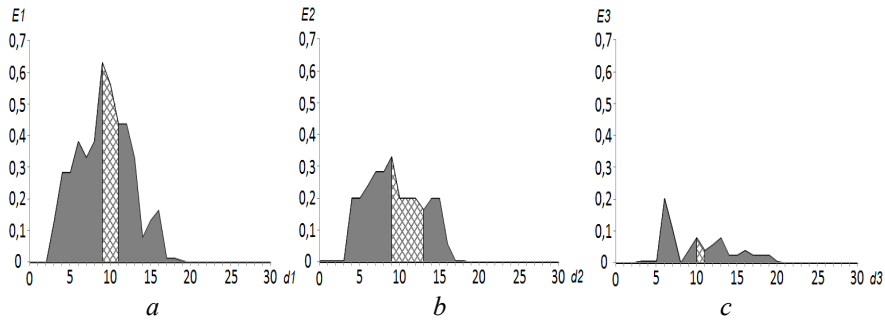


Fig. 6. Graphs of the dependence of the information criterion on the radiuses of containers of recognition classes: a – class X_1^o ; b – class X_2 ; c – class X_3^o

Analysis of Fig 6 shows that the optimal value of the radius of the container of the recognition class X_1^o is $d_1^* = 10$ (here and below in the code units), for class X_2 – $d_2^* = 9$ and for class X_3^o – $d_3^* = 11$.

The results of the physical simulation of the on-board recognition system in the exam mode showed that the full probability of correct recognition of the vector-realization of class X_1 is $P_t = 0,84$, class $X_2 - P_t = 0,80$ and class $X_3 - P_t = 0,78$.

6 Conclusion

The method of information-extreme machine learning of the ORS ground vehicle with optimization of the system of control tolerances on recognition signs is proposed. The formation of the input mathematical description was carried out on the basis of the results of processing images of vehicles in the polar coordinate system, which ensured the invariance of the deciding rules to the arbitrary position of the object of recognition in the frame of the interest zone. The results of machine learning were not allowed to construct non-error-based educational matrix deciding rules due to the high degree of intersection of classes of recognition in the space of signs. Therefore, in order to increase the functional efficiency of the ORS, it is necessary to increase the depth of machine learning by optimizing other learning parameters, including the parameters of image processing of terrestrial objects.

References

1. Jauregi, E., Lazkano, E., Martinez-Otzeta, J.M., Sierra, B.: Visual Approaches for Handle Recognition. Springer Tracts in Advanced Robotics. vol. 44. pp.313-322 (2008) DOI:10.1007/978-3-540-78317-6_32.
2. Huang, H., Lu, L., Yan, B., Chen, J.: A new scale invariant feature detector and modified SURF descriptor. Conference: Sixth International Conference on Natural Computation, ICNC 2010, Yantai, Shandong, China, 10-12 August 2010, pp. 3734-3738 (2010) DOI:10.1109/ICNC.2010.5583377.
3. Hastie, T.: The Elements of Statistical Learning: data Mining, Inference and Prediction. – 2nd ed. Springer Verlag, 2009, p. 746 (2009)
4. Muller, K. R., Mika, S., Ratsch, G.: An introduction to kernelbased learning algorithms, IEEE Transactions on Neural networks. vol. 12(2), pp. 181 – 202 (2001)
5. Subbotin, S.: The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition. Optical Memory and Neural Networks (Information Optics). vol. 22(2), pp. 97-103 (2013) DOI=10.3103/S1060992X13020082
6. Dovbysh, A.S., Martynenko, S. S., Kovalenko, A. S., Budnyk, M. M. Information-extreme algorithm for recognizing current distribution maps in magnetocardiography. Journal of Automation and Information Sciences, vol. 43(2), pp. 63-70 (2011) DOI:10.1615/JAutomatInfScien.v43.i2.60
7. Simonovskiy, J., Piatachenko, V., Myronenko, M. On-board Geographic Information System of Images' Identification. In Proceedings of the VI International Conference «Advanced Information Systems and Technologies – AIST-2018», Sumy, May 16 – 18, 2018, Sumy, Sumy State University, pp. 115 – 118 (2018)
8. Zheng, H.L., Fu, J.L., Mei, T., Luo, J.B. Learning Multi-Attention Convolutional Neural Network for Fine-Grained Image Recognition. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 5219–5227 (2017)

9. Ciresan, D., Meier, U. and Schmidhuber, J., Multi-column deep neural networks for image classification, *Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference on, pp. 3642-3649 (2012)
10. Dovbysh, A.S., Rudenko, M. S. Information-extreme learning algorithm for a system of recognition of morphological images in diagnosing oncological pathologies. *Cybernetics and Systems Analysis*, vol. 50, pp. 157-163 (2014) DOI:10.15587/1729-4061.2016.71930.
11. Dovbysh, A. S., Moskalenko, V. V., Rizhova, A. S. Information-Extreme Method for Classification of Observations with Categorical Attributes. *Cybernetics and Systems Analysis*, vol. 52(2), 224–231 (2016) DOI:10.1007/s10559-016-9818-1
12. Dovbysh, A., Velikodniy, D., Shelehov, I. Bibyk, M. Information-extreme machine learning of the control system over the power unit of a thermal power main line. *Eastern-European Journal of Enterprise Technologies*, vol. 5/4 (89), pp. 17-24 (2017) DOI:10.15587/1729-4061.2017.112121.

A Joint Application of Fuzzy Logic Approximation and a Deep Learning Neural Network to Build Fish Concentration Maps Based on Sonar Data

Dmitry Glukhov¹ [0000-0003-4983-2919], Rykhard Bohush¹ [0000-0002-6609-5810],
Juho Mäkiö² [0000-0001-9987-7600] and Tatjana Hlukhava¹

¹Polotsk State University, Blokhina st., 29, Novopolotsk, Republic of Belarus, 211440
{d.gluhov, r.bogush, t.gluhova}@psu.by

²University of Applied Sciences Emden/Leer, Constantiaplatz 4, Emden, Germany, D-26723
juho.maekioe@hs-emden-leer.de

Abstract. This paper proposes an effective method for obtain topographic lake map with fish concentration based on the results of an intelligent sonar data processing. Fuzzy logic special implementation for approximation of sonar data is used. The mathematics apparatus of fuzzy logic provides the possibility of flexible adjustment approximator under conditions of problem to be solved when working with data of high dimensionality. An algorithm for obtaining fish concentration maps based on the results of intelligent processing of the sonar data is also proposed. The algorithm is based on the following steps: input frame separation into overlapping blocks, blocks-processing using convolutional neural networks YOLO v2, and merging extracted bounding boxes around one object. Experimental results for fish detection and fish concentrations map building are presented.

Keywords: sonar data; fish concentration; maps of lakes; fuzzy logic; convolutional neural networks

1 Introduction

Modern tools for detecting underwater objects with application of ultrasound (sonars) have become widespread in solving various applied problems. A highly specialized class of sonars designed to study the relief of the lake bottom and search for fish are called echo sounders.

Currently, there is a wide range of sonars from different suppliers. The most famous sonars are produced by Lowrance, Raymarine and Humminbird. Moreover, most modern sonars have a GPS module. Such devices are called chartplotters. Chartplotters fix echogram data to X, Y coordinates of Mercator projection (WGS-84/UTM coordinate system). The peculiarity of echograms is that GPS data is updated much less frequently than ultrasonic sounding data. So, each individual act of acoustic sounding cannot be geographically fix.

The modern echograms formats contain the Mercator projection coordinates, changing stepwise after GPS data update. The most common echograms formats are SLG and SL2 developed by the Lowrance company. The SL2 format is used for a multi-beam sonar equipped with a DownScan bottom scan function and a Structure-Scan structural scan (455kHz or 800kHz beam) and the ability to probe Primary and Secondary beams (at frequencies 83kHz and 200kHz) simultaneously.

Now there are several software packages designed for processing sonar data: ReefMaster by ReefMaster Software Ltd., DrDepth (currently this project was purchased by Humminbird and on its basis created the program AutoCharts), Surfer, ArcGis, GlobalMapper and others. In addition to the high cost, the most of geographic information systems (GIS) ignore the acoustic echo information and analyze only the water depth data. This approach does not allow to create a fish concentration map or a presence of vegetation map or a large fish habitats map or other water body characteristics maps, indirectly extracted from the echolocation data.

Traditionally, various methods of spatial interpolation are used to construct a topographic map of the bottom from a discrete set of measurements. Geostatic estimation methods, such as kriging, require a large amount of computation, but allow us to obtain interpolations that are optimal in a certain sense.

When it comes to the processing of sonar data, it is important to note a feature that data is fragmentary, limited, and often inadequate to obtain reliable statistical estimates. The presence of uncertainty of this kind is an additional argument in favor of soft calculations. If we regard the unknown parameter as continuous, then we can draw a parallel between the conclusion about the value of the unknown parameter and the approximation of the function.

The idea of applying a fuzzy logical approximator for constructing a bottom topographic map follows from an analogy. The set of depth point measurements can be considered as a system of knowledge about the properties and structure of the water body. Each acoustic sounding can be described in terms of formal logic.

The deployment of the best tool for image recognition based on deep learning neural networks allows us to talk about the use of the echo sounder for solving new applied problems, such as tourist, ecological, nature protection, search tasks.

In [1] authors propose sliding window filters with contour detection to extract low-level features and fishes contours on echo images. This approach cannot adapt to various shapes of fish-schools and bottom artifacts, because filters kernels were not specialized for complicated forms.

In [2] presented approach which use sliding window filtering to extract objects from echo image. At first a median filter was used for noise removing, after a low pass filter with adaptive threshold was used to separate tracks with fishes from the background noise level, finally a perimeter filter was used to remove small regions with echo pulses from stochastic noise and bottom-structure. Described method can give false negative results in case complicated forms school of fishes.

Next algorithm, presented in [3] proposes convolutional neural networks (CNN) approach to extract information about fish localization on echo image. This algorithm uses the sonar images of moving agent obtained by forward-looking sonar. Authors used CNN Yolo to binary classification. But in case high-resolution images and

small-sized fishes this algorithm can give a lot of false negative results. Also, school of fishes can be missed, because it was not taken into account when learning CNN.

Traditionally, to surface a topology map building by a number of discrete measures, different kinds of interpolation methods are used. Geostationary estimation methods [4], like kriging, have large computational costs, but they can achieve optimal interpolation. In the case of echo data processing, it is important to take into account that the data are limited and frequently and the proper evaluation of static characteristics by these values is hard. GPS data updates rarely than the echo data that is why georeferencing is performing for group of echo sounding points. These groups are not equable located in water body and that is why we perform fuzzy logic for calculations.

We propose novel approach to generate fish concentration maps based on sonar data using CNN and that can adapt to different environment conditions. The presented approach to detect fishes or other objects on sonar images is based on the following steps: 1) separation of the input image into overlapping blocks; 2) blocks-processing using CNN YOLO v2, and 3) merging extracted bounding boxes around one object. After fish detection, to construct maps of the distribution of features along the lake, we propose a novel method for constructing the approximation of GPS-referenced CNN results based on the original implementation of fuzzy logic.

2 Fish detection using CNN

Deep machine learning systems provide perfect performance for object detection and classification challenges. Object detection systems have to dedicate following contributions: accuracy, precise extraction of regions of interest (RoIs) on images, and their classification with minimal deviation and speed. Usually typical image processing systems (optical character recognition system, fire detection video systems and others) include the following steps: preprocessing, features extraction, classification, and context processing [5, 6]. Machine learning systems simulating the human brain, can solve detection and classification problem as good as or even better than the human brain. At the same time, machine learning systems are faster in problem solving than the human brain. Currently, CNNs are increasingly used for image processing in various practical areas. Unlike traditional networks, CNNs provide a reduced number of extracting parameters and as an alternative of whole image processing and can process only extracted feature map, which takes into account the image topology and is stable to affine transformation.

We analyzed famous neural network architectures, like AlexNet[7], Faster R-CNN [8], CoogoleNet[9], ResNet[10] and etc. These networks process whole image as the feature map. This approach can make calculation faster, than the whole image processing. As stated in [11], YOLO is the fastest object detection system that works better than Faster R-CNN. In [12] the authors propose YOLO v2, YOLO9000, proposing modifications of YOLO. Better segmentation and classification was achieved by: 1) batch normalization from [13]; 2) high resolution classifier; 3) convolutional with anchor boxes; 3) use of k-means; 4) direct location prediction; 5) RPN usage; 6) fine-

grained features; 7) multi-scale training; and 8) novel classification model Darknet-19.

CNN Darknet-19 has 19 convolutional and five max-pooling layers. It can distinguish 9000 classes. At training time, instead of fixing the input image size, the network was changed in every few iterations. After ten batches YOLO v2 randomly chooses a new image dimension size. Since this model down-sampled by a factor of 32, was pulled from the following multiples of 32: {320, 352, ..., 608}. Input image resolution was resized to that dimension and continues training.

Since the sonar moves during the scanning of the lake along a complex trajectory with an alternating speed, it is necessary to perform the procedure of echogram normalizing. For this purpose, an algorithm to convert the echogram to metric coordinates along the length of the sonar track was developed. Due to the corresponding stretching/compression of the echogram, all objects of the acoustic echo may be represented on a single scale (fig.1).

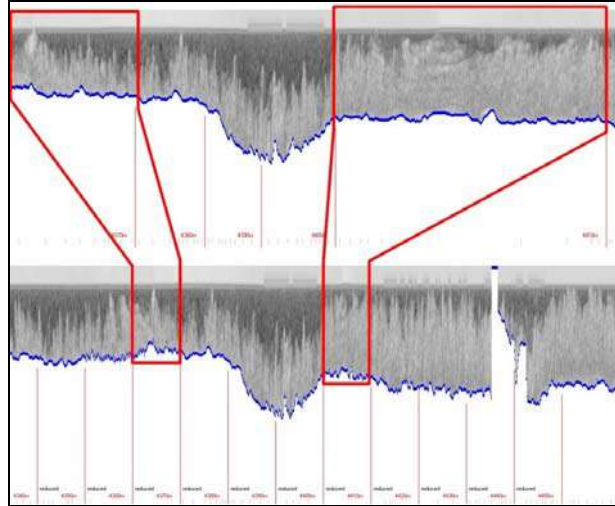


Fig. 1. Example for echogram normalization process

The input images scaled before CNN processing. This means that the tiny objects (fishes) can be missed. To solve this problem, we decided to process patches of echo image for precisely tiny object detection and subsequently concatenate the output results with performing post processing actions.

We propose an effective algorithm for fish detection on sonar images based on the following steps: input frame separating into overlapping blocks, blocks processing using CNN YOLO v2, merging extracted bounding boxes around one object.

Input image I with sizes $H \times W$ is divided into overlapping blocks $C_{i,j}$ with sizes $ch \times cw$, $i = 0, H / ch - 1$, $j = 0, W / cw - 1$. Overlap size can vary by input frame resolution and percentage ratio of minimal objects sizes.

Each block goes to CNN YOLO v2 [13] in which network predict objects are localized by using sequences of convolutional filters. YOLO v2 uses convolution with

anchor boxes, like in Faster R-CNN and run k-means ($k=5$) clustering for getting good priors for predicted objects. After YOLO v2 processing, we have bounding boxes in every block C_{ij} that are presented as top left corner coordinates $B_{ij}(x_1, y_1)$, bottom right corner coordinates $B_{ij}(x_2, y_2)$, object classification, and probability value.

In the next step - blocks post-processing - the neighbor RoIs, which have combined overlapped region located closer than 20% from blocks edge, are searched. If these blocks are found, we calculate *IoU* (Intersection over Union) which describes two regions overlapping:

$$IOU = \frac{B_1 \cap B_2}{B_1 \cup B_2}, \quad (1)$$

in which B_1 and B_2 are regions areas.

3 Maps building based on fuzzy logic

The idea of “fuzzy logic approximator” application to bottom topographical map building or for building feature map is based on following analogy: The assembly of separated depth measuring may be presented as knowledge-based systems including information about body water features and structure. The echo detection is described with formal logic as:

**IF coordinates X, Y AND time t THEN depth D,
water temperature T and other parameters.**

In this work the universal adaptive approximation is presented as fuzzy logic specific realization mathematical tools evolved in [14].

If distance between two points measuring operator is $L(p, p_i)$ then, as analogy of production system, water body fragment knowledge proliferates on neighbor fragments in accordance unimodal function has maximum value in specified point. We propose membership function in distance from nodal approximation point p_i to point p as:

$$\varphi(p, p_i) = \frac{1}{L(p, p_i)^n} \mid L(p, p_i) < L_{limit}, \quad (2)$$

In case application defuzzification by COG (Center Of Gravity) method, desired feature value for unknown point p can be calculated as:

$$p.z = \frac{\sum_i \varphi(p, p_i) p_i.z}{\sum_i \varphi_i}, \quad (3)$$

In addition, we present defuzzification methods modifications which overcome neighbor point's influence with help of space discretization. The replacement of rules group comes into discretization interval on one rule with maximum membership function in point p influence.

Fuzzy logic approximation, in contrast to traditional approximation methods, can take into account several predicates and build complex conditions.

For example, we can formulate approximation condition which allows not only depth but also information of structure of the lake bottom for correction abrupt depth change which may arise by bottom objects (artifacts).

IF coordinates X, Y **AND** bottom structure without artifacts, **THEN** depth D, water temperature T, and other parameters.

We also proposed modifications to the defuzzification methods to eliminate the disproportion of the influence of nearby points. This was done by space discretizing and replacing the influence of the rule group, included in a single discretization interval, on the influence of one rule with the maximum membership function at point p. An interesting way to eliminate the influence of the nodes location unevenness is the angular discretization.

We introduce the operator $angle(p, p_i)$ that returns the number of the circle sector into which the angle between the unknown point p and the approximation node p_i falls. Define the nearest point for each sector as follows:

$$PAngle(p, p_i) = p_i \mid \forall p_j \in P, p_j \neq p_i, angle(p, p_j) = angle(p, p_i), \varphi(p, p_i) > \varphi(p, p_j), L(p, p_i) < Llimit \quad (4)$$

The value of the unknown parameter, for example, the depth z, for an 8-sector split:

$$p.z = \frac{\sum_{k=1}^8 \varphi(p, PA_k) PA_k.z}{\sum_{k=1}^8 \varphi(p, PA_k)}. \quad (5)$$

If the angular discretization interval tends to 0, then the output by the center of gravity method will look as follows:

$$p.z = \frac{\int_{-\pi}^{\pi} \varphi(p, \alpha) z(\alpha) d\alpha}{\int_{-\pi}^{\pi} \varphi(p, \alpha) d\alpha}, \quad (6)$$

where $\varphi(p, \alpha)$ is piecewise-linear interpolation in polar coordinates of the values of the maximal membership functions of the nearest points in the direction to α points, and $z(\alpha)$ is a piecewise-linear interpolation in polar coordinates of the depth values of the points with the maximum membership function value.

The key difference between fuzzy logic approximation and traditional methods of approximation is the possibility of taking into account several predicates. For example, we formulate an approximation condition that should include both the depth information and the bottom structure information in order to eliminate the effect of depth jumps from bottomed artifacts.

In this case predicates value needs to be normalized and we propose the following membership function modification:

$$\varphi(p, p_i) = \left(1 - \frac{L(p, p_i)}{L_{limit}}\right)^n \mid L(p, p_i) < L_{limit} . \quad (7)$$

To define depth irregularity for point p as $R(p)$, we normalize this value and estimate “bottom without artifacts” model as:

$$\varphi_R(p) = \left(1 - \frac{R(p) - R_{min}}{R_{max} - R_{min}}\right)^m . \quad (8)$$

Then:

$$p.z = \frac{\int \max(\min(\varphi(p, p_i), \varphi_R(p))) z dz}{\int \max(\min(\varphi(p, p_i), \varphi_R(p))) dz} . \quad (9)$$

In conclusion, we get a classical logical output minimax representation about unknown parameter value and bottom approximation with low artifact influence.

4 Experimental Results

For YOLOv2 we build our own training set including about 80 000 objects. We selected ground truth bounding boxes around RoIs manually using VOTT (Visual Object Tagging Tool) software [15]. VOTT can make ground truth coordinates and convert them into Yolo format. Using this program, we additionally created annotations files. We predicted five classes of objects: “fish”, “grass”, “school of fish”, “predator”, “bottom fish”. Fig. 2 depicts ground truth boxes in VOTT.

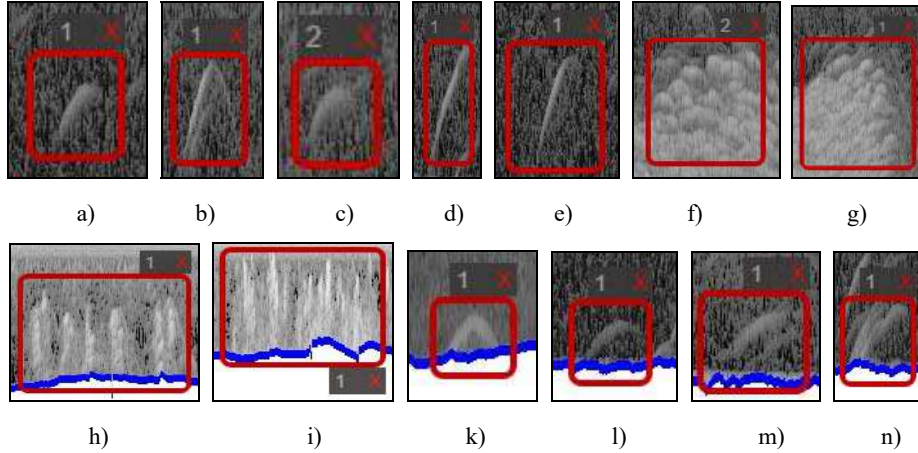


Fig. 2. Ground truth bounding boxes in VOTT: a,b,c) “fish”; d,e) “predator”; f,g) “school of fish”; h,i) “grass”; k,l,m,n) “bottom fish”

All images were taken from the river Western Dvina and lakes in the Republic of Belarus with a maximum depth for river of 12 meters and for lakes of 38 meters. Double-beam (200kHz and 450 kHz) echo sonar Lowrance HOOK 4 was used. Fig. 3 depicts the resulting classification after YOLOv2 processing.

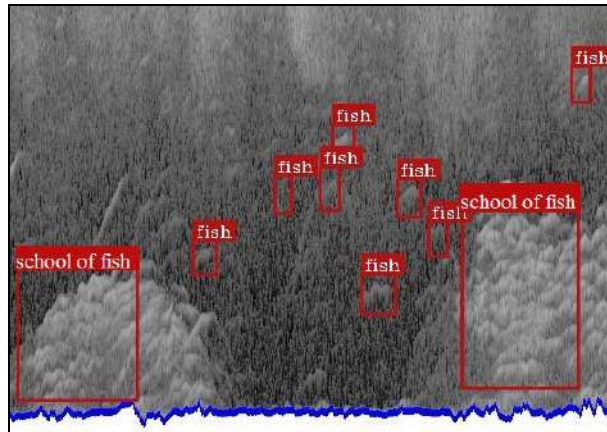


Fig. 3. Fish detection and classification results

Presented algorithm has an accuracy of 72.1% and a low percentage of false positive results in case of fish presence. However, our approach, as shown in Fig 4, cannot properly distinguish classes “grass” and “school of fish”, especially in case similar shapes.

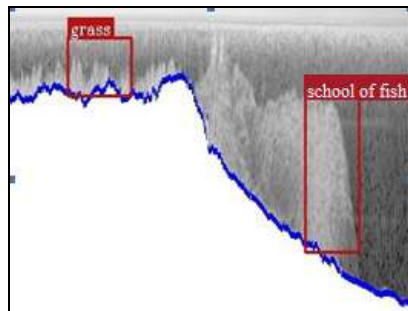


Fig. 4. Example for incorrect classification

Therefore, the accuracy of the approximation increases with increasing number and density of approximation nodes. As the number of approximation nodes increases, the accuracy of approximation increases. In this case it is important that the sonar passes over all the most complex and characteristic sections of the bottom.

A special type of track is the lake contour, imported in KMZ format from known GIS systems, such as Google Earth. The contour is a track with zero-depth points (Fig. 5). Similar contours are used to simulate contours of islands. It is possible to

delete incorrect points of the sonar tracks, as well as some points of the contours for modeling the open contour of the river bed. The same approach is used to construct the approximation by other features (Fig. 6).

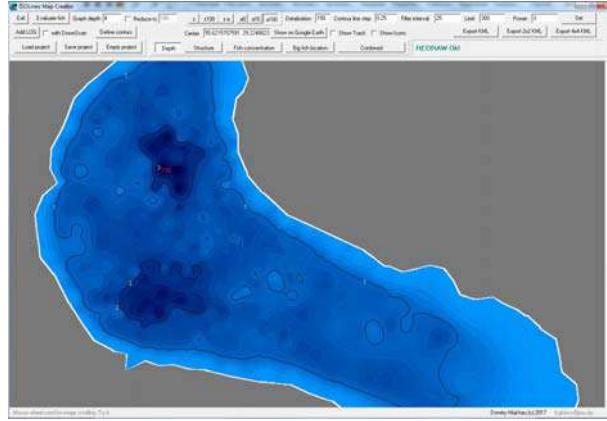


Fig. 5. Example for map building taking into account the contour of the lake

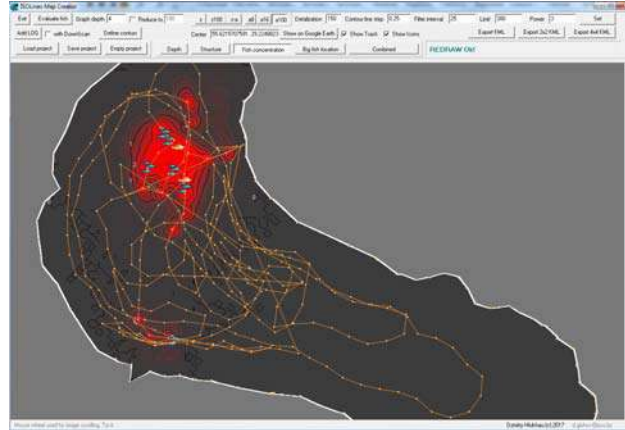


Fig. 6. Example for map of fish concentration

5 Conclusion

Method for obtaining topographic maps of lakes, maps of fish concentration and a map of predator location based on the results of intelligent sonar data processing is presented. The presented algorithm is based on sonar images for the detection of classes “fish”, “grass”, “school of fish”, “predator”, “bottom fish”. The algorithm includes following steps: input frame separating into overlapping blocks, blocks-processing using CNN YOLO v2, and merging extracted bounding boxes around one object, fish concentration map building. To construct maps of the distribution of fea-

tures along the lake, we propose a novel method for constructing the approximation of GPS-referenced CNN results based on the original implementation of fuzzy logic. Our method has an accuracy of 72.1% and has low percentage of false positive results in case of fish presence. To increase the accuracy, we need to significantly expand the dataset for CNN training.

References

1. Balk H., Lindem T. Improved fish detection probability in data from split-beam sonar. *Aquatic Living Resources*. 13(5): 297–303(2000) doi: 10.1016/S0990-7440(00)01079-2
2. Helge B., Torfinn L.: Improved fish detection probability in data form split-beam sonar: <https://slides.tips/improved-fish-detection-probability-in-data-form-split-beam-sonar.html>
3. Kim J., Yu, SC.: Convolutional neural network-based real-time rov detection using forward-looking sonar image. *Autonomous Underwater Vehicles (AUV)*, IEEE/OES. pp. 396–400. (2016) doi: 10.1109/AUV.2016.7778702
4. Krivoruchko, K.: *Spatial Statistical Data Analysis for GIS Users*. Redlands, Esri Press, (2011)
5. Demant, C., Garnica, C., Streicher-Abel, B. : *Industrial Image Processing: Visual Quality Control in Manufacturing*. Heidelberg, Springer (2013)
6. Shiping, Y., Zhican, B., Huafeng, C., Bohush, R. and Ablameyko, S.: An effective algorithm to detect both smoke and flame using color and wavelet analysis. *Pattern Recognition and Image Analysis*. 27(1):131-138 (2017) doi: 10.1134/S1054661817010138
7. Krizhevsky, A., Sutskever, I. and Hinton, G. E.: ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS'12)*, vol. 1, pp. 1097-1105 (2012)
8. Ren, Sh., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 39(6): 1137 - 1149 (2017) doi: 10.1109/TPAMI.2016.2577031
9. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 27-30 June 2016, pp. 2818–2826 (2016) doi:10.1109/CVPR.2016.308
10. He, K., Zhang, X., Ren, Sh., Sun, J.: Deep Residual Learning for Image Recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 27-30 June 2016, pp. 770–778 (2016) doi: 10.1109/CVPR.2016.90
11. Redmon, J., Divvala, S. K., Girshick, R. B., Farhadi, A.: You Only Look Once, Unified, Real-Time Object Detection. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 27-30 June 2016, pp. 779–788 (2016) doi: 10.1109/CVPR.2016.91
12. Redmon, J., Farhadi, A.: YOLO9000: Better, Faster, Stronger. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 21-26 July 2017, pp. 6517–6525 (2017) doi:10.1109/CVPR.2017.690
13. Ioffe, S., Szegedy, Ch.: Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning* Microtome Publishing, 6 -11 July 2015., pp. 448–456 (2015)
14. Glukhov, D.: Dynamic expert system by fuzzy inference rules to automations an examination of complex objects. *Budownictwo i Inzynieria, Srodowiska*, pp. 105–109 (1998)
15. Visual Object Tagging Tool: An electron app for building end to end Object Detection Models from Images and Videos: <https://github.com/Microsoft/VoTT>

Modification of the genetic method for neuroevolution synthesis of neural network models for medical diagnosis

Serhii Leoshchenko¹[0000-0001-5099-5518], Andrii Oliinyk²[0000-0002-6740-6078],
Sergey Subbotin³[0000-0001-5814-8268], Nataliia Gorobii⁴[0000-0003-2505-928X] and
Vadym Shkarupilo⁵[0000-0002-0523-8910]

^{1,2,3,4} Dept. of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia
69063, Ukraine

⁵ Dept. of Computer Systems and Networks, National University of Life and Environmental
Sciences of Ukraine, Kyiv 03041, Ukraine

sergleo.zntu@gmail.com, olejnikaa@gmail.com,
subbotin@zntu.edu.ua, gorobiy.natalya@gmail.com,
shkarupilo.vadym@nubip.edu.ua

Abstract. The main aim of the paper is researching the possibility of application artificial neural networks as the neural network models that can be used in medical diagnostics. One of the most problematic and complex issues of neural network models implementation is the initial stage of synthesis. The article presents a comparison of existing methods of synthesis, as well as a new method. The experiments confirm the effectiveness and expediency of the proposed method.

Keywords: artificial neural networks, synthesis, neuroevolution, genetic method, support vector machine.

1 Introduction

The diagnosis stage plays a crucial role in medicine. Timely accurate diagnosis facilitates the choice of therapy and significantly increases the probability of treatment of the patient. The using of neural networks is one of the ways to improve the efficiency of medical diagnosis [1].

The accuracy of the diagnosis and the speed with which it can be delivered depend on many factors: the patient's condition, the available data on the symptoms and signs of the disease and the results of laboratory tests, the total amount of medical information on the observation of such symptoms in a variety of diseases and, finally, the qualification of the doctor. A major role in this process is played by the human factor, which often leads to errors [1], [2].

Some of the specific medical diagnosis difficulties that need to be considered are listed below [3].

The basis for a reliable diagnosis is a wealth of practical experience, which can be reached only the middle of a doctor's career and is absent at the end of training, of

course. This is especially true for rare or new diseases, where experienced doctors are in the same situation as beginners.

The quality of diagnosis depends on the skill, knowledge and intuition of the doctor.

Emotional problems and fatigue adversely affect the work of the doctor.

Training of specialists is a long and expensive procedure, and therefore in many, even in developed countries, there is a lack of skills.

Medicine is one of the fastest growing and developing fields of science. New results disqualify the old ones, new drugs appear every day. The same applies to the diseases themselves, which take new forms.

These factors necessitate the search for new solutions and tools, for example, the use of artificial neural networks (ANNs) [2].

2 Using the ANN in medical diagnosis

The ANN technologies are designed to solve difficult-to-formalize problems, which, in particular, are reduced to many problems of medicine [4], [5]. This is primarily due to the fact that the researcher is often provided with a large number of heterogeneous factual material for which a mathematical model has not yet been created. In addition, it is necessary to present the results of the analysis in a form, which will be understandable to the specialist. So ANN is a powerful and flexible method of simulating processes and phenomena. Neural networks can be different in structure and form, but they have several common features. A distinctive feature of neural networks is their ability to training on the basis of experimental data of the subject area. With regard to medical subjects, experimental data are presented in the form of a set of initial features or parameters of the object and the diagnosis based on them. Training of ANN is an interactive process in which the neural network finds hidden nonlinear relationships between the initial parameters and the final diagnosis, as well as the optimal combination of weight coefficients of neurons connecting adjacent layers, in which the classification error tends to a minimum [6]. In the training process, the input of the neural network is fed a sequence of initial parameters along with the diagnoses that characterize these parameters. Careful formation of the training sample determines the quality of work, as well as the level of error of the neural network.

A number of difficulties are associated with the use of neural networks in practical problems. One of the main problems of application ANN technologies is a previously unknown degree of complexity of the projected ANN, which will be enough for a reliable diagnosis. This complexity can be unacceptably high and will require more complex network architecture. It is known, for example, that the simplest single-layer neural networks are able to solve only linearly separated problems [7]. This limitation can be overcome by using multilayer neural networks.

The basis of ANNs are neurons with a structure similar to biological analogues. Each neuron can be represented as a microprocessor with several inputs and one output. When neurons are joined together, a structure is formed, which calls a neural network. Vertically aligned neurons form layers: input, hidden and output. The num-

ber of layers determines the complexity and, at the same time, the functionality of the network, which is not fully investigated.

For researchers, the first stage of creating a network is the most difficult task. The following recommendations are given in the literature [8–10].

1. The number of neurons in the hidden layer is determined empirically, but in most cases the rule is used $N_h \leq N_i + N_o$, where N_h is the number of neurons in the hidden layer, N_i in the input and N_o output layers.
2. Increasing the number of inputs and outputs of the network leads to the need to increase the number of neurons in the hidden layer.
3. For the ANNs modeling multistage processes required additional hidden layer, but, on the other hand, the addition of hidden layers may lead to overwriting and the wrong decision at the output of the network.

Based on these recommendations, the number of layers and the number of neurons in the hidden layers is chosen by the researcher, based on his personal experience.

3 Review of the literature

The ANN are attractive from an intuitive point of view, because they are based on a primitive biological model of nervous systems. In this connection, there is an assumption that to improve it may be appropriate to apply another borrowing from nature, for example, evolutionary calculations and, in particular, neuroevolution. Neuroevolution in this paper refers to the automatic modification of neural networks using genetic algorithms. With this methodology, possible variations of neural networks with different topologies are grown, which with each iteration, called generation solve the problem better. Despite genetic programming, as well as evolutionary calculations in general, do not guarantee finding the optimal result, this approach eventually allows us to come to the results applicable to solving practical problems. However, it will take a reasonable amount of time to achieve such results. Thus, the level of complexity of the neural network that arises when it is necessary to create a neural network is significantly reduced, because when it is created, it is only necessary to select the parameter that evaluates the work of the neural network and provide a suitable set of data.

Despite the fact that most of the works devoted to the neuroevolutionary approach offer only a theoretical approach to solving problems of neural network optimization, it is possible to find several promising and noteworthy methods [11–14].

From the early works of noteworthy cellular Frederick Gruau method [15], [16] uses a special grammar for the representation of neural network structures. One individual represented an entire neural network, with each neuron considered as a biological cell, and the growth of the network was determined through the mechanisms of sequential and parallel "division" of neurons - cells. However, this method involves the implementation of a large number of specific operators that provide simulation of cell activity.

The Hierarchical SANE (Symbiotic, Adaptive NeuroEvolution) [17] method uses a different approach. It considers the development of two independent populations, one of which individuals are separate neurons, and the other contains information about the structures of an artificial neural network. The disadvantages of this method include the fact that the number of hidden neurons and connections is limited.

The ESP method [18] is a development of the sane method. Its main difference is that the network structure is fixed and is given a priori. The population of neurons is divided into subpopulations, in each of which the evolution is independent. Due to parallelization of the solution search, as well as simplification of the problem due to the rejection of the evolution of the artificial neural network structure, ESP works much faster than SANE, sometimes by an order of magnitude, but for the successful operation of the method it is required to choose the appropriate structure of the neural network [19].

One of the most potentially successful attempts to get rid of the disadvantages of direct coding while preserving all its advantages is the method proposed in 2002, called NEAT — Neural Evolution through Augmenting Topologies [15], [20]. Designed by Kenneth Stanley, the NEAT method allows to customize the structure of the network, and without restrictions on its complexity. The solution proposed by the authors is based on the biological concept of homologous genes (alleles), as well as on the existence in nature of the synapsis process — the alignment of homologous genes before the crossover. The technique assumes that two genes (in two different individuals) are homologous if they are the result of the same mutation in the past. In other words, with each structural mutation (gene addition), a new gene is assigned a unique number, which then does not change during evolution. The method uses a number of techniques, such as historical labels and specialization of individuals, to make the process of evolution significantly more efficient [21].

Summing up, it can be noted that the joint use of evolutionary methods and artificial neural networks allows us to solve the problems of configuration and training of artificial neural networks both individually and simultaneously. One of the advantages of this synthesized approach is largely a unified approach to solving a variety of problems of classification, approximation, control and modeling. The use of qualitative evaluation of the functioning of artificial neural networks allows the use of neuroevolutionary methods to solve the problems of the study of adaptive behavior of intelligent agents, the search for game strategies, signal processing. Despite the fact that the number of problems and open questions concerning the development and application of neuroevolutionary methods (coding methods, genetic operators, methods of analysis, etc.) is large, often for the successful solution of the problem with the use of neuroevolutionary method adequate understanding of the problem and neuroevolutionary approach, as evidenced by a large number of interesting and successful works in this direction [15].

4 Materials and methods

In the method, which is proposed to find a solution using a population of neural networks: $P = \{NN_1, NN_2, \dots, NN_n\}$, that is, each individual is a separate ANN $Ind_i \rightarrow NN_i$ [19–21]. During initialization population divided into two halves, the genes $g_{Ind_i} = \{g_1, g_2, \dots, g_n\}$ of the first half of the individuals is randomly assigned

$g_{Ind_i} = \{g_1 = \text{Rand}, g_2 = \text{Rand}, \dots, g_n = \text{Rand}\}$. Genes of the second half of the population are defined as the inversion of genes of the first half $g_{Ind_i} = \overline{\{g_1 = \text{Rand}, g_2 = \text{Rand}, \dots, g_n = \text{Rand}\}}$. This allows for a uniform distribution of single and zero bits in the population to minimize the probability of early convergence of the method ($p \rightarrow \min$).

After initialization, all individuals have coded networks in their genes without hidden neurons (N_h), and all input neurons (N_i) are connected to each output neuron (N_o). That is, at first, all the presented ANNs differ only in the weights of the interneuron connection w_i . In the process of evaluation, based on the genetic information of the individual under consideration, a neural network is first built, and then its performance is checked, which determines the fitness function ($f_{fitness}$) of the individual. After evaluation, all individuals are sorted in order of reduced fitness, and a more successful half of the sorted population is allowed to cross, with the best individual immediately moving to the next generation. In the process of reproduction, each individual is crossed with a randomly selected individual from among those selected for crossing. The resulting two descendants are added to the new generation $G = P' = \{Ind_1, Ind_2, \dots, Ind_n\}$. Once a new generation is formed the mutation operator starts working. However, it is important to note that the selection of the truncation significantly reduces the diversity within the population, leading to an early convergence of the algorithm, so the probability of mutation is chosen to be rather large ($p_{mut} = 15 - 25\%$) [22].

If the best individual in the population does not change for a certain number of generations (by default, it is proposed to set this number at eight), this individual is forcibly removed, and a new best individual is randomly selected from the queue. This makes it possible to realize the exit from the areas of local minima due to the relief of the objective function, as well as a large degree of convergence of individuals in one generation. The general scheme of the method demonstrated at Fig.1.

4.1 Using of genetic operators

It is obvious that the chosen method requires special genetic operators that implement crossover and mutation.

At crossover two parental individuals which produce two descendants are used. Common neurons and connections are inherited by both offspring, and the value of connections in the networks of descendants are formed by a two-point crossover. Elements of ANN, of distinct played out between generations.

An important feature is that neurons with the same indices are considered identical, despite the different number of connections and position in the network, as well as the fact that one of these neurons could have a different index, which changed as a result of correction of indices after mutation. For this purpose, three coefficients were introduced that regulate the size and direction of the network.

The first of them characterizes the degree of connectedness of neurons in the network and is calculated by the formula:

$$f_{con} = \frac{N_c}{2^{FB-1} [N_s(N_s-1) - N_i(N_i-1) - (1-FB)N_o(N_o-1)]} \quad (1)$$

where N_c is the number of connections in the network, N_i , N_o , N_s are respectively, the number of input, output neurons and the total number of neurons in the network, FB is a variable indicating the permitted occurrence of feedbacks ($FB = 1$) or not ($FB = 0$). It is worth noting that connections from hidden neurons to the output can appear in any case. Thus, the smaller f_{con} the more likely a new relationship will be added as a result of the mutation [23].

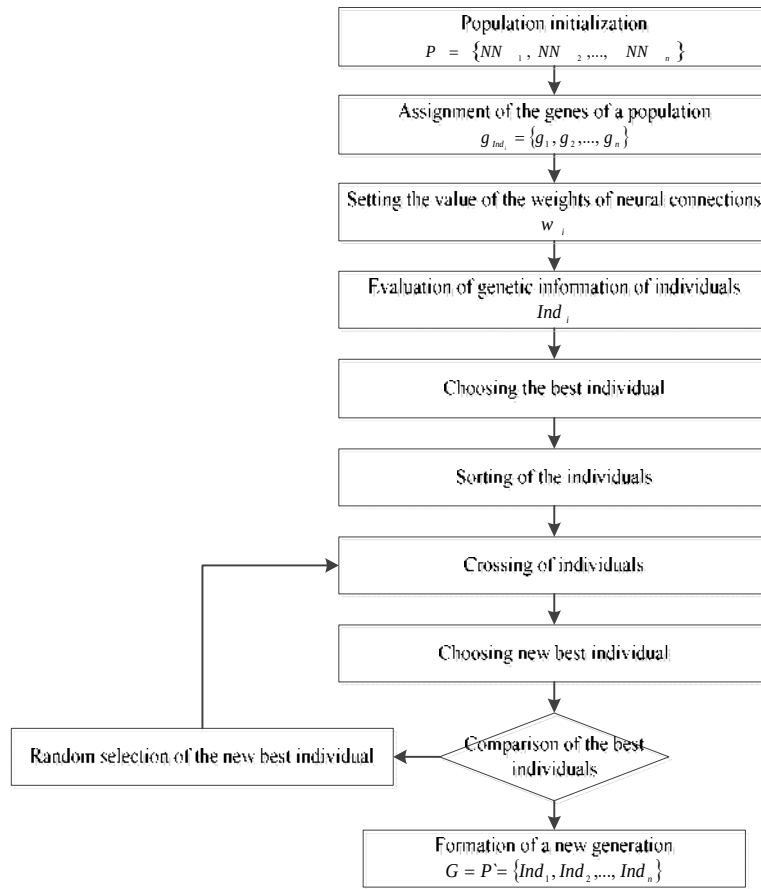


Fig. 1. The general scheme of the method

The use of the second coefficient is based on the assumption that the more elements in the sum of the input and output vectors of the training choice (the greater the total number of input and output neurons), which is probably a more complex network is necessary to solve the problem. The second coefficient is calculated by the formula:

$$f_{top.diff} = \frac{N_i + N_o}{N_s} \quad (2)$$

That is, the more neurons in the network, the less will be $f_{top.diff}$ and the less likely will be selected mutation that adds a new hidden neuron [23].

The third criterion is also based on the assumption that a more complex network should be used to solve more complex problems. However, this criterion characterizes the conditional complexity of the network. This criterion is based on the concept of cyclomatic complexity [24], [25].

$$f_{comp.diff} = \frac{N_i + N_o}{N_s} \quad (3)$$

For any of the described cases, the algorithm uses a ligament $f_{con} \cdot f_{top.diff} \cdot f_{comp.diff}$, because for use it is necessary to take into account the degree of connectivity of already existing neurons.

Thus, using mutations can be pointwise to change the parameters of the structure of the ins.

Chaotic the addition (removal) of neurons and connections can lead to situations where, for example, in a network of many neurons and few connections. It would be more logical to apply different types of mutations depending on the features of the network architecture represented by the mutating individual [26–28].

Removing links gives a side effect: there may be hanging neurons that have no incoming connections, as well as dead-end neurons, that is, without output connections [26], [27], [29]. In cases where the function of neuronal activation is such that at zero weighted sum of inputs its value is not equal to zero, the presence of hanging neurons makes it possible to adjust the neural displacement. It is worth noting that, on the other hand, the removal of links may contribute to the removal of some uninformative and uninformative input features.

4.2 Choosing the mutation type

Consider the dependence of the type of mutation on the values f_{con} , $f_{top.diff}$ and $f_{comp.diff}$. Adaptive mutation mechanism is one of the key features of the proposed method.

The choice of mutation type is determined based on the values of f_{con} , $f_{top.diff}$ and $f_{comp.diff}$. This approach, on the one hand, does not limit the number of hidden neurons from above, on the other hand, it prevents the immeasurable increase of the network, because the addition of each new neuron to the network will be less likely. The mutation of the weight of a random existing bond occurs for all mutating individuals with a probability of 0.5.

Let us consider in more detail how to choose the type of mutation. Fig. 2 shows the block diagram of the selection of the type of mutation. Here RV is a random variable, N_h is the number of hidden neurons in the mutating network.

Conventionally, the entire algorithm can be divided into two branches on the first conditional transition:

- branch increase f_c is carried out for the fulfilment of the conditions of transition;
- branch reduction f_c , performed if the transition condition is not met.

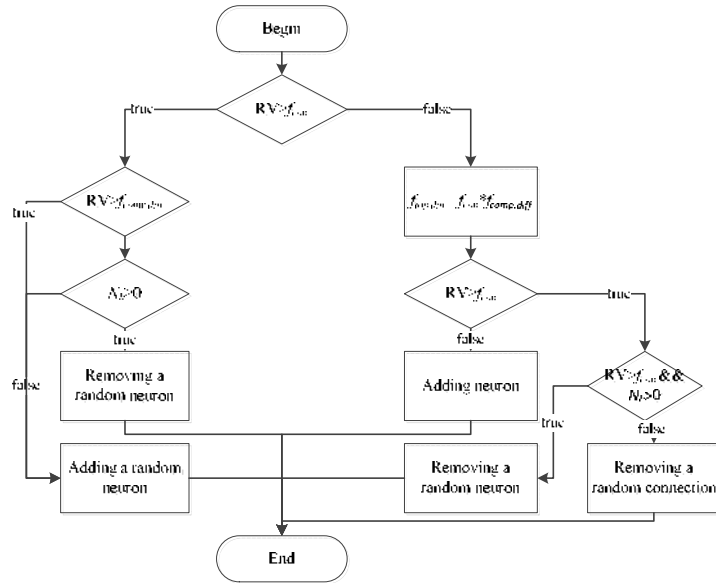


Fig. 2. The diagram of the selection of the type of mutation

Multiplication $f_{con} \cdot f_{comp.diff}$ is necessary in order to change the number of neurons adequately network topology, because the addition (removal) of neurons need information about the feasibility of changes. This information can be obtained indirectly from the value of the characteristic.

4.3 The calculation of the output layer of ANN

On condition using the support vector machine, the optimality criterion for calculating the output weights may not be specified. If the value of the mean square error is replaced by the criterion of the maximum separation of the support vectors, then the optimal linear weights of the output can be estimated using, for example, quadratic programming, as in the traditional method of support vectors, for this it is advisable to use the Evoke operator [30], by the formula:

$$y(t) = w_0 + \sum_{i=1}^k \sum_{j=0}^{l_i} w_{i,j} K(\phi(t), \phi^i(j)), \quad (4)$$

where $\phi(t) \in R^n$ is the output of a recurrent neural network $f(\cdot)$ at a time t ; $K(\cdot, \cdot)$ is a predefined kernel function; $w_{i,j}$ is weights corresponding to k training sequences ϕ^i , each length l_i , and are calculated using the support vector machine.

The value of the mean square error is replaced by the criterion of maximum separation of support vectors. In this case, the optimal linear weights can be estimated using quadratic programming, as in the traditional support vector machine.

One of the problems of neuroevolutionary method realization is the algorithm of ANN output calculation with arbitrary topology.

ANN can be represented as a directed planar graph. Based on the fact that the network structure can be any, loops and cycles containing any nodes are allowed in the graph, except for the nodes of the corresponding input neurons. Let denote the set of nodes of the graph by $V = \{v_i \mid i \in [0; N_v - 1]\}$, and a set of arcs through $E = \{e_j \mid j \in [0; N_e - 1]\}$, where N_v and N_e are accordingly, the number of nodes and arcs in the graph, and $N_v = N_s$, and $N_e = N_c$. The arc, which goes from node k to node 1 denote by an ordered pair $c_{k,l} = (v_k, v_l)$, the weight of the corresponding link will be denoted by $w_{k,l}$.

Give the index to the nodes of the graph as neurons, that is, the nodes that are the input neurons, called input, have an index out of range $[0; N_l - 1]$. By analogy, the indexes of outgoing nodes belong to the interval $[N_l; N_l + N_o - 1]$, and indexes for hidden nodes will be set in the interval $[N_l + N_o; N_v - 1]$.

Let introduce an additional characteristic for all nodes of the graph equal to the minimum length of the chain to any of the input nodes and denote it ch_i . Let's call ch_i the layer to which the i^{th} node belongs. Thus, all input nodes belong to the 0^{th} layer, not all input nodes that have input arcs from the input belong to the 1st layer, all other nodes with input arcs from nodes of the 1st layer will belong to the layer with index 2, etc. in this case, there may be situations when the node does not have input arcs, we will call it a hanging node with the layer number $ch_i = -1$.

For arcs, we also introduce an additional characteristic $b_{k,l}$ for the arc $c_{k,l}$, which is necessary to determine whether the arc corresponds to forward or reverse. It will be calculated as follows:

$$b_{k,l} = \begin{cases} 1, ch_l - ch_k > 0 \\ -1, ch_l - ch_k \leq 0 \end{cases} \quad (5)$$

That is, if the index of the layer of the end node of the arc is greater than the index of the layer of the beginning node, then we will consider such an arc as a straight line, otherwise we will consider the arc as an inverse.

Since each node of the graph represents a neuron, we denote by sum_i the value of the weighted sum of inputs, and through o_i is the value of the output (the value of the activation function of the i^{th} neuron-node). Then, $o_i = f_{fitness}(sum_i)$ where $f_{fitness}$ is the function of neuron activation.

Let's divide the whole process of signal propagation from the input nodes into stages, and during one such stage the signals manage to pass only one arc. The number of the stage is denoted by s . For the very first stage $s=1$. For short assumed that all arcs have the same length, and the signals are sewn on them instantly. We denote the feature that the output of node i was updated at this stage through a_i , that is, if $a_i=1$, then the output of the node at stage s is calculated, otherwise, if $a_i=1$ – not.

Let's introduce one more designation $X = \{x_i | i \in [0; N_l - 1]\}$ it is vector of input signals. Then the algorithm for calculating the ANN output is as follows:

1. $o_i = x_i$, $a_i = 1$, for all $i \in [0; N_l - 1]$;
2. $o_i = 0$, for all $i \in [N_l; N_s - 1]$;
3. $s=1$;
4. $sum_i = 0$, $a_i = 1$, for all $i \in [N_l; N_s - 1]$;
5. if $s=1$, than go to the step number 7;
6. calculation of the feedback network. For all input feedbacks $c_{j,k}$ node v_k , where $k \in [N_l; N_s - 1]$: $sum_k = sum_k + o_j$, if $ch_j < s$;
7. if $a_i = 0$, than $fn(i)$ for all $i \in [N_l; N_s - 1]$;
8. if the stop criterion is not met, than $s=s+1$ and go to the step number 4.

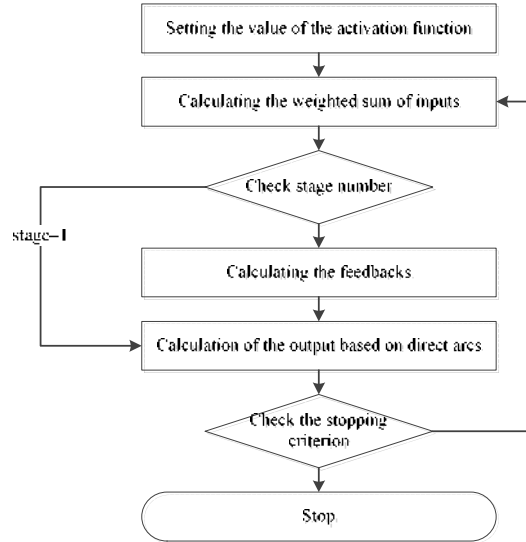


Fig. 3. The general scheme of the calculation of the output layer of ANN

Here $fn(i)$ is a recursive function that calculates the output of the 1^{st} node taking into account all straight arcs. Works on the following algorithm:

1. if $ch_i < 0$, than go to the step number 3;
2. for all input arcs $c_{k,l}$ node v_i : if $a_k = 1$, than $sum_i = sum_i + o_k$, else $fn(k)$;
3. $o_i = f(sum_i)$;
4. exit.

The stopping criterion of the ANN output calculation algorithm can be one of the following:

- stabilization of values at the output of ANN;
- s exceeds the set value.

It is more reliable to calculate the output until the values at the output of ANN do not change, but for the case when the network contains cycles and/or loops, its output may never become stable. Therefore, the required additional stopping criteria limiting the maximum number of stages of calculation of network output. For networks with no feedback ($FB = 0$) in many cases, allow the $\max(ch_i) + 1$ phases.

5 Experiments

During testing, the main task is to track the speed of the proposed method, quality and stability. Since synthesized ANN can be further used as diagnostic models for medical diagnosis, testing should be carried out on the relevant test data. Also, testing will be carried out in 2 stages: the first stage will consist in the synthesis of ANN only with the help of modified genetic algorithm, and the second – in additional processing of the initial layer by the support vector machine. This strategy will allow to know more clearly how useful the support vector machine is.

Data for testing were taken from the open repository – UC Irvine Machine Learning Repository. Data sample was used: Breast Cancer Coimbra Data Set [31]. Clinical features were observed or measured for 64 patients with breast cancer and 52 healthy controls. There are 10 predictors, all quantitative, and a binary dependent variable, indicating the presence or absence of breast cancer. The predictors are anthropometric data and parameters which can be gathered in routine blood analysis. Prediction models based on these predictors, if accurate, can potentially be used as a biomarker of breast cancer. Table 1 shows the main characteristics of the data sample.

Table 1. Main characteristics of the Breast Cancer Coimbra Data Set

Criterion	Characteristic	Criterion	Characteristic
Data Set Characteristics	Multivariate	Number of Instances	116
Attribute Characteristics	Integer	Number of Attributes	10

During the evaluation of the test results we will pay attention to the following criteria:

- the spent time, s ;
- average error of final network (E);
- standard deviation (SD).

The relative error value in this case will be calculated as the ratio of the classification error to the total sample size (number of instances).

$$E = \frac{error_{class}}{Number_{sampl}} \cdot 100\%, \quad (6)$$

where E is relative error; $error_{class}$ is classification error; $Number_{sampl}$ the number of instances in the sample.

Standard deviation gives an idea about how one or the other, the ANN accurately predicts the user's rating, since the estimation is calculated the difference between the result of the work of ANN's and known result. It is also important to know that this indicator can be calculated only with a sufficient amount of observations. Otherwise, the calculation of the SD will be uninformative and its use will not lead to improvement of the results of the ANN.

$$SD = \sqrt{\frac{1}{Number_{sampl}} \sum_{i=1}^{Number_{sampl}} (x_i - \bar{x})^2}, \quad (7)$$

where SD is standard deviation, x_i is i^{th} element of the set, $Number_{sampl}$ the number of instances in the sample, $\bar{x}$ is the mean value of these observations [32].

6 The results analysis

Table 2 shows the results of testing the modified genetic algorithm in comparison with the NEAT and ESP methods.

Table 2. Results of testing

	Time, s	E	SD
NEAT	468.013	4.40%	3.70
ESP	9389.55	3.07%	2.99
Modified GA	631.373	2.96%	3.05

As the table shows, the modified GA according to the time of fulfillment is ahead of ESP in time, however, inferior to the NEAT. However, it should be noted that in the analysis of error values, the proposed method is significantly ahead of existing methods.

Let's repeat testing, but now with additional use of the support vector machine. The results are shown in Table 3.

Table 3. Results of testing

	Time, s	E	SD
NEAT	468.013	4.40%	3.70
ESP	9389.55	3.07%	2.99
Modified GA (with support vector machines)	5326.326	2.36%	2.17

As can be seen from the table, the modified GA using the support vector machine is inferior to the opponents in terms of execution time. However, on indicators of minimum, maximum errors and average errors of the output ANN significantly better than their competitors. Therefore, we can conclude that the use of the support vector machine really significantly improves the results of the synthesis.

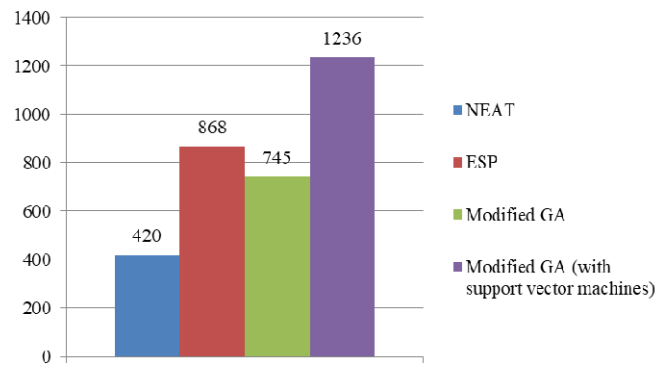


Fig. 4. The distribution of iterations in experiments

As you can see from the diagram, the modified genetic method was more iterative than the existing methods, but the time spent on iteration was less. That is, it can be concluded that iterations are not complex and for their reduction it is possible to resort to parallelization, which will significantly speed up the work even when using the support vector machine [33–35].

7 Conclusion

The problem of finding the optimal method of synthesis of ANN requires a comprehensive approach. Existing methods of ANNs training are well tested, but they have a number of nuances and disadvantages. The paper proposes a mechanism for the use of a modified genetic algorithm for its subsequent application in the synthesis of ANNs. Based on the analysis of the experimental results, it can be argued about the good

work of the proposed method. However, to reduce iterativity and improve accuracy, it should be continued to work towards parallelization of calculations.

Acknowledgment

The work was performed as part of the project “Methods and means of decision-making for data processing in intellectual recognition systems” (number of state registration 0117U003920) of Zaporizhzhia National Technical University.

References

1. Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, The National Academies of Sciences, Engineering, and Medicine: Improving Diagnosis in Health Care. In: Balogh, E., Miller, B., Ball, J. National Academies Press, Washington, USA (2015).
2. Lugovskaya, A. M.: Artificial neural networks in medical diagnostics [Iskusstvennyye neyronnyye seti v meditsinskoy diagnostike]. Computer systems and networks: proceedings of the 54th scientific conference of postgraduates, undergraduates and students 2018, vol. 1, pp. 182 – 183. BGUIR, Minsk (2018).
3. Goold, S.D., Lipkin, M.Jr.: The doctor-patient relationship: challenges, opportunities, and strategies. *Journal of General Internal Medicine* (1), 23–33 (1999).
4. Kolpakova, T., Oliinyk, A., Lovkin, V.: Improved method of group decision making in expert systems based on competitive agents selection. *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Institute of Electrical and Electronics Engineers, pp. 939–943, Kyiv (2017). doi: 10.1109/UKRCON.2017.8100388
5. Stepanenko, O., Oliinyk, A., Deineha, L., Zaiko, T.: Development of the method for decomposition of superpositions of unknown pulsed signals using the secondorder adaptive spectral analysis. *EasternEuropean Journal of Enterprise Technologies*, vol. 92, issue 2/9, pp. 48–54, (2018). doi: 10.15587/1729-4061.2018.126578
6. Callan, R.: *The Essence of Neural Networks (The Essence of Computing Series)*. Prentice Hall, Europe (1999).
7. Yasnitskiy, L.N.: *Introduction to artificial intelligence [Vvedenie v iskusstvennyiy intellekt]*. 3th edn. Academy, Moscow (2010).
8. Bondarenko, I.B., Gatchin, Yu.A., Geranichev, V.N.: Sintez optimalnykh iskusstvennykh neyronnykh setey s pomoschyu modifitsirovannogo geneticheskogo algoritma. *Nauchno-tehnicheskii vestnik informatsionnykh tehnologiy, mehaniki i optiki*, vol. 2 (78), 51–55 pp. (2012).
9. Lukichev, D.V., Usoltsev, A.A.: Sintez optimalnoy strukturyi neyrosetevykh ustroystv, 97–102 (2005).
10. Van Tuc, N.: Approximation contexts in addressing graph data structures. *University of Wollongong Thesis Collection*, 30–55 (2015).
11. Oliinyk, A., Skrupsky, S., Subbotin, S., Korobiichuk, I.: Parallel Method of Production Rules Extraction Based on Computational Intelligence. *Aut. Control Comp. Sci.* (2017) 51 pp. 215–223. DOI: 10.3103/S0146411617040058
12. Oliinyk, A., Subbotin, S., Lovkin, V., Leoshchenko, S., Zaiko, T.: Development of the indicator set of the features informativeness estimation for recognition and diagnostic model synthesis. *14th International Conference on Advanced Trends in Radioelectronics, Tele-*

communications and Computer Engineering (TCSET 2018), pp. 903–908, (2018). doi: 10.1109/TCSET.2018.8336342.

13. Oliinyk, A., Subbotin, S., Lovkin, V., Leoshchenko, S., Zaiko, T.: Feature Selection Based on Parallel Stochastic Computing," 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), pp. 347-351. Lviv (2018). doi: 10.1109/STC-CSIT.2018.8526729
14. Oliinyk A., Fedorchenko I., Stepanenko A., Rud M., Goncharenko D.: Evolutionary method for solving the traveling salesman problem. Problems of Infocommunications. Science and Technology : 5th International Scientific-Practical Conference PICST2018, Kharkiv, 9–12 October 2018, Kharkiv, Kharkiv National University of Radioelectronics, pp. 331-339 (2018). doi: 10.1109/INFOCOMMST.2018.8632033.
15. Tsoy, Y.R.: Evolutionary Algorithms Design: State of the Art and Future Perspectives. Proceedings of IEEE East-West Design and Test Workshop (EWDTW'06), Sochi, Russia, pp. 375-379 (2006).
16. Gruau, F.: Genetic synthesis of Boolean neural networks with a cell rewriting developmental process. In Proceedings of the International Workshop on Combination of Genetic Algorithms and Neural Networks (COGANN-92). Los Alamos, CA: IEEE Computer Society Press, pp. 55-74 (1992).
17. Moriarty, D., David, R., Miikkulainen, R.: Hierarchical evolution of neural networks. Evolutionary Computation Proceedings, pp. 428–433 (1998). doi: 10.1109/ICEC.1998.699793.
18. Greer, B., Hakonen, H., Lahdelma, R., Miikkulainen, R.: Numerical optimization with neuroevolution. Evolutionary Computation CEC '02(1), pp. 396–401 (2002). doi: 10.1109/CEC.2002.1006267.
19. Enforced Subpopulations (ESP) neuroevolution algorithm for balancing inverted double pendulum, <http://blog.otoro.net/2015/03/10/esp-algorithm-for-double-pendulum>.
20. Stanley, K.O., Miikkulainen, R.: Evolving Neural Networks through Augmenting Topologies. The MIT Press Journals, vol. 10, num. 2, pp. 99-127 (2002).
21. Whiteson, S., Whiteson, D.: Stochastic optimization for collision selection in high energy physics. Proceedings of the 19th national conference on Innovative applications of artificial intelligence, IAAI'07, vol. 2, pp. 1819-1825 (2007).
22. Hochreiter, S., Schmidhuber, J.: Long Short-Term Memory. Neural Computation, vol. 9, issue 8, pp. 1735–1780 (1997).
23. Tsoy Yu.R.: Development of a genetic algorithm for setting up an artificial neural network [Razrabotka geneticheskogo algoritma nastroyki iskusstvennoy neyronnoy seti]. Tomskiy politehnicheskii universitet (2004).
24. Papadimitriou, F.: Mathematical modelling of land use and landscape complexity with ultrametric topology. Journal of Land Use Science vol. 8(2), pp. 1-21 (2011).
25. Papadimitriou, F.: Artificial Intelligence in modelling the complexity of Mediterranean landscape transformations. Computers and Electronics in Agriculture, pp. 87-96 (2012).
26. Leoshchenko, S., Oliinyk, A., Subbotin, S., Gorobii, N., Zaiko, T.: Synthesis of artificial neural networks using a modified genetic algorithm. Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018), pp. 1-13 (2018). dblp key: conf/iddm/PerovaBSKR18
27. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Analysis of inrush currents of the unloaded transformer using the circuitfield modelling methods. Eastern-European Journal of Enterprise Technologies, vol. 3, № 5 (93), pp. 6-11 (2018). doi: 10.15587/1729-4061.2018.134248
28. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Enhancing the effectiveness of calculation of parameters for short circuit of threephase transformers using field simulation

- methods. *Eastern-European Journal of Enterprise Technologies*, vol. 4, № 5 (94), pp. 22-28 (2018). doi: 10.15587/1729-4061.2018.140236.
29. Shkarupilo, V., Skrupsky, S., Oliinyk, A., Kolpakova T.: Development of stratified approach to software defined networks simulation. *EasternEuropean Journal of Enterprise Technologies*, vol. 89, issue 5/9, pp. 67–73 (2017). doi: 10.15587/1729-4061.2017.110142
 30. Schmidhuber, J., Wierstra, D., Gagliolo, M., Gomez, F.: Training Recurrent Networks by Evolino. *Neural computation*. vol. 19(3), 757–779 pp. (2007). doi: 10.1162/neco.2007.19.3.757.
 31. Breast Cancer Coimbra Data Set, <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra>
 32. Bland, J.M., Altman, D.G.: Statistics notes: measurement error. *BMJ*. 312 (7047): 1654 (1996). doi:10.1136/bmj.312.7047.1654
 33. Oliinyk, A., Leoshchenko, S., Lovkin, V., Subbotin, S., Zaiko, T.: Parallel data reduction method for complex technical objects and processes. 9th International Conference on Dependable Systems, Services and Technologies (DESSERT'2018), 526–532 pp. (2018). doi: 10.1109/DESSERT.2018.8409184 IEEE Catalog number: CFP18P47-ART 978-1-5386-5903-8.
 34. Leoshchenko, S., Oliinyk, A., Subbotin, S., Zaiko, T.: Methods of semantic proximity extraction between the lexical units in infocommunication systems. 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T 2017), 7–13 pp. (2017). doi: 10.1109/INFOCOMMST.2017.8246137.
 35. Alsayaydeh, J.A., Shkarupilo, V., Hamid, M.S., Skrupsky, S., Oliinyk, A.: Stratified Model of the Internet of Things Infrastructure, *Journal of Engineering and Applied Science*, vol. 13, issue 20, pp. 8634-8638, (2018). doi: 10.3923/jeasci.2018.8634.8638.

Facial Expressions Analysis for Applications in the Study of Sign Language

Vladyslav Kuznetsov¹[0000-0002-1068-769X], Iurii Krak^{1,2}[0000-0002-8043-0785],
Olexander Barmak³[0000-0003-0739-9678], Anatolii Kulias¹[0000-0003-3715-1454]

¹Glushkov Cybernetics Institute, Kyiv, 40 Glushkov avenue, 03187, Ukraine,
kuznetsow.wlad@gmail.com, kulyas@nas.gov.ua

²Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska str., 01601, Ukraine,
krak@univ.kiev.ua

³National University of Khmelnytsky, 11, Instytutska str., 29016, Ukraine,
alexander.barmak@gmail.com

Abstract. Elements of information technology for analysis of facial expressions to apply in interactive study of sign language are described. The main elements of information technology, its structure and experimental implementation has been discussed. Analysis was carried out in order to identify how the classifier error rate depends on type of classifier, number of feature and teach set. Optimal constructs of classifiers were proposed, giving appreciable improvement of existing algorithms.

Keywords: information technology, sing language, modeling, identification, facial expressions

1 Introduction

Sign language is a tool for communication data transmission in communication of deaf and hard of hearing people [1]-[3]. According to statistics, significant percent of the world population has congenital hearing problems that make it difficult or impossible communication with hearing people in everyday speech. A lot of countries has a government program that determines the main directions of improving the conditions of deaf people, including a broad plan to study sign language among the general population that often in contact with deaf people. One step that aims to solve this problem is creation of educational programs using modern information technology.

The most complete solution to this problem is approached educational system [4] using animated 3D models, integrated tools and user interfaces of interactive communication with computer (see. Fig. 1).

Despite the fullness of the means implementing interactive learning sign language environment development of other means of reproduction and recognition of sign language, including facial expressions that indicate a signed language intonation, accent, logical and emotional meaning of sentence, is required. Modeling and recognition of elements of sign language [5],[6] is important because these elements in

spoken language hearing people passed by voice instead in signed language these aspects of speech transmitted in part by modifying the gesture (amplitude and smoothness of movements) that studies hearing people are not quite acceptable because gestures at high speed badly perceived. For ease of reading, the gesture is accompanied by facial expressions, which are typically used in communication deaf in a sign language (which is the same grammar speaking to the sequence of grammar inclusive), or in carriers of sign language with a very high level of sign language (there are analogue expression in spoken language).

It is important to obtain the characteristics of displays of emotion based on their accurate detection [7],[8] instead classification ones. In previous studies we have worked on the algorithm and methods that can be used in the information technology for analysis (identification) of facial expressions on a human face; in particular, we considered a sequence of receiving video clips of facial expressions, involving optical sensors (markers), mounted on a human face [4], the application of computer vision algorithms for numerical data describing the state change facial expressions over time numerical data analysis tools, including a single-layer perception and methods for reducing data dimensionality [9].

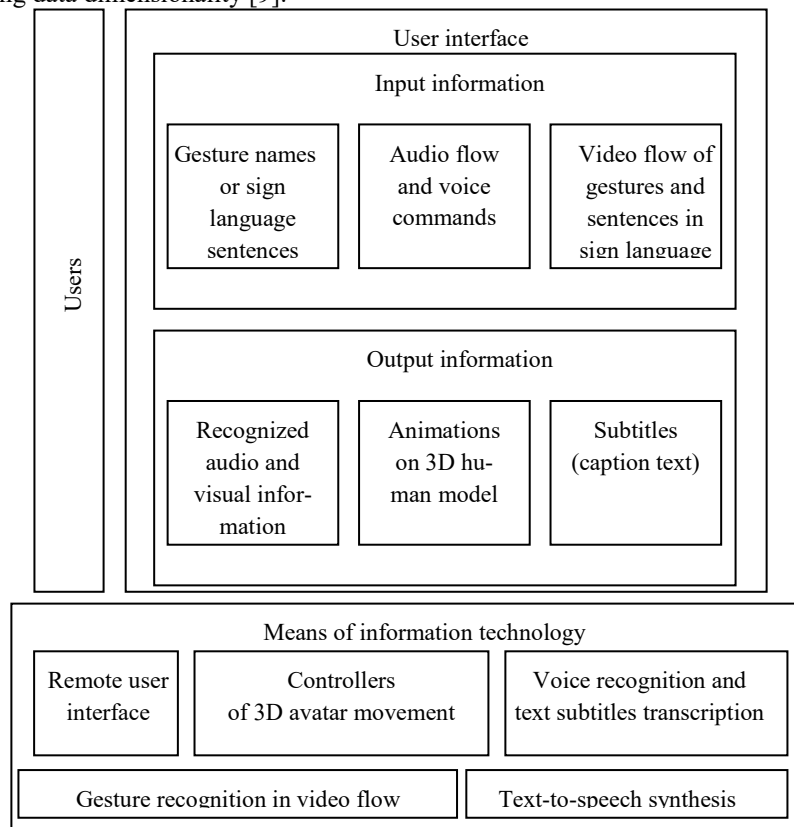


Fig. 1. Scheme information technology user interaction with the computer in the educational system to study sign language.

Results of the study [9] showed that there is a need to improve the efficiency of algorithms for problem of identification facial expressions changes over time; current implementation showed a low level of resolution of different classes of facial expressions. In order to improve current results suggest the following problem statement:

- propose a set of characteristics that can be used to identify facial expressions based on data obtained from the video stream;
- analyze existing means and methods of recognition of facial expressions on a set of characteristics that are applied in the analysis of facial expressions and other problems that reduce to identify changes in the characteristics of the object in time;
- develop a pilot information technology for identifying facial expressions and its software implementation;
- conduct a pilot test of the existing information technology to determine the best settings, algorithms, combinations thereof and the conditions under which it is possible to improve obtained previously results.

2 Methods of identification of the change over time characteristics

In the research [9] states that the problem of identifying the instantaneous state of the face does not require complex algorithms; perception or support vectors machines method are quite enough and give acceptable quality recognition - the order of 80-90% in the worst case - and uninformative small study sample small number of signs and to 99.5% at a rate sufficient training set (Table. 1). Instead, the task of examination of time dependent signal even for two classes of facial expressions is very complex and requires careful study. The same can be said when using the methods of reduction of data dimensionality while examining the single countdown (state of momentary) of facial expression. This indicates the similarity of sources of distortions that lead to identification errors.

Table 1 Identification accuracy for different samples of the algorithm in the existing technology.

Methods	Implementations and the number of samples in signals				
	Whole signal				1 countdown
	432	1080	2160	4320	36
PLA	64,8%	64,8%	64,8%	64,8%	98%
DCT+KLT	64,8%	64,8%	64,8%	64,8%	-
SVD+KLT	55%	58%	60%	62,5%	-
SVD	55%	58%	59%	61%	32%
KLT/FFT+KLT	52%	54%	57%	59,5%	-

In order to analyze changes of the object in time, during the duration of facial expression, the algorithm need to cope with varying intensity at the beginning and in the end, different amplitude, different waveform and signal distortion, different initial phase, the shape of the front and recession signal and varying duration.

In order to solve this problem (analysis of sequences of states) the most appropriate and expedient is to observe similar problems and algorithms for solving such problems which give a reasonable accuracy on a similar data.

Closest are several types of problems:

- phonograms-analysis (spectrograms);
- neural impulses, brain-waves, cardiograms, miohram analysis and so all;
- individual nerve cells process analysis.

In modern studies related to the analysis of phonograms [10] deep learning algorithms (with a teacher and without teacher), such as convolution neural networks, stacked denoised auto encoders, convolution auto encoders, deep belief networks, are used very often. Dynamic time warping method [11], which previously used to compare two close (similar) waveforms of varying duration, can also be taken in account as noticeable. Electrocardiograms and ion channels are analyzed using Karhunen-Loeve [12] and other integral transformations.

Methods of examination of time series studied in paper are:

- method of dynamic time warping, combined with some type of classifier (in particular, based on SVM, ANN and deep learning methods), using correlation between signals of corresponding sensors of different facial expressions and within correlation between signals from different sensors from each other in the same facial expression, as a net input. It also can help to test the hypothesis of generation of facial mimic signals by single nerve impulse (and their respective synchronization) through the evaluation of the correlation various mimic components;

- methods of analysis for spectrum signal including bringing output to the sound-track-like input of defining the parameters of spectral analysis, constructing spectral characteristics of sample signals and build classifier based on convolution neural network. This will test the possibility of signal convolution with fixed window dimension compared to dynamic time warping.

Methods which include integral transformation, especially Karhunen-Loeve expansion and discrete cosine transform [13] for signal analysis of fixed length (normalized in amplitude and time) will also be implemented in order to find out possibility/necessity of reduction of dimensionality of data and their applicability in connection with classification algorithms in task of analysis of the instantaneous state facial expressions.

3 Information Technology of the Facial Expressions Analysis

Information technology (Fig. 2) describes the processing of data containing the facial expression. In order to obtain a formal description of the process as a generalized block diagram we need to identify all the components of this process - the objects, properties and states of objects, relationships and dependencies at various stages of processing.

Input information. As input video objects are the examples of facial expressions (video stream) containing the sequence of states of facial expression (facial gesture or

mimic morpheme) during an interval of muscle or mimic activity (contraction - relaxation within a range of muscle activity).

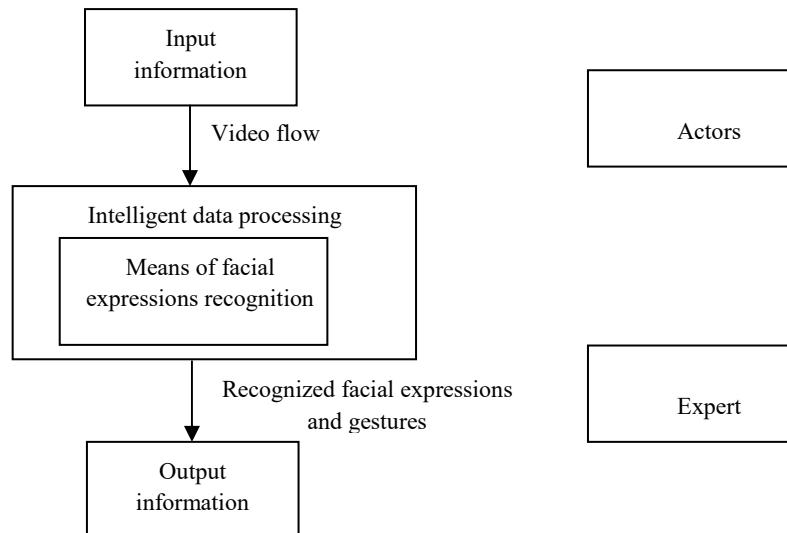


Fig. 2. Scheme information technology user interaction with the computer in the educational system to study sign language

Sampling frequency, signal number, information readout, signal number (ID), speaker identification are the main information signal characteristics.

The elements of input connections have a hierarchy: each countdown of a signal is part of a set of samples from a specific serial number, determining the order of appearance of a specific point in the video stream.

Processing the input data. The processing of the input data stream operates using algorithms to identify and track objects in video sequences to obtain coordinates of key points on each frame and the emerging key points coordinate data flow.

At this stage you must have to input constraints described by such terms, face orientation angle relative to the camera, sampling frequency, type of video compression, brightness of underlying objects on video, as well as object tracking algorithm restrictions, depending on the type and configuration of algorithms.

Getting the trajectory of the key points. The flow of data can be seen as a sampled characteristic points movement trajectory in some space formed by geometric transformations of space facial expressions with scale, rotation and shift of the base coordinate system.

Since each element (reference) of trajectory of movement of the key points is associated with a certain frame of video sequence, it is possible to group input on some characteristics derived by analyzing the trajectories of movement key points (for example, maximum activity, a minimum of activity, the mean interval values, etc.) on derived trajectory and original video sequence.

Data markup. This process describes the creation of some metadata that allow to group video frames and corresponding key points on the face. Automated processing

of video sequences and the corresponding data on the trajectory of motion was used to markup data. The output data is an ordered list of video sequence frame number and the corresponding notation (e.g. frame corresponding to the rest state and the frame corresponding to some activity) that allows to operate on specific instant face states that are most informative.

Normalization of data. For further processing of the ordered list of frames is necessary to bring all examples of movements to one scale. For this, all samples were ordered by a specific actor, thus data can be easily transferred to the new metric that will be comparable for all actors and allows changing scale of the mimic movements of each morpheme that is, obviously, specific to each actor.

Setting the functional converter. This process describes the iterative procedure for determining relationships (functional relationships) between input and some linguistic variables describing relation of inputs (video examples of facial expressions and its metadata) to a certain class. As for input, functional converter establish a priori class (beforehand known value of linguistic variable), which operates according to the motion trajectories key points on the face.

The most adaptive to this problem is classification algorithms and other algorithms that implement iterative procedure linkages.

Output data. The output data are linguistic variables corresponding to some input video sequence or reference throughout the interval samples. For such data we may need to set some attribute that describes the accuracy of functional dependence (identification) for some specific case (facial expression or mimic morpheme) by repeated execution of iterative procedure for multiple implementations of the data generated from the entire set of facial expressions video examples.

4 Experimental Realization of Information Technology for Identifying Facial Expressions

Database (DB) was created to implement storage means; a block diagram of this DB (Entity-Relationship model) is shown in Fig. 3.

We used database management system PostgreSQL [14] to implement database and the appropriate driver database to Oracle JVM (Java Virtual Machine) [15], which is connected to the executive files experimental program implementation.

Experimental software information technology was implemented in the environment IntelliJ IDEA [16] for Scala language [17] involving Java libraries and consists of several modules.

In order to provide the basic operations management we implemented by module and class **Main**. The module communicates with other modules to produce and output the converted data, including:

- input data** containing the trajectory of the markers on the actor's face;
- metadata** describing the characteristics of individual facial manifestations, including temporal range of activity, type of facial expression, identifier, actor who demonstrates mimic display, etc;

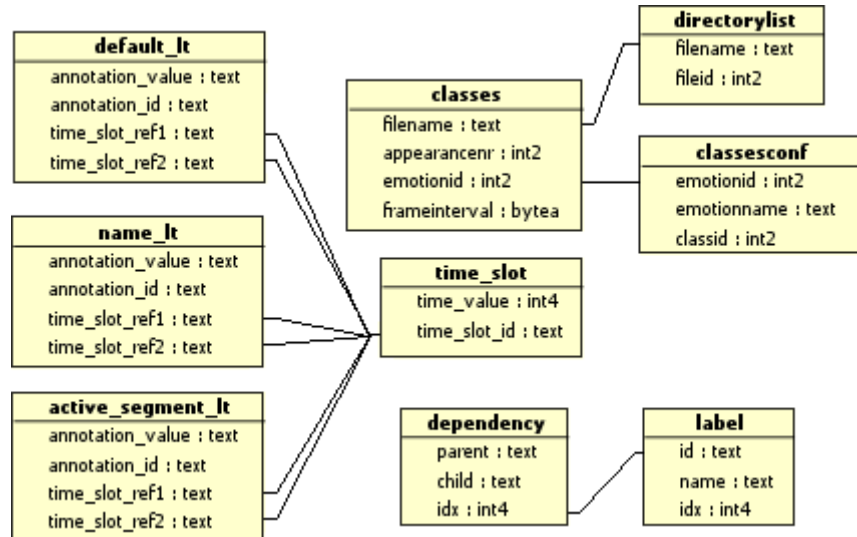


Fig. 3. Data base block diagram

- **transformed data** containing parameters obtained on the basis of input and removal of shift, scale and rotation distortions customized for each actor separately;
- **input and functional converter output data**, obtained by the rejection of the transformed data (may be rejected as fragments interval activity of facial expression, and the whole facial expression) and functional values of the dependent variable transformer;
- **coefficient setting of functional converter** that control facial expressions withdrawal rules that implements functional converter. To verify the configuration of functional converter (applicability coefficient settings for other data), its settings are checked for data that was rejected at the preliminary stage.

Object-oriented approach was offered to implement information technology; **InputReader** module, that implements reader input operations, consist of such classes: **Env** (file and means of reading the contents of the directory), **Marker** (information about the coordinates of each coordinate of the marker in the time and the means to import data containing the trajectory of motion markers), field **Data**, which contains an array unique samples of facial expressions, read from raw input, implemented as an array of objects of type **Marker**.

XML [16] metadata reading module (and their corresponding classes and methods) **ConfigReader** and **SQLQuery**, that implements JDBC-interface to the database PostgreSQL, were implemented.

Through the interface, implemented in the module, metadata video files stored in XML-container, and recorded in the database table. Database interface separately read two types of file metadata - time characteristics and a cloud of tags that describe the video files and structural links between them, respectively.

After reading metadata file, containing the time activity characteristics of facial expressions, the data are entered in the table (Fig. 3); data on time intervals is stored in the table **time_slot**, the names of basic facial expressions – in table **default_lt**, facial expressions names – in table **name**, time values (countdowns) of maximum activity (saturation of facial expressions) – in table **active_segment**. All three tables are combined in two key fields: **annotation_id** and **time_slot**.

After reading the tag cloud metadata file, module creates following tables: **item**, which contains information about the connections of individual tags from the tag cloud from video files, table **label**, containing the tag names, tables **dependency** and **hierarchy** containing links tags in the tag cloud respectively. All three tables are combined on a key field **id**, which contains a unique hex SHA-256 cloud tag hash id.

File structure that describes a list of video files is stored in table **directorylist**, links between basic and derivative mimic expressions – in table **classesconf**. This information can be manually adjusted from the outside interface that provides software implementation or database administration tools **pgAdmin** provided by ODBC PostgreSQL.

Database queries that perform union and intersection tables, aggregates data derived from two different sources XML-data. Aggregate table **classes** keep separate tracks, with such table labels as: time intervals of activity, the names of facial expressions, file names, names of actors, classes of facial expressions and other service information used in the experimental program implementation.

The next module (**ParameterMaker**) implements the basic methods for processing the coordinates of markers and form the parameters, namely "rejection" of affine distortion calculation movements markers algorithms determine the quantitative characteristics of secondary facial expressions and more. It also implements methods of calculating the average, maximum and minimum quantitative and qualitative characteristics of facial expressions.

Affine distortion rejection algorithm subtracts flipping movement of fixed point on the horizontal and vertical movement from relative (namely, non absolute) coordinates of other points. To calculate estimates of quantitative characteristics of facial expressions, **ParameterMaker** also implements appropriate methods of calculating them based on the distance between the eyes of the actor. The result is a dimensionless ratio of this distance in the range from 0 to 1.

Appropriate scaling is performed to determine the quantitative characteristics of facial expressions. In order to take into account the differences in facial expressions forms in different people, each mimic expression determined as a percentage of the maximum values for the individual.

Data obtained by the module **ParameterMaker** and corresponding metadata transmitted to the input module **TrainTestDataMaker**. Depending on the type of transducer and the type of study facial expressions, of the total volume data is shown in part comprising representatives of facial expression, a particular actor or interval activity mimic activity (e.g. muscle contraction or retraction), which are divided into two sets of data - set that is used to configure the functional converter and to test coefficients set-functional converter. Elements from the original dataset are selected using

a random number generator; the amount of data in each of the data sets is defined by a certain constant that controls random number generator.

Next (***TrainTestParametersMaker***) module implements functional transformation rules that describe the output of facial expressions based on input data from the module ***TrainTestDataMaker***. The transformation is implemented using algorithms library MLLib, with the involvement of these constructs: single- and multilayer perceptron (PLA) with hypotheses based on linear and radial basis functions (RBF), stacked denoising autoencoders (SAE), means reducing the dimension (convolution) data, including: cosine transform (DCT) and Karhunen-Loeve (KLT). The software implementation use different combinations of constructs, which were limited to the following algorithm: 1) single or multistep reduction the dimensionality of data and 2) reduced dimension processing by using the methods of linear and nonlinear classification.

Built-in library of algorithms MLLib was used to set the coefficients properly. After setting the coefficients, in order to verify coefficients of each functional configuration of the converter, its settings are checked for test data.

The influence of distortion of scale of a small amount of data was checked by performing iteratively a number of trials on test data applying coefficients of functional transformers on train on test data, and comparing values of errors, that describe the output on trained data comparing with known (a priori) output value.

5 Experimental Tests of Information Technology for Identifying Facial Expressions

Input data. Test information technology tests were carried out on a set of 300 facial expressions taken from five different faces of actors, which included from 2 to 14 representatives of basic facial expressions. In addition, for a number of experiments that included the study of a single countdown, we did not used whole interval of mimic activity of facial expressions; these intervals covered the mimic expression in a state of saturation.

Samples of facial expressions, or single samples of the proposed sets during processing module ***TrainTestDataMaker*** were broken at a ratio of 20% and 80% using the integrated random number generator.

Description of trials. Variants of functional implementations converters were studied under different conditions. There were 5 different tests conducted on experimental implementation.

During the *first trial* we had to reveal if there is some specific information, that appear only when observing time changes of mimic expression. This test was conducted on individual reference data, and on the interval of activity of facial expressions over time. Tests showed that the available data on individual samples (countdowns) and key expansion coefficients, obtained by reducing the dimensionality of data (direct cosine transform), correlated. Therefore, for the data type "countdown" we observed relatively low value of the error for both proposed methods. However, these algorithms have limits in creation of complex hypotheses based on the input data, that,

possibly may incorrectly perceive facial expression of low intensity (e.g. 10-30% of power in respect to “weakest” sample in teach set).

During the *second trial* we had been tested whether the reduction of the dimension of the data (both single point, and at intervals of integral activity of facial expressions over time) is applicable. This test was necessary to establish the causes of great recognition errors obtained in previous studies [9], while using methods of dimensionality reduction of the data.

Tests showed (Fig. 4): 1) in order to classify the data obtained from the dimensionality reduction methods using classification methods, they require a large number of major expansion coefficients (much greater than 3), indicating that the eigenvectors with low energy have significant informational contribution applied to classification algorithms (Fig. 4.a); 2) individual samples (countdowns) require a smaller amount of additional vectors than for the entire interval of mimic activity - at least 4 vs 20-40 respectively.

This indicates that 1) the effectiveness of classification algorithms strongly depends on the number of expansion coefficients and methods of reduction of dimensionality of data respectively 2) large variation in the data and its phase-frequency characteristics makes unclear the dependency of the value of error on quantity (and even order) of basis functions.

During the *third trial* we conducted an experiment in order to figure out the minimum amount of sample data (teach data) that is used to configure the function converter and, in particular, to test the effectiveness of each constructs (e.g. DCT+SVM, SVD+ANN and so all) only case of "single signal" (as showed the previous test, the individual samples values and key ratios of discrete cosine transform correlated and, therefore, the inputs of classifiers were relatively homogeneous).

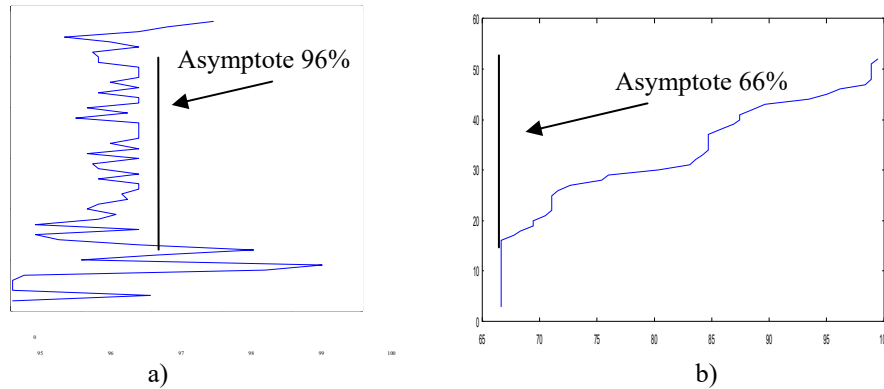


Fig. 3. Dependency of value of error on the number of basis functions for: SVM method - a) and single-layer neural network - b). Here vertical axis depict the number of basis functions (3-63), horizontal axis - the size of error (lower limit for errors 99.5%)

Tests have shown that even for the 103 order of data, one can find abnormal emissions (Fig. 5), reflecting one of the many bad implementations (e.g. iterations) of training on specific algorithm on exact data set. One can argue that the size of the

training set, sufficient to obtain results of satisfactory accuracy recognition (95%) in the range of 70-700 items, with no sample size is crucial in size accuracy, and most items are included in each class training and test samples and adaptability features to the elements of learning sample, which builds hyper resolution. This indicates a significant variance within the class that causes class transitions errors (e.g. element of one class appears to be element of another class).

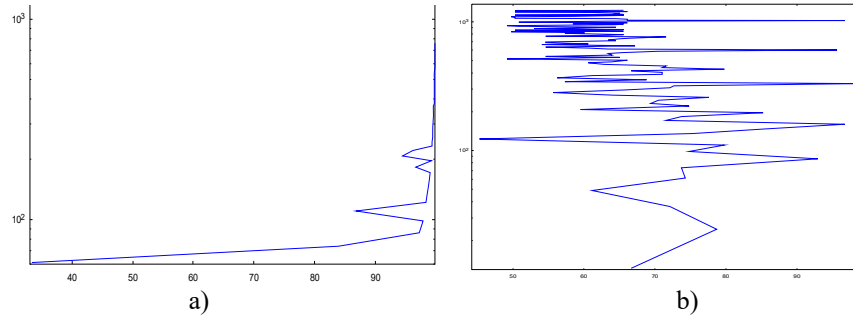


Fig. 4. The learning curve for the two implementation of classifiers: single layer neural network - a) and SVM based on tanh kernel - b). Logarithmic scale: horizontal axis depicts the volume of training set (101-103), the vertical axis - the value of errors (right asymptote ~ 99.5%, left asymptote ~ 40%)

During the *fourth trial* we had to suggest one way to improve the recognition accuracy of the neural network (which showed a relatively small number of emission value errors) on net input, with not applied dimension reduction. One good solution is good initial weights selection; deep learning methods answer this requirements. As a result of testing, we had found that the stacked autoencoders are able to improve learning accuracy of neural network with one hidden layer neurons compared with vanilla neural network of same structure.

During the *fifth trial* we had to figure out how to reduce the abnormal emissions in learning curves of learning algorithms enrolled on data, obtained after reduction of data dimension. One way to improve the algorithms is eliminating the time scale, frequency and phase mismatch between signals from different samples of facial expressions, without removing information about temporal characteristics. In order to do this, we compared the characteristics of each of the facial expressions together using Dynamic Time Warping (DTW). Values of independent pair DTW-correlation form (i.e. fill) the characteristic vector signal in the form of a square matrix (Fig. 5 a)). If necessary, the thinning is applied on matrix (Fig. 6, b). As a result of experiments, we had found that the most applicable algorithm that can be used in conjunction with thinned characteristic DTW vectors is SVM.

Therefore, as a result of all trials described above, we obtained the optimal set of constructs (combinations of algorithms) that give relatively high rates of accuracy of identification. Table 2 lists these constructs and optimistic value of precision out-of-sample for the best scenario (the largest sample size, the best set of basis functions,

the best implementation of a specific algorithm, etc.); numerator depicts the accuracy, denominator – number of eigen functions and size of teach vector respectively.



Fig. 5. Matrix of DTW pair correlation for samples of different facial expressions without thinning - a) and two-dimensional thinning using DCT - b).

Table 2 Identification accuracy for various implementations of the algorithms

Classifier type	Signal dimensions	
	1x20 (1 countdown)	20,30,60,120,...x20 (whole signal)
SVM (RBF)	99,5%/-	-
ANN MLP	99,5%/-	75%/-
KLT(SVD) + SVM (RBF)	99,5%/1x4	99,5%/1x18
DCT + SVM (RBF)	-	99,5%/3x20
SDAE (400x20x2) + ANN MLP	-	99,5%/-
DBN (400x20x2)	-	75%
Convolutional NN 12c-2s-6c-2s	-	75%
DCT+SVM (RBF)	-	99,5%/7x20
DCT+SVM (Linear)	-	99,5%/2x20
DTW + SVM	-	87,5%/-
DTW + ANN/DBN/SDAE/CNN	-	87,5%/-
DTW + DCT + ANN/DBN/SDAE/CNN		87,5%/10x19
DTW + DCT + SVM (RBF)	-	99,5%/7x19
DTW + DCT + SVM (Linear)	-	99,5%/2x19

6 Conclusion

As a result of experiments we greatly improved existing technology [9], using relatively the same data examples - facial expressions and facial expression components captured by means of optical motion capture of control points using optical markers, particularly for methods of identification changes facial expressions over time. Several different structures or combinations of the most effective methods of feature extraction and classification of facial expressions were applied, making possible to implement an effective system for identifying facial expressions. We analyzed the impact of the number of inputs, the training set on amount and type of classification errors of 1 and 2 race, which allowed to offer new constructs of multiple classifiers.

In particular, the proposed algorithms and the neural network in connection with stacked denoising autoencoders, using eigenvalues of decomposition of output data, in particular, in task of identification of the temporal changes of micro expressions allowed to achieve higher efficiency (of the order of 99.5%) .

Research, that was carried out on the proposed algorithms, helped to improve the preliminary results and to established causes of poor quality recognition. Some solutions were proposed as well - a set of structures of multiple classifiers, that are alternative for methods for dimensionality reduction, based on three simple steps: 1) calculating DTW inner-correlation matrix; 2) applying thinning; 3) using classifier on thinned matrix, which can be applied in classification of other discrete time-dependent and correlated multiple channel signals that have different duration (number of countdowns).

References

1. Stokoe. W.C. Jr.: Sing language structure: An outline of the visual communication systems of the American deaf. Univ. of Buffalo (1960)
2. Sing Languages. Ed. by D. Brentari. Cambridge: Cambridge University Press (2010)
3. Smith, R., Morrissey, S., Somers, H.: HCI for the Deaf community: developing human-like avatars for sign language synthesis. In: Proceedings of the 4th Irish Human Computer Interaction Conference. Dublin (iDCI2010), Dublin, Ireland, 2-3 September 2010, pp.129-136 (2010)
4. Kryvonos, I.G., Krak, I.V.: Modeling human hand movements, facial expressions, and articulation to synthesize and visualize gesture information. *Cybernetics and Systems Analysis*. 47(4): 501-505 (2011)
5. Kryvonos, I.G., Krak, I.V., Barmak, O.V., Shkilniuk, D.V.: Construction and identification of elements of sign communication. *Cybernetics and Systems Analysis* 49(2): 163-172 (2013)
6. Krak, I.V., Kryvonos, I.G., Barmak, O.V., Ternov, A.S.: An Approach to the Determination of Efficient Features and Synthesis of an Optimal Band-Separating Classifier of Dactyl Elements of Sign Language. *Cybernetics and Systems Analysis* 52(2): 173-180 (2016)
7. Brunelli, R., Poggio, T.: Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(10): 1042-1052 (1993)

8. Martinez, A., Du, S.: A model of perception of facial expressions of emotion by human: research overview and perspectives. *Journal of Machine Learning Research* 13: 1589-1608 (2012)
9. Kryvonos, I.G., Krak, I.V., Barmak, O.V., Ternov, A.S., Kuznetsov, V.O.: Information Technology for the Analysis of Mimic Expressions of Human Emotional States. *Cybernetics and Systems Analysis* 51(1): 25-33 (2015)
10. Lee, H., Largman, Y., Pham, P., Ng, A.Y.: Unsupervised Feature Learning for Audio Classification Using Convolutional Deep Belief Networks, In: *Proceedings of Conference on Neural Information Processing Systems (NIPS 2009)*, Vancouver, Canada, 7-10 December 2009, pp. 1096-1104 (2009)
11. Al-Naymat, G., Chawla, S., Taheri, J.: Sparse DTW: A novel approach to speed up Dynamic Time Warping, *Australasian Data Mining*, Melbourne, Australia, ACM Digital Library, pp. 117-127 (2009)
12. Chumakov, A.G., Kurashov, V.N.: Karhunen-Loeve Basis Synthesis for Grate Capacity Signal Performance in Optical Processors, *SPIE Proc.*, 21-08, pp. 338-342 (1993)
13. Loeffler, C., Ligtenberg, A., Moschytz, G.: Practical Fast 1-D DCT Algorithms with 11 Multiplications, In: *International Conference on Acoustics, Speech, and Signal Processing*, Glasgow, UK, 23-26 May 1989, pp. 988-991 (1989)
14. PostgreSQL: The World's Most Advanced Open Source Relational Database, <https://www.postgresql.org>, last accessed 2019/04/09.
15. Java SE at a Glance <http://www.oracle.com/technetwork/java/javase/overview/index.html>, last accessed 2019/04/10.
16. JetBrains, <https://www.jetbrains.com/idea/>, last accessed 2019/04/10.
17. The Scala Programming Language, <http://www.scala-lang.org>, last accessed 2019/04/10.

Fingerspelling Alphabet 3D Modeling and Recognition Base on CNN Technology for Cross Platform Applications

Serhii Kondratiuk¹[0000-0002-5048-2576], Iurii Krak^{1,2}[0000-0002-8043-0785], Olexander
Barmak³[0000-0003-0739-9678], Anatolii Pashko¹[0000-0001-6944-8477]

¹Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska str., 01601, Ukraine
{sergey.kondrat1990,aapashko}@gmail.com, krak@univ.kiev.ua

²Glushkov Cybernetics Institute, Kyiv, 40 Glushkov avenue, 03187, Ukraine
krak@nas.gov.ua

³National University of Khmelnytsky, 11, Institutes str., 29016, Ukraine
barmakov@khnu.km.ua

Abstract. The technology, which is implemented with cross platform tools, is proposed for modeling of gesture units (fingerspelling alphabet) of sign language, animation between states of gesture units with a combination of gestures (words). Implemented technology simulates sequence of gestures using virtual spatial hand model and performs recognition of dactyl items from camera input. With the cross platform means technology achieves the ability to run on multiple platforms without re-implementing for each platform.

Keywords: cross platform, sing language, modeling, recognition, convolutional neural network.

1 Introduction

Modern hardware is able to collect information fast and almost without restriction, process data both in cloud computing (model, which provides a universal, easy access on demand through the network to the virtual cluster computing resources) [1] and locally on the device, and through data channel processing results are returned to the user. All this is also true for sign language [2], [3]. Sings can be stored and reproduced via a variety of devices and platforms, stationary or mobile, high performance or energy efficient. The actual problem is the reproduction of sign language on all these platforms, for further usage by people with hearing disabilities in particular and everyone in general. Deployment of a single unified technology on various platforms (android, ios, windows, linux, web) without need to port it or to implement it under each platform is a major problem.

One way of solving the stated problem of visualization and reproduction of sign language is cross platform software development. Unlike single-platform technologies

that operate only on a specific platform under which they were developed, “cross platform software provides the ability to perform on more than one platform with identical (or nearly identical) functionality” [4]. The term “platform” in this context may refer to one of or a combination of several definitions: 1) the type of operating system (such as Microsoft Windows, Mac OS X, Linux, Solaris, Android, iOS); 2) processor type (such as x86, PowerPC, ARM); 3) the type of hardware (e.g., main-frame, workstation, personal computer, mobile device) [4]. Cross-platform technologies are on a par with the platform independent technologies (those that can operate on any platform, such as Web application) [4] and cross-platform virtual machines (technologies that support individual processes or systems, depending on the level of abstraction at which is virtualization) [5].

In this article the proposed solution of the problem is via cross platform development, taking into account characteristics of different classes of devices (such as hardware, CPU power, amount of memory, presence on the Internet) and setting the number of polygons of the three-dimensional hand model and gesture animation step. Given paper is a progress of author's investigations [2], [6], [7]. Gesture modeling and gesture recognition is performed via cross platform means as a part of proposed communication technology.

2 Approaches for modeling and recognition of sign language and problem statement

Sign modeling is a problem that is considered both independently and as part of the problem of modeling and recognition of gestures and thus as a technology learning and evaluating sign language. One of the systems to display the sign language is American Sing Language Online Dictionary [8], which consists of a video database of words and phrases displayed via sign language. These developments were involved in a number of commercial agencies [7], but the systems they propose are configured to pre-determined number of gestures, and therefore do not solve the problem of modeling sign language. Also all of them lack functionality of gestures recognitions, thus not allowing to evaluate quality of sing language performed by a user.

Creating a model hand is the first step in the task of sing language modeling. In their work [10], authors analyze existing approaches of hand modeling, which are divided into two main groups: spatial and temporal. Former consider the characteristics of different positions for the hand gestures, while latter refer to the description of the dynamics of gestures. Modeling hands in the spatial area can be completed in two and three dimensions.

In [11] proposed system by authors is able to simulate sign animations for a given text. As a part of this system a statistical model is used to analyze input text and generative algorithm is used when creating the appropriate simulated kinematics of sign animations. Within the article, the authors have provided ANVIL tools for input text annotation, gesture generator NOVA, and DANCE library developed in [12] is used for gesture animation. The system is built on the Microsoft Windows platform and

x86 processor. In [11]-[18] authors discuss the modeling of virtual character for spatial reproduction of sign language. The training system is based on Microsoft Windows platform and x86 processor. Gesture recognition for mobile platforms is developed in [19], but gesture modeling on mobile devices is not performed.

The proposed technology should perform modeling of sign units [13], [14] of sign language, and reproduce animation of gestures structures (words, sentences) via state transitions between shown units using spatial virtual model hand. The proposed technology should perform recognition of sign language based on camera input from the device in order to evaluate sign language performed by user. The technology should be a combined solution for learning sign language via gesture modeling and recognition.

Technology should solve the problem of running on existing platforms using cross platform development without implementing the functionality for each platform separately. The effectiveness of the proposed approach is shown in building cross platform technology for modeling and recognition of Ukrainian dactyl (fingerspelling) alphabet.

3 Proposed methods for gestures modeling

To address the modeling of sign language and perform animation of sign structures using spatial virtual model hand the cross platform technology based on cross platform framework Unity3D [20] is proposed. Cross platform framework Unity3D is also used for the user interface, both libraries and technology are implemented with programming language C#. Proposed tools can solve the problem of running the technology on multiple existing platforms. The novelty of the proposed technology is that it is cross platform and has customizable level of polygons for three dimensional hand model and animation step for gesture transitions. This allows to run proposed technology without changes on multiple platforms (different types of processors, operating systems and hardware).

Advantage of cross platform technology over technologies developed for a single platform is that there is no need to modify or re-implement the functionality already available for other platforms (porting) [4], which speeds up the process of developing and deploying technologies, and increases the number of potential users. The advantage of cross platform technology over cross platform virtual machine emulators is performance speed and absence of necessity to install additional software (software dependencies).

The core of the technology is composition of three cross-platform modules (Fig. 1): three dimensional hand model (which is implemented with cross platform framework Unity3D), user interface (implemented also with cross platform framework Unity3D) and gesture recognition module (implemented with cross platform framework Tensorflow [21]). Core functionality is implemented with C# and Python and runs on desktop OS (MacOS, Linux, Windows) and on mobile OS (Android, iOS).

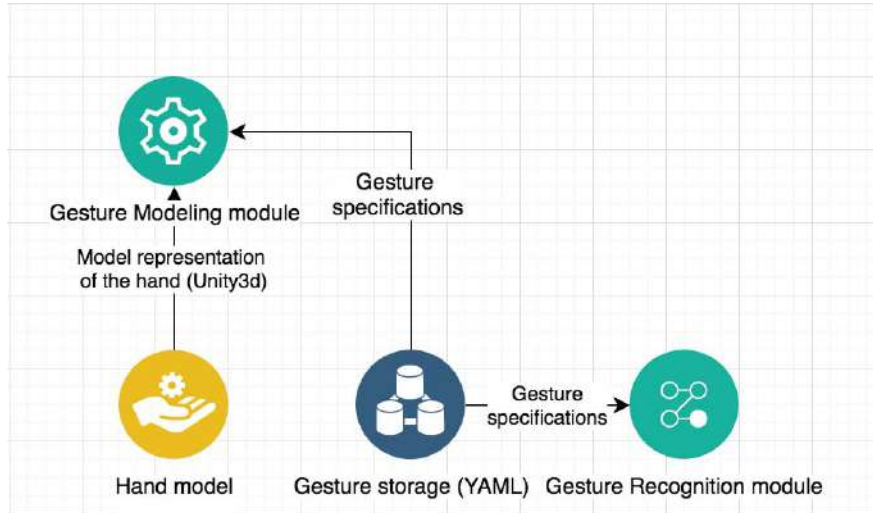


Fig. 1. Diagram of infological model of cross platform gesture communication technology

Hand model module is cross platform and provides hand model representation for gesture recognition module. Hand renderer receives hand model representation and gesture specifications from gesture storage module, and provides a high-polygon rendered hand model. Gesture learning module and gesture modification module are implemented with cross platform Unity3D, both taking as input results of hand model renderer. Gesture modification module provides updated gesture specifications and transmits them to gesture storage. Gesture recognition module is proposed to be implemented with Tensorflow framework and receives as input hand model, gesture specifications and input from camera

The hand model which is built in gesture modeling module has 27 bones, 8 of the bones are in wrist, 3 are in the thumb (one metacarpal and 2 phalanx) and 4 metacarpus and 12 phalanges are in other fingers. Each bone is connected to the other through different types of joints.

Designing your own cross platform engine for simulating the hand is non-trivial task, thus as the core technology for modeling three-dimensional hand model and gesture animations between morphemes cross platform framework Unity3D was selected. Unity3D framework is able to effectively reproduce a realistic hand model which consists of more than 70,000 polygons (Fig. 2).

Based on the anatomy of the hand within Unity3D hand model was developed with 25 degrees of mobility, four of them located in the metacarpal-carpal joint, to the little finger and thumb to provide movement palm. The thumb has 5 degrees of freedom, middle and index fingers have 4 degrees of freedom (metatarsophalangeal joint with two degrees of mobility, and the distal and proximal interphalangeal joints each have one). To preserve the gesture YAML format was selected [22]

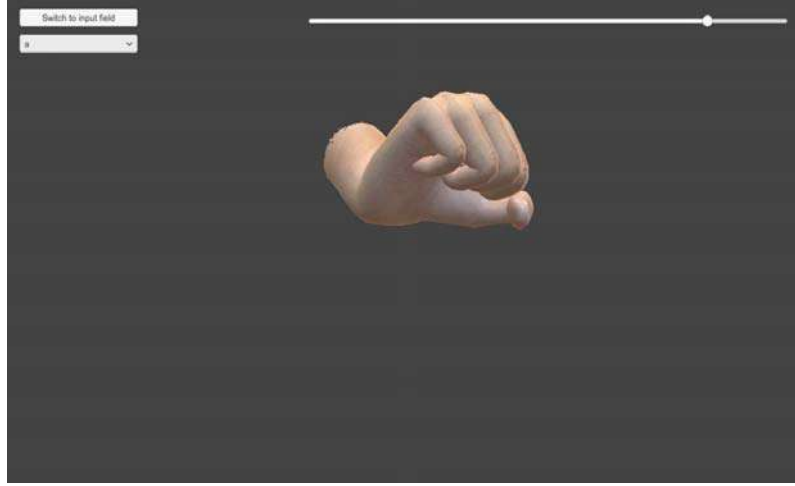


Fig. 2. 3D model of gestures under iOS platform

4 Proposed methods of gestures recognition

Gesture learning and gesture recognition modules, developed with cross platform tools (frameworks based on Python, C++) can be embedded into information and gesture communication cross platform technology. Multiple approaches were considered as an approach for gesture recognition. Automatic sign language recognition can be approached similarly to speech recognition, with signs being processed similar to phones or words. Conventionally, sign language recognition consists of taking an input of video sequences, extracting motion features that reflect sign language linguistic terms, and then using pattern mining techniques or machine learning approaches on the training data. For example, Ong et al. propose a novel method called Sequential Pattern Mining (SPM) that utilizes tree structures to classify signs [23].

Convolutional Neural Networks (CNNs) have shown robust results in image classification and recognition problems, and have been successfully implemented for gesture recognition in recent years. In particular, deep CNNs have been used in researches done in the field of sign language recognition, with input-recognition that utilizes not only pixels of the images. With the use of depth sense cameras, the process is made much easier via developing characteristic depth and motion profiles for each sign language gesture. Multiple existing researches done over various sign languages show that CNNs achieve state-of-the-art accuracy for gesture recognition [24],[25].

Convolutional neural networks have such advantages: no need in hand crafted features of gestures on images; predictive model is able to generalize on users and surrounding not occurring during training; robustness to different scales, lightning conditions and occlusions. Although, selected approach has couple of disadvantages, which may be overcome with a relatively big dataset (1,000 images for each gesture, among

more than 10 people of different age, sex, nationality and images taken under different environment conditions and scales): need to collect a rather big and labeled gesture images dataset; black-box approach which is harder to interpret. Usage of cross platform neural network framework such as Tensorflow allows to implement gesture recognition as a cross platform module of proposed technology and serve trained recognition model on server or transfer it to the device.

For experiment there was collected a dataset with Ukrainian dactyl language letters (fingerspelling alphabet). Each gesture consists of 1000 sample images, and 50 different people hands were showing gestures, with distribution of 70% male and 30% female hands. Different light conditions were used (with distribution of 20 % images in bad light conditions, 30% in mediocre light conditions and 50% in good light conditions). About 10% of images were distorted with noise and blur.

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
5×	Conv dw / s1 $3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
	Conv / s1 $1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

Fig. 3. MobileNet body architecture [26]

MobileNet[26] architecture (see Fig. 3) was used as a CNN architecture. It has multiple advantages, such as good trade-off on accuracy and performance, especially on mobile devices, which are aimed to use, as the technology is cross-platform. The MobileNet model is based on depth wise separable convolutions which is a form of factorized convolutions which factorize a standard convolution into a depth wise convolution and a 1×1 convolution called a point wise convolution. For MobileNets the depth wise convolution applies a single filter to each input channel. The point wise

convolution then applies a 1×1 convolution to combine the outputs the depth wise convolution. A standard convolution both filters and combines inputs into a new set of outputs in one step. The depth wise separable convolution splits this into two layers, a separate layer for filtering and a separate layer for combining. This factorization has the effect of drastically reducing computation and model size.

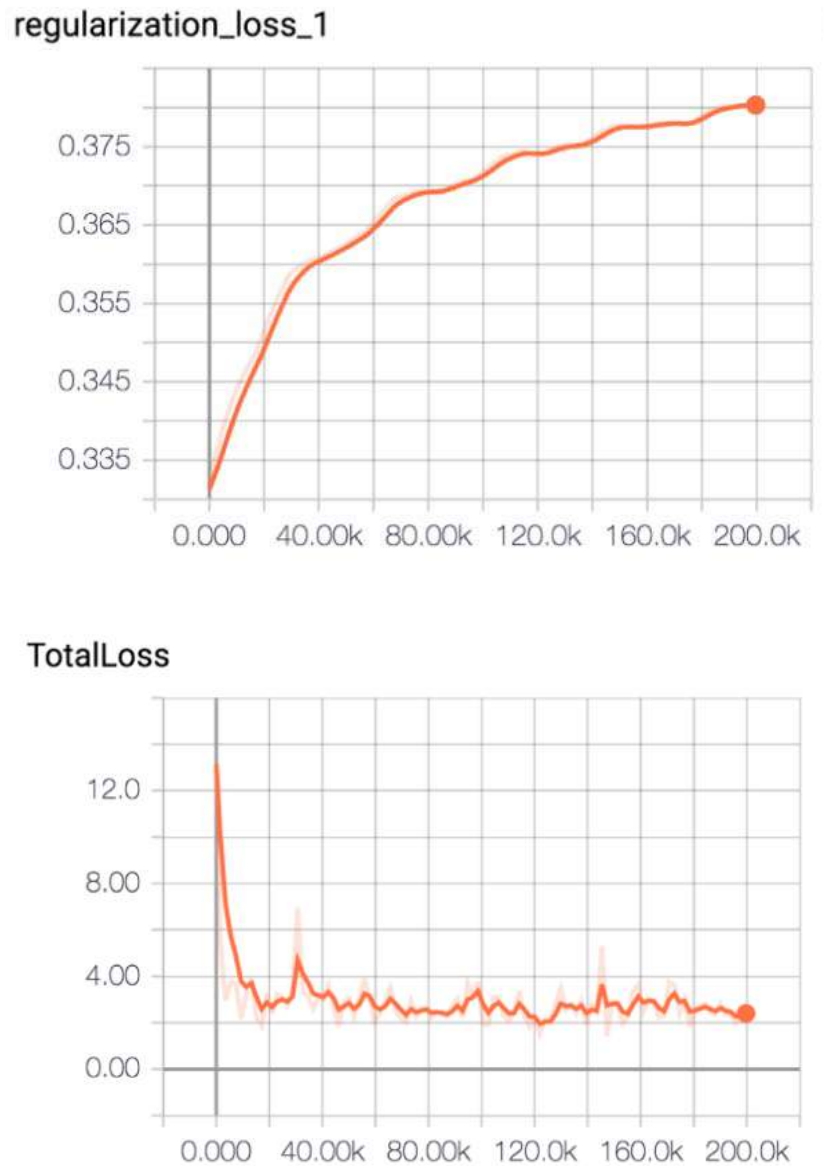


Fig. 4. Curves of the neural network optimization

Process of training MobileNet network for gesture detection takes ~ 200.000 iterations, which is approximately 10 epochs. On the Fig. 4 shows curves of how the neural network is optimized.

On the Fig. 5 shows example of how the proposed technology detects specific gesture from Ukrainian dactyl and draws a bounding box over a detected gesture.



Fig. 5. Example of fingerspelling letter (b)

5 Application of cross platform tools

Due to selected cross platform implementation tools, the proposed technology solves the problem of executing on multiple platforms without the implementation under each platform separately.

Software offered and used in the implementation of information technology is cross-platform and operates unchanged regardless of operating system (Windows, Linux, Android, iOS), CPU type (x86, arm), and the type of hardware (mobile or stationary device).

With its cross platform build system Unity3D it is possible to create applications for each platform without porting or changing the original code.

As there are no specific hardware requirements for information technology for modeling sign language, there are objective obstacles for performance speed of older generations devices. To overcome this problem, the following adaptive approach to information technology was proposed[6].

Further modules implementation will leverage from existing cross-platform technology. Gesture learning and gesture recognition modules, developed with cross platform technologies (Python, Tensorflow) will be embedded into information and ges-

ture communication cross-platform technology. In case of the mobile app (iOS, Android) or application on the device with a stationary operating system (Windows, Linux), during installation on the device, information technology analyzes the existing hardware and, depending on its capacity, conducts a series of adjustments: 1) number of polygons of the hand model changes to priority for performance speed; 2) during rotation hand model changes pitch angle at which it rotates, with priority for speed. If the available hardware does not meet the minimum requirements of information technology, the user is given the recommendation to choose “online” mode, in which the calculation is not performed on hardware.

6 Conclusion

The proposed technology is built with cross platform tools for gesture modeling, gesture transitions animation and gesture recognition. The technology uses virtual spatial model of hand. With the help of cross platform development, the technology solves the problem of execution on the existing multiple platforms without implementing functionality under each platform separately. Thus, it was shown the effectiveness of the technologies built using cross platform tools, for example modeling and recognition elements of dactyl Ukrainian alphabet sign language. Information and gesture communication technology was developed with further scaling capabilities in mind for gestures of other languages alphabets.

To implement this idea, the validation mechanism of new gestures to the common database can be applied. Cross platform information and communication technology and standardized protocol and data format (YAML) allows a range of solutions for remote computing using cloud computing, Web servers, local servers using a single sign database PostgreSQL [27]. The gesture communication technology can be augmented with other cross platform modules, such as gesture recognition and gesture learning modules.

References

1. Mell, P., Grance, T.: The NIST Definition of Cloud Computing (Technical report). National Institute of Standards and Technology: U.S. Department of Commerce. Special publication 800-145 (2011) doi:10.6028/NIST.SP.800-145.
2. Kryvonos, I.G., Krak, I.V., Barmak, O.V., Kulias, A.I.: Methods to Create Systems for the Analysis and Synthesis of Communicative Information. *Cybernetics and Systems Analysis* 53(6): 847-856 (2017)
3. Stein, D., Dreuw, P., Ney, H., Morrissey, S., Way A.: Hand in hand: automatic sign language to English translation. In: *Proceeding of the 11th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI 2007)*, Skovde, Sweden, 7-9 September 2007, pp. 214-220 (2007)
4. The Linux Information Project, Cross-platform Definition [<http://www.linfo.org/cross-platform.html>] (2005)
5. Smith, J., Nair, R.: The Architecture of Virtual Machines. *Computer* 38(5): 32-38 (2005)

6. Krak, I., Kondratiuk, S.: Cross-platform software for the development of sign communication system: Dactyl language modelling. In: Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2017, Lviv, 05-08 September 2017, vol. 1, pp. 167-170 (2017)
7. Kryvonos, Iu.G., Krak, Yu.V., Barchukova, Yu.V., Trotsenko, B.A.: Human hand motion parametrization for dactylemes modeling. *Journal of Automation and Information Sciences* 43 (12): 1-11 (2011)
8. ASL Sing language dictionary [<http://www.signasl.org/sign/model>] (1999)
9. Apple Touchless Gesture System for iDevices [<http://www.patentlyapple.com/patently-apple/2014/12/apple-invents-a-highly-advanced-air-gesturing-system-for-future-idevices-and-beyond.html>] (2014)
10. Rafiqul, Z. K., Noor, A. I., Natarajan, M., et al : Comparative study of hand gesture recognition system: SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 06, pp. 203-213 (2012)
11. Neff, M., Kipp, M., Albrecht, I., Seidel, H.-P.: Gesture Modeling and Animation by Imitation. MPI-I-2006-4-008 (2006)
12. Shapiro, A., Chu, D., Allen, B., Faloutsos, P.: Dynamic Controller Toolkit [http://www.arishapiro.com/Sandbox07_DynamicToolkit.pdf] (2005)
13. Kryvonos, Iu.G., Krak, Iu.V.: Modeling human hand movements, facial expressions, and articulation to synthesize and visualize gesture information. *Cybernetics and Systems Analysis* 47(4): 501-505 (2011)
14. Kryvonos, Iu.G., Krak, Iu.V., Barmak, O.V., Shkilniuk, D.V.: Construction and identification of elements of sign communication. *Cybernetics and Systems Analysis* 49(2): 163-172 (2013)
15. The 3D Biomechanics Data Standard [<https://www.c3d.org/>]
16. Holte, M.B., Moeslund, T.B.: Gesture recognition using a range camera, Technical Report. CVM-07-01 (2007)
17. Krak, Yu.V., Golik, A.A., Kasianiuk, V.S.: Recognition of dactylemes of Ukrainian sign language based on the geometric characteristics of hand contours defects. *Journal of Automation and Information Sciences* 48(4): 90-98 (2016)
18. Krak, I.V., Kryvonos, I.G., Barmak, O.V., Ternov, A.S.: An Approach to the Determination of Efficient Features and Synthesis of an Optimal Band-Separating Classifier of Dactyl Elements of Sign Language. *Cybernetics and Systems Analysis* 52(2): 173-180 (2016)
19. Raheja, J. I., Sadab, A.S., Chaudhary, A.: Android based portable hand sign recognition system, Science Gate Publishing, USA, pp. 1-18 (2015) doi: 10.15579/gcsr.vol3.ch1
20. Unity3D framework [<https://unity3d.com/>]
21. Tensorflow framework documentation [<https://www.tensorflow.org/api/>]
22. YAML - The Official YAML Web Site [<http://yaml.org/>]
23. Ong, E.-J., et al.: Sign language recognition using sequential pattern trees. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, 16-21 June 2012, pp. 2200-2207 (2012) doi: 10.1109/CVPR.2012.6247928
24. Garcia, B.: American Sign language: Real-time American Sign Language Recognition with Convolutional Neural Networks. Stanford University, Stanford, CA (2015)
25. Bobic, V., Tadic, T., Kvascev, G.: Hand gesture recognition using neural network based techniques, In: NEUREL 2016: 2016 13th Symp. on Neural Networks and Applications (NEUREL), Belgrade, 22-24 November 2016, pp.35-38 (2016)
26. Howard, A.G., Zhu, M., Bo C., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, [<https://arxiv.org/pdf/1704.04861.pdf>] (2017)
27. PostgreSQL official web site [<https://www.postgresql.org/>]

Implementing automaton behavior with fuzzy controllers

Mykhailo Poliakov^[0000-0002-7772-3122]

Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine
polyakov@zntu.edu.ua

Abstract: Structural diagrams of control automata using fuzzy controllers for building control systems are proposed, types of behaviors of such automata that differ from the traditional ones by the possibility of parametric and structural adaptation of behavior, the formation of parallel activity relative to the state activity interval are described. Examples of the use of the proposed functional structures are given.

Keywords: Control systems, control machines, fuzzy controllers.

1 Introduction

Since the middle of the last century, finite automata have been used as a model of control devices for discrete and logical control [1 - 5]. These automata are defined as a tuple [6]

$$A = \langle S, X, Y, s_0, \delta, \lambda \rangle, \quad (1)$$

where S is a finite nonempty set (of states); X is a finite non-empty set of inputs (input alphabet); Y is a finite nonempty set of exits (output alphabet); $s_0 \in S$ - initial state; $\delta: S \times X \rightarrow S$ - transition function; $\lambda: S \times X \rightarrow Y$, $\lambda: S \rightarrow Y$ - functions of the outputs of the Mealy and Moore automata, respectively.

Such machines have limited behavior in the control system. They lack the mechanisms of structural adaptation of the automaton; the activity of the exits is rigidly "tied" to the time of the activity of the state. The need for an extended automaton behavior arises in the design of integrated and cognitive control systems [7 - 10]. An integrated system is such a system that combines several interconnected subsystems, for example, built according to the principle "control device at the i - level is the object of control at the $(i + 1)$ - level of control". In the cognitive system, knowledge and elements of cognitive behavior, such as perception, judgment, planning, learning, and others, are used to achieve management goals.

The behavior proposed in [11] is possible when an automaton is defined as a tuple

$$A = \langle X, Y, S, s_0, C, c_0, F \rangle, \quad (2)$$

where C is the set of controls; c_0 is the initial control; F is the set of functions of the automaton in its states. The element f_i of the set F for the i -th state is defined by the functions of this state

$$f_i = \langle \mu_i, \lambda_i, \sigma_i \rangle, \quad (3)$$

where μ_i is the activation function, λ_i is the output, σ_i is the structure function.

2 Purpose of research

On the basis of definitions (2) and (3), automata with non-binary elements forming the automaton sets X , Y , S and C can be built. At the same time, the author-accessible publications lack the functional structures and mechanisms for realizing the behavior of automata, built on the basis of tuples (2) and (3), which complicates their practical application.

The purpose of this work is to study the functional structures and mechanisms for implementing the extended behavior of automata constructed using Fuzzy controllers.

3 Analysis of publications

Fuzzy controller is a functional block that implements one of the fuzzy inference algorithms [12 - 15]. In the functional structures of the fuzzy control systems, the controller takes the place of the PID controller in the loop “outputs of the control object — sensors — controller — actuators — inputs of the control object” or is used in conjunction with the PID controller to improve its characteristics. With the help of the fuzzy controller, continuous object control is implemented.

The IEC 61131-7 [16] standard describes a unified FCL (Fuzzy Control Language) language for building a fuzzy controller. Such a description is made in the form of a functional block, which is included in the controller program of the facility control. These programs are written in one of the languages recommended by the IEC 61131-3 standard [17, 18].

These standards note the advantages of fuzzy controllers compared to other types of controllers, including the simplicity of the implementation of the parametric adaptation of the controller, but there are no recommendations for using fuzzy controllers to build control machines with both binary and non-binary elements of automaton sets.

4 Materials and methods

First consider the use of a fuzzy controller for constructing a classic control automaton based on a tuple (1) with binary elements of the sets X , Y , and S . The block diagram of the control automaton based on the fuzzy controller is shown in Fig. 1.

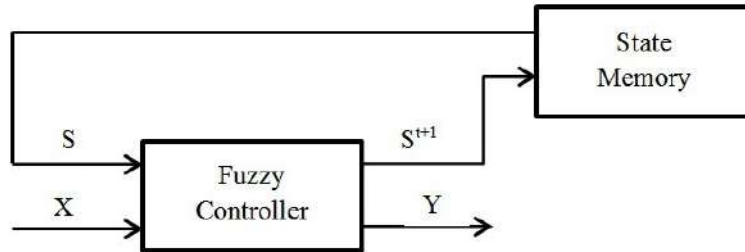


Fig. 1. Structural diagram of the control machine based on fuzzy controller

To build the control automaton, the fuzzy controller was used together with the state memory block of the automaton. The controller has two groups of inputs and two groups of outputs. Inputs X of the controller are connected to the outputs of the control object. The current state S of automata is fed to the inputs S of controller. A new state value is formed at the outputs S^{t+1} , into which the control automaton passes to the transition. And outputs Y are connected to executive mechanisms or to inputs of output operating automatic machines.

The controller's operation will be described in the traffic light control system, the block diagram of which is shown in Fig.2.

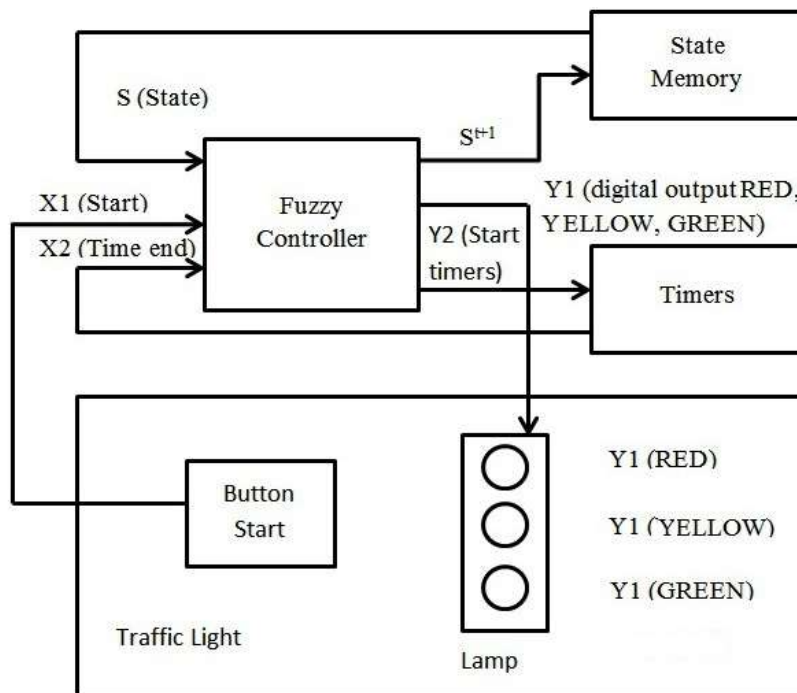


Fig. 2. Block diagram of the traffic light control system

The system implements a state machine whose graph is shown in Fig. 3.

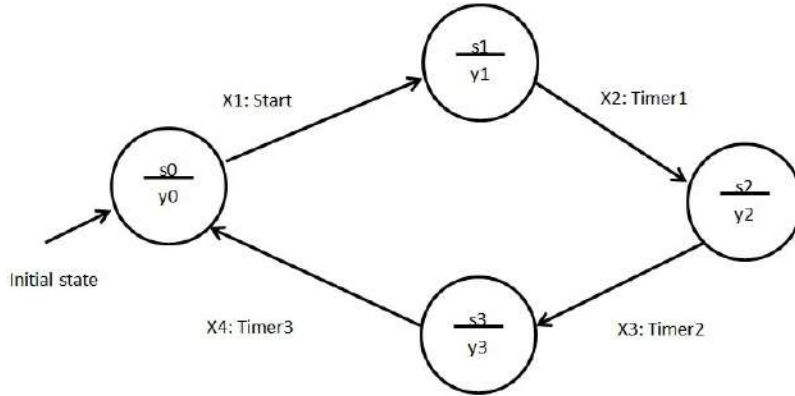


Fig. 3. Graph automatic simplified control system of traffic lights

The control device has four states ($s0$ - Green, $s1$ - Yellow1, $s2$ - Red, $s3$ - Yellow2), four transitions, and four outputs ($y0$ - Green; $y1$ - Yellow, Start Timer 1; $y2$ - Red, Start Timer 2; $y3$ - Yellow, Start Timer 3). One of transitions ($x1$) is an event transition, the rest ($x2 - x4$) are time transitions, which is counted using timers.

The states of the automaton S at the input and output of a fuzzy controller are described as fuzzy variables with membership functions of the “Singletons” type [15]. They are described only for a single linguistic term. In fig. 4 examples of terms are given.

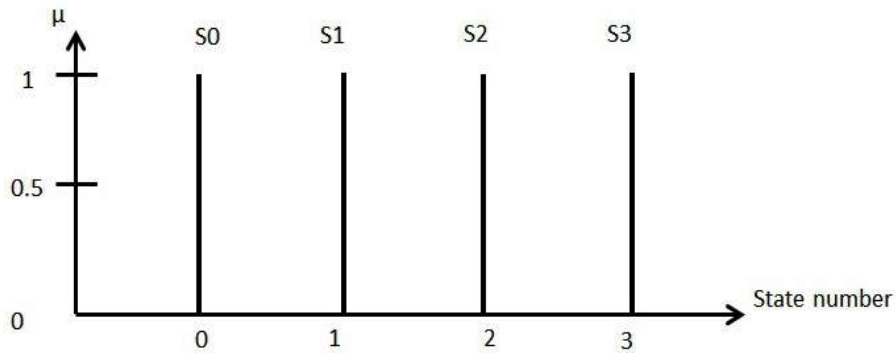


Fig. 4. The membership functions of the states on the output and the input of a fuzzy controller

The structure of the control device graph defines the control rules that are stored in the fuzzy controller's rule base. These rules are given in table 1.

Table 1. Traffic Light Control Rules

The rule	Comments
1. IF <i>Reset</i> IS <i>true</i> THEN <i>S</i> IS <i>s0</i>	Initial State
2. IF <i>S</i> IS <i>s0</i> THEN <i>Y</i> IS <i>y0</i>	Lamp «green»
3. IF <i>S</i> IS <i>s1</i> THEN <i>Y</i> IS <i>y1</i> AND Timer1 IS <i>Start</i>	Lamp «yellow» and Timer1 work
4. IF <i>S</i> IS <i>s2</i> THEN <i>Y</i> IS <i>y2</i> AND Timer2 IS <i>Start</i>	Lamp «red» and Timer2 work
5. IF <i>S</i> IS <i>s3</i> THEN <i>Y</i> IS <i>y3</i> AND Timer3 IS <i>Start</i>	Lamp «yellow» and Timer3 work
6. IF <i>S</i> IS <i>s0</i> AND <i>x</i> IS <i>Start</i> THEN S^{t+1} IS <i>s1</i>	Transition from <i>s0</i> to <i>s1</i>
7. IF <i>S</i> IS <i>s1</i> AND timer1 IS <i>end</i> THEN S^{t+1} IS <i>s2</i>	Transition from <i>s1</i> to <i>s2</i>
8. IF <i>S</i> IS <i>s2</i> AND timer2 IS <i>end</i> THEN S^{t+1} IS <i>s3</i>	Transition from <i>s2</i> to <i>s3</i>
9. IF <i>S</i> IS <i>s0</i> AND timer3 IS <i>end</i> THEN S^{t+1} IS <i>s0</i>	Transition from <i>s3</i> to <i>s0</i>

Rule 1 sets the initial state, rules 2 - 5 describe actions in the states, and rules 6 - 9 describe transitions from one state to another.

The fuzzy traffic light control model can be written in FCL notation, as shown in Figure 5.

Next, we consider the use of fuzzy controller for implementing control devices in accordance with the tuples of sets (2) and (3). Note that the activation, outputs and structures functions provided by the tuple (3) must be implemented for each state of the control unit's automaton. To do this, the control device must have a network of fuzzy controllers that interact with each other and with the operating machines of the control device through the memory of parameters, states, controls and the knowledge base. If for some states, the use of fuzzy controllers is redundant, then conventional (FSM) controllers are used to implement state functions. A generalized block diagram of a control system based on networks of operational machines that control fuzzy and FSM controllers is shown in Fig. 6.

The following types of controls are possible in this structure:

- Continuous control in the loop Control Object - Sensors - Input operating machines - Intermediate operating machines 1 - Output operating machines - Actuators. The operating machines involved in the loop perform a PID or other controller.

<pre> FUNCTION _BLOCK traffic light VAR_INPUT state, reset, start, end_timer :real END_VAR VAR_OUTPUT n_state, lamp, timer_start :real END_VAR FUZZIFY state Term S0:=0; Term S1:=1; Term S2:=2; Term S3:=3; END_FUZZIFY FUZZIFY reset Term norm: = 0; Term reset: = 1; END_FUZZIFY FUZZIFY start Term wait: = 0; Term start: = 1; END_FUZZIFY FUZZIFY end_timer Term timer1: = 0; Term timer2: = 1; Term timer3: = 2; END_FUZZIFY DEFUZZIFY n_state Term S0:=0; Term S1:=1; Term S2:=2; </pre>	<pre> Term S3:=3; ACCU: MAX METHOD: COGS DEFAULT: 0 END_DEFUZZIFY DEFUZZIFY lamp Term green: = 0; Term yellow: = 1; Term red: = 2; ACCU: MAX METHOD: COGS DEFAULT: 0 END_DEFUZZIFY DEFUZZIFY timer_start Term timer1: = 0; Term timer2: = 1; Term timer3: = 2; ACCU: MAX METHOD: COGS DEFAULT: 0 END_DEFUZZIFY RULE_BLOCK number1 RULE 1: IF state IS s0 THEN lamp IS green; RULE i: IF state IS s0 AND start IS 1 THEN n_state IS 1; END_RULE_BLOCK END_FUNCTION_BLOCK </pre>

Fig. 5. The program of control of the traffic light in FCL notation

- Event (logical) control using the FSM controller. Input operation machines transform analog signals from sensors into an event (binary signal) at the input of the FSM controller. As a result of this controller operation, its new state and output are determined. The output operation machines initiate the activation of the actuators or count down the time intervals necessary for control.
- Hybrid control using automatic machines and FSM controller [19]. There is more than one continuous control option, each of which is activated by its controller

FSM output. At each moment of time, the active circuit of the automatic machines that is selected by this controller is active.

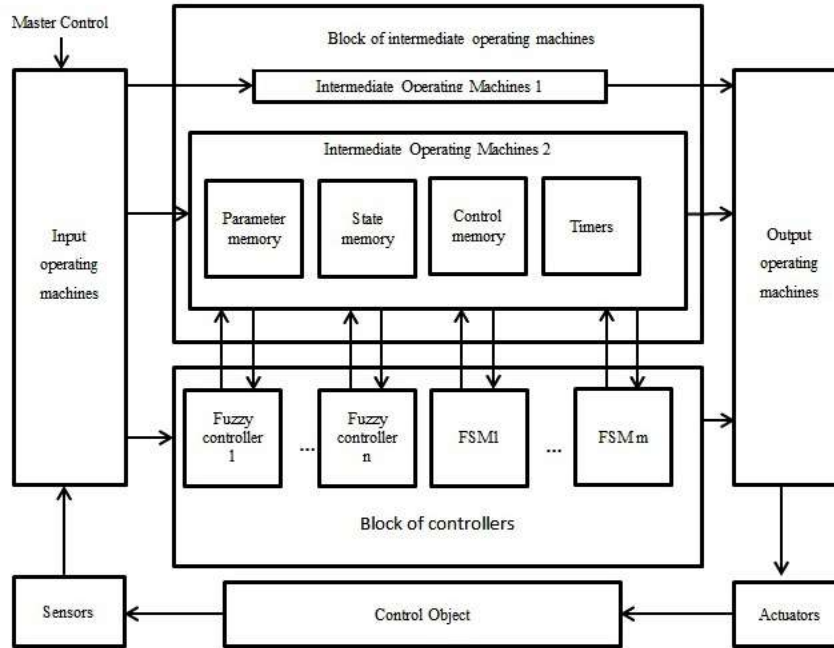


Fig. 6. A generalized block diagram of a control system based on networks of operating machines that control fuzzy and FSM controllers

- Continuous control using a fuzzy controller in the loop Control Object - Sensors - Input operating machines - fuzzy controller - Output operating machines - Actuators. The inputs of the fuzzy controller receive analog signals from the sensors. These signals are fused, processed using the fuzzy products rule base, defuzzified, and in the analog format are sent to the actuators inputs.
- Event fuzzy control with a fuzzy controller. This controller interacts with the elements of the block of intermediate operational automata 2. The inputs of a fuzzy controller are the inputs of analog and (or) discrete signals from sensors of the control object, timers, state memory and control. The processing of this data uses the appropriate membership functions and fuzzy product rules. The outputs of the fuzzy controller are the outputs of the control machine in the active state and the signals for memorizing its new active state. This control implements the activation functions μ_i (only for the active state), the outputs λ_i , and the structure σ_i . The form of these functions is described in [11].
- Advanced event fuzzy control with a fuzzy controller. This type of control differs from the previous one in that the activation function sets nonzero values of states in the vicinity of the active state. As a result, it becomes possible to activate the out-

puts in these states along with the outputs of the active state in order to post and preprocess tasks of the corresponding state.

Consider the features of the implementation of such controls by example. The fragment of the state graph of the controlling automaton with non-binary elements of the sets is shown in fig. 7.

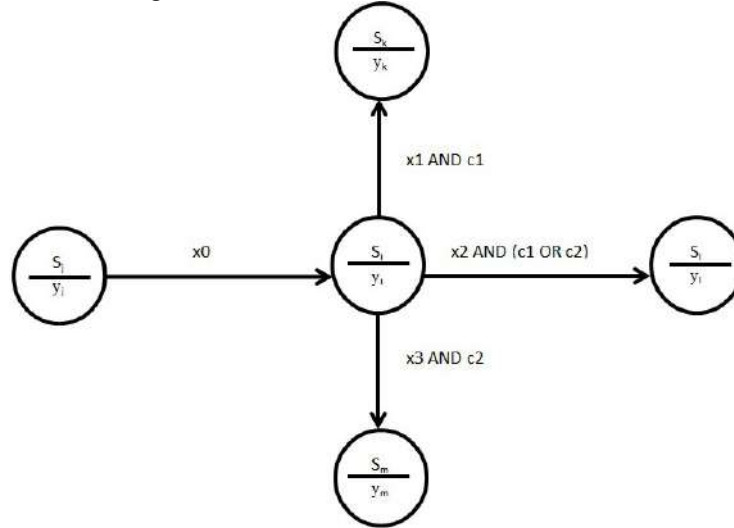


Fig. 7. Fragment of the state graph of the controlling automaton with non-binary elements of sets

The elements of the state set s_j, s_b, s_k, s_l, s_m take the values: active, post-active, pre-active or passive. The elements of the set of inputs $x0 - x3$ and outputs y_j, y_b, y_k, y_l, y_m can be logical, numerical or fuzzy type variables.

Let the state s_i be active at the current time. The transition to state s_i was made from state s_j . The function of the structure sets the control $c2$ in state s_i . Note that the control $c2$ does not allow the transition from the state s_i to the state s_k . Activation function μ_i calculate the value of states in the vicinity of the active state: s_j is post-active, s_i is active, s_k is passive, s_b, s_m are pre-active. State values are used by the output functions of these states to determine the level and / or type of activity of the outputs. Post-activity can be used to obtain processing and storage of secondary data that were relevant in the s_j state. The pre-activity of a certain state is intended to prepare the necessary data, information and knowledge to perform actions after the activity is received by these states.

The behavior described above can be implemented with the help of fuzzy controllers like control of a traffic light. For example, to calculate the pre-activity in the s_l state of the automaton Fig.7 rules like "IF s_i IS "active" AND $x2$ IS "probable" AND "structure" IS $c2$ THEN s_l IS "pre-active"" are used. The variables $s_i, x2, \text{Structure}, s_l$, presented in the rule, are linguistic variables with terms and membership functions given in Table. 2.

Table.2 Description of linguistic variables

Variable name	Terms	Type of membership function
si, sl	Active	S-shaped
	Pre-active	Π -shaped
	Post-active	Π -shaped
	Passive	Z-shaped
x2	Probably	S-shaped
	Average probability	Π -shaped
	Low probability	Z-shaped
Structure	c1	S-shaped
	c2	Z-shaped

5 Conclusion

The capabilities of control automata in the form of FSM are not enough to effectively solve the problems of constructing adaptive, integrated, and cognitive control systems.

The decision proposed in [11] about introducing into the definition of an automaton a set of controls and expanding the number of values of the automaton sets led to the need to move from the functions of the automaton as a whole to the set of state functions.

It follows that the various states of the controlling automaton may have different functions and devices for their realization. This paper describes the use of fuzzy controllers as devices for implementing state functions. The advantages of fuzzy controllers in control tasks are the ability to use the experience of experts, qualitative factors, and the availability of methods for the automatic synthesis of such controllers by training neural networks.

The technique of using fuzzy controllers is presented on examples of traffic control and a generalized control system combining fuzzy controllers and traditional FSM with a set of input, intermediate and operating machines.

Control devices based on fuzzy controllers implement a wide range of control types - continuous, discrete, hybrid, which differ in parametric and structural adaptation capabilities, are easily integrated into hierarchical control structures, and implement nonlinear controllers.

The proposed methods for constructing control automata are proposed to use cognitive monitoring systems for power transformer parameters [20] and remote laboratories for studying cognitive control systems in studies at Zaporizhzhia National Technical University.

References

1. Glushkov, V.M.: Synthesis of digital machines [Sintez tsifrovyykh avtomatov]. - M.: GIFML (1962).

2. Huffman, D. A.: The synthesis of sequential switching circuits. J. Franklin Inst. 257: 3–4, 161–190 and 275–303 pp. (1954).
3. Mealy, G. H.: A method for synthesizing sequential circuits. The Bell System Technical Journal, 34, 5, 1045-1079 pp., September (1955).
4. Moore, E. F.: Gedanken - experiments on sequential machines, pp. 129-153 in Shannon and McCarthy (eds.), Automata Studies, Annals of Mathematics Studies, Number 34, Princeton University Press (1956).
5. Hennie, F. C.: Finite-State Models for Logical Machines, Wiley, New York (1968).
6. Karpov, Ju. G.: Automata theory [Teoriya avtomatov] – Saint Petersburg, Piter, (2002).
7. Miller, G. A.: The cognitive revolution. Trends in Cognitive Sciences, 7 (2003).
8. Vernon, D.: Artificial Cognitive Systems. Carnegie Mellon University Africa (2014).
9. Polyakov, M.: Set-theoretic model of the integrated control system. [Teoretiko - mnozhestvennyye modeli integrirovannykh sistem upravleniya]. Dnipropetrovsk, System technologies, Regional interuniversity collection of scientific papers, №4, 131-137 pp. (2009).
10. Poliakov, M. A.: Cognitive control based on a dynamic set of goals: structures and models. [Kognitivnoye upravleniye na osnove dinamicheskogo kompleksa tseley: struktury i modeli]. Electrotechnical and computer systems, № 28, 127-133 pp. (2018).
11. Poliakov, M. A., Andrias, I.A.: Finite automata with non-binary elements of sets. [Konechnyye avtomaty s ne binarnymi elementami mnozhestv]. Dnipro, Systems technology, Regional interuniversity collection of scientific papers, №2, 85-94 pp. (2019).
12. Zadeh, L.A.: Fuzzy sets. Information and Control, vol. 8, 338-353 pp. (1965).
13. Ross, T.J.: Fuzzy logic with engineering applications.-McGraw-Hill (1995).
14. Feng, G.: Analysis and Synthesis of Fuzzy Control Systems: A Model-Based Approach. 1st Edition CRC Press Published (2017).
15. Leonenkov, A.V.: Fuzzy modeling in the MATLAB environment and fuzzyTECH. [Nechetkoye modelirovaniye v srede MATLAB i fuzzyTECH]. –SPb.: BHV-Petersburg (2005).
16. IEC 61131-7. Programmable controllers – Part 7: Fuzzy control programming. International Standard. First Edition 2000-08.
17. IEC 61131-3:1993, Programmable controllers – Part 3: Programming languages.
18. Parr, E. A.: Programmable Controllers. An engineer's guide. Third edition. Oxford: Newnes (2003).
19. Poliakov, M. A., Andrias I.A.: Set-theoretic models of functional structures of hybrid automata of control systems [Teoretiko-mnozhestvennyye modeli funktsional'nykh struktur gibridnykh avtomatov sistem upravleniya]. Dnipro, Systems Technologies. Regional interuniversity collection of scientific papers, № 3, 146-152 pp. (2018).
20. Poliakov, M. A., Andrias, I.A., Konohrai, S. P., Vasilevsky, V. V: Cognitive control of the life cycle of the insulation of the windings of an oil-filled power transformer [Kognitivnoye upravleniye zhiznennym tsiklom izolyatsii obmotok maslonapolnennogo silovogo transformatora]. The Bulletin of the National Technical University KhPI. Seriya: "Electric Machines and Electric Mechanisms of Energy Transformation". - Kharkov: NTU "KhPI", № 5, 90-96 pp. (2018).

Hybrid immune algorithms in the gene regulatory networks reconstruction

Fefelov A.O.¹[0000-0003-1140-0985], Lytvynenko V.I.¹[0000-0002-1536-5542],
Taif M.A.¹[0000-0002-3449-6791], Savina N.B.²[0000-0001-8339-1219],
Voronenko M.A.¹[0000-0002-5392-5125], Lurie I.A.¹[0000-0001-8915-728X],
Boskin O.O.¹[0000-0001-7391-0986]

¹Kherson National Technical University, 24, Beryslavskoe highway, Kherson, Ukraine, 73008
fao1976@ukr.net, immun56@gmail.com, taifmohamedali@gmail.com,
mary_voronenko@i.ua, aandre.lenoge@gmail.com

²National University of Water Management and Environmental Engineering, 11, Soborna
Street, Rivne, Ukraine, 33000
n.b.savina@nuwm.edu.ua

Abstract. The reconstruction is the process of identifying the structural and dynamic properties of the system on the basis of observations of its behavior and certain knowledge in the relevant subject area. Today, reconstruction plays a crucial role in biology as one of the main tools for modeling biological systems and their interactions, which is key to understanding the mechanisms of their functioning. One of the most widely used applications of reconstruction methodology is the gene regulatory networks identification. In this paper, we propose hybrid methods for the gene regulatory networks reconstruction, which allow us to increase the speed and accuracy of solving the S-system parametric identification problem. Experimental studies of the hybrid algorithm's parameters influence the quality of solving the S-system identification problem have been carried out on test examples. A comparative analysis of developments with other similar computational methods was carried out.

Keywords: gene regulatory networks, reverse engineering, gene expression, S-system, clonal selection algorithm, differential evolution.

1 Introduction

The modern world is characterized by the enormous complexity of the systems that fill it. The examples of such systems are biological organisms. Modeling of biological systems and their interactions is of key importance for understanding their functioning mechanisms, as evidenced by the numerous studies results that have been particularly intensive in the last few decades. During this time, a lot of tools and technologies have been developed that allow building and researching models of biological sys-

tems and processes. One of the fundamental modern science areas, the specialization of which is to decipher biological information at the genetic level, is molecular biology. Using gene expression profiles obtained using DNA microarrays [1], sets of genes involved in a particular biological process are identified. Interactions between different genes or their groups within the same process are modeled using the gene regulatory network (GRN).

Detection of the presence and nature of the interactions between the genes in the GRN is essential for the new drugs creation. The reconstruction helps researchers find answers to a number of questions, among which are the following: what body's processes are regulated by the gene under study; what genes influence the gene under study; how genes interact; which genes are responsible for a particular type of disease; which drugs will have an effect in case of a disease, etc. The recent interest in system biology, in the GRN's reconstruction, is primarily due to the intensive development of high technologies and the emergence of effective methods for measuring gene expression levels to produce impressive amounts of information available for processing. Expression data can be static with the ability to study the network structure. Expression data can be dynamic, which allows modeling of the complex behavior of the system over time. GRN reconstruction methods can be divided into several groups, the main of which are: relevant networks, Bayesian networks and differential equations. Relevant networks [2] are simplified models that use correlation or mutual information as a gene interaction measure. Due to the low computational complexity, these methods are suitable for the GRN structure reconstruction consisting of thousands of genes, however, they do not allow modeling the gene networks dynamics. Bayesian networks and dynamic Bayesian networks [3] are more complex models based on the fundamental concepts of probability theory and mathematical statistics. In these models, the relationship between genes is given in the form of conditional probability distributions. In addition, dynamic Bayesian networks allow simulating the behavior of a GRN in discrete time.

Models based on ordinary differential equations (SODE) systems have the ability to accurately reproduce the system evolution in continuous time [4]. Among the many linear and non-linear SODE models, the S-system is most widely accepted as a compromise between accuracy and mathematical flexibility. It is non-linear and therefore capable of reproducing different types of GRN behavior. On the other hand, the S-system structure features make its interpretation and analysis quite simple. The main difficulty of the GRN reconstruction, presented in the SODE form is the high computational complexity of the problem. It is necessary to take into account the fact that mathematical models, as a rule, have their own structure and a number of parameters that must be adjusted (identified). Moreover, the data used for the gene networks reconstruction are characterized by the highest dimensionality of the attribute space. Matrix gene expression profiles include tens of thousands of genes. In [5,6,7], the authors presented the studies results on the reduction and step-by-step clustering of gene expression profiles for the GRN reconstruction. However, it should be noted that this problem currently does not have a unique solution. Also unresolved is the problem of creating new, fast and accurate model identification methods that allow the use of expression data to reconstruct the architecture and behavior of GRN.

2 Formal problem statement

The GRN reconstruction problem involves two main aspects: the choice of a gene network model, and the development of a method for identifying the selected model. For a regulatory network consisting of N genes, the S-system is represented by the following SODE [8]:

$$\frac{dx_i}{dt} = \alpha_i \prod_{j=1}^N x_j^{g_{ij}} - \beta_i \prod_{j=1}^N x_j^{h_{ij}}, i = 1 \dots N \quad (1)$$

where $x_i(t)$ – state variable expressing the expression level of the i -th gene at time t ; α_i, β_i – nonnegative numbers called rate constants; g_{ij}, h_{ij} – kinetic orders that determine the direction and regulatory action degree, i.e. stimulation or suppression.

Minuend $\alpha_i \prod_{j=1}^N x_j^{g_{ij}}$ in the right part of equality (1) corresponds to the GRN elements cumulative effect, leading to an increase in the concentration x_i , and deducted

$\beta_i \prod_{j=1}^N x_j^{h_{ij}}$ reflects all effects resulting in reduced concentrations x_i . With increasing concentration, positive values g_{ij} imply gene j activation on gene i (enhances gene production), while negative values g_{ij} indicate the gene j suppressive effect on gene i (weakens gene production). In the decreasing concentration phase, the amplification and weakening effects of degradation correspond to positive and negative values of the parameter h_{ij} . Values $g_{ij} = 0$ ($h_{ij} = 0$) indicate the relationship absence between genes i and j .

The S-system identification consists in finding the optimal parameters values from the set $\nu = \{\alpha, \beta, g, h\}$. The complexity of solving this problem is due to its high dimensionality. The number of parameters to be found is determined by the expression $2N(N+1)$. That is, for a GRN consisting of only four genes, the search space dimension is 40. For this reason, the reconstructing S-system task is almost impossible to solve by analytical methods, especially with a large number of genes participating in the GRN. It is known that in solving such problems, population methods, such as genetic algorithms or artificial immune systems, have proven themselves well. But on large dimensions, the convergence rate of population methods can be very low. Thus, the study purpose is to develop new, fast and accurate methods for optimizing the S-system parameters, taking advantage of the population-based approach and hybridization technology.

3 Review of the literature

Evolutionary algorithms (EA), as well as artificial immune systems (AIS), are powerful optimization tools often used to solve GRN reconstruction tasks. EA is a computational paradigm based on methods that simulate the natural evolution processes of living beings. AIS models are based on the theoretical immunology studies results. According to these studies, the natural immune mechanisms of higher beings possess features characteristic of pattern recognition systems. Immune mechanisms are adaptive and resistant to small changes in the recognizable antigens characteristics (viruses and bacteria). They are able to develop their skills in recognizing alien agents. Despite the differences in paradigms, EA and AIS have a lot in common. Both those and others use a population of individuals (solutions), which from generation to generation is exposed to evolutionary or immune operators. In general, these operators are mutation and selection. The population of solutions, initially generated randomly, under the mutation and selection influence, as well as other specific operators, in each new generation improves its characteristics. The population is quantified in quantitative terms by assigning to each individual values of suitability for EA or affinity for AIS. Based on the estimates, individuals can leave the population or move on to a new generation.

Currently, several varieties of EA and AIS have been proposed for the GRN reconstruction. A combined algorithm using genetic programming (GP) and the least squares method (LSM) was proposed in [9]. Here, the ODE system general form is chosen as the GRN model, and the individuals evaluation is performed by means of the rms discrepancy between the observed and experimentally obtained data of gene expression time series. In [10], a two-stage GRN reconstruction method was proposed, where, at the first stage, using GP, approximation of the expression data with obtaining differentiable functions is performed, and at the second stage, the derivative values calculated after differentiating the functions are used by the hybrid evolutionary algorithm to evaluate individuals and search for structure and parameters GRN model presented in the S-system form. This approach allows you to get rid of complex calculations of the integration stage and increase the calculations overall speed. In [11], an evolutionary algorithm called trigonometric differential evolution (TDE) was used to reconstruct the S-system [12]. The reconstruction process consists of three phases: structural identification, parametric identification and synchronization. In addition, each next phase uses the results obtained in the previous phase. In [11], by differentiating the expression data time series, the S system is translated into an untied form. In an isolated ODE system, every equation that is included in it can be identified independently of the others, which reduces the search space dimension. Then TDE method is applied to each equation to search for its structure and parameters. In [13], a method for reconstructing an S-system based on multi-purpose optimization was proposed. Individuals multipurpose EA consist of two parts: binary and real. The binary part encodes the GRN structure, namely the presence or absence of a link between the genes, and the real part contains the S-system parameters values. For each of the parts formulated its own objective function. The result of the algorithm is a set of Pareto-optimal alternatives from which the final decisions are formed using the automatic selection procedure. In work [14], AIS in the classical clonal algorithm

form was used to optimize the S-system parameters, which was later improved in [15] by hybridization with one of the quasi-Newton methods. In the resulting hybrid, AIS performs the global optimizer functions, while the quasi-Newtonian method deals with local search. In [16], the clonal algorithm is applied to the S-system linearized model, where the exponents are replaced by two weight matrices: the activation matrix and the suppression matrix. This allows you to reduce the optimized parameters number, and increase the algorithm convergence rate by eliminating significant dependencies. In [17], the authors used a clonal selection hybrid method and wavelet neural network for the GRN reconstruction, represented by the ODE system general form. In this case, the neural network, as a universal approximator, is chosen to describe the functional dependencies of the ODE system right part, and the clonal algorithm is used to identify the neural network structure and parameters.

4 Materials and methods

4.1 Clonal selection

In [2], the immune system is considered from the point of view of the clonal selection mechanism. On the basis of the clonal selection principle in [18], the optimization algorithm CLONALG was proposed, which is widely used today as one of the AIS varieties. In the clonal algorithm, affinity values express the individual's proximity measures to the optimal solution and are calculated based on the objective function of the problem. In CLONALG, depending on the problem type, you can use different ways of representing solutions. The most frequently used binary and real representations. Also, the conditions and goals of the task are decisive when choosing how to represent the immune operators, the affinity function type, the algorithm parameters values.

When calculating the main population affinity, conditions are created for the selection of those cells that most fully (at this stage) interact with the antigen, i.e. form the objective function minima. In the activation process, the selected antibodies increase their representation in the solution space due to cloning. Cells whose affinity is higher create more clones but are less susceptible to mutations. Mutation in CLONALG has a high intensity because is the main driving force of evolution. In the replacement process, cells with low affinity are removed from the main population, and new, randomly generated individuals take their place. Theoretically, this allows you to avoid local extremes, and explore the entire target surface.

In [19] and [20], using comparative experiments, it was shown that the clonal selection algorithm, in its classical version, is unable to solve the problem of optimizing the S-system parameters with the required accuracy. To improve the results, the authors apply hybridization technology.

4.2 Differential evolution

The differential evolution algorithm (DE) is one of the evolutionary algorithms types [21]. Possessing high efficiency, DE has found application in many subject areas, as a global optimization method. There are several versions of the DE, differing in the evolutionary operators' implementation details. In this paper, the variant presented in [21] is used. A brief description of the DE method is given below.

The problem of minimizing the objective function is considered:

$$f(\bar{x}) \rightarrow \min, \bar{x} = (x_1, \dots, x_n) \quad (2)$$

where $\bar{x}$ – the parameters vector of the problem, on the basis of which individuals are built decision populations $\bar{x}_i^G, i = 1, \dots, P$, P – decision population size; G – current generation.

The DE algorithm main difference from other evolutionary algorithms is the mutation operator implementation. DE mutation is as follows:

$$\bar{v}_i^{G+1} = \bar{x}_{r_3}^G + F(\bar{x}_{r_1}^G - \bar{x}_{r_2}^G) \quad (3)$$

где $\bar{v}_i^{G+1}, i = 1, \dots, P$ – the individual that was received as a mutation result; $r_1, r_2, r_3 \in \{1, \dots, P\}$ – indices of individuals that are randomly selected from a population of solutions in the current generation are such that $r_1 \neq r_2 \neq r_3 \neq i$; F – scale factor ($F \geq 0$).

Components of individuals $\bar{x}_i^G$ partially replaced by the corresponding components of the vectors $\bar{v}_i^{G+1}$ with the formation of a candidates population $\bar{u}_i^{G+1} = (u_{i1}^{G+1}, \dots, u_{in}^{G+1})$. Vector formation $\bar{u}_i^{G+1}$ occurs according to the following expression:

$$u_{ij}^{G+1} = \begin{cases} v_{ij}^{G+1}, & \text{if } randEvent(p_{DE}) = 1 \vee j = k \\ x_{ij}^G, & \text{otherwise} \end{cases}, j = 1, 2, \dots, n \quad (4)$$

where $k \in \{1, \dots, n\}$ – random parameter index, selected once for each individual, the meaning of which is to ensure the transition of at least one component of the vector $\bar{v}_i^{G+1}$ to vector $\bar{u}_i^{G+1}$; p_{DE} – transition probability of j -th component of the vector $\bar{v}_i^{G+1}$ to vector $\bar{u}_i^{G+1}$; $randEvent()$ – binary function generating a random event with a given probability.

Since expression (4) is associated with the crossover operator in evolutionary algorithms, p_{DE} it is called the crossover probability.

The next generation population is formed from the current generation population and the candidate population through the DE-selection:

$$\bar{x}_i^{G+1} = \begin{cases} \bar{u}_i^{G+1}, & \text{if } f(\bar{u}_i^{G+1}) \leq f(\bar{x}_i^G) \\ \bar{x}_i^G, & \text{otherwise} \end{cases} \quad (5)$$

that is, each individual in the candidate population is compared with the corresponding individual in the current population. If the candidate has a smaller value of the objective function, he moves to a new generation. Otherwise, the current individual enters a new generation.

In [12], the DE algorithm was supplemented by the introduction of a new operator, called the trigonometric operator of mutation (TOM). Studies show that TOM is a local search tool that provides the DE algorithm with additional stability and accelerated convergence. In [20], the CLONALG and DE hybrid proposed in [20] was supplemented with the TOM operator, for the description of which the following relation is used:

$$\begin{aligned} \bar{v}_i^{G+1} = & (\bar{x}_{r_1}^G + \bar{x}_{r_2}^G + \bar{x}_{r_3}^G) / 3 + (p_2 - p_1)(\bar{x}_{r_2}^G - \bar{x}_{r_1}^G) + (p_3 - p_2)(\bar{x}_{r_2}^G - \bar{x}_{r_3}^G) + \\ & + (p_1 - p_3)(\bar{x}_{r_3}^G - \bar{x}_{r_1}^G); p_1 = |f(\bar{x}_{r_1}^G)| / p'; p_2 = |f(\bar{x}_{r_2}^G)| / p'; \\ & p_3 = |f(\bar{x}_{r_3}^G)| / p'; p' = |f(\bar{x}_{r_1}^G)| + |f(\bar{x}_{r_2}^G)| + |f(\bar{x}_{r_3}^G)|, i = 1 \dots P \end{aligned} \quad (6)$$

where $f(\bar{x}_i^G)$ – the affinity function value of the i -th individual of the population in the current generation.

In each iteration, TOM is applied with probability m_{TDE} , and the operator DE-mutation, respectively, with probability $1 - m_{TDE}$. As in the case of the DE mutation, the parameter p_{TDE} , used in TOM is analogous p_{DE} , forming the final configuration of the solution vector by expression (4).

4.3 Hybrid solutions

The proposed hybrid optimization algorithms are based on the method of expanding the mutation phase of the clonal selection algorithm with operators taken from the differential evolution algorithm. The block diagram of the hybrid clonal selection and differential evolution algorithm is shown in the figure. 1.

In the second case, a TOM is introduced into the hybrid, working with the DE-mutation in the mutually exclusive mode. The block diagram of the hybrid clonal selection algorithm and trigonometric differential evolution is shown in the figure 2.

The description of the developed hybrids main blocks is presented below.

Randomly create an antibodies initial population Ab^0 . An initial version of the possible size decisions basic population is created N_{Ab} . Population Individuals are encoded as strings of real numbers, which can acquire values in the range from 0.0 to 1.0. Each line contains the S-system parameters full set (Fig. 3). Each element of the initial population individual line is generated by generating a random number in the specified interval.

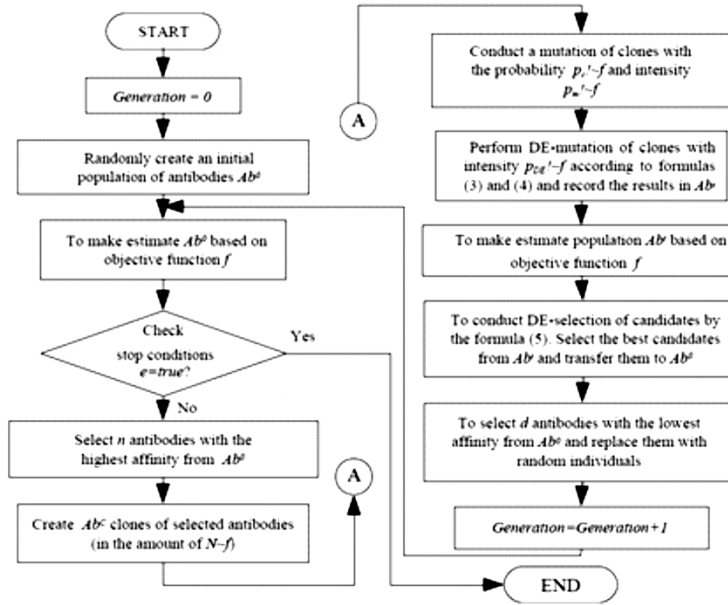


Fig. 1. Block diagram of the clonal selection hybrid algorithm and differential evolution

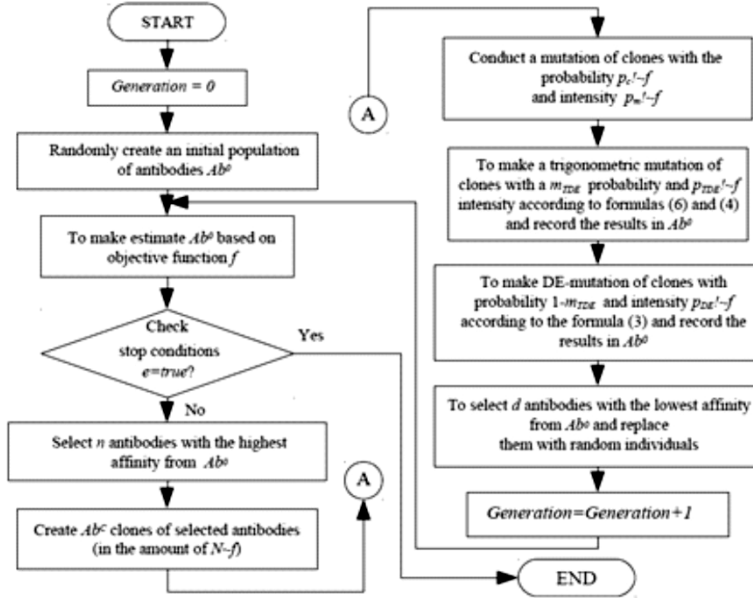


Fig. 2. Block diagram of the hybrid clonal selection algorithm and trigonometric differential evolution

Estimate Ab^0 based on the objective function f . When evaluating the next solution, the individual line values are recalculated into new ranges, in accordance with the change permissible intervals of one or another the S-system parameter, which are set before starting the algorithm.

α_I	β_I	g_{II}	...	g_{IN}	h_{II}	...	h_{IN}	...	α_I	β_I	g_{II}	...	g_{IN}	h_{II}	...	h_{IN}	...	α_N	β_N	g_{NI}	...	g_{NN}	h_{NI}	...	h_{NN}

Fig. 3. Antibody coding for clonal hybrid algorithm

The problem objective function, used as an affinity measure of the antibodies to the antigen, is represented by the model error on the gene expression data time series. In this paper, the following expression is used to calculate the error [22]:

$$f = \min \sum_{i=1}^N \sum_{j=1}^T \left(\frac{x_i^M(t_0 + j\Delta t) - x_i(t_0 + j\Delta t)}{x_i(t_0 + j\Delta t)} \right)^2 \quad (7)$$

where t_0 – time to start measuring gene expression levels; Δt – time step between successive measurements; T – number of measurements; $x_i^M(t_0 + j\Delta t)$ – the expression level values of the i -th gene, obtained using the model, i.e. SODE (1) solution (in this case, the fourth-order Runge-Kutta method); $x_i(t_0 + j\Delta t)$ – measured expression levels of the i -th gene.

Check break conditions. As a condition for stopping the algorithm, various indicators can be used. For example, the achievement of a given minimum value of the objective function, maximum operating time, cessation of the best antibody affinity growth over several generations, etc. In this paper, the stopping criterion is the established number of generations.

Select n antibodies with the highest affinity of Ab^0 . In the CLONALG algorithm, only the antibodies with the highest affinity are selected for further cloning. Since affinity is represented by a scalar value, selection in a population is easy. You can simply sort the population and select the first n antibodies, you can use the roulette method, etc. In the present work, selection is made using a tournament, when k individuals are randomly selected from the population (k -size of the tournament) and a winner is selected by further comparison of their affinities, which proceeds to the next stage. The tournaments themselves are repeated n times..

Create clones Ab^C of selected antibodies. From the individuals selected in the previous step, a clones population of size N_{Ab^C} is created. When cloning, each individual is duplicated N times. The number of clones is directly proportional to the affinity of the antibody.

Conduct mutation of clones with probability $p_C \sim f$ and intensity $p_m \sim f$. The first phase of the mutation is represented by a simple scheme, which is used in the CLONALG algorithm. According to this scheme, the value of each element of the

antibody string varies randomly (with a given probability p_m) by the following formula:

$$Ab_{ij} = \begin{cases} randInit(), & \text{if } randEvent(p_m) = 1 \\ Ab_{ij}, & \text{otherwise} \end{cases}, j = 1, 2, \dots, L \quad (8)$$

where Ab_{ij} – j -th element of the i -th individual in the population; L – individual line length; p_m – probability of changing the line element, called here the mutation intensity; $randInit()$ – random element initialization function; $randEvent()$ – binary function of generating a random event with a given probability.

In this case, the probability p_C determines the fact of applying the formula (8) to an individual. Also, according to the mechanism of clonal selection, mutation of antibodies with higher affinity is performed with less probability and with less intensity, i.e. there is an inverse relationship.

This is followed by blocks of diagrams representing operators taken from the differential and trigonometric differential evolution algorithm. Their detailed description is given above. Here Ab^l corresponds to the population of candidates, which is formed by the expression (4).

Select d antibodies with the lowest affinity of Ab^0 and replace them with random individuals. In this case, using the tournament selection, d individuals with the lowest affinity are selected from the population. Instead, new antibodies are placed in the population, which are created randomly. Thus, the algorithm provides the necessary population diversity level.

4.4 Search space transformation

Denote by v any of the parameters of the S-system α, β, g or h . To search for the optimal value of v by an evolutionary or immune algorithm, it is necessary that v be specified in the interval $v \in \mathcal{R} : a \leq v \leq b$. The ends of the intervals a and b depend on the parameter type and are set on the basis of the subject area preliminary knowledge before the start of the optimization process. Since the search algorithms types considered here belong to combinatorial optimization tools, the values set $V, v \in V$ is finite.

When using binary coding of individuals, the power V is defined as $|V| = 2^{m_v}$, where m_v – the length of the individual string segment (the bits number) allocated to encode the parameter v . The total number of solutions, among which the algorithm must find the optimal one, is equal to $2^{m_v^{2N(N+1)}}$. Obviously, the reduction m_v can significantly reduce the computational load and improve the performance of the search algorithm. However, here it is necessary to take into account the fact that all solutions are placed in the solution space not at random, but strictly in the nodes of the multidimensional grid formed by the values from the sets of $V_i, i = 1, \dots, 2N(N+1)$ parameters of the S-

system (Figure 4, *a*). If the global optimum P_{opt} does not coincide with any of this grid the nodes, then with this m_v search algorithm it is not able to find the optimal solution. Moreover, if it is small enough, then the best possible solutions, such as P_1 , will be too far from the global optimum and thus will not meet the requirements for accuracy of the result.

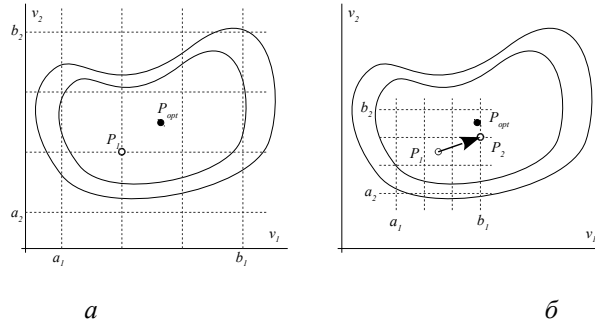


Fig. 4. *a* – initial placement of solutions in a two-dimensional search space; *b* – the transformation result of the search space relative to the solution P_1

The essence of this proposal is that, in the process of optimization, periodically, when certain conditions are reached, transform the search space by recalculating the possible values intervals of v with their simultaneous centering relative to the current best solution found (Figure 4, *b*). As can be seen in the figure, after receiving the solution P_1 the ends of the intervals of possible changes in parameters v_1 and v_2 were recalculated, which led to the transformation of the search space. As a result, without changing the values m_v , a more frequent grid is created, which allowed at the next iteration to get the best solution P_2 , which is closer to P_{opt} , than the solution P_1 . Then space is transformed relative to the solution P_2 , and the process is repeated until the required accuracy of the result is achieved.

A positive feature of this approach is the ability to work with small values m_v , which significantly reduces the solution options number and speeds up the convergence of the search algorithm. The transformation involves not only compressing the space, as shown in Figure 1, *b*, but also a possible expansion if the best solution in the next iteration is at the ends of the gap of one or more model parameters. At the same time, the transformation can include simultaneous compression of space in one dimension and expansion in another, which makes the proposed approach more flexible.

Formally, the process of the solution space transformation can be represented as follows:

$$a^{j+1} = v_p^j - \frac{s^{j+1}(b^0 - a^0)}{2}; \quad b^{j+1} = v_p^j + \frac{s^{j+1}(b^0 - a^0)}{2} \quad (9)$$

where a^{j+1}, b^{j+1} – the ends of the parameter values interval v for the next transformation; a^0, b^0 – initial interval of parameter values v ; v_p^j – best parameter value v , obtained using the optimization algorithm in the current (j -th) iteration; s^{j+1} – the scaling factor of the parameter values interval v for the next transformation, which is calculated by the formula:

$$s^{j+1} = \begin{cases} s^j k_g, & \text{if } (v_p^j - a^j) \leq \varepsilon \text{ or } (b^j - v_p^j) \leq \varepsilon \\ s^j k_s, & \text{otherwise} \end{cases} \quad (10)$$

where s^j – parameter scaling factor v in the current (j -th) iteration; $k_g (k_g > 1)$ – span expansion factor; $k_s (0 < k_s < 1)$ – gap compression ratio; a^j, b^j – the ends of the parameter values interval v in the current (j -th) iteration; ε – threshold value that records the fact that the parameter v coincides with the left or right end of the gap.

It should be noted that expressions (9) and (10) are valid for both binary and real methods of encoding individuals of the search algorithm.

Figure 5 presents a flowchart showing the operation of the proposed method. In this scheme, v_p it implies the best value of the S-system parameter v obtained for the entire time of the algorithm.

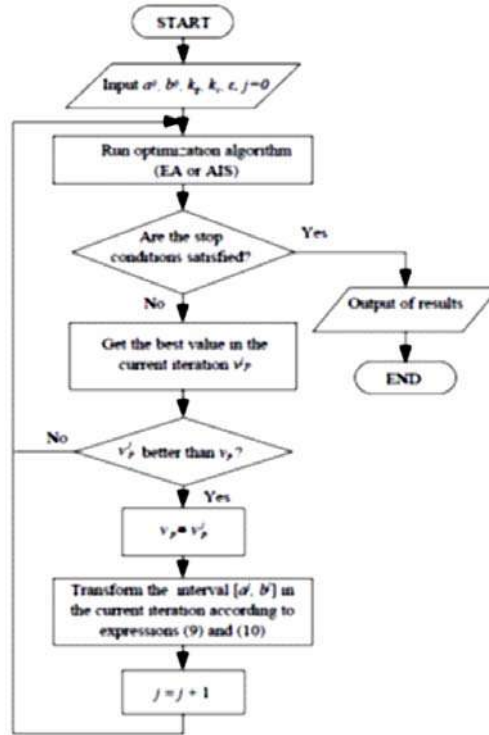


Fig. 5. Block diagram of the search space transformation

The work of the proposed method and algorithm can be described as follows. After a single launch of an evolutionary or immune optimization algorithm with a fixed number of generations, the best solution obtained as a result of this start is compared with the best solution obtained for all previous launches. The comparison is made based on the approximation error values of the observed gene expression data. If the new solution turns out to be better than all previous ones, the solution space is transformed according to the formulas (9) and (10). Next, the optimization algorithm is restarted and the process is repeated until the stopping conditions are satisfied (the required number of iterations j or some minimum error value is reached). Thus, in each subsequent iteration j EA or AIS works with a new solution space, which is gradually compressed in the vicinity of the global optimum. As an optimization method, it is recommended to use either classic AIS or hybrid, discussed above in this paper.

5 Experiments

To test the proposed approach effectiveness, the authors chose two tasks to identify the S-system parameters. In both tasks, the same artificial GRN is used [10], consisting of four genes. Reconstruction is performed for a network with a fixed and non-fixed structure. The fixed network structure, in contrast to the non-fixed, implies the availability of information about the links between the genes (that is, the matrix of connections is considered constant), which simplifies the solution, since the algorithm searches only the weights of the existing links.

The parameters of the S-system for the regulatory network consisting of four genes ($N=4$), are presented in table 1. To build a table of observed expression data needed for the experiments, system (1) is solved with the parameters taken from table 1. Solution S -systems were carried out by the fourth-order Runge-Kutta method with the number of time samples $T=26$. The following initial conditions were selected in the experiments $(x_1(t_0), x_2(t_0), x_3(t_0), x_4(t_0)) = (1.4, 2.7, 1.2, 0.4)$.

Table 1. S-system parameters for GRN consisting of four genes

i	α_i	β_i	g_{i1}	g_{i2}	g_{i3}	g_{i4}	h_{i1}	h_{i2}	h_{i3}	h_{i4}
1	12.0	10.0	0.0	0.0	-0.8	0.0	0.5	0.0	0.0	0.0
2	8.0	3.0	0.5	0.0	0.0	0.0	0.0	0.75	0.0	0.0
3	3.0	5.0	0.0	0.75	0.0	0.0	0.0	0.0	0.5	0.2
4	2.0	6.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.8

A time series of changes in gene expression levels was obtained, its graph is shown in the figure. 6.

The developed algorithms parameters, selected for research, are presented in Table 2. The settings for simple mutation, DE and TDE mutations are set according to the results of studies conducted in [19,20,23], where at these values the maximum solution quality was obtained.

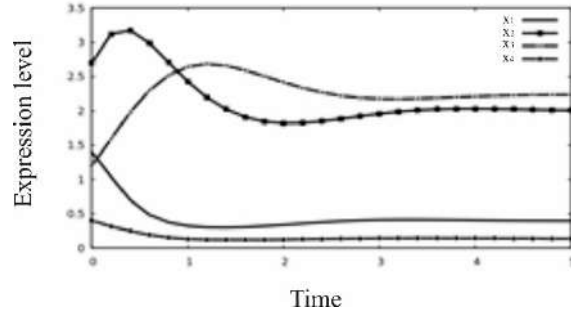


Fig. 6. The gene expression time series for experimental GRN

Table 2. Parameters of the proposed reconstruction algorithms

Parameter name	CLONALG + DE	CLONALG + TDE	CLONALG + transformation
Initial intervals of values α and β	[0.0, 15.0]	[0.0, 15.0]	[0.0, 15.0]
Initial intervals of values g and h	[-1.0, 1.0]	[-1.0, 1.0]	[-1.0, 1.0]
Coding method of individuals	real	real	binary
Parameter representation accuracy (m_v)	—	—	4 bit
Size of main population (N_{Ab})	300	300	300
Clone population size (N_{Ab^c})	900	900	900
Number of generations	3000	3000	50
Selection coefficient (n/N_{Ab})	1.0 ($n=300$)	1.0 ($n=300$)	0.7 ($n=210$)
Type of selection	tournament	tournament	tournament
Tournament size (k)	4	5	4
Probabilit of simple mutation (p_c)	0.1	0.1	0.01
Intensity of simple mutation (p_m)	0.05	0.05	—
Intensity of DE-mutation (p_{DE})	1.0	1.0	—
Probability TOM (m_{TDE})	—	0.6	—
Intensity TOM (p_{TDE})	—	0.7	—
Scale factor (F)	0.8	0.8	—
Span expansion factor (k_g)	—	—	1.5
Gap compression ratio (k_s)	—	—	0.8
Threshold value ε	0.0	0.0	0.0
Stop condition (number of space transformations)	—	—	50

For the experiments, the GRN reconstruction information system was used, the object-oriented architecture of which is presented in [24]. The results of each experiment were averaged on the basis of ten runs of the algorithm.

6 Results and Discussion

Comparison of the results of three AIS, based on clonal selection, as applied to the GRN reconstruction task, consisting of four genes with a fixed structure, is presented in the table 3.

Table 3. Comparative table of model errors in the study of three methods based on clonal selection for GRN with a fixed structure

Model error (f)	CLONALG + DE	CLONALG + transformation	CLONALG
Minimum	$1.634 \cdot 10^{-10}$	0.00112	0.03
Maximum	0.00068	0.0089	1.021
Average	0.000272	0.00466	0.6

It should be noted that in this test, out of ten starts, in 60% of cases, the CLONALG + DE algorithm converged to the minimum possible approximation error of $1.634 \cdot 10^{-10}$, and in 40% of cases, to error 0.00068, after which stagnation began. Stagnation is a differential evolution phenomenon characteristic that requires additional research that goes beyond the scope of this work.

Table 4. Comparative table of model errors in the study of three reconstruction methods for GRN with non-fixed structure

Model error (f)	CLONALG + TDE	CLONALG + DE	TDE
Minimum	0.00034	0.011	0.45
Maximum	0.0015	1.882	2.366
Average	0.00068	0.472	0.889

Table 4 presents comparative information on the GRN reconstruction results with a non-fixed structure. At the same time, the settings for the classical TDE algorithm are taken from the published research results and on the basis of recommendations found in the literature.

As can be seen from the table, the developed algorithms, thanks to the hybridization technology, provide a higher quality identification of the test SODE than the individual components that comprise them. In particular, the search space transformation procedure and the hybrid clonal selection and differential evolution algorithm give a smaller model error compared to the classical clonal algorithm. The

clonal selection hybrid algorithm and trigonometric differential evolution shows better results than the classical differential evolution.

7 Conclusion

The paper proposes hybrid GRN reconstruction methods that allow increasing the rate of convergence and accuracy of the optimization algorithm in solving the identifying S-system problem. The methods are based on hybridization technology, combining an artificial immune system in the clonal selection algorithm form, differential evolution algorithms, and also a solution space transformation, in which space is compressed in the global optimum vicinity. To test the proposed approaches effectiveness, experiments on the artificial GRN reconstruction were carried out. During the experiments, the various parameters influence of hybrid algorithms on the solution to the S-system reconstruction problem quality was studied, and also a comparison was made with other computational methods. Experiments have shown the presence of a negative effect from the use of differential evolution operators, such as selection and crossing over. On the other hand, the operator DE-mutation showed a significant positive effect. The comparative experiments results confirmed the advantages of the developed hybrid methods and algorithms over similar classical computational methods.

Obviously, the proposed methods do not eliminate the search algorithm from the falling into local optima possibility, therefore the abilities development to choose from them is one of the priorities for further research. In addition, in the future we plan to test the development effectiveness on real biological data.

References

1. De Risi, J.: Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*. 1997, vol. 278, pp. 680-686 (1997)
2. Burnet, F.: A modification of jerne's theory of antibody production using the concept of clonal selection. *CA: a cancer journal for clinicians*. 1976, vol. 26(2), pp. 119-121 (1976)
3. Friedman, N.: Using Bayesian networks to analyze expression data. *Journal of Computational Biology*. 2000, vol. 7(3-4), pp. 601-620 (2000)
4. Chen, T.: Modeling gene expression with differential equations. 4th Pacific Symposium on Biocomputing, PSB '99 - Proceedings, Big Island of Hawaii, Hawaii, USA, January 1999, pp. 29-40 (1999)
5. Babichev, S., Lytvynenko V., Gozhyj, A., Korobchynskyi, M., Voronenko, M.: A fuzzy model for gene expression profiles reducing based on the complex use of statistical criteria and shannon entropy (2019) *Advances in Intelligent Systems and Computing*, 754, pp. 545-555 (2019). doi: 10.1007/978-3-319-91008-6_55
6. Babichev, S., Krejci, J., Bicanek, J., Lytvynenko, V.: Gene expression sequences clustering based on the internal and external clustering quality criteria (2017) *Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2017*, 1, art. no. 8098744, pp. 91-94 (2017) doi: 10.1109/STC-CSIT.2017.8098744

7. Babichev, S., Lytvynenko, V., Skvor, J., Korobchynskyi, M., Voronenko, M.: Information Technology of Gene Expression Profiles Processing for Purpose of Gene Regulatory Networks Reconstruction (2018) Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018, art. no. 8478452, pp. 336-341 (2018) doi: 10.1109/DSMP.2018.8478452
8. Savageau, M.: Introduction to S-systems and the underlying power-law formalism. *Mathematical and Computer Modelling*. 1988, vol. 11, pp. 546-551 (1988)
9. Sakamoto, E.: Inferring a system of differential equations for a gene regulatory network by using genetic programming. Congress on Evolutionary Computation, COEX 2001—Proceedings, Seoul, Korea, 27-30 May 2001, pp. 720-726 (2001)
10. Kim, K.: Multi-stage evolutionary algorithms for efficient identification of gene regulatory networks. Applications of evolutionary computing: EvoWorkshops—Proceedings, Budapest, Hungary, April 2006, pp. 45-56 (2006)
11. Noman, N.: Reverse engineering genetic networks using evolutionary computation. *Genome Informatics*. 2005, vol. 16(2), pp. 205-214 (2005)
12. Fan, H.: A trigonometric mutation operation to differential evolution. *Journal of Global Optimization*. 2003, vol. 27(1), pp. 105-129 (2003)
13. Chen, Y.: Inference of biochemical S-systems via mixed-variable multiobjective evolutionary optimization. *Computational and mathematical methods in medicine*. 2017, vol. 20. — 9p. (2017) id 3020326
14. Jereesh, A.: A clonal based algorithm for the reconstruction of genetic network using S-system. *International Journal of Research in Engineering and Technology (IJRET)*. 2013, vol. 2, pp. 44-50 (2013)
15. Jereesh, A.: Clono-hybrid algorithm for the reconstruction of gene regulatory network using S-system. *International Journal of Pure and Applied Bioscience*. 2014, vol. 2(6), pp. 241-248 (2014)
16. Jereesh, A.: Two weight matrix model based GNET reconstruction using clonal algorithm. *Journal of Advanced Bioinformatics Applications and Research*. 2014, vol. 5(3), pp. 121-133 (2014)
17. Fefelov, A.: A hybrid approach in the reconstruction of gene regulatory networks. *Control systems and machines*. 2017, Kiev, № 3, pp. 63-72 (2017)
18. De Castro, L.: Learning and optimization using the clonal selection principle. *IEEE Transactions on Evolutionary Computation*. 2002, vol. 6(3), pp. 239-251 (2002)
19. Fefelov, A.: Parametric identification of the S-system using a modified clonal selection algorithm. *Control systems and machines*. 2017, Kiev, № 5, pp. 43-53 (2017)
20. Fefelov, A.: Reconstruction of the S-system by a hybrid algorithm of clonal selection and differential evolution. *Control systems and machines*. 2017, Kiev, № 6, pp. 41-51 (2017)
21. Storn, R.: Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*. 1997, vol. 11(4), pp. 341-359 (1997)
22. Tominaga, D.: Efficient numerical optimization algorithm based on genetic algorithm for inverse problem. Genetic and Evolutionary Computation Conference, GECCO '00 — Proceedings, Las Vegas, Nevada, USA, July 2000, vol. 251, pp. 251-258 (2000)
23. Fefelov, A.: Reconstruction of the Gene Regulatory Network by Hybrid Algorithm of Clonal Selection and Trigonometric Differential Evolution. IEEE 38th International Conference on Electronics and Nanotechnology, ELNANO 2018— Proceedings, Kyiv, Ukraine, April 24-26, pp. 305-309 (2018)

24. Fefelov, A.: Object-oriented architecture of the information system for the reconstruction of gene regulatory networks. Control systems and machines. 2017, Kiev, № 4, pp. 67-75, 82 (2017).

Genetic method of image processing for motor vehicle recognition

Ievgen Fedorchenko¹[0000-0003-1605-8066], Andrii Oliinyk²[0000-0002-6740-6078], Alexander Stepanenko³[0000-0002-0267-7440], Tetiana Zaiko⁴[0000-0003-1800-8388], Anton Svyrydenko⁵[0000-0003-0114-8632] and Dmytro Goncharenko⁶[0000-0001-8973-3742]

^{1,2,4,5,6} Dept. of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia 69063, Ukraine

evg.fedorchenko@gmail.com, olejnikaa@gmail.com,
alex@zntu.edu.ua, nika270202@gmail.com,
antonsvyrydenko@gmail.com, gdink99@gmail.com

Abstract. An analysis of scientific articles on similar topics was conducted; it was determined that the best recognition accuracy rate is achieved by using convolutional neural networks. Modifications of a simple genetic algorithm (Alfa-Beta, Alfa-Beta Fixed, Fixed) were developed. The implementation of the program for recognition of road users (cars, bicycles, pedestrians, motorcycles, trucks, etc.) was developed. Also, a comparison was made between the use of modifications of a simple genetic algorithm and the best approach for solving the problem of road user recognition. The purpose of the research conducted was to find an optimal approach for solving the problem of road user recognition, since the system that hasn't been implemented yet can recognize road users accurately. It was found that the improved Alpha-Beta modification is the best approach from the considered ones used to solve the problem. This modification allowed getting the best accuracy rate in less time selection, in comparison with other base modifications and simple genetic algorithm. The obtained results have a high practical value, since the developed modification allows optimizing the process of selection of values in other subject areas.

Keywords: Pattern Recognition, Genetic Algorithm, Evolutionary Algorithm, Neural Networks, Python, OpenCV, Keras.

1 Introduction

The theory of pattern recognition exists as a branch of informatics and related disciplines. It develops methods for the classification and identification of objects of various nature: signals, situations, objects that are characterized by a comprehensive number of certain features [1]. The problem of object recognition can also be a purpose of interdisciplinary researches - including work on creation of artificial intelligence. It is also often used in solving practical problems in the field of computer vision. When setting the classical problem of object recognition, it is necessary to apply a mathematical language based on logical thinking and mathematical principles. In

contrast to this approach, there are methods for recognizing objects using machine learning and 6 artificial neural networks, formed not so formalized approaches to recognition, and show not worse, but in some cases, a much better result.

One of these areas is road safety, which includes many narrow areas: number-plate recognition, traffic congestion determination, recognition of road users in self-driving cars, etc.

However, for today, problems of recognition quality are not completely solved. For example, the vulnerability was detected in the Tesla Autopilot system: the system in most cases cannot recognize a bicycle. It identifies a bicycle as a person or a small car; accordingly, the system can make decisions that may endanger cyclist's life [1].

Vulnerabilities in recognizing road users in systems such as Mazda Smart City Brake Support [2], Waymo Autopilot [3] are also known.

It can be said that recognizing bicycles and other road users is a topical task. Recognition of road users can be used in the system of control and / or traffic safety or it can be used to study the reasons of constructing specialized bicycle lanes, etc.

2 Analysis of literary data and problem statement

This work [4] presents a study of solving the problem of bicycle recognition on a video stream by combining the following approaches: Histogram of Oriented Gradients (HOG) [5, 6], Support Vector Machine (SVM) [7], cascade classifier (Viola-Jones method) [8, 9].

Presented results of performing bicycle recognition in different weather conditions, and it is argued that the developed approach can be used in systems of real time recognition.

However, using of a combination of HOG and SVM methods is not optimal, since it allows you to recognize large enough objects that are not always present in video in real time and due to the fact that it takes a lot of time to calculate complex parameters. Using a cascaded classifier is more optimal in terms of recognition time. However, this method has errors in recognition - because of the application of different scales and the size of the scanning window, one object can be recognized as two.

In general, low image quality, weather conditions and various lighting conditions make precise recognition more complicated for the methods used. The approach developed is complex and complicated, and each method requires sufficient computational resources that effect on results.

An expedient approach to the implementation of the bicycles recognition would be the use of neural networks.

Thus, in [10], the study of the analysis of the flow of machines on the CCTV (video surveillance systems) images, based on the solution of the problem of recognition of machines with the help of neural networks Faster R-CNN (the faster region-based convolutional neural network) [11] and SSD (single shot multibox detector) [12].

In general, SSDs are capable of quickly recognizing, but are often mistaken for small objects, while Faster R-CNNs are slower than SSD, but can provide more precision.

Work aim is to test the specified trained neural networks on images of low quality and under different weather conditions. As a result of the study, it was found that the considered neural networks are not capable of recognizing all the machines, and in the case of bad weather conditions (rain, snow, fog), the accuracy indicator falls significantly. Therefore, the study of neural networks to detect vehicles should use more diversity training set. In addition, the study implements the recognition of only cars, although the concept of "traffic" covers several types of transport (motorcycles, trucks, etc.).

In [13], the solution of the machine recognition problem by studying a convolutional neural network [14] is considered, which is based on reinforced learning mistake and samples of error-prone probes. That is, in the training of the convolutional neural network, it is proposed to use the training errors of the current stage as training data in the next stage. A comparison of histogram oriented gradient recognition with convolutional neural networks, one of which was taught in the standard way, and the second - by the proposed authors, was made, which resulted in the proposed approach showed better performance.

However, there are no examples of car recognition, because of the small amount of study sample used in the study, one cannot say with complete certainty that a significant improvement in the training of the firewall neural network has been achieved.

In [15] the decision of the recognition problem among the crowd of each pedestrian separately is considered. The main purpose of the study is to recognize pedestrians in black and white images in various situations, the pedestrian's identification of the image, even with partial overlap, and with high accuracy, localize it. Also, the goal is to determine how many pedestrians are presented in the image.

The approach developed in this article is based on the scale-independent expansion of the Implicit Shape Model (ISM).

Authors succeeded in realizing all goals, but the accuracy of pedestrian recognition is 71.3% and is not satisfactory.

Also, it's worth noting that the developed method works only with images. However, the method has a great potential and can, as a result of further development, develop towards image recognition and tracking.

In [16] we consider the solution of the problem of pedestrian recognition by combining several neural networks.

The proposed system consists of two subsystems: the main pedestrian detection system for generation of all detections and a system of semantic segmentation to improve the results. Pedestrian Detection system in turn consists of a generator "candidates" for pedestrians and classification system. In order to provide more information about the coordinates of the object of the classification network, it is proposed to use the new method of soft tokens, which relates the object to all classes. For the implementation of the classification system the idea of "combined training" was implemented. Also, the technique of "soft rejection" is proposed for combining the conclusions of all networks of classification and the network of semantic segmentation with the conclusion of the network of "candidates" generation.

Pedestrian recognition has been tested on four popular image sets: Caltech Pedestrian set, INRIA set, ETH set, KITTI set. The first three sets received the highest recognition accuracy. Also, it indicates the significant speed of the developed approach.

In [17] the implementation of pedestrian recognition is considered, using the combination of the method of histogram oriented gradients, the method of reference vectors and the convolutional neural network.

The purpose of work is to construct a semantic areas of interest in order to get the foreground object is to reduce the errors associated with error detection background.

First, using a convolutional neural network trained by the Caltech Pedestrian set, the thermal map is generated by the input image. Further, semantic interest regions are extracted from the thermal map by morphological image processing. In the rest, semantic interest regions share the whole image on the background and foreground, to facilitate the work of the detectors who make the decision.

It is noted that with the help of semantic regions of interest, the work of detectors to varying degrees improves.

However, the use of a combination of HOG and SVM methods is not optimal, as it allows you to recognize large enough objects that are not always present in video in real time and due to the fact that it takes a lot of time to calculate complex parameters.

In work [18] the implementation of the system of recognition of vehicles belonging to 7 classes (motorcycle, car, pickup, bus, truck, truck with a trailer, truck with several trailers) is considered on the CCTV video. The Deep Convolutional Neural Network (DCNN) was used in the implementation.

The proposed system algorithmically consists of two tasks: localization and classification. Initially, the localization is performed, by generating regions independent of the class, for each frame, and then using a deep convolutional neural network to produce feature descriptions for each region. Finally, using the Support Vector Machines (SVM) method, for each region, the resulting description descriptions are compared with the templates, and a conformity assessment and classification are performed.

The accuracy of vehicle recognition differs for each class, but in general, the accuracy is within the range of 92 - 95%. The accuracy of the recognition significantly decreases when downloaded traffic, rain and fog, with low video quality.

The disadvantage of this development is the use of a deep convolutional neural network, since its calculation requires a lot of computational resources, which is not acceptable. The use of this neural network is possible with the presence of a powerful CPU (CPU) or graphics processor (NVIDIA GPU), in conjunction with such artificial intelligence environments as CUDA, Caffe, Torch.

Also, the system detects recognition errors on a small scale, then the object of one class is recognized as an object of another class, for example, the "truck" is recognized as a "truck with a trailer", and so on.

The analysis [4, 10, 13, 15-26] in the subject area suggests that conducting research on the implementation of the system traffic participant's recognition and increasing the accuracy of recognition is a rather topical task.

3 Purpose and objectives of the research

The purpose of the study is to develop a program for solving the problem of road users recognition - cars, bicycles, pedestrians, motorcycles, trucks. Also, it is necessary to develop an evolutionary method that would allow to obtain the best indicators

of recognition accuracy when selecting the parameters of teaching neural networks at lesser time.

To achieve this goal, the following tasks were solved:

1. Conduct an analysis of existing methods for solving the pattern recognition problem.
2. To develop methods of evolutionary search for selection of parameters for teaching neural networks.
3. To implement the program realization of developed methods of evolutionary search.
4. To conduct the study of the effectiveness of the developed method.

4 Materials and methods of studying the influence of parameters in the training of neural networks

4.1 Equipment used in the experiment

The study of the influence of the values of parameters in the training of neural networks on the accuracy and time of training was carried out on a computer with an Intel Core i5 7400 processor with a clock speed of 3 GHz, with a memory capacity of 8 GB.

To study neural networks, a study sample of 8,000 images was divided into 5 classes and 3 sub-sets: educational (1000 images), test (200 images), verifying (400 images).

4.2 Methods of determining the influence of parameters of teaching neural networks

The study of the effect of values of the parameters of training neural networks was carried out using a simple genetic algorithm and its modifications.

The sequence of actions of a simple genetic algorithm is as follows: generation of an initial population P_s consisting of N individuals occurs:

$$P_s = (I_1, \dots, I_N), \quad (1)$$

where I - the individual.

Each individual I (its genes) is encoded by certain values from the set of all parameters of the set Q . Let the parameters of the neural networks have a set of parameters Q :

$$Q = (n_j, m, f_{act}, f_{opt}), \quad (2)$$

where n - number of neurons of a certain layer,

$j = 1, \dots, m$,

m - number of layers,

f_{act} - activation function

f_{opt} is a function of optimization.

Then the individual I_N of the initial population P_s can be encoded by formula (3).

$$\begin{aligned} Q(n_i) &= R[32, 64, \dots, n_{max}], \\ Q(m) &= R[1, 2, \dots, m_{max}], \end{aligned} \quad (3)$$

$$Q(f_{act}) = R[relu, elu, tanh, sigmoid],$$

$$Q(f_{opt}) = R[adam, sgd, ada\ max, nadam, rmsprop, adagrad, adadelta],$$

where $i = 1, \dots, m$,

R - function of choice of a random variable,

n_{max} - maximum number of neurons,

m_{max} - maximum number of layers.

$$I_N = R(Q). \quad (4)$$

Next, to evaluate the optimality of each individual I of the population of the P_s , the fitness function f_{fit} is calculated:

$$f_{fit}(P_s) = \max(I_i), \quad (5)$$

where $i = 1, \dots, N$.

The next step is to check if the conditions for completing the search for the optimal solution are not fulfilled. Such conditions can be the number of generations N , the value of the fitness function f_{fit} , runtime, and so on. In the example of selecting parameters for teaching neural networks, the value of fitness function is the accuracy of the trained neural network. If the conditions are fulfilled, it is necessary to complete the execution of the algorithm and output the results, the most optimal generation of the individual.

Then the selection of specimens is performed in a new population P_n : individuals are ranked according to the fitting indicator, 40% of the most adapted individuals (T) are selected in a new population. Also randomly selected 10% less fitting individuals (L). Another 50% of the new population P_n is obtained by crossing the selected individuals (K).

$$P_n = T + L + K. \quad (6)$$

Intersection (C) of individuals has the following algorithm: subsets of B and W randomly selected individuals on M and F having a specific set of genes:

$$M = ((n_1^l, \dots, n_m^l), m^l, f_{act}^l, f_{opt}^l), \quad (7)$$

$$F=((n_1^2, \dots, n_m^2), m^2, f_{act}^2, f_{opt}^2).$$

Then the crossing can be expressed as follows:

$$C=(M \times F). \quad (8)$$

The result of the crossing of individuals is two descendants K_1, K_2 , which can be written as follows:

$$K_1=(R[(n_1^1, \dots, n_m^1), (n_1^2, \dots, n_m^2)], R[m^1, m^2], R[f_{act}^1, f_{act}^2], R[f_{opt}^1, f_{opt}^2]), \quad (9)$$

$$K_2=(R[(n_1^1, \dots, n_m^1), (n_1^2, \dots, n_m^2)], R[m^1, m^2], R[f_{act}^1, f_{act}^2], R[f_{opt}^1, f_{opt}^2]),$$

where R - function of the choice of a random variable.

This approach of crossing individuals is called uniform cross-linking, since the coding of the offspring randomly selects the values for each gene from each ancestor.

For the received descendants the mutation operator μ is applied according to (3):

$$\mu(K)=R[R(n_i), R(m), R(f_{act}), R(f_{opt})]. \quad (10)$$

Then, after applying the mutation operator, the descendant can be written as follows:

$$K_\mu=((n_i, \dots, n_m), m, f_{act}, f_{opt}). \quad (11)$$

After applying the mutation operator descendants included the new generation P_n . The bursts and mutations are held until a new generation of P_n of size N is created:

$$P_n=((T+L), K_j), \quad (12)$$

where $j = 1, \dots, (N - (T + L))$.

Next, the f_{fit} fitness function for each new generation P_n (5) is calculated.

It is checked whether the conditions for completing the search for an optimal solution were not fulfilled. If the conditions are fulfilled, it is necessary to complete the execution of the algorithm and output the results, the most optimal generation of the individual. If the conditions are not fulfilled - continue the implementation of the algorithm, proceed to the formation of the next generation P_{n+1} based on the generation P_n .

The proposed modifications to a simple genetic algorithm are as follows:

1) modification Alfa-Beta: for crossing in each generation, a different number of couples for crossing is selected, in which one individual refers to the most suitable one and the other to the least adapted individuals. Also, random mutations (base and double) or one mutation (base) can occur: Monte Carlo method generates a random number 0 or 1. If 0 falls, then one mutation arises if there is one mutation - there are two mutations.

The sequence of actions of this modification of a simple genetic algorithm is

similar to its basic version, but it has some differences. At the stage of selection of individuals to a new population of P_n individuals are ranked according to the fitness parameter, then the number of pairs is determined randomly - a certain number of the most adapted individuals, and the same number of the least adapted.

The most suitable individuals form a subset of B , the least adapted - subset of W . Both subsets are included in the set of pairs V . The number of individuals that can be chosen in pairs is in the range of 20-60% of the total number of individuals. The rest of the new population P_n is obtained by crossing the selected individuals (K).

$$P_n = (V, K_j), \quad (13)$$

where $j = 1, \dots, (N-V)$.

Also, in the proposed modification, two ($\mu 1, \mu 2$) or one mutation (μ) may occur randomly, whereas in the basic version of the algorithm, one mutation arises randomly. Moreover, in a situation when there are two mutations - one gene can mutate twice. After applying the mutation operator descendants included the new generation P_n . Crossing and mutation are held as long as we create the next generation of P_n size N (13).

2) modification Alfa-Beta fixed: the number of couples for crossing is selected in each generation, with one person referring to the most suitable one and the other to the least adapted individuals. Also, random mutations (base and double) or one mutation (base) can occur: Monte Carlo method generates a random number 0 or 1. If 0 falls, then one mutation arises if there is one mutation - there are two mutations. A fixed point of crossing is established - the first half of the genes participate in the crossing - the genes responsible for the number of neurons in the layers, the values of other genes are always passed on to the descendants of one of the individuals.

The sequence of actions of this modification is similar to the sequence of Alfa-Beta modifications, however, the crossing operation is excellent: according to formula 7, subsets B and W are randomly selected by the individual M and F having a certain set of genes. Intersection can be expressed as:

$$C^* = \left(\left(\frac{1}{2} M \times \frac{1}{2} F \right), \frac{1}{2} F \right). \quad (14)$$

The result of the crossing of individuals is two descendants K_1, K_2 , which can be written as follows:

$$K_1 = (R[(n_1^1, \dots, n_m^1), (n_1^2, \dots, n_m^2)], m^2, f_{act}^2, f_{opt}^2), \quad (15)$$

$$K_2 = (R[(n_1^1, \dots, n_m^1), (n_1^2, \dots, n_m^2)], m^2, f_{act}^2, f_{opt}^2),$$

where R is a random variable selection function,

$m^2, f_{act}^2, f_{opt}^2$ - the value of the genes transmitted from the individual F .

3) modification Fixed: a fixed point of crossing is established - half of the genes are involved in the crossing - the genes responsible for the number of neurons in the layers, the values of other genes are always passed on to the descendants of one of the

individuals. Also, in the mutation stage randomly, there are two mutations (base and double) or one mutation (base): the Monte Carlo method generates a random number, 0 or 1. If 0 falls, then there is one mutation, if there is 1 - occurs two mutations. The sequence of actions of this modification is similar to the sequence of Alfa-Beta modification, but there is a difference: the selection of individuals to a new population occurs both in the basic version of the genetic algorithm (6), and the crossing occurs as in the second modification (14), (15).

5 Experiments and results of research on the use of modifications of a simple genetic algorithm

There were 4 experiments in which the genetic algorithms and the simple simple genetic algorithm were considered. All experiments were carried out with the same parameters: the number of generations - 5, the size of the population - 20, the mutation rate - 0.01. Input data is a sample of 5-digit images (pedestrian, bicycle, motorcycle, auto, truck), in which each class has 1600 images.

To modify Alpha-Beta and a simple genetic algorithm, the elite coefficient was also set at 0.4 and the randomly selected individuals were 0.1.

In the figures below, graphs of parameter selection parameters are presented in the training of the neural network by the proposed modifications of the genetic algorithm. After selecting the parameters of learning by the genetic algorithm and its modifications, the recognition accuracy for the best neural network and the selection time were obtained. These results are shown in the Table 1. Figure 1 shows a graph of the mean accuracy for each generation of selection parameters for training neural networks. Figure 2 shows a chart of indicators for the 5 best individuals of the last generation of parameter selection for each neural network.

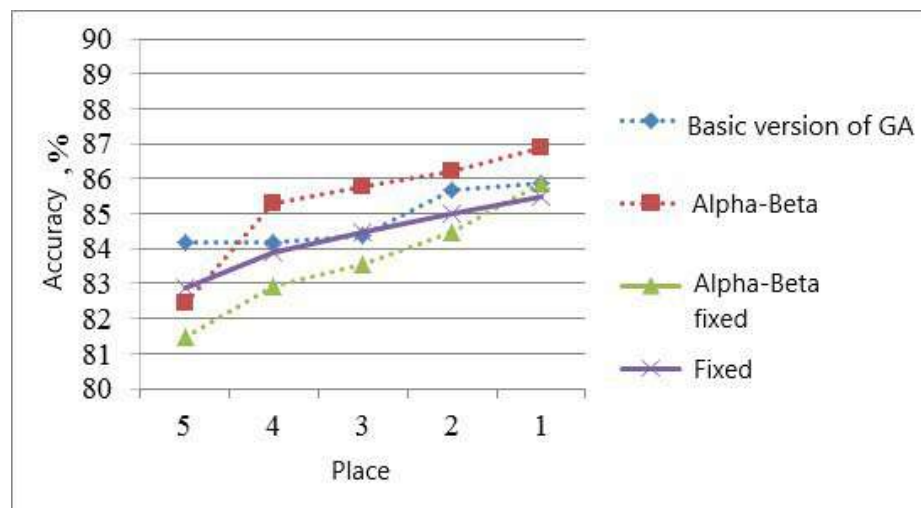


Fig.1. Schedule of mean accuracy for selection of parameters for teaching neural networks

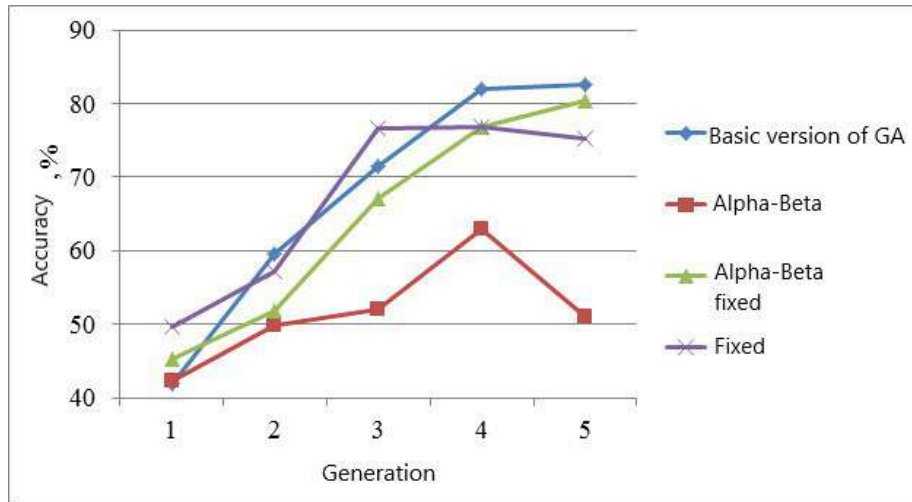


Fig.2. Schedule of accuracy indicators of the last generation when selecting parameters for teaching neural networks

Table 1. Results of selection of parameters using a simple genetic algorithm

GA version	Time of training	Accuracy on test data	Accuracy of auto-testing
Basic	8 days, 8 minutes, 39 seconds	79.45 %	85.89 %
modification Alpha-Beta	7 days, 7 hours, 16 minutes, 43 seconds	82.12 %	86.90 %
modification Alpha-Beta fixed	10 days, 12 hours, 4 minutes, 58 seconds	80.02 %	85.89 %
modification Fixed	5 days, 22 hours, 27 minutes, 5 seconds	78.48 %	85.48 %

Based on the results presented in table 4, we can conclude that the use of the Alpha-Beta modification of the genetic algorithm is the best approach to achieving higher recognition accuracy in less time.

6 Discussion of the results of research on the use of modifications of a simple genetic algorithm

In the Fig. 4 and 5 are graphs of the execution time (in minutes) of the proposed method on computer systems, which depends on the number of cores involved. It can be seen from the graphs that the proposed method has an acceptable degree of parallelism and is effectively performed on both MIMD and SIMD systems. This way, the IPME cluster was able to reduce the method execution time from 1565 minutes (on one core) to an acceptable 147 minutes on 16 cores. On the ZNTU the computing system, the method execution time was reduced from 1268 minutes on a single core to 110 minutes on 16 cores. The differences in the performance of the systems are due to their architectural features: in the cluster cores are connected by means of the Infini-Band communicator, and in the multi-core computer they are located on a single chip, which explains the smaller impact of overhead (transfers and synchronizations). In addition, the processor in multi-core computer supports Turbo Boost technology [25-32], making the time of the method execution on the single core much less than the execution time on the core of the cluster that does not support this technology.

As a result of the study, a modification of a simple genetic algorithm - Alfa-Beta - was developed. Using this modification allows you to speed up the selection of the parameters of training neural networks, and increase the accuracy. It has been shown that the increase in the number of mutations and selection in pairs of different individuals provides a greater variety of gene combinations, which leads to better performance in less time.

Fig. 3-7 shows examples of recognition using a trained neural network.

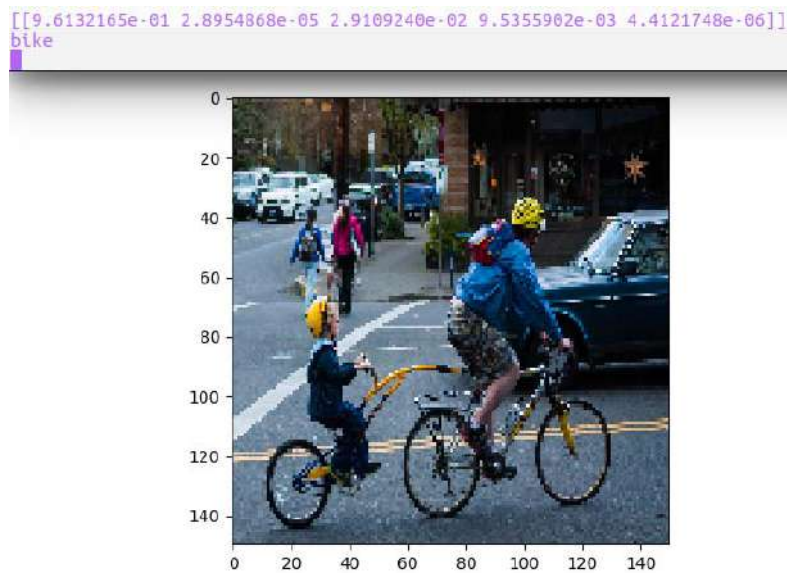


Fig. 3. Example of recognizing the object "bicycle"

```
[[5.3577603e-16 9.9984324e-01 1.5912704e-07 1.3068089e-13 1.5657884e-04]]
```

car

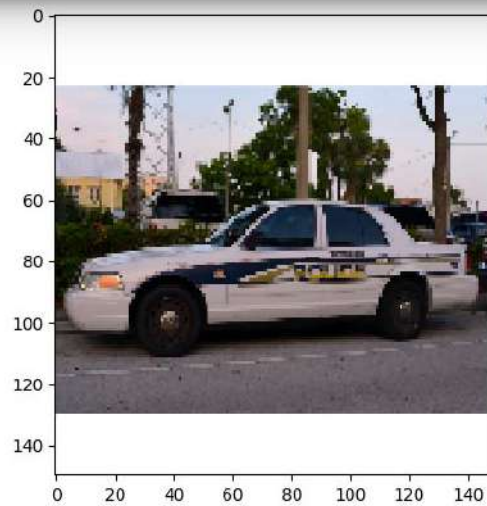


Fig. 4. Example of recognizing the object "car"

```
[[0.01507533 0.00409241 0.52248806 0.45667475 0.00166949]]
```

moto



Fig. 5. Example of recognizing the object "motorcycle"



Fig. 6. Example of recognizing the object "pedestrian"

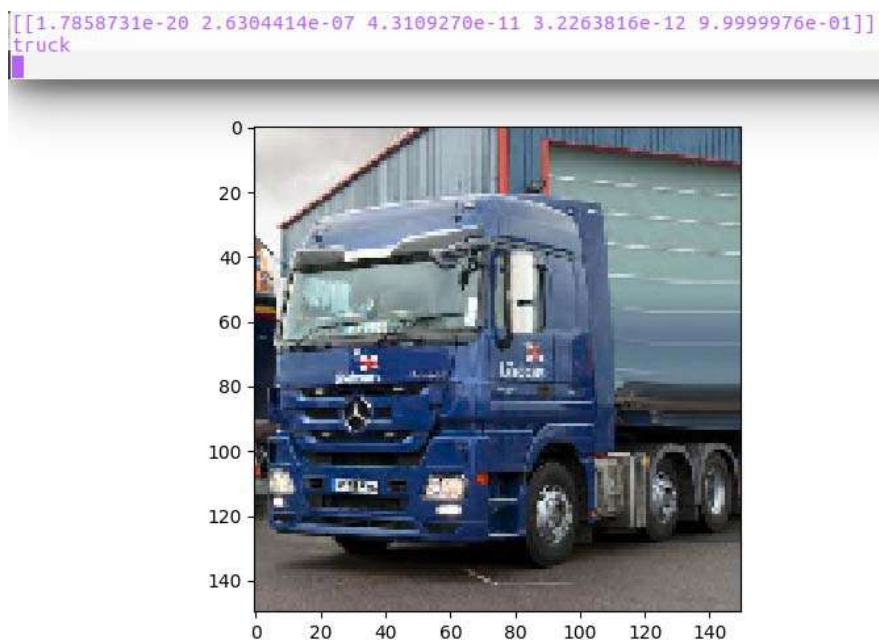


Fig. 7. Example of the recognition of the object "truck"

Based on the recognition results, we can conclude that a trained neural network is able to confidently determine the relevance of an object to a particular class.

7 Conclusion

An analysis of existing methods of pattern recognition from scientific articles on similar problems was carried out. It was determined that the best recognition accuracy rate is achieved using convolutional neural networks.

Three modifications of a simple genetic algorithm (Alfa-Beta, Alfa-Beta fixed and Fixed) were developed for selecting parameters of training neural networks.

A program using the developed modifications of a simple genetic algorithm for selecting parameters of training neural networks and recognition of road users was implemented. When applying the developed modifications, the following results were obtained:

- Alpha Beta - accuracy 86.90%, selection time - 7 days, 7 hours, 16 minutes, 43 seconds;
- Alpha Beta fixed - accuracy 85.89%, selection time - 10 days, 12 hours, 4 minutes, 58 seconds;
- Fixed - accuracy 85.48%, selection time - 5 days, 22 hours, 27 minutes, 5 seconds.

Comparison of the rates of a simple genetic algorithm was conducted; it was determined that the Alpha-Beta modification is the best approach to recognize road users since this modification allowed better accuracy (86.90%) for shorter selection time (7 days, 7 hours, 16 minutes, 43 seconds).

Acknowledgment

The work was performed as part of the project “Methods and means of decision-making for data processing in intellectual recognition systems” (number of state registration 0117U003920) of Zaporizhzhia National Technical University.

References

1. Lugovskaya, A. M.: Artificial neural networks in medical diagnostics [Iskusstvennyie neyronnyie seti v meditsinskoy diagnostike]. Computer systems and networks: proceedings of the 54th scientific conference of postgraduates, undergraduates and students 2018, vol. 1, pp. 182 – 183. BGUIR, Minsk (2018).
2. Goold, S.D., Lipkin, M.Jr.: The doctor-patient relationship: challenges, opportunities, and strategies. *Journal of General Internal Medicine* (1), 23–33 (1999).
3. Van Tuc, N.: Approximation contexts in addressing graph data structures. University of Wollongong Thesis Collection, 30–55 (2015).
4. Barkoulas, J. T., Baum, Ch. F., Travlos, N.: Long Memory in the Greek StockMarket. *Applied Financial Economics*, vol. 10, no. 2, 177–184 pp. (2000).
5. Dalal, N. Histograms of Oriented Gradients for Human Detection. *CVPR*, 886–893 (2005).

6. Schmidhuber, J., Wierstra, D., Gagliolo, M., Gomez, F.: Training Recurrent Networks by Evolino. *Neural computation*. vol. 19(3), 757–779 pp. (2007). doi: 10.1162/neco.2007.19.3.757.
7. Bland, J.M., Altman, D.G.: Statistics notes: measurement error. *BMJ*. 312 (7047): 1654 (1996). doi:10.1136/bmj.312.7047.1654
8. Viola, P., Jones M.J. Rapid Object Detection using a Boosted Cascade of Simple Features. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 1–9 (2001).
9. Viola, P. Robust real-time face detection. *International Journal of Computer Vision*, 57, 2, 137–154 pp. (2004).
10. Peppas M. V., Bell D., Komar T., Xiao W. Urban traffic flow analysis based on deep learning car detection from CCTV image series. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 499–506 pp. (2018).
11. Ren S., He K., Girshick R., Sun J.. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 6, 1137–1149pp. (2017)
12. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., Berg A. SSD: Single Shot MultiBox Detector. *Computer Vision ECCV*, 21–37 (2016).
13. A CNN Vehicle Recognition Algorithm based on Reinforcement Learning Error and Error-prone Samples, <http://iopscience.iop.org/article/10.1088/1755-1315/3/032052/pdf>.
14. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Analysis of inrush currents of the unloaded transformer using the circuitfield modelling methods. *Eastern-European Journal of Enterprise Technologies*, vol. 3, № 5 (93), pp. 6-11 (2018). doi: 10.15587/1729-4061.2018.134248
15. Boden M.: A guide to recurrent neural networks and backpropagation. Halmstad University, pp. 1–10 (2001).
16. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Enhancing the effectiveness of calculation of parameters for short circuit of three-phase transformers using field simulation methods. *Eastern-European Journal of Enterprise Technologies*, vol. 4, № 5 (94), pp. 22-28 (2018). doi: 10.15587/1729-4061.2018.140236.
17. Whitley D. Genetic Algorithms and Neural Networks. *Genetic Algorithms in Engineering and Computer Science*, pp. 203-216, (1995).
18. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Analysis of inrush currents of the unloaded transformer using the circuitfield modelling methods. *Eastern-European Journal of Enterprise Technologies*, vol. 3, № 5 (93), pp. 6-11 (2018). doi: 10.15587/1729-4061.2018.134248
19. Oliinyk, A., Subbotin, S., Lovkin, V., Leoshchenko, S., Zaiko, T.: Development of the indicator set of the features informativeness estimation for recognition and diagnostic model synthesis. *14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET 2018)*, pp. 903–908, (2018). doi: 10.1109/TCSET.2018.8336342.
20. Leoshchenko, S., Oliinyk, A., Subbotin, S., Zaiko, T.: Methods of semantic proximity extraction between the lexical units in infocommunication systems. *4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T 2017)*, 7–13 pp. (2017). doi: 10.1109/INFOCOMMST.2017.8246137.
21. Leoshchenko, S., Oliinyk, A., Subbotin, S., Gorobii, N., Zaiko, T.: Synthesis of artificial neural networks using a modified genetic algorithm. *Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018)*, pp. 1-13 (2018). dblp key: conf/iddm/PerovaBSKR18

22. Stepanenko, O., Oliinyk, A., Deineha, L., Zaiko, T.: Development of the method for decomposition of superpositions of unknown pulsed signals using the second order adaptive spectral analysis. *Eastern European Journal of Enterprise Technologies*, vol. 92, issue 2/9, pp. 48–54, (2018). doi: 10.15587/1729-4061.2018.126578
23. Alsayaydeh, J.A., Shkaruplyo, V., Hamid, M.S., Skrupsky, S., Oliinyk, A.: Stratified Model of the Internet of Things Infrastructure, *Journal of Engineering and Applied Science*, vol. 13, issue 20, pp. 8634–8638, (2018). doi: 10.3923/jeasci.2018.8634.8638.
24. Shkaruplyo, V., Skrupsky, S., Oliinyk, A., Kolpakova T.: Development of stratified approach to software defined networks simulation. *Eastern European Journal of Enterprise Technologies*, vol. 89, issue 5/9, pp. 67–73 (2017). doi: 10.15587/1729-4061.2017.110142
25. Kolpakova, T., Oliinyk, A., Lovkin, V.: Improved method of group decision making in expert systems based on competitive agents selection. *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Institute of Electrical and Electronics Engineers, pp. 939–943, Kyiv (2017). doi: 10.1109/UKRCON.2017.8100388
26. Oliinyk A., Fedorchenko I., Stepanenko A., Rud M., Goncharenko D.: Evolutionary method for solving the traveling salesman problem. *Problems of Infocommunications. Science and Technology : 5th International Scientific-Practical Conference PICST2018*, Kharkiv, 9–12 October 2018, Kharkiv, Kharkiv National University of Radioelectronics, pp. 331–339 (2018). doi: 10.1109/INFOCOMMST.2018.8632033

Automated Software Vulnerability Testing Using In-Depth Training Methods

Alexandr Kuznetsov ¹[0000-0003-2331-6326], Oleksiy Shapoval ¹[0000-0003-4478-3193],

Kyrylo Chernov ¹[0000-0003-2417-3793], Yehor Yeromin ¹[0000-0001-9720-746X],

Mariia Popova ¹[0000-0002-9751-1717], Olga Syniavska ²[0000-0002-7507-3541]

¹V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, 61022, Ukraine
kuznetsov@karazin.ua, alex.shapoval@protonmail.com,
kirillfilippsky@gmail.com, suvenick2@gmail.com,
mariia.popova26@gmail.com

²Sumy State University, Rymskogo-Korsakova st., 2, Sumy, 40007, Ukraine
o.syniavska@uabs.sumdu.edu.ua

Abstract. The article provides a view on modern technologies, which are used for automatic software vulnerability testing in critically important systems. Features of fuzzing realization (which is based on making many inputs with different mutated data) are also studied. As a result, testing algorithm picks input data that is more likely to cause a fail or incorrect work of software product. Deep learning algorithms are used to decrease the computational complexity of testing process. The use of simple fuzzer and Deep Reinforcement Learning algorithm shows that the amount of mutations necessary to find vulnerabilities decreases by 30%.

Keywords. Fuzzing, Testing, Reinforcement Learning, Q-Learning, Software Security.

1 Introduction

The development of modern computer technology leads to the emergence of new high-quality information services and their implementation in all spheres of human activity [1-5]. The development of the IT-industry has led to the construction of global computer networks, extensive data warehouses, automated control systems, including critical infrastructures, Smart Grid and much more [6-8].

In the age of the Internet and global implementation of information technologies, information security is more important for critical infrastructures [9, 10]. Complex solution for problems related to informational security is connected with solving different objectives in cryptography [11-14], computations optimization, technical and physical security as well as many others [10, 15-17].

This paper focuses on the problem of automated software vulnerability scanning [18-22]. As practice shows, computer programs are the most exposed fragment of

modern IT infrastructure. Failure or configuration errors and undeclared operations can lead to disastrous consequences. Development and research of methods and tools for automated software vulnerability scanning is extremely important and relevant task.

2 Known Fuzzers and Their Analysis

Barton Miller, who is a professor at the University of Wisconsin, Madison, was the first to introduce the term “fuzzing” together with his students in 1989 [23]. Henceforward, the development of automated testing continued and the creators of fuzzing used this method to search for vulnerabilities in software using different operating systems including UNIX, Windows and Mac OS [23-26]. Today fuzzing is most commonly used in Software Quality Assurance. It is also one of the key steps of Microsoft Security Development Lifecycle (SDL). Software experts from Microsoft consider that deliberate input of wrong or random data is a sufficient and non-costly way to detect potential errors before product release [18-22].

Fuzzing testing has several advantages among which [23-26]:

- high speed (usually much higher than manual code review);
- no need to involve human work;
- fuzzer does not need to be controlled, while human capabilities are finite;
- scalability, i.e. if there is a need to find more vulnerabilities, the only thing needed is more fuzzers.

However, fuzzing has several disadvantages. For example, when fuzzing is used, it is very difficult to discover deep errors, such as business logic errors etc. [27]. It is relevant to conduct comparative analysis of different fuzzing methods and experimental researches using the most common software products as an example.

This work presents the results of analysis and comparative research of automated vulnerability search technologies. Particularly, the following common fuzzing utilities are reviewed [28-32]: American Fuzzy Lop, MiniFuzz, and Peach. Experimental researches are conducted for such everyday software products as Google Chrome; Notepad++; Winamp; Microsoft Paint.

American Fuzzy Lop (AFL) is an open-source fuzzer, which was developed by Polish computer security expert Michał Zalewski [29, 30]. The program uses genetic algorithms to automatically look for test cases. This fuzzer’s main goal is to cause unexpected behavior of target programs by changing or shifting input channel bytes.

MiniFuzz was developed by Microsoft. This fuzzer is intended for simple and routine usage [31]. Its operating principle is forming data beforehand, then passing them to the target and catching errors. This utility belongs to dumb fuzzers category which conduct fuzzing randomly.

Peach is more advanced tool for “intellectual” fuzzing developed by Michael Eddington [32]. It supports not only mutation mode but also fuzz-file generation. Since program needs to know the structure of target files, specialized XML-documents are used as input. Peach can fuzz applications, servers, network protocols, drivers, internal protocols, devices, systems and so on.

During experimental research of automated vulnerability search effectiveness MiniFuzz was used. Fuzzing process was conducted for several common desktop applications, including Google Chrome; Notepad++; Winamp; Microsoft Paint.

For Google Chrome testing several dozens of different html-files were selected. Aggressiveness (how much of input data is mutated) was alternately set to 5%, 15%, 25%, 35%. Testing results for this case are shown in Table 1. As can be seen from the table, fuzzing testing can help in discovering “File Not Found” type errors. Apparently, higher aggressiveness leads to more errors of this type.

Table 1. Expected value of “File Not Found” error and confidence interval (for significance level $\alpha = 0,01$ sample size $N = 1000$)

Error	Aggressiveness Level			
	5%	15%	25%	35%
“File Not Found”	$0,239 \pm 8,23 \cdot 10^{-4}$	$0,269 \pm 2,6 \cdot 10^{-4}$	$0,295 \pm 3,49 \cdot 10^{-4}$	$0,320 \pm 8,28 \cdot 10^{-4}$

After several hours of fuzzing for Notepad++ no errors occurred even with 100% aggressiveness. This shows high level of stability and security of this application. Testing of Winamp led to similar results.

To test Microsoft Paint images of different formats and sizes were selected. Testing results are shown in Table 2. As it can be seen from the results, selective testing with fuzzing revealed errors “File Not Found” and “Wrong Format”. Frequency of these are almost similar and increase with higher aggressiveness.

Table 2. Expected value for errors and confidence interval (for significance level $\alpha = 0,01$ sample size $N = 1000$)

Error	Aggressiveness Level			
	5%	15%	25%	35%
“File Not Found”	$0,13 \pm 2,6 \cdot 10^{-4}$	$0,15 \pm 2,58 \cdot 10^{-4}$	$0,18 \pm 2,64 \cdot 10^{-4}$	$0,22 \pm 1,58 \cdot 10^{-4}$
“Wrong Format”	$0,09 \pm 2,96 \cdot 10^{-4}$	$0,16 \pm 3,77 \cdot 10^{-4}$	$0,2 \pm 3,58 \cdot 10^{-4}$	$0,23 \pm 4,57 \cdot 10^{-4}$

Conducted tests show that the majority of common applications are secured and cannot be crashed using primitive fuzzing. In the case of Google Chrome, for example, it is not a surprise, because software engineers and testing professionals at Google use much more complex fuzzing during the development cycle of their products [28]. The same goes for Microsoft.

Thus, the analysis of different fuzzers in the area of automated testing shows that this approach to software vulnerability search can vary depending on the goal, tester’s skills, data format and other factors. Some applications have privilege separation system, which depends on user level. Using fuzzer as a tool for automated vulnerability search, it is possible to find errors in software products, which let attacker gain full or

partial control over the system. Some low-lever errors are very similar to each other so it is possible to use the same logic to find vulnerabilities in more applications.

Relying on conducted experimental research for selective testing it can be said that fuzzing is quite promising method of automated vulnerabilities search. We were able to find errors even in reliable and tested applications, like Google Chrome and Microsoft Paint, as we managed to discover random input data, which would cause errors. However, these are not critical for application functioning and/or operating system, their handling is correct that, apparently, is stipulated behavior for such kind of corrupted input data.

It is worth mentioning that fuzzing has some limitations when it comes to practical use and have not gained wide popularity for automated vulnerability testing yet. However, considering the fact big companies, such as Google and Microsoft are using fuzzing as a part of their methodology and vigorously work on its development, it can be safely said that fuzzing has quite strong potential [29-32].

The most promising direction of future development of automated vulnerability search methods is fuzzing intellectualization [18-22]. This is about using deep learning methods to improve computational procedures for automated vulnerability search. It is believed that such approach can significantly improve the process of selecting input data, which will cause failure or errors in the target application.

3 Intellectual Fuzzing Algorithm

Fuzzing is a method of software and security vulnerabilities testing which is conducted by making multiple tests using mutated input data [21]. Repeated testing is performed with random mutation, and usually testing time is far from optimal. This article considers the problem of intellectual fuzzing and tries to find a solution [22]. The main goal is to develop a technology that can be guided and will make decisions based on the experience it gained during testing. The solution to this question lies in machine learning and reinforcement learning based on the deep Q-learning algorithm [33]. It uses maximum possible rewards, which are defined during development process by analyzing program source data and available rewards. This allows to apply optimal input data mutations. Thus, agent gets an opportunity to learn to formulate an optimal action policy for obtaining maximum reward. In this paper, we propose an algorithm and a computer model of the in-depth training as well as research of automated vulnerabilities testing effectiveness in comparison with the random mutation test. During testing, the "black box" method is used. It means that the available information represents only results of the program's work and the input data that it needs to perform [21].

During research, we realized there is a serious problem with randomized fuzzing: if it works with randomly generated input data, the time will not be optimal, since the process is performed blindfold. There may be a lot of testing rounds resulting in huge amounts of mutations that do not provide any progress. The process of fuzzing is an execution of a task cycle in certain defined program, where input is a sequence that was changed by some mutation. The ideal solution to solve such problem is machine-

learning technology called reinforcement learning. The best example of using this algorithm is the AlphaGO developed by Google DeepMind in 2015. It becomes the world's first program to win the game of "Go" with a top-ranked professional Lee Sedol [34].

As a combination of fuzzing and reinforcement training, a system capable of changing the rules for selecting specific mutation was created. It sends mutated data to input and, depending on the program's source data, generates a reward in order to rely on its own experience and select optimal mutations for particular case upon further testing. Thus, the amount of mutations does not contribute to the testing process significantly reduces. This makes testing process faster. Schematic diagram of the developed system is shown in Fig. 1. Testing process begins by defining the original non-mutated input data. Its format depends on fuzzing kind. Packages are submitted to the program's input channel and the response of the program is determined using special debugging software. From obtained data (this may be the program execution time, code coverage, code completion, etc.) system's State is formed. It will be pre-processed and presented to the Deep Q-learning model's input. The system then decides what Action should be taken next.

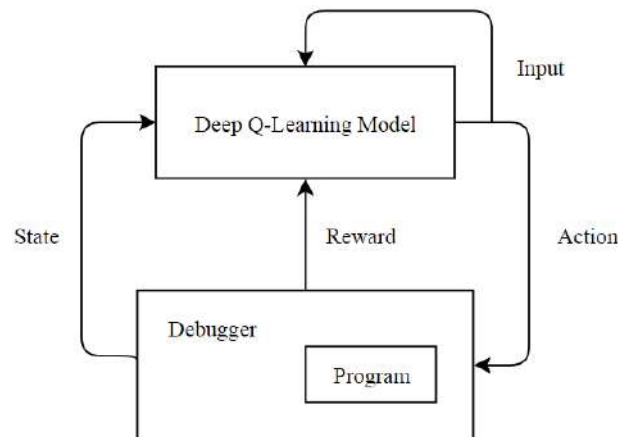


Fig. 1. Simplified diagram of the intellectual fuzzing algorithm

At the same time, depending on the selected action and the obtained program state, system forms the Reward for the algorithm. Using this reward, algorithm understands assigned task and determines optimal behavior for its implementation. In addition, algorithm remembers which actions brought it to the maximum reward (finding a mistake or a program failure, etc.) and, in the following testing rounds, decides what action should be taken based on its already gained experience. To calculate the next step, the previous inputs are mutated according to the action that was selected. Program input receives new mutated data. This procedure is repeated until the algorithm reaches its goal. Developed model uses Markov decision-making process – deep Q-learning [35].

4 Reinforcement Learning

This section provides necessary data about the algorithm of reinforcement training. Reinforcement learning is a computational approach to understanding and automating targeted learning and decision making. It stands out among other machine learning algorithms by the fact that agent learns directly by interacting with the environment without referring to examples [35]. This algorithm is primarily aimed to solve problems that arise during interaction with the environment to achieve long-term action. It uses formal structure of Markov decision-making process [35], defining the interaction between agent and environment in terms of states, actions and rewards. These features include understanding of causes and effects, as well as presence of clear goals. The notion of value and function of value is the main features of reinforcement training methods.

As noted earlier, the interaction between agent and environment can be described using Markov decision-making process $M = (S, A, P)$, where S – a set of system's states, A – action's set, P – transition probabilities set. For each state-action pair $(s, a) \in S \times A$, P is a set of probabilities $P(s' \vee s, a)$, where s' corresponds to the next system's state. Agent considers possible system states for the selected action, where each transition has its own reward $r(s, a)$, and studies the optimal behavior for maximizing the reward.

During the learning process, the main goal is to maximize the finite amount of rewards:

$$R = \sum_{t=0}^{\infty} \gamma^t r_{t+1},$$

where $\gamma \in (0, 1)$ – discount rate, which determines the remuneration priority over time. Action a_t at state s_t is determined by the policy of action $a_{t\pi}(\vee s_t)$. Policy π attaches considered possible states to action, which in turn determines agent's behavior. Expected cumulative reward for the policy-maker π is defined as:

$$Q^\pi(s, a) = E \left[\sum_{t=0}^{\infty} \gamma^t r_{t+1} \vee s_0 = s, a_0 = a \right].$$

The problem of finding the optimal value $Q_\pi(s, a)$ can be reduced to the procedure of function approximation. To achieve this $Q_\pi(s, a)$ needs to be updated after each iteration of receiving the award [36]. It is defined as

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha,$$

where α – learning rate. The entire procedure can be stated in the following sequence: agent gets the status s_t , takes action

$$a_t = \arg \max(Q(s_t, a)),$$

which defines the reward r_t , and causes the system to go to the state $s_t + 1$.

After receiving a reward r_t and state $s_t + 1$, agent determines the best possible effect

$$a_t + 1 = \arg \max(Q(s, a)).$$

Then it updates the value $Q(s_t, a_t)$. To approximate the function $Q(s_t, a_t)$, deep neural networks are used (it defines the name of deep Q-learning), which in turn aim to minimize the loss function:

$$L = (r + \gamma \cdot \max(Q(s_t + 1, a_t)) - Q(s_t, a_t))^2.$$

5 Simulation Results

To conduct an experiment simple fuzzer was used. Its main function is to implement input data selection for automated software vulnerability search. Software launch process was simulated and special logic tests were developed. The program could return an error code when certain data is received to compare fuzzing with and without artificial intelligence. Possible states of the system are presented in the form of data generated after the completion of the program. This data is transmitted by neural network, which consists of one input layer, two hidden layers with 50 neurons each and activation functions (Rectified Linear Unit): $f(x) = \max(0, x)$.

The output of the neural network has 45 elements, representing the number of possible mutations. Complete scheme of neural network for function approximation $Q(s_t, a_t)$ is shown in Fig. 2.

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 50)	150
activation_1 (Activation)	(None, 50)	0
dense_2 (Dense)	(None, 50)	2550
activation_2 (Activation)	(None, 50)	0
dense_3 (Dense)	(None, 45)	2295
activation_3 (Activation)	(None, 45)	0
Total params: 4,995		
Trainable params: 4,995		
Non-trainable params: 0		

Fig. 2. Neural network scheme Network training is performed using Adam optimization algorithm [36]

Let $f(\theta)$ be noisy target function: a stochastic scalar function is differentiated relatively to the parameter θ . The expected cost of this feature, $E[f(\theta)]$ relative to θ must be minimized. Using $f_1(\theta), \dots, f_T(\theta)$ the implementation of a stochastic function in the next steps $1, \dots, T$ is denoted. Stochasticity can be based on the estimation of random subsets (mini-groups) of data points or noise of a function. At $g_t = \nabla \theta f_t(\theta)$ the gradient is denoted – the vector of partial derivatives f_t relatively to θ , which is estimated by time t .

The algorithm updates exponential moving average values of the gradient (m_t) and the square of the gradient (v_t), where hyper parameters $\beta_1, \beta_2 \in [0, 1]$ control the exponential velocity of decomposition for these moving averages. The most moving averages are the estimations of the first moment (mean value) and the second moment (uncentered dispersion) of the gradient. However, these moving averages are initialized as zero vectors, which lead to the estimation of moments moving in zero direction, especially in the initial time steps and when expansion rates are small (for example, the value β_s close to 1). The good news is this bias initialization can be easily prevented by getting bug fixes m_{bt} and v_{bt} .

The algorithm itself has the following form:

Required input data:

- α : Learning speed
- $\beta_1, \beta_2 \in [0, 1]$: Exponential decay rates for moment estimates (standard settings $\beta_1 = 0.9, \beta_2 = 0.999$)
- $f(\theta)$: Stochastic function with parameter θ
- θ_0 : Vector of initial parameters
- Algorithm:
 - $m_0 \leftarrow 0$ (Initialize the first moment vector);
 - $v_0 \leftarrow 0$ (Initialize the second moment vector);
 - $t \leftarrow 0$ (Initialize the time);
 - while θ_t not converged:
 - o $t \leftarrow t + 1$;
 - o $g_t \leftarrow \nabla \theta f_t(\theta_t)$ (Take the gradient relative to the stochastic function during t);
 - o $m_t \leftarrow \beta_1 m_{t-1} + (1 - \beta_1) g_t$ (Update rejected first moment);
 - o $v_t \leftarrow \beta_2 v_{t-1} + (1 - \beta_2) g_{2t}$ (Update the second estimation for a rejection);
 - o $m_{bt} \leftarrow \frac{m_t}{1 - \beta_{t_1}}$ (Calculate the corrected bias estimation of the first moment);

- o $y b_t \leftarrow \frac{y_t}{1 - \beta_{t_2}}$ (Calculate the corrected bias estimation of the second moment);
- o $\theta_t \leftarrow \theta_t - 1 - \alpha \frac{m b_t}{\sqrt{v b_t}}$ (Update parameters);

end while

- return θ_t (Calculated parameters).

Input mutations were selected based on the standard list: increasing and decreasing line length, integer insertion, adding special characters (for example, "%s", which may also cause errors). In sum, 45 functions were created and placed in a dictionary for further use.

The system receives an award if execution time was greater than the previous or if there were errors occurred during testing. When an error occurs, the algorithm finishes its work. While forming this type of award, there was a problem when algorithm already found one error and started calling it repeatedly to get maximum reward. To avoid this problem, two constants must be set: $\gamma \in (0,1)$ – discount rate and $\varepsilon \in (0,1)$ – intelligence speed. The first constant was already being mentioned before. The second one determines how the algorithm is capable in terms of discovering new solutions. Random action will be chosen with probability ε , and the most profitable action will be selected with probability $1 - \varepsilon$. The hypothesis is a scientific assumption that made to explain any phenomenon and requires testing on theoretical basis in order to become a reliable scientific theory [37]. Statistical hypothesis is any assertion (assumption) concerning the type or distribution parameters of a certain feature of the objects being studied [37].

The following sequence of actions was necessary to test the hypotheses:

1. Make calculations of certain statistics, the distribution of which is known.
2. Find the P-value for the calculated results.
3. Make appropriate conclusions depending on the significance criterion and P-value.

A special test for identifying an error was developed. The hypothesis of the experiment is Q-learning fuzzing works faster than randomized one. The discount rate is set at 0.9, and the exploration speed is equal to 0.5. The latter parameter decreases 0.99 times after each era. The Student's t-test was used [38] to test the hypothesis.

Student's t-test – the general name for the class of methods for statistical criteria testing. It is based on comparison with the Student's distribution. The most common application of the criteria is related to checking the equality of mean values in two samples [38]. To use this criterion, some conditions must be met: the initial data must have normal distribution and dispersion must be equal.

The first group contains the results of testing using developed algorithm, while the second one presents the results of random mutations. Testing was performed in the following sequence: generate 15 experiments, and record the amount of mutations

necessary to find an error at the end of each. The results of the experiments are shown in Table 3.

The results of Student's t-test calculation: $t = -12.40$. The number of degrees of freedom:

$$v = 2n - 2 = 2 \times 15 - 2 = 28$$

Considering significance criteria

$$\alpha = 0.01 \text{ and } P\text{-value} = 2.763,$$

while $t < P\text{-value}$, the hypothesis is valid.

The result is statistically significant for a given criterion, if the probability of accidental occurrence of the same or extreme result is less than the given level (0.01) under the condition of loyalty of the null hypothesis.

Table 3. Experimental Results

№	Deep Q-learning model	Random selection of mutations
1	3671	3686
2	1191	1897
3	1879	3164
4	1640	3233
5	1966	10446
6	1585	5358
7	1135	1134
8	4877	752
9	2465	2157
10	3266	3684
11	1895	2026
12	2093	2993
13	1150	295
14	1181	3381
15	1153	358

The testing time of developed algorithm is better than the time of random mutations testing considering the fact the algorithm did not learn before. On average, developed algorithm finds an error after 2076 mutations, whereas random testing needs 2832 mutations. Represented algorithm finds error 30% faster.

This result proves the direction of research was chosen right. Particularly, it shows that the main problem of fuzzing (large amount of mutations) can be solved using artificial intelligence methods. Actually, if input data is changed directionally depending on previous results, it can speed up mutation process and receive results (which

are finding the vulnerability in certain software product) after less amount of mutations.

6 Conclusions

Due to conducted research, one of promising automated software testing methods in critically important systems was analyzed. Fuzzing bases on multiple input of different (mutated) data to find parameters, which will cause failure or incorrect functioning of software. Repeated testing is usually carried using randomized mutations and the time of testing is very high in the most cases. This article researches the problem of intellectual fuzzing – the technology, which uses previous testing experience to make choices, related to mutation, and reduce testing time.

Deep Reinforcement Learning algorithm was used to implement intellectual fuzzing. With the use of simple fuzzing app, it becomes possible to prove that testing time decreases by 30%. This result was received because fuzzer used previous experience to adjust mutations.

This research may continue in other spheres. For example, Intrusion Detection and Prevention Systems [39-42] are also can be built using some elements of artificial intelligence. Critically important information systems in different spheres, including banking, industrial facilities management and Smart Grids are especially interesting for further research [43-50].

References

1. Bitner, M.J., Zeithaml, V.A., Gremler, D.D., Maglio, P.P., Kieliszewski, C.A., Spohrer, J.C.: Technology's Impact on the Gaps Model of Service Quality" in Handbook of Service Science - Service Science: Research and Innovation in the Service Economy, US, Boston, MA: Springer (2010)
2. Brocke, J., Thurnher, B., Winkler, D.: Improving IT-Service Business Processes by Integrating Mobility: The Role of User Participation-Results from Five Case Studies. In: International Conference on eLearning e-Business Learning e-Business Enterprise Information Systems and e-Government (EEE 2008) (2008)
3. Budzik, J., Hammond, K.: Watson: Anticipating and Contextualizing Information Needs. In: Proc. of the 62 Annual Meeting of the American Society for Inform. Science, 1999, vol. 36, pp. 727–740 (1999)
4. World Manufacturing Production 2012. United Nations Industrial Development Organization, 2013
5. Structural Change in the World Economy. United Nations Industrial Development Organization, 2009
6. Bush, S.F., Goel, S., Simard, G.: IEEE Vision for Smart Grid Communications: 2030 and Beyond. In: IEEE Vision for Smart Grid Communications: 2030 and Beyond, 2014, pp. 1492–1499. (2014) doi:10.1109/IEEESTD.2013.6690098
7. Albarakati, A., Moussa, B., Debbabi, M., Youssef, A., Agba, B.L., Kassouf, M.: Open-Stack-Based Evaluation Framework for Smart Grid Cyber Security. In: 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart

- Grids (SmartGridComm), Aalborg, 2018, pp. 1–6. (2018) doi:10.1109/SmartGridComm.2018.8587420
8. Hariri, M.E., Youssef, T., Habib, H.F., Mohammed, O.: A Network-in-the-Loop Framework to Analyze Cyber and Physical Information Flow in Smart Grids. In: 2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), Singapore, 2018, pp. 646–651 (2018)
9. Sokolov, S.S., Glebov, N.B., Antonova, E.N., Nyrkov, A.P.: The Safety Assessment of Critical Infrastructure Control System. In: 2018 IEEE International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), St. Petersburg, 2018, pp. 154–157. (2018) doi:10.1109/itmqls.2018.8524948
10. Krasnobayev, V., Kuznetsov, A., Koshman, S., Moroz, S.: Improved Method of Determining the Alternative Set of Numbers in Residue Number System. In: Recent Developments in Data Science and Intelligent Analysis of Information. ICDSIAI 2018. Advances in Intelligent Systems and Computing, Springer, Cham, 2019, vol. 836, pp. 319–328. (2019) doi:10.1007/978-3-319-97885-7_31
11. Cybersecurity for Smart Grid Systems (25 June 2018). <https://www.nist.gov/programs-projects/cybersecurity-smart-grid-systems>
12. Andrushkevych, A., Gorbenko, Y., Kuznetsov, O., Oliynykov, R., Rodinko, M.: A Prospective Lightweight Block Cipher for Green IT Engineering. In: Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control, Springer, Cham, 2019, vol. 171, pp. 95–112. (2019) doi:10.1007/978-3-030-00253-4_5
13. Ferguson, N., Schneier, B.: Practical Cryptography. John Wiley & Sons (2003)
14. Kuznetsov, O., Potii, O., Perepelitsyn, A., Ivanenko, D., Poluyanenko, N.: Lightweight Stream Ciphers for Green IT Engineering. In: Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control, Springer, Cham, 2019, vol. 171, pp. 113–137. (2019) doi: 10.1007/978-3-030-00253-4_6
15. Yinghua, H., Yanchun, M., Dongfang, Z.: The Multi-join Query Optimization for Smart Grid Data. In: 2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA), Nanchang, 2015, pp. 1004–1007. (2015) doi:10.1109/ICICTA.2015.255
16. Du, Z.: Intelligence Computation and Evolutionary Computation. Springer-Verlag Berlin Heidelberg (2013)
17. Tong, C., Gao, Y., Tong, M., Li, J., Wang, Q., Chen, K.: Application of lightning detection in Smart Grid dynamic protection. In: 2012 International Conference on Lightning Protection (ICLP), Vienna, 2012, pp. 1–13. (2012) doi:10.1109/ICLP.2012.6344209
18. Choi, S.Y., Lee, H.Y.: Toward Automated Scanning for Code Injection Vulnerabilities in HTML5-Based Mobile Apps. In: 2016 International Conference on Software Security and Assurance (ICSSA), St. Polten, 2016, pp. 24–24. (2016) doi:10.1109/ICSSA.2016.11
19. Peng, H., Shoshitaishvili, Y., Payer, M.: T-Fuzz: Fuzzing by Program Transformation. In: 2018 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, 2018, pp. 697–710. (2018) doi:10.1109/SP.2018.00056
20. Zhao, J., Pang, L.: Automated Fuzz Generators for High-Coverage Tests Based on Program Branch Predications. In: 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), Guangzhou, 2018, pp. 514–520. (2018) doi:10.1109/DSC.2018.00082
21. Sutton, M., Greene, A., Amini, P.: Fuzzing: Brute Force Vulnerability Discovery. 1st ed. Boston, MA, USA: Addison-Wesley Professional (2007) <http://bxi.es/Reversing-Exploiting/Fuzzing%20Brute%20Force%20Vulnerability%20Discovery.pdf>

22. Böttinger, K., Godefroid, P., Singh, R.: Deep Reinforcement Fuzzing. Submitted on 14 Jan 2018. (2018) <https://arxiv.org/abs/1801.04589>
23. Miller, B.P., Fredriksen, L., Barton, B.S., Miller, P.: An Empirical Study of the Reliability of UNIX Utilities (1989) ftp://ftp.cs.wisc.edu/paradyn/technical_papers/fuzz.pdf
24. Miller, B.P.: Fuzz Revisited: A Re-examination of the Reliability of UNIX Utilities and Services (1995) ftp://ftp.cs.wisc.edu/paradyn/technical_papers/fuzz-revisited.pdf
25. Forrester, J.E., Miller, B.P.: An Empirical Study of the Robustness of Windows NT Applications Using Random Testing. Computer Sciences Department University of Wisconsin Madison, WI 53706–1685 (2000) ftp://ftp.cs.wisc.edu/paradyn/technical_papers/fuzz-nt.pdf
26. Miller, B.P., Cooksey, G., Moore, F.: An Empirical Study of the Robustness of MacOS Applications Using Random Testing. Computer Sciences Department University of Wisconsin Madison, WI 53706–1685 (2006) ftp://ftp.cs.wisc.edu/paradyn/technical_papers/Fuzz-MacOS.pdf
27. René, F.: The Art of Fuzzing (2017) https://www.sec-consult.com/wp-content/uploads/files/vulnlab/the_art_of_fuzzing_slides.pdf
28. Guided in-process fuzzing of Chrome components. Google Security Blog. <https://security.googleblog.com/2016/08/guided-in-process-fuzzing-of-chrome.html>
29. American fuzzy lop. <http://lcamtuf.coredump.cx/afl/>
30. Böck, H.: How Heartbleed could've been found (2015) <https://blog.hboeck.de/archives/868-How-Heartbleed-couldve-been-found.html>
31. Dallman, J., Pulikkathara, G.: MiniFuzz File Fuzzer Overview (2011) <https://msdn.microsoft.com/en-us/biztalk/gg675011>
32. Peach Fuzzer Community Edition. <http://www.peach.tech/resources/peachcommunity/>
33. Li, Y.: Deep Reinforcement Learning: An Overview. Submitted on 25 Jan 2017 (v1), last revised 26 Nov 2018 (this version, v6) (2018) <https://arxiv.org/abs/1701.07274>
34. AlphaGo Games. <https://deepmind.com/research/alphago/match-archive/alphago-games-english/>
35. Sutton, R.S., Barto, A.G.: Reinforcement learning: An introduction. MIT press Cambridge (1998) <http://incompleteideas.net/book/bookdraft2017nov5.pdf>
36. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization”. Submitted on 22 Dec 2014 (v1), last revised 30 Jan 2017 (this version, v9) (2017) <https://arxiv.org/abs/1412.6980>
37. Statistical hypotheses and their verification (in Ukrainian). <http://stat.org.ua/statclasses/hypotheses-testing/>
38. t-criterion for Student (in Ukrainian). <http://fpo.bsmu.edu.ua/static/t-kryteriy-styudenta>
39. Muller, S.: Risk Monitoring and Intrusion Detection for Industrial Control Systems. Doctor de l'Université du Luxembourg en Informatique, University of Luxembourg, Luxembourg (2018)
40. Kuznetsov, A.A., Smirnov, A.A., Danilenko, D.A., Berezovsky, A.: The statistical analysis of a network traffic for the intrusion detection and prevention systems. Telecommunications and Radio Engineering, 2015, vol. 74(1), pp. 61–78. (2015) doi: 10.1615/TelecomRadEng.v74.i1.60
41. Jahan, S., Habiba, R.: An analysis of smart grid communication infrastructure & cyber security in smart grid. In: 2015 International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, 2015, pp. 190–193. (2015) doi:10.1109/ICAEE.2015.7506828
42. Wang, Y., Zhang, B., Lin, W., Zhang, T.: Smart grid information security - a research on standards. In: 2011 International Conference on Advanced Power System Automation and Protection, Beijing, 2011, pp. 1188–1194. (2011) doi:10.1109/APAP.2011.6180558

43. Rassomakhin, S., Kuznetsov, A., Shlokin, V., Belozertsev, I., Serhiienko, R.: Mathematical Model for the Probabilistic Minutia Distribution in Biometric Fingerprint Images. In: 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 2018, pp. 514–518. (2018) doi:10.1109/DSMP.2018.8478496
44. Borkar, A., Donode, A., Kumari, A.: A survey on Intrusion Detection System (IDS) and Internal Intrusion Detection and protection system (IIDPS). In: 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 949–953. (2017) doi:10.1109/ICICI.2017.8365277
45. Moskovchenko, I., Pastukhov, M., Kuznetsov, A., Kuznetsova, T., Prokopenko, V., Kropyvnytskyi, V.: Heuristic Methods of Hill Climbing of Cryptographic Boolean Functions. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 1–6. (2018) doi:10.1109/INFOCOMMST.2018.8632017
46. Lei, H., Chen, B., Butler-Purry, K.L., Singh, C.: Security and Reliability Perspectives in Cyber-Physical Smart Grids. In: 2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), Singapore, 2018, pp. 42–47. (2018) doi:10.1109/ISGT-Asia.2018.8467794
47. Kuznetsov, A., Kavun, S., Panchenko, V., Prokopovych-Tkachenko, D., Kurinniy, F., Shoiko, V.: Periodic Properties of Cryptographically Strong Pseudorandom Sequences. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 129–134. (2018) doi:10.1109/INFOCOMMST.2018.8632021
48. Kavun, S.: Conceptual fundamentals of a theory of mathematical interpretation. *Int. J. Computing Science and Mathematics*, 2015, vol. 6(2), pp. 107–121. (2015) doi:10.1504/IJCSM.2015.069459
49. Kavun, S.: Indicative-geometric method for estimation of any business entity. *Int. J. Data Analysis Techniques and Strategies*, 2016, vol. 8(2), pp. 87–107. (2016) doi:10.1504/IJDATS.2016.077486
50. Kovtun, V., Kavun, S., Zyma, O.: Co-Z Divisor Addition Formulae in Homogeneous Representation in Jacobian of Genus 2 Hyperelliptic Curves over Binary Fields. *International Journal Biomedical Soft Computing and Human Sciences*, 2012, vol. 17(2), pp. 37–43. (2012) doi:10.24466/ijbschs.17.2_37

Method of data control in the residue classes

Victor Krasnobayev ^[0000-0001-5192-9918], Alexandr Kuznetsov ^[0000-0003-2331-6326],

Anna Kononchenko ^[0000-0002-8101-6500], Tetiana Kuznetsova ^[0000-0001-6154-7139]

V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, 61022, Ukraine

v.a.krasnobaev@gmail.com, kuznetsov@karazin.ua,
akononpro@gmail.com, kuznetsova.tatiana17@gmail.com

Abstract. Methods of data control in the residue classes are considered in the article. The main advantage of non-positional notation in the residue classes lays in the possibility of an organization of the process of quick implementation of modular arithmetic operations of addition, subtraction and multiplication. The base of the method is the procedure of generating and using the positional indication of non-positional code. That allows increasing efficiency of the procedure of data control granted by the residue classes.

Keywords. Methods of data control, the system of the residue classes, the positional indications of the non-positional code, computer systems and components.

1 Introduction

It is well known, that the main advantage of non-positional notation in the residue classes (NRC) is laid in the possibility of quick process of organization of modular arithmetic operations of addition, subtraction and multiplication [1, 2].

However, in computer systems for common purpose it is needed to perform so-called non-modular (position) operations except the above listed arithmetic operations. The following operations belong to these:

- Arithmetic and algebraic congruence of numbers and its absolute values;
- Sign of number definition;
- Definition of existing of overflow of bit grid of computer system (CS);
- Rounding of value of result of operation;
- Computing the absolute value;
- Division and multiplication of fraction
- Translation of data from code of NRC to positional notation and back;
- Expanding of original NRC (it is an informational process in which familiar remainders $\{a_i\}$, that correspond to basis $\{m_i\}$, define value of remainders with the same code structure by other additional basis);

- Control, diagnostic and correction mistakes of NRC data, etc.

Generally, all positional operations come down to the procedure of definition of an index of j numerical $[jm_i, (j+1)m_i)$ interval of entering (detecting) of number $A=(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$. It is appropriate to use so-called positional indications of non-positional code (PINC) to define an index of j numerical interval of detecting A . The following indications are the most frequently used in NRC (within existing PINCs) [3, 4]:

- Indications based on the procedure of number's translating from NRC to PINC;
- Indications based on the procedure of nulevization (reduction to zero polynomial) (definition of the value of y_{n+1});
- Indications based on the procedure of expanding given system of basis of NRC;
- Rank r of number A , etc.

There are disadvantages of the above-mentioned PINC. At first, it is technical and time complexity of PINC generating (developing) for given code structure $A=(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$. At second, it is no mean time of implementation having applied the existing positional indications, non-modular operations in NRC, specifically, operations of control, diagnostic of correction of data mistakes [3, 4].

It becomes obvious that the research of developing methods of generating new PINC in NRC is important; because of them, we can use operatively non-modular operations. It should be pointed out, population (sequence) of defined modular and non-modular operations, which are implemented by PINC, can realize any non-modular operation.

Beforehand we will consider the general requirements to PINC, on base of which in article will be developed method of increasing of operational efficiency of data control:

- To define exactly correctness and incorrectness of number A in NRC (to define fact of detecting/not detecting of number A in an informational numerical $[0, M]$

interval, where $M = \prod_{i=1}^n m_i$) by used (chosen, developed, generated) PINC;

- Simplicity of generating PINC for given code structure of data $A=(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$;
- Simplicity of using of generated indication for data control in NRC (in general case for implementing positional operations);
- Indication has to have clear and understandable physical meaning;
- Indication must be analytically described by simple mathematical relator;
- It is possible to technically implement process of data control in NRC by using PINC;
- Using the chosen indication of non-positional code has to provide the given fidelity of data control in NRC;

- PINC using has to provide the possibility of exception from the control procedure, diagnostic and correction of mistakes in NRC the most difficult positional operations.

It is reasonable to develop and research the method of data control in NRC on the base of PINC using, resulting from the above-mentioned information.

2 Method of generating of positional indication of non-positional code structure of data in NRC

Looking upon method of generating of PINC (Fig. 1).

The indication n_A forms proceeding from representation of initial code structure $A=(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$, which is represented as NRC with basis $\{m_i\}, i = \overline{1, n+1}$, in the way of creating of so-called uniserial code (UC).

In standard form, UC

$$K_N^{(n_A)} = \{Z_{N-1}^{(A)} Z_{N-1}^{(A)} \dots Z_1^{(A)} Z_0^{(A)}\}$$

represents sequence of binary bits, which contains ones and a zero, index of which is at n_A place (the count is performed from right to left, from digit bit $Z_0^{(A)}$ to digit bit $Z_{N-1}^{(A)}$). Indication of PINC n_A defines an index of j numerical interval $[jm_i, (j+1)m_i)$ of detecting of number $A=(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$. Mathematically PINC n_A represents a positive integer, which points out the location of zero binary bit in the record of UC $K_N^{(n_A)} (Z_{n_A}^{(A)} = 0)$.

The procedure of generating UC $K_N^{(n_A)}$ lies in the following. A constant

$$KH_{m_i}^{(A)} = (a_1', \dots, a_{i-1}', a_i, a_{i+1}', \dots, a_{n+1}') \quad (1)$$

defines in the block of constants of nuvelization (BCN). It is performed by the definition of remainder a_i of number $A=(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$ for chosen integer m_i of NRC.

Further, using the chosen constant of nuvelization $KH_{m_i}^{(A)}$, we displace number A on the left edge of the interval $[jm_i, (j+1)m_i)$ by implementing operation

$$\begin{aligned} A_{m_i} &= A - KH_{m_i}^{(A)} = (a_1, a_2, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_{n+1}) - (a_1', a_2', \dots, a_{i-1}', a_i, a_{i+1}', \dots, a_{n+1}') = \\ &= [a_1^{(1)}, a_2^{(1)}, \dots, a_{i-1}^{(1)}, 0, a_{i+1}^{(1)}, \dots, a_{n+1}^{(1)}]. \end{aligned}$$

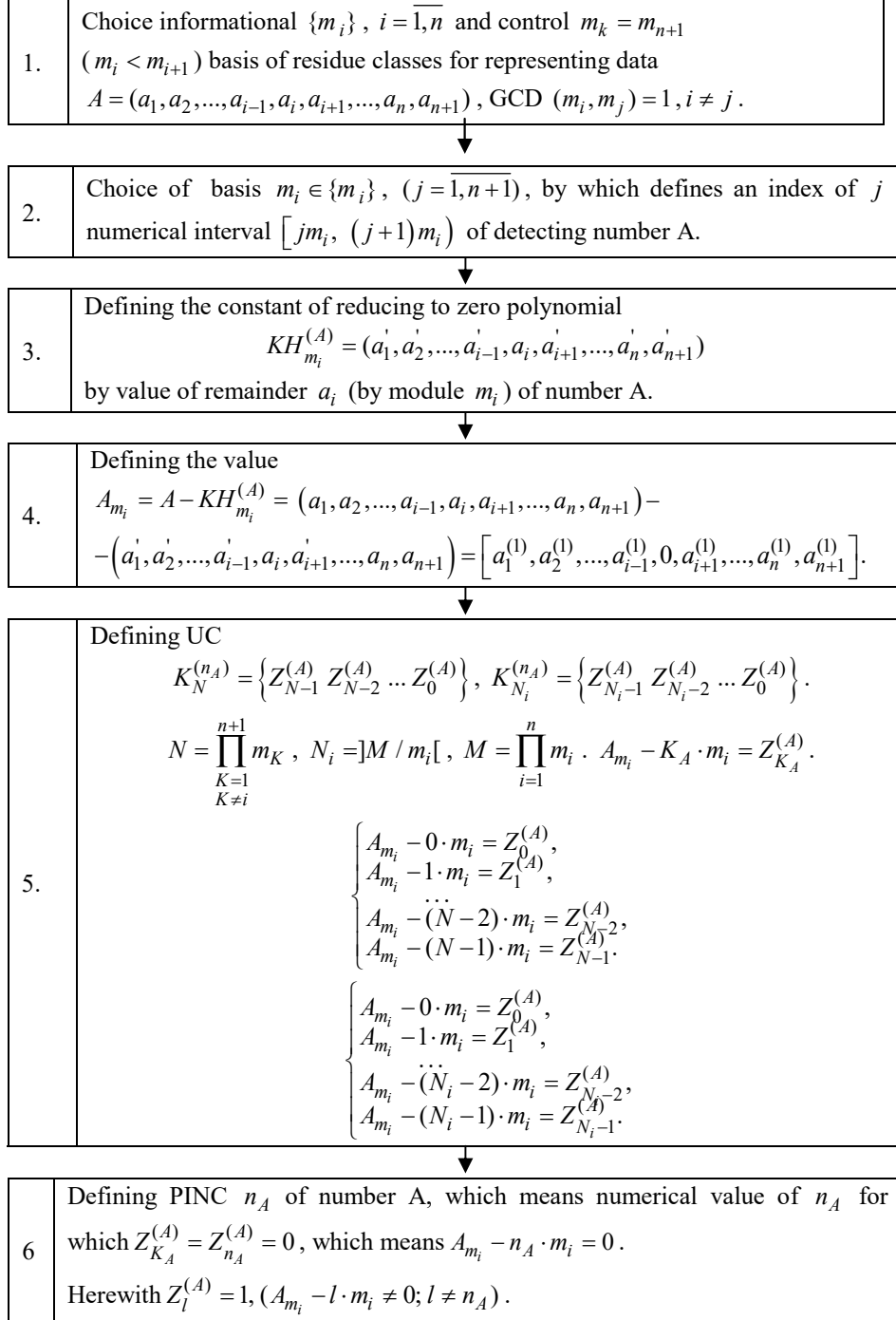


Fig. 1. The method of generating PINC in NRC

It is obvious, that number A is aliquot to value of module m_i of NRC.

It is known, that an indication of the correctness of number A in NRC defines by it's entering or not entering in its numerical informational interval $[0, M)$. If number A is out of the interval $(A \geq M)$, A is considered as garbling (wrong). In this case, PINC n_A has to define the fact of entering or not entering initial number A in the interval $[0, M)$. We have to implement operation

$$A_{m_i} - K_A \cdot m_i = Z_{K_A}^{(A)} \quad (1)$$

to define the fact of detecting number in numerical informational interval $[0, M)$.

The operation (1) is carried out simultaneously using the population of N constants

$$K_A \cdot m_i \left(K_A = \overline{0, N-1} \right),$$

where $N = \prod_{\substack{K=1 \\ K \neq i}}^{n+1} m_K$:

$$\begin{cases} A_{m_i} - 0 \cdot m_i = Z_0^{(A)}, \\ A_{m_i} - 1 \cdot m_i = Z_1^{(A)}, \\ A_{m_i} - 2 \cdot m_i = Z_2^{(A)}, \\ \dots \\ A_{m_i} - (N-2) \cdot m_i = Z_{N-2}^{(A)}, \\ A_{m_i} - (N-1) \cdot m_i = Z_{N-1}^{(A)}. \end{cases} \quad (2)$$

In this case, UC represents in the view

$$K_N^{(n_A)} = \{ Z_{N-1}^{(A)} Z_{N-2}^{(A)} \dots Z_1^{(A)} Z_0^{(A)} \} \quad (3)$$

The only value n_A from (1) exists in total (2) analytical rations. For n_A there is

$$Z_{K_A}^{(A)} = Z_{n_A}^{(A)} = 0 \quad (K_A = n_A),$$

which means

$$A_{m_i} - n_A \cdot m_i = 0.$$

The rest of values (2) equal

$$Z_l^{(A)} = 1 \quad (A_{m_i} - l \cdot m_i \neq 0; l \neq n_A).$$

In the general case, the amount of N binary bits in the record UC $K_N^{(n_A)}$ equal to value

$$N = \prod_{\substack{K=1 \\ K \neq i}}^{n+1} m_K .$$

It should be pointed out, that there is no necessity to have the whole sequence of values $Z_{K_A}^{(A)}$ from (3) to define only the fact of the garbling of number $(A \geq M)$. It is enough to have UC $K_{N_i}^{(n_A)}$ with length only $N_i =]M / m_i[$ of binary bits (where value $]M / m_i[$ describes the lesser integer of number M / m_i).

In this case values of variables of numerical intervals $[jm_i, (j+1)m_i)$, which are out of informational interval $[0, M)$, make no matter for establishing the fact of correctness control of number A.

3 Method of operational data control in NRC on base of using positional indication of non-positional code

The procedure of generating of positional indication of non-positional code (fig. 1) laid in the base of method of operational data control in residue classes. In that way, the essence of the method of data control in residue classes lies in the following. For the controlled code structure $A = (a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n, a_{n+1})$, which is represented in residue classes, developed (defined) PINC n_A by generating UC

$$K_{N_i}^{(n_A)} = \{Z_{N_i-1}^{(A)} Z_{N_i-2}^{(A)} \dots Z_1^{(A)} Z_0^{(A)}\}$$

in view of sequence of N_i binary bits. Choosing of basis m_i of NRC is performed by special-purpose in accordance with defined measures. The constant of nuvelization

$$KH_{m_i}^{(A)} = (a'_1, a'_2, \dots, a'_i, \dots, a'_n, a'_{n+1})$$

is selected depending on a result of the value of remainder a_i of number A. Further implementation operation is carried out:

$$\begin{aligned} A_{m_i} &= A - KH_{m_i}^{(A)} = (a_1, a_2, \dots, a_i, \dots, a_n, a_{n+1}) - (a'_1, a'_2, \dots, a'_i, \dots, a'_n, a'_{n+1}) = \\ &= [a_1^{(1)}, a_2^{(1)}, \dots, 0, \dots, a_n^{(1)}, a_{n+1}^{(1)}]. \end{aligned}$$

Using N_i constants $K_A \cdot m_i$ ($K_A = \overline{0, N_i - 1}$) simultaneously the operations of subtractions $A_{m_i} - K_A \cdot m_i$ are carried out, in the result of which appears the values of binary bits $Z_{K_A}^{(A)}$, so the UC $K_{N_i}^{(n_A)}$ forms.

The values of PINC n_A defined from the equation $A_{m_i} - n_A \cdot m_i = 0$.

If $n_A > N_i$, number A is considered as a wrong. In the opposite case ($n_A \leq N_i$) number A is correct. In common view, method of data control is represented on Fig. 2.

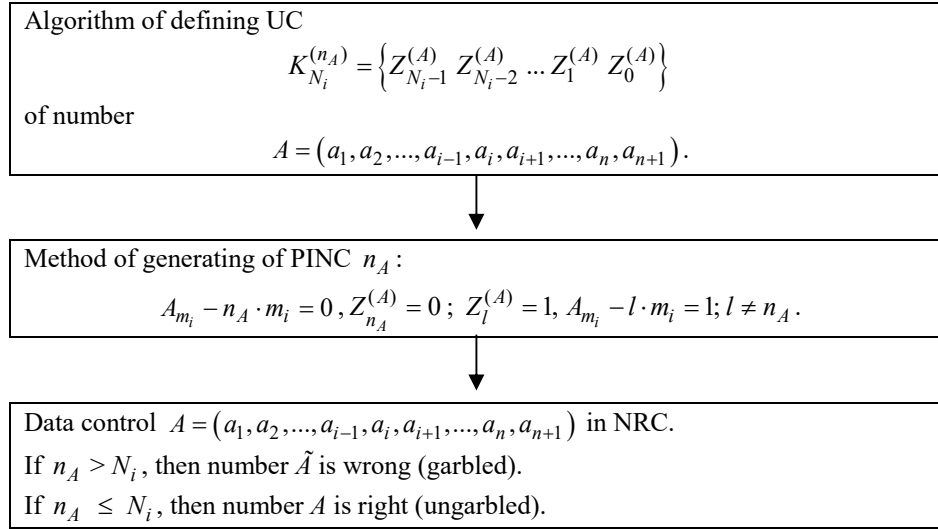


Fig. 2. Method of data control in NRC

Examine now the examples of implementing of method of control for specific NRC, which is given with basis $m_1=3, m_2=4, m_3=5, m_4=7$ and $m_k = m_{n+1} = m_5 = 11$. The NRC provides the data handling in single-byte (1=1) bit grid of CS. Herewith

$$M = \prod_{i=1}^4 m_i = 420, M_0 = M m_{n+1} = 4620;$$

$$N_i = N_{n+1} = \lceil M / m_i \rceil = \lceil M / m_{n+1} \rceil = \lceil 420 / 11 \rceil = \lceil 38,18 \rceil = 39.$$

Contents of the block of constants of nuvelization (BCN) concerning basis $m_k = m_{n+1} = 11$ are given in Table 1.

Example 1. Perform control of data that represented in view $A = (01, 00, 000, 010, 0001)$.

Constant of nuvelization

$$KH_{m_{n+1}}^{(A)} = (01, 01, 001, 001, 0001)$$

is selected by values of remainders $a_K = a_{n+1} = a_5 = 0001$ of number A in BCN CS (Table 1). Then we define value

Table 1. Constants $KH_{m_{n+1}}^{(A)}$ of nuvelization BCN

Remainder $a_k = a_{n+1}$	Constant of nuvelization				
	$m_1 = 3$	$m_2 = 4$	$m_3 = 5$	$m_4 = 7$	$m_k = m_5 = 11$
	a'_1	a'_2	a'_3	a'_4	a_5
0000	00	00	000	000	0000
0001	01	01	001	001	0001
0010	10	10	010	010	0010
0011	00	11	011	011	0011
0100	01	00	100	100	0100
0101	10	01	000	101	0101
0110	00	10	001	110	0110
0111	01	11	010	000	0111
1000	10	00	011	001	1000
1001	00	01	100	010	1001
1010	01	10	000	011	1010

$$A_{m_{n+1}} = A - KH_{m_{n+1}}^{(A)} = (00,11,100,001,0000) .$$

By implementation of ratio (2) we form UC

$$K_{N_i}^{(n_A)} = K_{39}^{(9)} = \{11...110111111111\} .$$

Resulting from the view of UC and using formula

$$A_{m_{n+1}} - n_A \cdot m_{n+1} = 0 ,$$

we define that $n_A = 9$

$$(A_{m_{n+1}} - n_A \cdot m_{n+1} = 99 - 9 \cdot 11 = 0) ,$$

meaning $Z_{n_A}^{(A)} = Z_9^{(A)} .$

Because of $N_i = 39 > n_A = 9$ it means that there is no mistake in data.

Check: $A = 100 < M = 420$ (number A is right).

Example 2. Perform control of data $A = (00,10,000,010,1010) .$

The constant

$$KH_{m_{n+1}}^{(A)} = (01,10,000,011,1010)$$

is selected by value of $a_5 = 1010$ in BCN (Table 1).

We deduce that

$$A_{m_{n+1}} = A - KH_{m_{n+1}}^{(A)} = (10,00,000,110,0000) .$$

Because

$$A_{m_{n+1}} - n_A \cdot m_{n+1} = 440 - 44 \cdot 11 = 0$$

then UC has view $K_{N_i}^{(n_A)} = K_{39}^{(40)} = \{11...11...11\}$ and $n_A = 40$.

Because of $N_i = 39 < n_A = 40$ it means that there is a mistake in data.

Check: $A = 450 > M = 420$ (number A is wrong).

Example 3. Perform control of data $A = (01,11,010,000,1001)$.

The constant

$$KH_{m_{n+1}}^{(A)} = (00,01,100,010,1001)$$

is selected by value of $a_5 = 1001$ in BCN (Table 1). We deduce that

$$A_{m_{n+1}} = A - KH_{m_{n+1}}^{(A)} = (01,10,011,101,0000) .$$

Because

$$A_{m_{n+1}} - n_A \cdot m_{n+1} = 418 - 38 \cdot 11 = 0$$

then UC has view

$$K_{N_i}^{(n_A)} = K_{39}^{(38)} = \{011...11...11\}$$

and $n_A = 38$.

Because $n_A = 38 < N_i = 39$ of it means that number A is right (ungarbled). Though the check $A=427 > M = 420$ shows us that number A is wrong (Fig. 3).

The represented method can be used for improving promising computer systems and their components and in the other practically important applications [5-9]. In particular, mathematical transformation in the system of the residue classes can be successfully adopted for optimization of calculations of cryptographic methods of data protection [10-15, 24-27], also in code theories and complicated discrete signals [16-21], in authentication and steganography [22, 23].

4 Conclusions

Thus, the method of data control in the system of residue classes is presented in the article. The procedure of forming and using of position indication of non-positional code is the base for a method of operational control of data in a residue class. Use of PINC allows to increase efficiency of the procedure of data control provided to NRC.

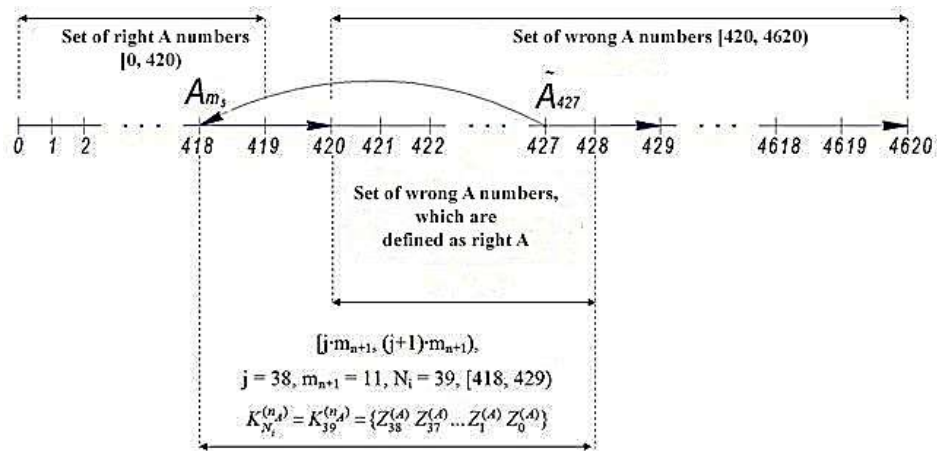


Fig. 3. Example of implementing operation of data control in NRC for $m_{n+1} = m_5 = 11$

It should be pointed out, that any non-modular operation can be implemented by set (sequence) of defined modular and non-modular operations, which are implemented by PINC. Using PINC in the method provides the possibility of exception of the most complicated positional operations from the procedure of control, diagnostic and correction of mistakes in NRC.

References

1. Akushsky, I.Y., Yuditsky, D.I.: Machine modular arithmetic. M.: Sov. Radio (in Russian) (1968)
2. Kolyada, A.A., Pak, I.T.: Modular structures of pipeline processing of digital information. Minsk: University (in Russian) (1992)
3. Krasnobayev, V., Kuznetsov, A., Koshman, S., Moroz, S.: Improved Method of Determining the Alternative Set of Numbers in Residue Number System. In: Recent Developments in Data Science and Intelligent Analysis of Information. ICDSIAI 2018. Advances in Intelligent Systems and Computing, Springer, Cham, 2018, vol. 836, pp. 319–328. (2019) doi:10.1007/978-3-319-97885-7_31
4. Krasnobayev, V.A., Yanko, A.S., Koshman, S.A.: A Method for arithmetic comparison of data represented in a residue number system. Cybernetics and Systems Analysis, 2016, vol. 52(1), pp. 145–150. (2016) doi:10.1007/s10559-016-9809-2
5. Xiao, H., Wu, D.: A Component Model for Network Processor Based System. In: 2007 IEEE/ACS International Conference on Computer Systems and Applications, Amman, 2007, pp. 47–50. (2007) doi:10.1109/AICCSA.2007.370863
6. Andrushkevych, A., Gorbenko, Y., Kuznetsov, O., Oliynykov, R., Rodinko, M.: A Prospective Lightweight Block Cipher for Green IT Engineering. In: Kharchenko V., Kondratenko Y., Kacprzyk J. (eds) Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control, Springer, Cham, 2019, vol. 171, pp. 95–112. (2019) doi:10.1007/978-3-030-00253-4_5

7. Kuznetsov, O., Potii, O., Perepelitsyn, A., Ivanenko, D., Poluyanenko, N.: Lightweight Stream Ciphers for Green IT Engineering. In: Kharchenko V., Kondratenko Y., Kacprzyk J. (eds) *Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control*, Springer, Cham, 2019, vol. 171, pp. 113–137. (2019) doi:10.1007/978-3-030-00253-4_6
8. Wang, A., Zhang, A., Chen, S.: Study On simulation and design of data links human computer interface. In: 2011 International Conference on Computer Science and Service System (CSSS), Nanjing, 2011, pp. 4066–4069. (2011) doi:10.1109/CSSS.2011.5974893
9. Irfan, S., Ghosh, S.: Optimization of information retrieval using evolutionary computation: A survey. In: 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, 2017, pp. 328–333. (2017) doi:10.1109/CCAA.2017.8229837
10. Moskovchenko, I., Pastukhov, M., Kuznetsov, A., Kuznetsova, T., Prokopenko, V., Kropyvnytskyi, V.: Heuristic Methods of Hill Climbing of Cryptographic Boolean Functions. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 1–6. (2018) doi:10.1109/INFOCOMMST.2018.8632017
11. Kuznetsov, A., Kavun, S., Panchenko, V., Prokopovych-Tkachenko, D., Kurinnii, F., Shoiko, V.: Periodic Properties of Cryptographically Strong Pseudorandom Sequences. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 129–134. (2018) doi:10.1109/INFOCOMMST.2018.8632021
12. Gorbenko, I., Kuznetsov, A., Tymchenko, V., Gorbenko, Y., Kachko, O.: Experimental Studies of the Modern Symmetric Stream Ciphers. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 125–128. (2018) doi:10.1109/INFOCOMMST.2018.8632058
13. Rodinko, M., Oliynykov, R.: An Approach to Search for Multi-Round Differential Characteristics of Cypress-256. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 659–662. (2018) doi:10.1109/INFOCOMMST.2018.8631904
14. Ivanov, O., Ruzhentsev, V., Oliynykov, R.: Comparison of Modern Network Attacks on TLS Protocol. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 565–570. (2018) doi:10.1109/INFOCOMMST.2018.8632026
15. Gruschka, N., Mavroeidis, V., Vishi, K., Jensen, M.: Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR. In: 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 5027–5033. (2018) doi:10.1109/BigData.2018.8622621
16. Kuznetsov, A., Serhiienko, R., Prokopovych-Tkachenko, D.: Construction of cascade codes in the frequency domain. In: 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2018, pp. 131–136. (2017) doi:10.1109/INFOCOMMST.2017.8246366
17. Kuznetsov, A., Kiyan, A., Uvarova, A., Serhiienko, R., Smirnov, V.: New Code Based Fuzzy Extractor for Biometric Cryptography. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 119–124. (2018) doi:10.1109/INFOCOMMST.2018.8632040
18. Gorbenko, I.D., Zamula, A.A.: Cryptographic signals: requirements, methods of synthesis, properties, application in telecommunication systems. *Telecommunications and Radio En-*

- gineering, 2017, vol. 76(12), pp. 1079–1100. (2017) doi:10.1615/TelecomRadEng.v76.i12.50
19. Kuznetsov, A., Pushkar'ov, A., Kiyan, A., Kuznetsova, T.: Code-based electronic digital signature. In: 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT), Kyiv, Ukraine, 2018, pp. 331–336. (2018) doi:10.1109/DESSERT.2018.8409154
 20. Gorbenko, I.D., Zamula, A.A., Semenko, Ye.A.: Ensemble and correlation properties of cryptographic signals for telecommunication system and network applications. *Telecommunications and Radio Engineering*, 2016, vol. 75(2), pp. 169–178. (2016) doi:10.1615/TelecomRadEng.v76.i17.40
 21. Gorbenko, I.D., Zamula, A.A., Semenko, A.E., Morozov, V.L.: Method for synthesis of performed signals systems based on cryptographic discrete sequences of symbols. *Telecommunications and Radio Engineering*, 2017, vol. 76(17), pp. 1523–1533. (2017) doi:10.1615/TelecomRadEng.v76.i17.40
 22. Kuznetsov, A., Shekhanin, K., Kolhatin, A., Mikheev, I., Belozertsev, I.: Hiding data in the structure of the FAT family file system. In: 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT), Kyiv, Ukraine, 2018, pp. 337–342. (2018) doi:10.1109/DESSERT.2018.8409155
 23. Gorbenko, Y., Svatovskiy, I., Shevtsov, O.: Post-quantum message authentication cryptography based on error-correcting codes. In: 2016 Third International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T), Kharkiv, 2016, pp. 51–54. (2016) doi:10.1109/INFOCOMMST.2016.7905333
 24. Kavun, S.: Conceptual fundamentals of a theory of mathematical interpretation. *Int. J. Computing Science and Mathematics*, 2015, vol. 6(2), pp. 107–121. (2015) doi:10.1504/IJCSM.2015.069459
 25. Kavun, S.: Indicative-geometric method for estimation of any business entity. *Int. J. Data Analysis Techniques and Strategies*, 2016, vol. 8(2), pp. 87–107. (2016) doi: 10.1504/IJDATS.2016.077486
 26. Zamula, A., Kavun, S.: Complex systems modeling with intelligent control elements. *Int. J. Model. Simul. Sci. Comput.*, 2017, vol. 08(01). (2017) doi:10.1142/S179396231750009X
 27. Kavun, S., Zamula, A., Mikheev, I.: Calculation of expense for local computer networks. In: Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 2017 4th International, Kharkiv, Ukraine, 2017, pp. 146–151. (2017) doi:10.1109/INFOCOMMST.2017.8246369

Intelligent System for Determining the Sufficiency of Metric Information in the Software Requirements Specifications

Tetiana Hovorushchenko ¹[0000-0002-7942-1857] and Olga Pavlova ²[0000-0003-2905-0215]

Khmelnytsky National University, Institutska str., 11, Khmelnytsky, 29016, Ukraine

¹ tat_yana@ukr.net

² olya1607pavlova@gmail.com

Abstract. The paper is devoted to the developing the intelligent system for determining the sufficiency of metric information in the software requirements specifications (SRS), which provides on the basis of the natural language processing of SRS: conclusion about the sufficiency of metric information (indicators for metrics calculation) in SRS, numerical assessment of the sufficiency level of metric information in the SRS, visualization of missing indicators for metrics calculation. The developed intelligent system provides an increase in the sufficiency of information by 12.71-50.28% for the SRS of the information and analytical system for the accounting of therapeutic and diagnostic activities provided to the wounded during transportation. In general, the developed system provides an increase in the sufficiency of metric information in the SRS to 100% – if it's necessary (for critical software) or at the customer's request. The developed intelligent system for determining the sufficiency of metric information in the SRS can be used during the software development for government agencies, military formations and law enforcement agencies, commercial organizations (both for organizations-developers of software and for organizations-customers of software).

Keywords: Software Metrics, Software Requirements Specification (SRS), Sufficiency of Metric Information, Intelligent Agent (IA), Intelligent System (IS).

1 Introduction

Given the fact that in today's world software development has become one of the most expensive industries, and any bottlenecks in the technological process of its creation can lead to unwanted results, one of the main requirements of users to modern software is its high quality and low complexity. According to ISO 25010 [1], ISO 25030 [2], ISO 19759 (SWEBOK) [3], the quality of software is the ability of software to meet the claimed and predictable customer's needs during its use under certain conditions. Software quality in ISO 9000 [4] and ISO 9001 [5] is the degree of accordance of software to the requirements. Definition of quality from standards [4, 5] doesn't take into account the fact that requirements may not fully reflect the customers' needs, then meeting the requirements isn't meeting the customers' needs,

so such software cannot be considered qualitative (in fact, there will only be formal quality satisfaction). The structural complexity of software is the number of interacting components, the number of links between components and the complexity of their interaction [6].

Today, there are a number of models that provide the calculation of the software quality and complexity, but the multiplicity of interpretation of these characteristics complicates such calculations. Most models are based on the use of different software metrics. According to ISO 24765 [7], the software metric is a measure, which gives the numeric value of a certain software feature as a weighted arithmetic mean taking into account the values of the indicators of this metric and their weights.

The using of quantitative metrics had to help in solving a number of practical tasks [8]: 1) predicting the number of bugs in the software from the beginning of the project; 2) predicting the level of software quality, software complexity and its maintenance on the basis of analysis of the results of the design stage; 3) predicting the level of complexity and quality of the testing processes and the number of unidentified bugs based on analysis of the source code; 4) predicting the final size of the source code based on the analysis of the complexity assessments of the design stage; 5) determining the impact of certain features of source code on the quality of the ready software; 6) control of project development stages; 7) analysis of explicit and hidden defects; 8) identifying the best methods and technologies of development on the basis of experimental comparison.

The modern software industry has accumulated a large number of metrics, that assess the certain production and operating features of software, but modern software is not ideal in terms of its quality [9], and in the field of evaluation and prediction of software characteristics based on the analysis of metrics remains a series of unsolved issues [8, 10]: 1) there are no common standards for metrics, which leads to a subjective choice of metrics – more than a thousand metrics has created, each developer of the "measuring" system offers its own methods and metrics for assessing the quality; 2) there is a problem of the complexity and subjectivity of the interpretation of the values of metrics – the value of metrics, which were obtained using "measuring" systems, are non-informative or little-informative for the user, for the customer, and often for the programmer; 3) tools of automating the calculation of metrics are aimed at metrology of the finished source code, and aren't aimed at prediction and calculation of software metrics at the design stage, when there is no source code, and there are only informational, functional and behavioural models of requirements analysis; 4) there is a problem of low level of automation of analysis and processing of software metrics – only the processes of collecting, registering and calculating the metric information are currently automated; 5) all known automated tools are aimed, in general, at quality assessment, but none of them is aimed at ensuring the adequate level of quality.

The complexity of the justification of the choice and interpretation of metrics during making the production decisions, and ignoring the stages of the software lifecycle don't allow the full use of metrics for evaluating and predicting the software characteristics at the early stages of the software lifecycle. At present, *the actual task* is calculating the metrics' values for estimating and predicting the quality and complexity of software at the early stages of the lifecycle.

2 State-of-the-Art

Let's consider the known methods and tools for evaluating the values of complexity and quality metrics.

The German National Research Center of Information Technology (GMD) has developed the ProcePT project [11], which includes information models of software development processes, and quality assessment methods based on the metrics, in particular, the SMV component aims at measuring software metrics and quantifying the software characteristics ; the SPM component aims to assess the quality of development processes based on quantitative indicators. These products are designed to evaluate the finished source code.

IBM corporation offers the methodology for creating complex software systems, called Cleanroom Software Engineering [12]. The Cleanroom tool for automated testing and evaluation of software reliability is the Cleanroom Certification Assistant environment, that uses statistical results of testing to calculate software reliability metrics by mathematical methods.

The Logiscope package [13] is a set of software tools (TestChecker, RuleChecker, Audit) that conduct comprehensive software testing and improve its quality. The basis of the package is the idea of analyzing the source code. The Logiscope package is designed to qualitatively evaluations of source code and identification of code's fragments, where the occurrence of bugs is most likely. After analyzing the code, Logiscope generates a set of various metric information in the form of quantitative metrics (over 200 types of metrics) about the code, its positive and negative features, generates a complete report that provides the conclusions about the quality of the code.

Pure Software company, a leading developer of automated software tools for creating high-quality software, offers the Purify system [14], which provides to identify a variety of software bugs and to calculate metrics of software complexity. The Purify system is mainly oriented on source code, rarely it can work without code, but it is not designed to work with the requirements and the software requirements specifications (SRS).

The Hindsight tool of IntegriSoft [15] analyzes the source code, measures the source code and calculates the values of software product metrics for their use in quality assessment. The cyclomatic complexity, data complexity, Halstead's metrics, complexity of software architecture are calculated.

The EzCover tool [15] measures the software and calculates the following metrics: cyclomatic complexity, modified complexity, data complexity, number of empty lines, number of commented lines, number of executable lines.

Metricate tool [15] probes practically all aspects of software companies activity: efficiency of technological processes, quality of source code, level of project management, cost of execution of different stages, productivity of the developed system, productivity of developers' work and quality of finished products.

DMS tool [16] provides a calculation of a number of metrics based on the analysis of the source code. CAST Application Intelligence Platform by CAST [17] provides detailed, audience-specific dashboards to measure quality and productivity. ConQAT [18] provides the continuous quality assessment toolkit that allows flexible

configuration of quality analyses and dashboards. GrammaTech CodeSonar [19] calculates the software metrics for C, C++, Objective-C, and Java source code. Moose [20] is the software analysis platform with many tools to manipulate, assess or visualize software; it can evolve to a more generic data analysis platform. Parasoft [21] provides static analysis (pattern-based, flow-based, in-line, metrics) for C, C++, Java, .NET. SideCI [22] is the static code analysis based automated code review tool; checks style, quality, dependencies, security and bugs. Sonargraph [23] provides the dependency analysis, automated architecture check, metrics and the ability to add custom metrics and code-checkers. SonarQube [24] provides tracks code complexity, unit test coverage and duplication. CppCheck [25] is the open-source tool for analysis of the source code and calculation of the software complexity metrics.

In the paper [26] the tool is developed, named as SWMetrics, using Microsoft Visual Studio C# to compute a metrics of LOC, SLOC and complexity based the cyclomatic complexity metric for quality measurement for many format languages of the source of code.

The paper [27] discusses how the Expert System can be used to automate the selection and implementation process of software quality assurance and provides recommendations for building an expert system for software quality assurance which can be used to help software-producing organizations in selecting the most suitable models to be adopted according to their properties and needs.

Authors of the paper [28] present a software quality support tool, a Java source code evaluator and a code profiler based on computational intelligence techniques that represent a new approach to evaluate and identify inaccurate source code usage and transitively, the software product itself. The aim of this project is to provide the software development industry with a new tool to increase software quality by extending the value of source code metrics through computational intelligence.

Authors of [29] discuss software measurement and metrics and their fundamental role in the software development life cycle, with focus on software test metrics, discuss their key role in software testing process and also classify and systematically analyze the various test metrics.

The proposed in [30] BornBaby model is a whole new dimension of Artificial Intelligence for software engineering domain: attempts are being made to come up with a software synthesis program, which is able to write programs on its own, like human programmers; defines how a program must learn in order to come up with solutions; emphasizes on how the program must learn the data based on natural concepts of living beings and the implementation is generic.

A fuzzy logic reputable paradigm is proposed in [31] for predicting software defect density on individual phases of the software development lifecycle.

The thesis [32] presents methodological investigations using search-based techniques, which are relevant to the task of software quality measurement and prediction, using search-based techniques in large-scale projects during verification, validation and testing of software.

The general disadvantages of the above tools of quality assessment are: a subjective dependence of the choice of metrics that tools of automating the calculation of metric information will calculate; the subjectivity of metric

interpretation, because the exact (standard) values of the metrics are absent; the focus of automated "measuring" tools is the testing and metrology of the finished source code, but not the prediction and calculation of software metrics at the design stage.

In [8] the neuronet method for software quality evaluation and prediction is proposed. The authors of [8] selected 24 metrics of software quality and complexity, which can be calculated already at the design stage (with exact or predicted values). The values of such metrics are analyzed by the artificial neural network, which, based on this analysis, provides assessments of the complexity and quality of the software project, and the predicted assessments of the complexity and quality of the future software. The obtained assessments are processed according to the developed production rules (which were formed on the basis of the empirically obtained threshold values), and as a result the user gets conclusions on the level of complexity and quality of the software project, and the conclusions-predictions on the level of complexity and quality of the future software. The developed method focuses not on the source code, but on the SRS, but is based on the analysis of the ready values of 24 metrics of complexity and quality (which depend on 72 indicators, including 42 different indicators) and doesn't take into account the possibility or impossibility of calculation such metrics on the basis of the information of the SRS.

Then, *the task* of assessing the sufficiency of metric information at the early stages of the software lifecycle, in particular, in the SRS (as the possibility of obtaining the indicators for calculation of metrics' values), *is actual*. Although completeness of software requirements is desirable, determination of completeness of the set of requirements is not realistic as was proved in [33]. The insufficiency of information on indicators in requirements will lead, accordingly, to the impossibility of calculating the values of certain metrics and hence to the reduction of the validity of assessments of software quality and complexity at the early stages of the lifecycle. Today the assessment of sufficiency of software safety requirements is known [33, 34].

Then, *the aim of this study* is the developments of the intelligent system for determining the sufficiency of metric information in the SRS, which, based on the natural language processing of SRS, will provide conclusions about the sufficiency of metric information in the SRS (of indicators for calculation of the chosen in [8] metrics).

3 Intelligent System for Determining the Sufficiency of Metric Information in the Software Requirements Specifications

Intelligent system (IS) for determining the sufficiency of metric information in the SRS is developed as an agent-oriented system, which consists of two intelligent agents (Fig. 1): the intelligent agent (IA) for parsing the SRS on the search of metric information (indicators for calculation of metrics) and the IA for determining the sufficiency of metric information in the SRS.

Both intelligent agents are built on the basis of the ontological approach. IAs use the early developed base ontology of the subject domain "Software Engineering" (part "Quality and complexity of software. Metric analysis") as the known knowledge.

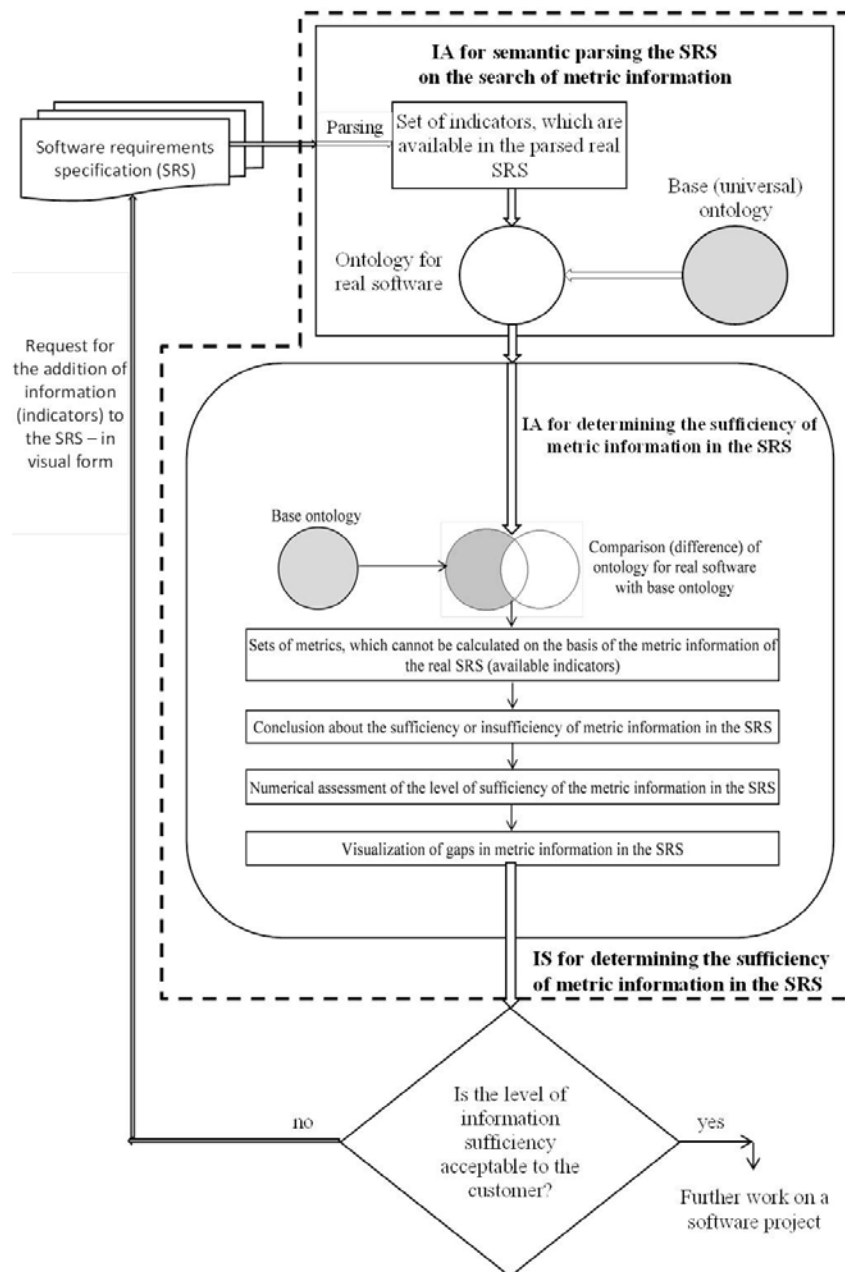


Fig. 1. Intelligent system for determining the sufficiency of metric information in the SRS.

The method of activity of the IA for semantic parsing the SRS on the search of metric information (indicators for the calculation of metrics) consists of the following steps:

1. Search of each indicator of the base ontology "Software Engineering" (part "Quality and complexity of software. Metric analysis") in the SRS for the real software (such the base ontology is contained in the agent's knowledge base).
2. If $\langle \text{indicator}_i \rangle$ is in the SRS, then $\langle \text{indicator}_i \rangle$ belongs to the set of available indicators, $i=1..42$ (since, according to [8], there are 42 different indicators, that effects on metrics of quality and complexity of software).
3. If $\langle \text{indicator}_i \rangle$ is not in the SRS, then $\langle \text{indicator}_i \rangle$ belongs to the set of missing indicators, $i=1..42$.
4. All indicators of the set of missing indicators are deleted from the base ontology "Software Engineering" (part "Quality and complexity of software. Metric analysis").
5. Checking that all indicators of the set of available indicators are in the ontology after its modification in the previous step.
6. Saving the made changes – creation of a real ontology "Software Engineering" (part "Quality and complexity of the software. Metric analysis").

In [35], we have developed *method of activity of ontology-based intelligent agent for evaluating the initial stages of the software lifecycle*, which, on the basis of comparison of the obtained real ontology with the base ontology for non-functional characteristics of the software, forms the conclusion about the sufficiency or insufficiency of the information, and numerical assessments of the sufficiency level of the SRS information for defining the non-functional characteristics. Now we use this method for development of the IA for determining the sufficiency of metric information in the SRS. But the base ontology for such IA is the ontology "Software Engineering" (part "Quality and complexity of the software. Metric analysis"). In addition, the numerical assessment of the level of sufficiency of metric information in the SRS will be calculated by the formula (1):

$$D = \frac{k - \sum_{j=1}^k \frac{qmi_j}{qni_j}}{k}, \quad (1)$$

where k – quantity of metrics ($k = 24$, because in [8] 24 metrics of quality and complexity, which are available at the early stages of lifecycle, were chosen), qmi_j – quantity of missing in the real SRS indicators for j -th metric, qni_j – quantity of necessary indicators for j -th metric (according to base ontology "Software Engineering" (part "Quality and complexity of the software. Metric analysis"))).

Thus, the developed IS for determining the sufficiency of metric information in the SRS performs the parsing (semantic analysis) of the natural language SRS with the purpose of the search of the indicators, which are necessary for calculation of the software metrics, and also forms the conclusion about the sufficiency or insufficiency of metric information in the SRS, assesses the level of sufficiency of metric information and visually shows missing indicators with the distribution by the metrics, for which they are used.

4 Experiments with Intelligent System for Determining the Sufficiency of Metric Information in the Software Requirements Specifications

Intelligent system for determining the sufficiency of metric information in the SRS has implemented in the form of free software, which is available by the link – <https://olp-project.herokuapp.com>.

The user of the IS for determining the sufficiency of metric information in the SRS upload the SRS in pdf-format. IA for parsing the SRS on the search of metric information (indicators for calculation of metrics) performs parsing of the specification and generates the ontology for real software as a .owl file, which the user can download (Fig. 2).

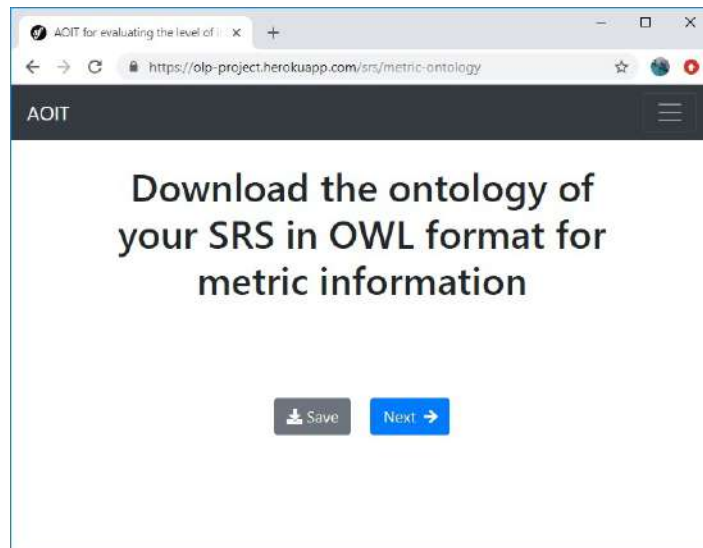


Fig. 2. Possibility of download of ontology for real SRS in OWL-format, which is provided by the intelligent system for determining the sufficiency of metric information in the SRS (results of IA for semantic parsing the SRS on the search of metric information).

Next, the IA for determining the sufficiency of metric information in the SRS performs a comparison of the ontology for real software with the base ontology and concludes about the sufficiency of metric information, namely: Count of missing indicators (without considering the number of uses of each indicator), Count of missing indicators (considering the number of uses of each indicator), Percent of missing indicators (without considering the number of uses of each indicator), Percent of missing indicators (considering the number of uses of each indicator), Total numerical evaluation of information sufficiency level for all metrics in SRS, and visualized list of missing indicators with distribution by the metrics.

For the experiments, two SRS of the information and analytical system for the accounting of therapeutic and diagnostic activities provided to the wounded during transportation were analyzed, which were developed by two different software companies at Khmelnytsky.

The level of sufficiency of metric information in the SRS1 was 49.72% when the 7 indicators without considering the number of uses of each indicator and 31 indicators considering the number of uses of each indicator were absent – Fig. 3.

The level of sufficiency of metric information in the SRS2 was 75.28% when the 14 indicators without considering the number of uses of each indicator and 20 indicators considering the number of uses of each indicator were absent – Fig. 4.

Obviously, the sufficiency of metric information in SRS2 is higher than the sufficiency of metric information in SRS1, and the number of missing indicators considering the number of uses of each indicator in SRS2 is lower than in SRS1 (while the number of missing indicators without considering the number of uses of each indicator in SRS2 is twice as high as in SRS1). These results are due to the fact that more important and priority are indicators that affect more than one metric.

On Fig. 5 the visualization of the missing indicators for determining the metrics of complexity with exact values at the design stage are represented with the distribution by metrics is represented.

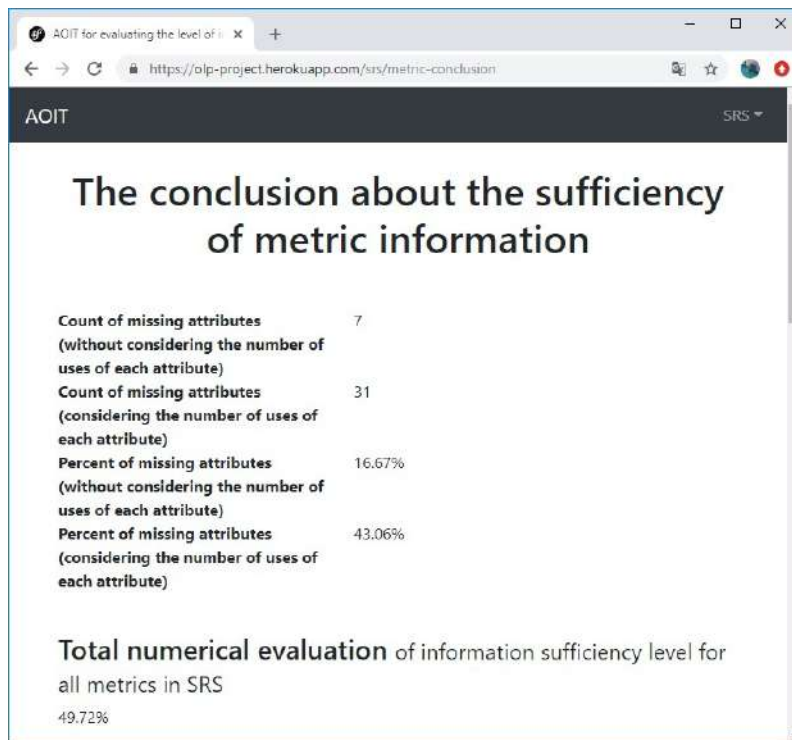


Fig. 3. Conclusion about the sufficiency of metric information in SRS1, which is provided by the intelligent system for determining the sufficiency of metric information in the SRS.

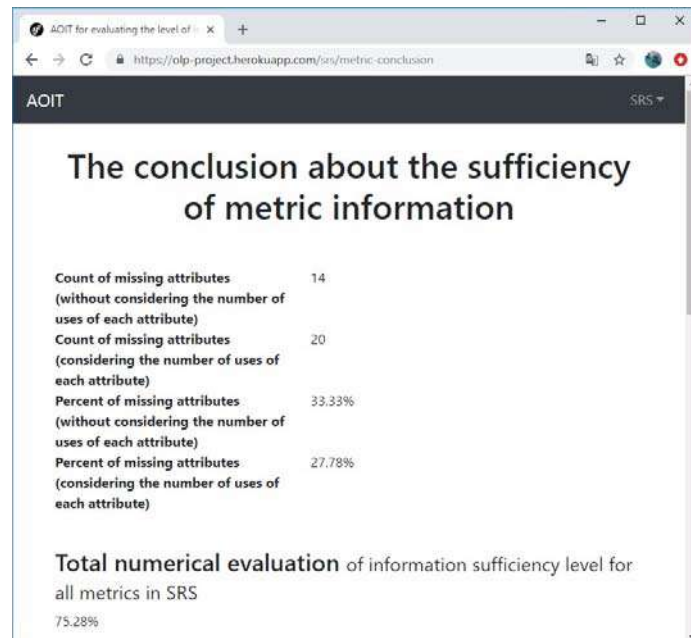


Fig. 4. Conclusion about the sufficiency of metric information in SRS2, which is provided by the intelligent system for determining the sufficiency of metric information in the SRS.

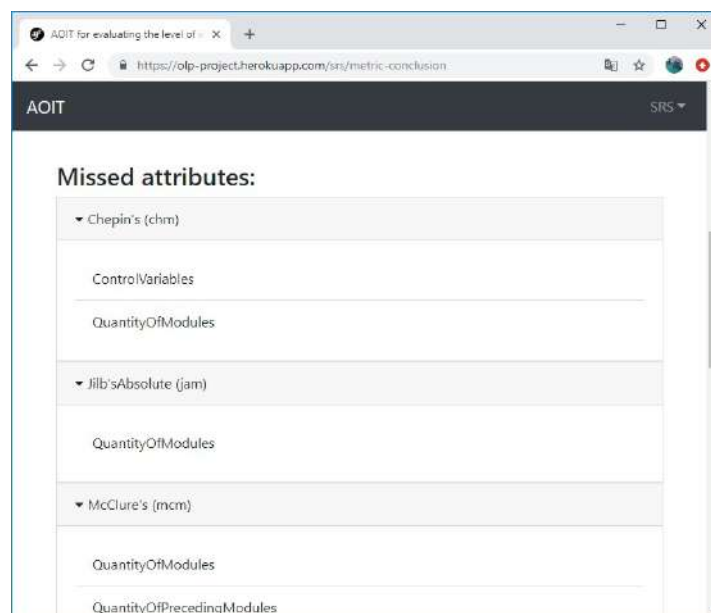


Fig. 5. Visualization of missing indicators (in SRS1) for determining the software complexity metrics with the exact values at the design stage, which is provided by the intelligent system for determining the sufficiency of metric information in the SRS.

The customer of the software of the information and analytical system for the accounting of therapeutic and diagnostic activities provided to the wounded during transportation considered the level of sufficiency equal 95% is acceptable (since medical software is the critical software). So the customer demanded the re-work of both SRS by their developers. After re-work, the SRS were again analyzed by the developed IS – Fig. 6, Fig. 7.

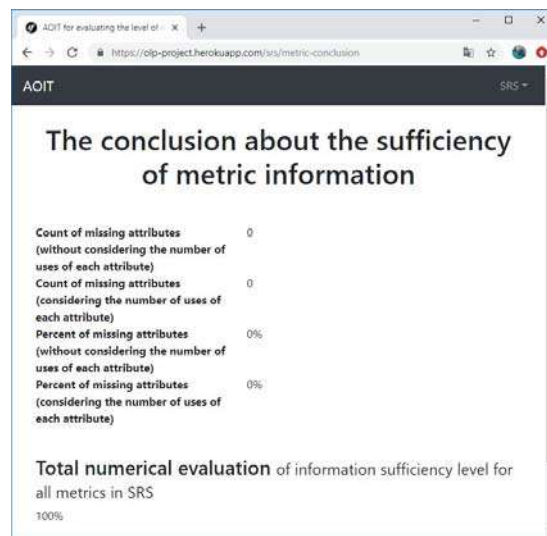


Fig. 6. Conclusion about the sufficiency of metric information in SRS1 (after re-work), which is provided by the intelligent system for determining the sufficiency of metric information in the SRS.

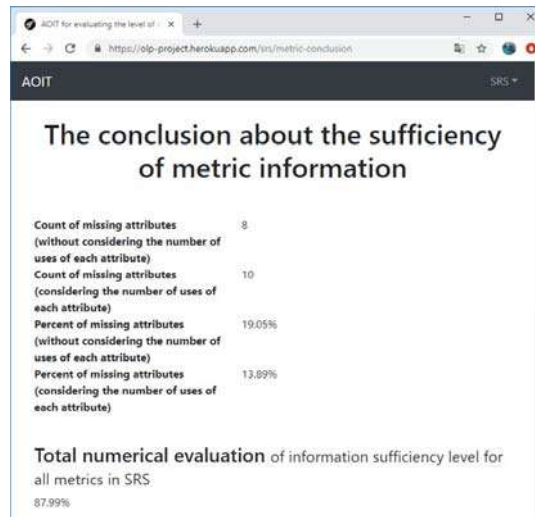


Fig. 7. Conclusion about the sufficiency of metric information in SRS2 (after re-work), which is provided by the intelligent system for determining the sufficiency of metric information in the SRS.

The level of sufficiency of the metric information in SRS1 after re-work is 100% (the SRS1 has all the necessary indicators for calculation of the metrics), and the level of sufficiency of the metric information in the SRS2 after re-work is 87.99% when 8 indicators without considering the number of uses of each indicator and 10 indicators considering the number of uses of each indicator are absent.

The customer of the software of the information and analytical system for the accounting of therapeutic and diagnostic activities provided to the wounded during transportation chose SRS1 with 100%-th level of sufficiency of metric information for further work on a software project.

The developed intelligent system for determining the sufficiency of metric information in the SRS has increased the sufficiency of information by 50.28% for SRS1 and by 12.71% for SRS2.

5 Conclusions

Nowadays the actual task is the calculation of the metrics' values for evaluating and predicting the quality and complexity of software at the early stages of the lifecycle. The conducted analysis of known methods and tools for performing metric analysis showed that most of these methods and tools are focused on the calculation of various metrics based on the analysis of the finished source code. Known methods and tools, which are aimed at early stages of lifecycle, are based on the analysis of the ready values of metrics and don't consider the possibility or impossibility of calculating such metrics on the basis of the available information in the SRS. Then, the actual task is assessing the sufficiency of metric information at the early stages of the software lifecycle, in particular, in the SRS.

In this paper the intelligent system for determining the sufficiency of metric information in the SRS is developed. This IS, based on the natural language processing of SRS, provides: the conclusion about the sufficiency of metric information in SRS, the numerical assessment of the sufficiency level of metric information in the SRS, visualization of missing indicators for metrics calculation.

The developed IS provides the increase in the sufficiency of information by 12.71-50.28% for the SRS of the software of the information and analytical system for the accounting of therapeutic and diagnostic activities provided to the wounded during transportation. Generally developed IS provides the increase in the sufficiency of metric information in the SRS to 100% – if it's necessary (for critical software) or at the request of the customer.

The developed intelligent system for determining the sufficiency of metric information in the SRS can be used in the process of software development for government agencies, military formations and law enforcement agencies, for commercial organizations (for organizations-customers of the software – with the purpose of the assessment of the level of implementation of the initial stages of the lifecycle by the developers and with the purpose of the grounded choice of the SRS with the highest level of sufficiency of information). For organizations-developers,

the developed IS can also be used – with the purpose of automation of the process of verifying the sufficiency of information in the SRS.

The limitation of the developed IS is the possibility of assessment of the sufficiency of only metric information in the SRS – as the sufficiency of the indicators in the SRS for determining the 24 selected software metrics, which are available at the design stage. The further efforts of the authors will be directed for elimination of this limitation (in particular, for expanding the set of metrics and, respectively, indicators).

References

1. ISO/IEC 25010:2011. Systems and software engineering. Systems and software Quality Requirements and Evaluation (SQuaRE). System and software quality models (2011).
2. ISO/IEC 25030:2007. Software engineering. Software product Quality Requirements and Evaluation (SQuaRE). Quality requirements (2007).
3. ISO/IEC TR 19759:2015. Software Engineering. Guide to the software engineering body of knowledge (SWEBOK) (2015).
4. ISO 9000:2015. Quality management systems. Fundamentals and vocabulary (2015).
5. ISO 9001:2015. Quality management systems. Requirements (2015).
6. Maedche, A., Botzenhardt, A., Neer, L.: Software for people: fundamentals, trends and best practices (Management for professionals). Springer-Verlag Berlin Heidelberg, Berlin (2012).
7. ISO/IEC/IEEE 24765:2010. Systems and software engineering. Vocabulary (2010).
8. Pomorova, O., Hovorushchenko, T.: Research of Artificial Neural Network's Component of Software Quality Evaluation and Prediction Method. In: The 2011 IEEE 6-th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications Proceedings. Prague (2011).
9. Pomorova, O., Hovorushchenko, T.: The Way to Detection of Software Emergent Properties. In: The 2015 IEEE 8-th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications Proceedings. Warsaw (2015).
10. Pincioli, F.: Improving Software Applications Quality by Considering the Contribution Relationship Among Quality Attributes. *Procedia Computer Science*. 83, 970-975 (2016).
11. Graphical Development Process Assistant: ProcePT, <http://www.scope.gmd.de/projects/ProcePT/>, last accessed 2019/03/19.
12. Cleanroom Software Engineering, <https://www.tutorialride.com/software-engineering/cleanroom-software-engineering.htm>, last accessed 2019/03/19.
13. Automatic Code Analysis with Logiscope Products, <https://www.kalimetrix.com/logiscope>, last accessed 2019/03/19.
14. PurifyPLUS > Dynamic Software Analysis, <https://www.almttoolbox.com/purify.php>, last accessed 2019/03/19.
15. Morse, J.: Expressive and efficient bounded modelchecking of concurrent software. University of Southampton, Southampton (2015).
16. DMS® Software Reengineering Toolkit™, <https://www.semanticdesigns.com/Products/DMS/DMSToolkit.html>, last accessed 2019/03/19.

17. Application Intelligence Platform, <https://www.castsoftware.com/products/application-intelligence-platform>, last accessed 2019/03/19.
18. ConQAT end of life, <https://www.cqse.eu/en/blog/conqat-end-of-life/>, last accessed 2019/03/19.
19. GrammaTech CodeSonar: Delivering resiliency for today's IoT devices, <https://www.grammatech.com/products/codesonar>, last accessed 2019/03/19.
20. Moose is a platform for software and data analysis, <http://www.moosetechnology.org/>, last accessed 2019/03/19.
21. Automated Software Testing Tools for Creating High Quality Software, <https://www.parasoft.com/>, last accessed 2019/03/19.
22. SideCI: Automated Code Review for GitHub, <https://alternativeto.net/software/sideci/>, last accessed 2019/03/19.
23. Sonargraph Product Family, <https://www.hello2morrow.com/products/sonargraph>, last accessed 2019/03/19.
24. SonarQube: The leading product for continuous code quality, <https://www.sonarqube.org/>, last accessed 2019/03/19.
25. CppCheck: A tool for static C/C++ code analysis, <http://cppcheck.sourceforge.net/>, last accessed 2019/03/19.
26. Zakariya, S., Belal, M.: Software quality management measured based code assessments. *International Journal of Computer Science Trends and Technology*. 3, 4, 263-268 (2015).
27. Shouman, M., Eldrandaly Kh., Tantawy, A.: Software quality assurance models and expert systems. In: *The 9-th International Conference on Production Engineering, Design and Control Proceedings*. Alexandria (2009).
28. Agüero, M., Madou, F., Esperón, G., López De Luise, D.: Artificial intelligence for software quality improvement. *International Journal of Computer and Information Engineering*. 4, 3, 399-404 (2010).
29. Farooq, S. U., Quadri, S., Ahmad, N.: Software measurements and metrics: role in effective software testing. *International Journal of Engineering Science and Technology*. 3, 1, 671-680 (2011).
30. Gururaj, H. L., Ramesh, B. BornBaby model for software synthesis: A program that can write programs. *Data Analytics and Learning: Lecture Notes in Networks and Systems*. 43, 403-412 (2019).
31. Ravi Kumar, T., Srinivasa Rao, T.: Software Defects Prediction based on ANN and Fuzzy logic using Software Metrics. *International Journal of Applied Engineering Research*. 12, 19, 8509-8517 (2017).
32. Azfal, W.: Search-based prediction of software quality: evaluations and comparisons. *Blekinge Institute of Technology, Blekinge* (2011).
33. Cruickshank, K. J.: A validation metrics framework for safety-critical software-intensive systems. *Naval Postgraduate School, Monterey* (2009).
34. Michael, J. B., Shing, M.-T., Cruickshank, K. J., Redmond, P. J.: Hazard Analysis and Validation Metrics Framework for System of Systems Software Safety. *IEEE Systems Journal*. 4, 2, 186-197 (2010).
35. Hovorushchenko, T., Pavlova, O.: Method of Activity of Ontology-Based Intelligent Agent for Evaluating the Initial Stages of the Software Lifecycle. *Advances in Intelligent Systems and Computing*. 836, 169-178 (2019).

Implementation of Audio Navigation for Smart Campus

O. Petrova^{1[0000-0002-6499-6017]}, G. Tabunshchyk^{2[0000-0003-1429-5180]},
P. Arras^{3[0000-0002-9625-9054]}

^{1,2}₁Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063,
Ukraine

petrovaoa353@gmail.com, galina.tabunshchik@gmail.com

³KU Leuven, Jan De Nayerlaan 5, 2860 Sint Katelijne Waver, Belgium
peter.arras@kuleuven.be

Abstract. The article deals with the task of in-door navigation of visually impaired people. The authors have carried out an analysis of audio navigation software programs such as Google Assistant, Siri and Cortana. Authors suggest a model of the voice navigator, which helps a person to conveniently find the location and build the desired route. Developed software is integrated into the smart-campus solution, which improve the infrastructure of the university.

Keywords: audio navigation, SMART-CAMPUS, BLE, voice navigator, in-door-positioning

1 Introduction

According to statistics nowadays in the European Union, people with disabilities make out about 1/6 of all citizens of working age. In Ukraine, the amount of persons with disabilities is 6.1% of the total population [1]. A person with disabilities faces many problems that are unknown to other people. Mainly this is caused by the restriction of the access of persons with disabilities to social benefits known to the majority of the population, such as shops, pharmacies, underground stations, stations, hair-dressers, educational establishments, et cetera [2, 3]. This is due to the fact that in such places there are no special devices for assistance to people with disabilities.

Ukraine is trying to adapt public buildings to this reality by constructing ramps and buttons for the disabled people. At the legislative level, the Government makes changes to the laws that regulate the rights of persons with disabilities, namely, the Laws "On the Basis of Social Protection of Persons with Disabilities in Ukraine" [4], "On Amendments to Some Laws of Ukraine on Increasing Access to the Blind, persons with visual impairments and persons with dyslexia to works published in a special format "[5], " On Amending Certain Legislative Acts of Ukraine on the Protection of the Rights of Persons with Disabilities "[6], " On Amendments to Certain Laws of Ukraine on Education on the organization of inclusive n the voice "[7]. However, there should pass a long period before our country achieves the result that we can see in Europe today. Therefore, the development of audio navigation systems to improve social adaptation of people with visual disabilities is a very important task.

The idea of smart campus based on BLE 4.0 where objects could talk to the students, staff and visitors were described in an number of publications [8,9]. The use of voice for navigation systems could allow visually impaired people: to connect many objects and events; to provide access to the information in the navigation systems; to support new systems of interaction with users, sensors, mobile devices, devices and applications [10].

2 Problem definition

For correctly detection of the location inside the building, it is necessary to determine the current coordinates, compare the position with the cartographic representation, update the location in real time, and check the compliance of the current position with respect to the planned route [11].

Further we will consider a system that uses data from beacons based on BLE 4.0 to identify the current location [112]:

$$S = \langle X, B, R, Z, K \rangle \quad (1)$$

where X - input data (x_1 - data from sensors x_2 - accelerometer readings, x_3 - gyro readings, x_4 data from beacons, x_5 - voice commands), B - cartographic representation (map represented as matrix $[M, N]$, where M - X , N is the number of points along the Y -axis), R is information about the decisions taken ($r_1, r_2 \dots r_n$), Z -output devices (z_1 - camera, z_2 - audio recording, z_3 -phone), K - robot mode (k_1 - autonomous, k_2 controlled).

Positioning methods were developed for this class of systems [13]. However, the task of integrating audio navigation in Smart-Campus systems has not been solved yet. Solving this problem will allow the existing system to be adapted for people with disabilities.

The aim of the work was to develop a voice navigator and integrate it into the indoor positioning and navigation system.

3 An analysis of existing approaches to the implementation of voice navigation

Voice Navigator should help a person to navigate in the building using only voice. However, for the correct information exchange between the user and the application, there should be developed a module which could recognize speech signals.

Problems of voice navigation and speech recognition were investigated by D. Shpakov [14], E. A. Vereshchagina [15], Jen-Tzung Chien [16], Shinji Watanabe [17], Mohamed Afify [18], Chia-Yu [19], Mark D. Skowronski [20].

Automated speech recognition systems can be classified according to many features: by type of language, by a set of dictators, by volume and completeness of the vocabulary that needs to be recognized. By type, the language is divided into discrete and continuous [21]. Discrete language is a language in which pauses between words

are much longer than natural pauses inside words. In continuous speech, there are no significant pauses between words. The natural human mode of communication is a continuous language.

Each person has a unique voice, but from a phonetic point of view language consists of many different sounds that have articulation differences. In general, these sounds are called phonemes. But in different words one and the same phonemes may be exaggerated, so there is the notion of allophone - phonemes [22].

For successful speech recognition, areas of the audio signal are considered in a few tens of milliseconds, which are called frames [22]. The difficulty is that some phonemes are quite similar to one another, but one can solve this problem in terms of "probabilities". Some phonemes are more likely for a given signal, others - less. An acoustic-on-model is being built, which is a function that receives an area of a small audio signal (frame) at the input and outputs the distribution of the probabilities of different phonemes on this frame. On the basis of the acoustic model, one can say with a certain degree of confidence that it was said.

The acoustic model can be built on the basis of such methods and algorithms as neural networks, a model of Gaussian mixtures, dynamic programming [23-26]. In practice, hidden mark models are widely used in practice [27].

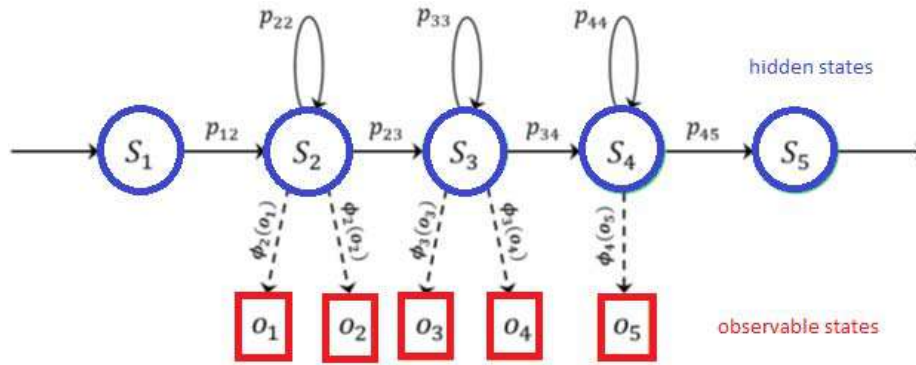


Fig. 1. - Acoustic model [13]

In the system discussed previously, incoming data can come in the form of voice messages. In this case, the task of recognizing audio events will look like this: an audio signal arrives at the audio event detector input, represented by the sequence:

$$\Omega = \{o_1, o_2, \dots, o_M\} \quad (2)$$

where: oi - the value of the sound signal parameter (one of M) taken by the detector at

the i^{th} moment of time. The segments of time in which the detector takes off these

parameters are states $S = \{s1, s2, ..., sN\}$ of the model $\lambda = (P, \Phi, \pi)$. Each of these

models corresponds to different types of audio events, such as certain words. In order for the system to be able to select the audio event that corresponds most to the initial segment of the audio signal (in other words, to recognize the word), it is necessary to

find the fidelity of the appearance of the sequence $\Omega = \{o1, o2, ..., oM\}$ for each

available models $\lambda = (P, \Phi, \pi)$. In this way, there is a set of observed states (speech signal) and a probabilistic model that conveys a hidden state (phonemes) and observable quantities.

Thus, the processing of a voice message occurs in a few steps:

Step 1. The input of the system for identifying the current location S is the input data X . One of the input parameters is a voice message $x5$.

Step 2. A voice message Ω arrives at the audio event detector, which starts with one of the keywords: start navigation, build route, cancel, stop the starting position, destination.

Step 3. The resulting sequence falls into the audio processing block where we get the λ model.

Step 4. This step defines a specific audio event in a probabilistic way. That is, the record is divided into frames and each frame is skipped through the acoustic model. System with machine learning, defines variants of spoken words and context. The accuracy of the results depends on the completeness of the phonetic alphabet of the system. For each sound, a complex statistical model is first constructed that describes the pronunciation of this sound in the language. The system of recognition compares the incoming speech signal with phonemes, and from them they collect words.

Step 5. In this step, the data fall into the next level of the system as a text for decision. The main teams will be: In what building am I? What floor? I need room №? What room do I have a couple of? How do I get to the room №?

Step 6. After receiving the request, the commands will be mapped to the source data, which include: schedule, group lists, placement and maps of the building and each floor, the list of audiences.

Step 7. Next, using the integrated method will determine the current position on the map of the room.

Step 8. In this step, the data is verified using a neuro-fuzzy method of verification [12].

Step 9. After processing the system, we receive z_2 messages and the route is being built [28].

4 Realization of the subsystem of voice indoor navigation

Within the Smart-Campus application, the ability to display the current position of the user inside the building and the search for the shortest path to the specified beacon [9] was implemented. The next step is to modify the Smart-Campus subsystem of voice navigation.

The Smart-Campus, is a system with Bluetooth Low Energy devices and a back-end database with dedicated content management system (CMS). The idea is to find the location from one beacon to the others, for an interactive tour around the campus or to guide visitors to their specific location of interest. To provide navigation, first a map of the building should be provided or developed. Next is showing the appropriate path to another beacon location. This is why the newly developed solution consists of two parts: a map editor and path detection.

The map editor allows creating a map of a floor. You can use a background picture of a known area or develop it from scratch with the easy-to-use editor. The app-user is the client of information related to a certain beacon at a certain location and our solution allow user to get this information in an attractive way on his or her smart phone through a dedicated application. The app itself fetches the information from the server, related to the unique user ID (UUID) the beacon broadcasts on regular basis. On this server the information is added and edited by the beacon owners through the developed CMS. The users can decide on groups of beacons which are allowed to display their information [8].

The voice navigator will help a person find the location of the audience and the body in which it is located. After the audience is found, the navigator will answer the question about the building in which the audience is located, on which floor and construct the map-device from the current position of the user to the required body. Also, the mobile add-on will provide the user with the opportunity to create and manage their class schedules. The timetable will be displayed for the week and the current day. From the schedule, the user will be able to build a route to the required building.

Let us consider software which contains similar functionality: Google Assistant, Siri, Cortana. For analysis following characteristics were selected: dependence on the Internet, speed of operation of the recognizer, understanding the request, number of satisfactory answers to questions, construction of the route, vocabulary, number of supported languages,. The summary is presented in the table 1.

After analyzing the applications, the main characteristics that should have a voice navigator have been highlighted. The voice navigator, for integration into the Smart-Campus must have the following features:

- to record a voice sentence to get an audience that the user is looking for;
- to recognize vocal sentences and convert them to text;

- to formulate a response to the user;
- to issue a voice message about the user's request;
- to determine the location of the user;
- to build a route from the current position of the user to the required body;
- to display the schedule of occupations of the user;
- to add classes to the schedule;
- to enter the name of the class not only through the virtual keyboard but also through speech recognition;
- to edit or delete selected classes from the schedule;
- to get the route to the chosen lesson;
- to display the schedule for the current day;
- to display the list of recent queries.

The interaction diagram for audio navigation is shown at fig.2.

Table 1. Mobile applications comparison

Parameters	Google Assistant	Siri	Cortana	Voice Navigation
Dependence on the Internet	+	+	+	+-
Speed of operation of the recognizer	+	++	+	+
Understanding the request, %	97	92	86	90
Number of satisfactory answers to questions, %	85,5	78,5	52,4	74,3
Construction of the route	+-	++	+	++
Vocabulary	+	++	+	++
Number of supported languages,	9	34	8	8

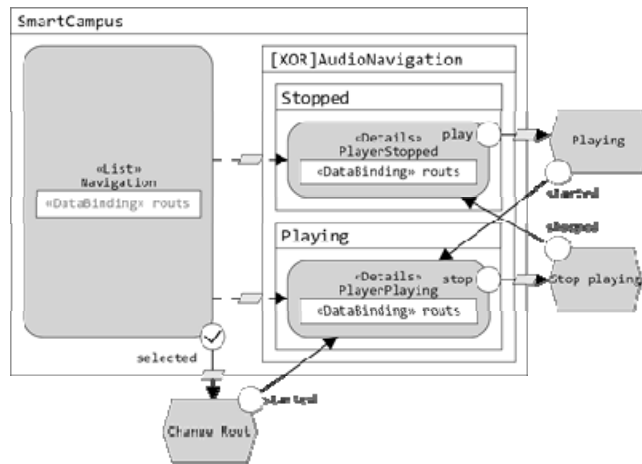


Fig. 2. - Interactions in the voice navigator

For development of the speech recognition, the frame Speech.framework was selected.

First the application is trained with commands which are stored at the local databases.

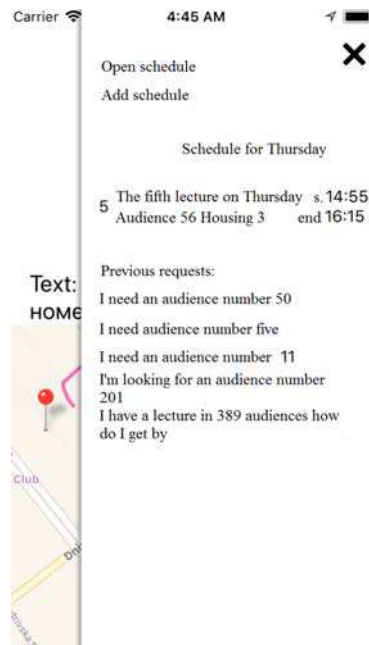


Fig. 3. – Menu

With each beacon there is connected the voice identification of the location. After the final location is recognised the path is built according to the shortest path algorithm [11]. One of the options is that the user can see the previous voice requests (fig.3).

5 Conclusion

For the in-door navigation system was designed a voice navigator. Integrating the audio navigator into the Smart-Campus system is improving the social adaptation of visually impaired people. Usage of the voice for navigation systems allow user to provide access to information in navigation systems; to connect many objects and events among themselves; to support new user interaction systems, sensors, mobile devices, devices and applications

6 Acknowledgment

The work was partly done within the framework of Erasmus+ [BIOART] Also, the work was carried out within the framework of the agreement on scientific and technical cooperation Agency # 417/156 / 1.4917 dated May 4, 2017 between ZNTU and Limited Liability Company Infocom LTD.

7 References

1. Gnibidenko I.F., Kravchenko M.V., Koval O.M., Novikova O.F.: Social defencing of the population of Ukraine: higher possibilities, per community. K.: View at NAPA; View of "Phoenix". p. 212 (2010)
2. Arras, P., Van Merode, D., Tabunshchyk, G.: Project Oriented Teaching Approaches for E-learning Environment. IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS). pp.317-320 (2017) [https://doi: 10.1109/idaacs.2017.8095097](https://doi.org/10.1109/idaacs.2017.8095097)
3. Tabunshchyk, G, Parkhomenko, A, Morshchavka, S, Luengo, D.: Engineering Education for HealthCare Purposes: A Ukrainian Perspective. The XIV-th International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), Lviv, Poliana, 18-21 April, pp 245 - 249 DOI: [10.1109/MEMSTECH.2018.8365743](https://doi.org/10.1109/MEMSTECH.2018.8365743)
4. On the basis of social protection of persons with disabilities in Ukraine: Law of Ukraine 19.12.2017 2249-VIII . (2017)
5. On Amendments to Some Laws of Ukraine on the Expansion of Access to the Blind, Visually Impaired, and Dyslexic Individuals for Works Published in a Special Format: Law of Ukraine 25.12 2015 № 927-VIII . (2015)

6. On amendments to certain legislative acts of Ukraine concerning the protection of the rights of persons with invalidity: Law of Ukraine 18.06 2014 № 1519-VII . (2014)
7. On amendments to some laws of Ukraine on education regarding the organization of inclusive education: Law of Ukraine 5.06.2014 № 1324-VII . (2014)
8. Tabunshchyyk, G., Van Merode, D.: Intellectual Flexible Platform for Smart Beacons. In book: Edit by M. Auer, D. Zhutin Online Engineering and Internet of Things, Springer International Publishing, pp. 895-900 (2017) https://doi.org/10.1007/978-3-319-64352-6_83
9. Tabunshchyyk, G., Van Merode, D., Goncharov, Y., Patrakhalko, K.: Smart-campus infrastructure development based on BLE4.0. J. Electrotechn. Comput. Syst. 18(94), 17–20 (2015)
10. Speech Recognition. Available at: <http://buchuk.domen.ua/index.php?id=realspeaker>
11. Petrova, O., Tabunshchyyk, G.: Modelling of location detection for indoor navigation systems. IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), pp. 961-964. (2017) <https://doi.org/10.1109/IDAACS.2017.8095229>
12. Petrova, O., Tabunshchyyk, G., Van Merode, D.: Method for determining the current location in positioning systems and indoor navigation. Electrotechnical and Computer Systems, № 25, pp. 270-278. (2017)
13. Petrova, O., Tabunshchyyk, G., Kaplienko, T., Kaplienko, O.: Fuzzy Verification Method for Indoor-Navigation Systems. In: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018 – Proceedings, Slavske, 20–24 February 2018, pp. 65 – 68 (2018) <https://doi.org/10.1109/TCSET.2018.8336157>
14. Shpakov, D.V.: Voice Recognition in the Sphere of Information Technologies. Young Scientist, №29, pp. 8-11. (2017)
15. Kolesnikova, D. S, Rudnichenko, A. K, Vereshchagina, E.A, Fominova, E.R.: The application of modern speech recognition technologies in the creation of a linguistic simulator to enhance the level of linguistic competence in the field of intercultural communication. Internet journal "Naukovedenie", vol. 9, No. 6. (2017)
16. Chien, J.-T.: Linear Regression Based Bayesian Predictive Classification for Speech Recognition. IEEE Transactions on Speech and Audio Processing, vol. 11, no. 1 January (2003)
17. Watanabe, Sh.: Variational Bayesian Estimation and Clustering for Speech Recognition. IEEE Transactions on Speech and Audio Processing, vol. 12, no. 4 (2004)
18. Afify, M., Liu, F., Jiang, H.: A New Verification-Based Fast-Match for Large Vocabulary Continuous Speech Recognition. IEEE Transactions on Speech and Audio Processing, vol. 13, no. 4 (2005)
19. Chia-Yu: Histogram-based quantization for Roboust and / or Distributed speech recognition. IEEE Transactions on Audio, Speech And Language Processing, Vol.16, Jan. 1, 2008. (2008)
20. Skowronski, M.D.: Noise Robust Automatic Speech Recognition using a predictive Echo state Network. IEEE transactions on Audio, Speech and Language processing, Vol.15, No.5, June 2007. (2007)
21. Alborova, Zh.V., Rubtsov V.I.: Algorithm and Methods of Speech Recognition. youth scientific and technical weight №FS77 51038, (2016)
22. Rabiner, L. R., Schafer, R. V.: Digital processing of speech signals. Radio and communication, . 496 p. (1981)
23. Subbotin, S.A.: Opt. Mem. Neural Networks 19: 126. (2010) <https://doi.org/10.3103/S1060992X10020037>

24. Oliinyk, A., Skrupsky, S., Subbotin, S.A.: Parallel Computer System Resource Planning for Synthesis of Neuro-Fuzzy Networks. In: Szewczyk R., Kaliczyńska M. (eds) Recent Advances in Systems, Control and Information Technology. SCIT 2016. Advances in Intelligent Systems and Computing, vol 543. Springer, Cham (2017) https://doi.org/10.1007/978-3-319-48923-0_12
25. Rabcan, J., Rusnak, P., Subbotin, S.: Classification by fuzzy decision trees inducted based on Cumulative Mutual Information. In: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018 - Proceedings, Slavske, 20-24 February 2018, pp. 208-212 (2018)
26. Leoshchenko, S., Oliinyk, A., Subbotin, S., Zaiko, T.: Using Modern Architectures of Recurrent Neural Networks for Technical Diagnosis of Complex Systems. International Scientific-Practical Conference on Problems of Infocommunications Science and Technology, PIC S and T 2018 – Proceedings (2018)
27. Hidden Markov Models, available at: http://www.machinelearning.ru/wiki/images/8/83/GM12_3.pdf

Intellectual Support of Control System Human-Machine Interface Designers

Zhanna Kaminska^[0000-0003-4029-1538]

Dept. of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia 69063,
Ukraine
kamzhana@gmail.com

Abstract. The problem of intellectual support of human-machine interface (HMI) designers for Process Control Systems is considered. A method for solving this problem in the form of a set of models, technologies and knowledge, forming a new information technology for Control System HMI design is proposed. Software in the form of an Expert System that implements the proposed technology is developed. The use of an Expert System in the HMI design process will improve the efficiency, quality, reliability and safety of Control System operator's activity.

Keywords: expert system, SCADA system, man-machine interface, mnemonic scheme, ergonomic design, intellectual support

1 Introduction

The introduction of new information technologies in production processes, improvement of automation equipment leads to the fact that the human-operator becomes the "narrowest" link in modern Control Systems. It was shown [1] through operational experience of modern Control Systems that human mistakes or erroneous behavior cause 42% abnormal situations in industrial systems used. Accidents of 45% for nuclear plants, 60% - aviation crashes, 80% - sea disasters and 90% - car accidents [2, 3] happened due to human psychophysiological capabilities limitations and faults. In fact, Cochran and Bullemer [4] have estimated that abnormal situations cost the petrochemical industry \$10B (US) annually in preventable losses. The Chemical Manufacturers Association estimates that 80 to 85% of human errors in chemical production result from poor design of the working environment [5]. European Agency for Health and Safety at Work, besides poor design of HMI, notes [6] also such reasons as multi-factorial risks (combined effects of poor ergonomic design, poor work organisation, mental and emotional demands) and complexity of new technologies, new work processes and HMI leading to increased mental and emotional strain. This demonstrates the relevance and importance of research in the domain of ergonomic design of Control Systems operator's HMI.

For today Automatic Control System development is mostly powered by Supervisory Control And Data Acquisition (SCADA) technology with wide variety of

special component libraries as well as graphic editors used. However, they include no modules or add-ons to provide intellectual support for HMI ergonomic design. Hereinafter HMI determines Control System HMI developed by SCADA technology. The main purpose of our research is intellectual support of Control System HMI designers.

2 Literature review

The use of intelligent systems as a component of SCADA systems or in conjunction with them is known [e.g. 7–9]. Such integration provides intelligent support of operator in the analyzing data process, diagnosing the state of technological equipment, forecasting the development of situations and making appropriate decisions in normal and emergency situations.

Operator is often remote from objects of observation and control (OOC) and performs remote control using the information model (IM) in modern Control Systems. There are many types of IM. A special case of IM is the HMI of the operator developed by SCADA technology.

There are a lot of techniques for ergonomically driven the IM design process for Control System operators [e.g. 10]. Their disadvantage is the insufficient formalization of the design stages. The analysis [11], as well as further research, showed the following:

- software is being developed for displaying operational dispatch information on the Internet, as well as archived data [12]. A language for describing mnemonic diagrams was developed based on the XML standard. The mnemonic description file contains information about the appearance of the mnemonic diagram and how to react to various events or interactive actions of the user. The disadvantage of this approach is the rejection of the presentation of mnemonic diagrams in the SVG format. As a result, the possibility of drawing mnemonic diagrams through graphic editors developed by such companies as COREL DRAW, VISIO etc. is lost;
- basic requirements for the organization of HMI for process engineers of NPPs with VVER-1000 reactors and a set of HMI elements that is sufficient for organizing the control and diagnostics of a power unit is defined [13];
- software development of the system for improving the quality of interfaces of computer simulation systems is under development [14]. A new method has been developed to automate the HMI quality assessment process. The method is based on analyzing the factors that most affect the quality of HMI computer-based training systems on mnemonic diagrams;
- currently there are domestic standards that regulate the requirements for ergonomic design of such types of IM as a mnemonic diagram [15, 16]. The closest to the problem under consideration is the international standard ANSI/ISA-101.01-2015 [17]. This standard addresses HMI for equipment and automated processes. If the standard, recommended practices and methodology are followed, the result should enable the users to be more effective yielding improved safety, quality, productivity and reliability. The standards ISO-9241, ISO-10075, ISO-13407, ISO-14915 provides

recommendations for ergonomic human-system interaction design, operator's mental workload, and human-centered design of the HMI, respectively.

There are also many HMI design guides for designing of task-based interface (Ecological Interface) (e.g. [18, 19]), human-centred design (e.g. [20]) and HMI for Process Control Applications (e.g. [21]), but they give only general guidelines for interfaces designing.

On the basis of the foregoing, we can conclude that there are practically no work in the Control System HMI design automation field. It should also be noted that the design and configuration tasks of the HMI can significantly increase the cost and complexity of the Control System design project. According to [22], designers spend \$10,000 or more to develop each page of HMI graphics for their process operations. Therefore, the actual task is to develop a method of intellectual support for the design of the Control System HMI.

3 Information technology of Control System HMI design

The base of the information technology for designing the Control System HMI were suggested by the author in [23] and require further development as part of this work.

The object-oriented approach, which simplifies the implementation of projects in SCADA-tools, is applied for modeling the subject domain "HMI design". The conceptual model of the subject domain "HMI design", made in the UML notation, is presented in Fig. 1.

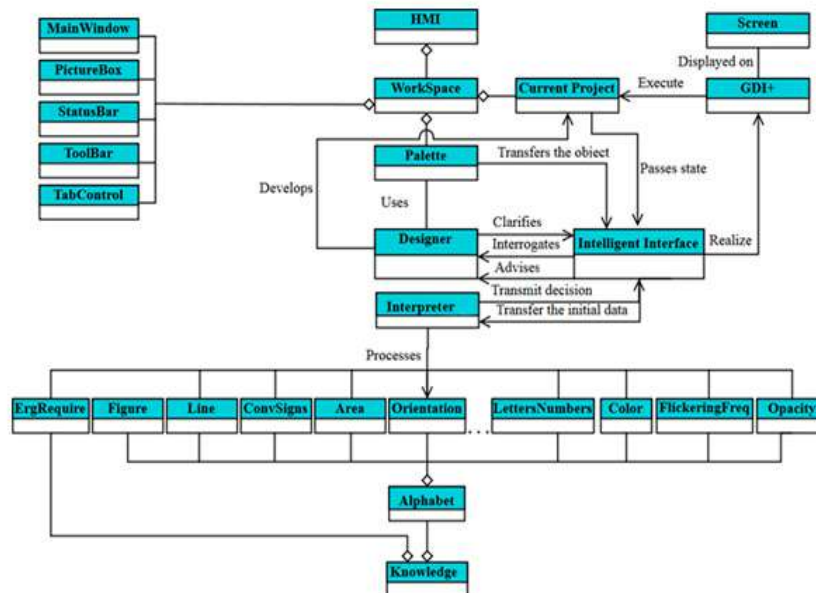


Fig. 1. The conceptual model of the subject domain "HMI design"

The internal hierarchical structure of the subject domain reveals the aggregate association. The nature of the other relations between classes in the model is set using the named directional associations.

According to Fig. 1, the designer (the class "Designer"), using a set of standard interface elements ("Palette"), is developing the project HMI ("Current Project") step by step. The current HMI project is carried out in the workspace ("WorkSpace"), which includes a set of tools, templates, and methods and means of using them within the project. The final state of the current project (the product of the designer's labor) is the Control System HMI ("HMI").

The designer fills the workspace with objects that are implemented using various program classes. The class "WorkSpace" is a container for the following classes: MainWindow (the main window of the HMI development tool); PictureBox (represents a Windows picture box control for displaying an image); StatusBar (represents a Windows status bar control); ToolBar (a toolbar that is a container for a group of commands or controls); TabControl (represents a control that contains multiple items that share the same space on the screen).

Intellectual support of the designer is carried out using certain Expert System (ES). A typical ES consists of the following main components: a inference engine (interpreter); working memory (data base) intended for storage of initial and intermediate data of a current problem; knowledge base; knowledge acquisition component; explanatory component; dialog component. To display them in the conceptual model (Fig. 1), the following classes are introduced:

- "Interpreter" for inference engine;
- "Knowledge" for knowledge and data base (KB);
- "Intelligent Interface" (subsystem "Intelligent Interface") for the remaining components.

As a result of the interaction of the designer with the subsystem "Intelligent Interface" is his intellectual support in the form of recommendations generated by ES.

ES provides intellectual support to the designer by generating advice during the design process. The process of advice inference is as follows. The input of the subsystem "Intelligent Interface" receives information: about the current state of the project (from the "Current Project"); about the typical interface element chosen by the designer (from the "Palette"), as well as the results of the dialogue with the designer (from the "Designer").

This subsystem transfers to the "Interpreter" the results of processing input information in the form of some initial judgments (prerequisites) that describe the current problem design task TS_i . The "Interpreter", using the initial data and knowledge from the KB of expert system ("Knowledge"), forms a sequence of rules (advice to the designer), which lead to the solution of the problem design task TS_i .

KB of N -th Man-Machine System (MMS) is represented as the tuple [24]:

$$KB(N) = \langle KBC(N), KBO(N) \rangle, \quad (1)$$

where $KBC(N)$ – KB about N -th MMS components; $KBO(N)$ – KB about N -th MMS organization, i.e. information about the relations of the MMS components.

The knowledge base of ES consists of a set of local KB. Knowledge about components $KBC(ES)$ and organizations $KBO(ES)$ contained in the KB of Expert System is discussed further. In Fig. 1 presents two local KB "ErgRecuire" (denote it $KBEV$) and "Alphabet" ($KBAL$) such as $KBEV, KBAL \in KBO(ES)$.

The first KB "ErgRecuire" contains the following knowledge of the ergonomic support for the HMI designer: psycho-physiological characteristics of the operator; rules for developing IM and HMI; code arrangement rules; rules and recommendations for the design of instrument scales and graphs.

The second KB, "Alphabet", represents knowledge about the coding methods (alphabets) of information about OOC of the technological process (TP). Qualitative and quantitative characteristics of OOC are coding in different ways (alphabets, codes). 17 coding methods are known [25].

Fig. 1 presents the following types of alphabet: "Figure" - coding with abstract geometric figures; "Line" - type of line; "ConvSigns" - conventional symbols; "Area" - area of geometric figures; "Orientation" - orientation of the figure or line in space; "LettersNumbers" - letters, punctuation marks and numbers; "Color" - color; "FlickeringFreq" - flicker frequency, "Opacity" - opacity.

Subsystem "Intelligent Interface" received the results of the logical inference from the "Interpreter" and in accordance with them modifies the current HMI project. Thus, we get a solution to the current problem task TS_i .

As it is known, GDI+ ("GDI+" in Fig. 1) is the Windows operating system graphic device interface, which enables functions to transfer graphic objects to a monitor. Access to GDI+ functions is implemented through a set of managed classes. The interface of the managed classes associated with images on a screen ("Screen" in Fig. 1) is part of the .NET Framework.

3.1 Analysis of Control System HMI structure

An analysis of HMI of modern Control Systems made it possible to reveal their typical structures that determine knowledge about the components of HMI (i.e. $KBC(ES)$). The following typical formats for presenting information in the Control System HMI are defined: mnemonic diagrams of three levels (generalized, group and individual); indications (sets of analog and discrete information); current events associated with the change of states of technological systems and processes; alarms; graphs of analog parameters (trends); instructions (for displaying the text of operational instructions); free formats that are used to obtain reference information in the operator's dialogue mode with an intelligent SCADA-system module.

Each information format is practically implemented as a specific HMI zone on the screen. Interface elements that can be located by the designer in the zones were combined into the conceptual class "Palette". The domain model "Palette" in the form of a class diagram of the UML language is shown in Fig. 2. The diagram displays two types of relations - aggregation and generalization, that is, the R_2 relation of the type "IS a". The relation R_2 , as applied to the class diagram, describes the hierarchical structure of classes and the inheritance of their properties and behavior.

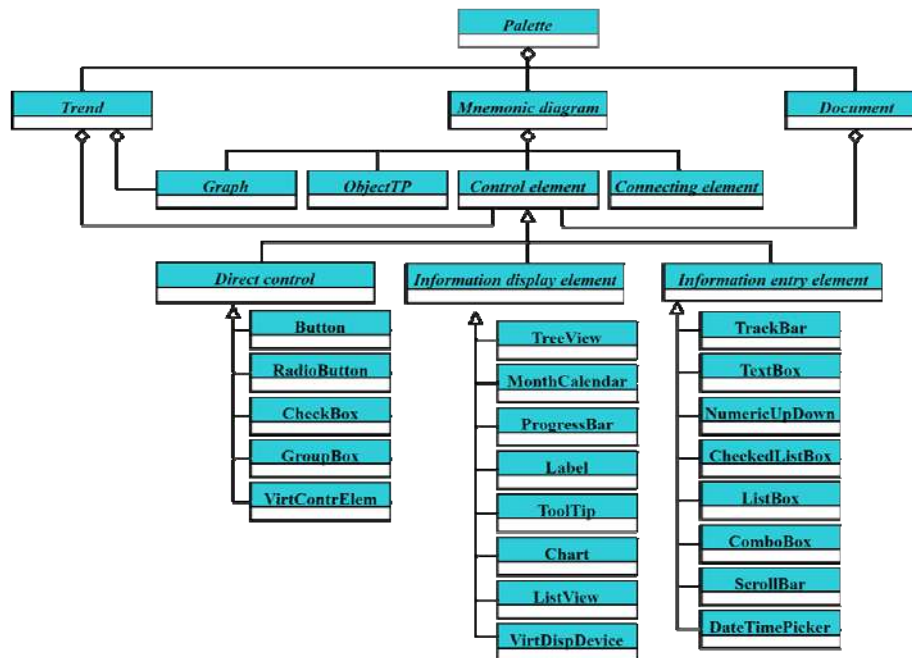


Fig. 1. Model of the subject domain "Palette"

Zones of all formats are presented in the form of three classes: "Mnemonic diagram", "Trend" and "Document" in Fig. 2. The purpose of the first two classes is obvious. The "Document" class was introduced to represent all the other formats listed above.

In accordance with the model, for example, the following elements can be placed in the zone of the mnemonic diagram: "Graph"; "ObjektTP" (OOC of technological process); "Control element" (HMI Controls) and "Connecting element" (connectors on a mnemonic diagram).

The "Control element" class is the parent for the following classes: "Direct control" (direct control elements); "Information display elements" and "Information entry element" (information input elements). Each of these classes, in turn, has child classes. Most of them are typical elements of the interface "Button", "TreeView", "ComboBox" and others. HMI may also contain virtual elements that simulate real controls and devices. To describe them, the model has introduced the classes "VirtContrElem" and "VirtDispDevice" respectively.

3.2 Knowledge about technology to transfer graphic objects to a monitor

The *KBO(N)* organizational knowledge of any MMS can be represented as a complex of morphological structures. The components of these structures are the many goals of the MMS, the conditions, tasks, functions, plans, means of technology and laws of the MMS control [23].

To achieve the "HMI design" goal, the designer must solve a lots of tasks $TS = \{TS_i \mid i = 1 \div K_i\}$, the nomenclature of which is due to known methodological techniques of IM and HMI design [10, 18-21, 25]. Knowledge about plans and technologies for their implementation (*KBTS*) is a subset of *KBEV* knowledge. They can be represented as a tuple [24]:

$$KBTS = \langle Fn_i, MTS(TS_i), MTc_i, P_{mm}(KM_p, MTc_i) \rangle, \quad (2)$$

where $Fn_i = \{Fn_{ij} \mid j = 1 \div K_j\}$ – set of functions Fn_{ij} , that must be implemented to solve the i -th task; $MTS(TS_i)$ – morphological-time structure of functions Fn_{ij} that shows a planned logical-time sequence of functions necessary for solving the TS_i task; MTc_i – set of technological ways to implement the function Fn_{ij} ; $P_{mm}(KM_p, MTc_i)$ – the predicate "The implementation of the function Fn_{ij} is possible in a set of technological methods MTc_i with a combination of conditions KM_p ".

Windows Forms graphic elements can be divided into three main categories: two-dimensional (2-D) vector graphics; drawings; typographic markup.

The managed class interface associated with transfer graphic objects to the monitor in the .NET Framework 4.5 and 4.6 consists of the following namespaces: System.Drawing; System.Drawing.Drawing2D; System.Drawing.Imaging; System.Drawing.Text; System.Drawing.Configuration and System.Drawing.Design.

The parent System.Drawing namespace contains types that support the basic graphical functions of GDI+. The basis of the GDI+ interface is the Graphics class. It directly performs the drawing of straight and curved lines, geometric figures, the output of drawings and text.

The analysis of the functional purpose of the System.Drawing namespace classes, the associated methods of the Graphics class, as well as the types of the alphabet, and the possibility of their use for coding the Windows Forms graphic elements was carried out.

We introduce four sets:

- A_1 – information transferred to a screen;
- A_2 – System.Drawing namespace;
- A_3 – Graphics class methods;
- A_4 – types of the alphabet.

In order to determine the knowledge required by the "Intelligent Interface" subsystem to implement the transfer and coding functions of HMI elements, we introduce the tetradic relation R_3 "To transfer and code a_1 , use a_2, a_3, a_4 " on the $A_1 \times A_2 \times A_3 \times A_4$ Cartesian product.

The relation R_3 can be matched to a 4-place predicate. A tuple (a_1, a_2, a_3, a_4) belongs to a relation if and only if the predicate of this relation is $P(a_1, a_2, a_3, a_4)=1$.

All elements of the R_3 relation are of the same type of tuples. Tuples of the same type make it possible to consider them as analogs of rows (i) in a simple table, that is, in a table where all rows consist of the same number of cells and the same data types

are located in the corresponding cells. An example of the R_3 relation representation in the form of a table is presented in the Table 1.

Table 1. An example of the R_3 relation representation

Displayed information	Classes Namespace System.Drawing	Class methods Graphics	Type of alphabet								
			Geometric figures	Symbols	Area of figures	Orientation in space	Hatching	Alphanumeric	Color, brightness	Flicker frequency	Opacity
2-D vector graphics	Brush	FillClosedCurve							1		
		FillEllipse					1		1		
		FillPolygon					1		1		
		FillRectangle					1		1		
	Pen	DrawArc	1			1			1		
		DrawClosedCurve	1	1	1	1			1		
		DrawEllipse	1		1	1			1		
		DrawLine				1			1		
		DrawPolygon	1		1	1			1		
Raster images	Image	DrawImage	1	1	1	1	1	1	1	1	
	Icon	DrawIcon		1	1				1		
Characters and Strings	StringFormat	DrawString						1			

The number of columns $j=4$ in the Table 1 is determined by the number of sets of Cartesian products. Table 1 for the convenience of presenting information (reducing the number of rows) has the following features:

- the 4-th column is presented in the form of 18 elements of the set A_4 (in the example of Table 1, nine alphabets are represented);
- tuples with common elements from $j=1\div 3$ columns and various elements of A_4 set are combined in one line;
- the following condition must be met for the x_{ij} elements of the table

$$x_{ij} = \begin{cases} 1, & \text{if } R_3 = \text{true}, \\ 0, & \text{if } R_3 = \text{false} \end{cases} \quad (3)$$

Such codes as letters, punctuation, and numbers are represented by one "Alphanumeric" alphabetic in Table 1. Due to the fact that the computer implementation of the "Color" and "Brightness" codes is associated with the same RGB color model, in Table 1 they are presented in one column.

After formalization and recording of the knowledge given in Table 1 in KB, the expert system will have knowledge about the transfer graphic objects to a screen technology and selection of codes for Windows Forms graphic elements. For example, when solving the problem of transfer raster images, the "Intelligent Interface" subsystem will propose to use the Image class AND/OR Icon class of the System.Drawing space and the corresponding methods of the Graphics class. If designer chooses the Image class, all kinds of codes will be recommended as possible. If he chooses Icon class, the following types of codes will be recommended: "Symbols" AND/OR "Area of figures" AND/OR "Color" AND/OR "Brightness".

The combination of the developed models, technologies and knowledge forms a new information technology of Control System HMI design.

4 Implementation

The developed information technology has become a theoretical basis to develop an ES, which could deliver a methodology of the design and set of functions as the intellectual tool to support technical designers whose task is to create Control System HMI. The general scheme of the ES demonstrated at Fig. 3.

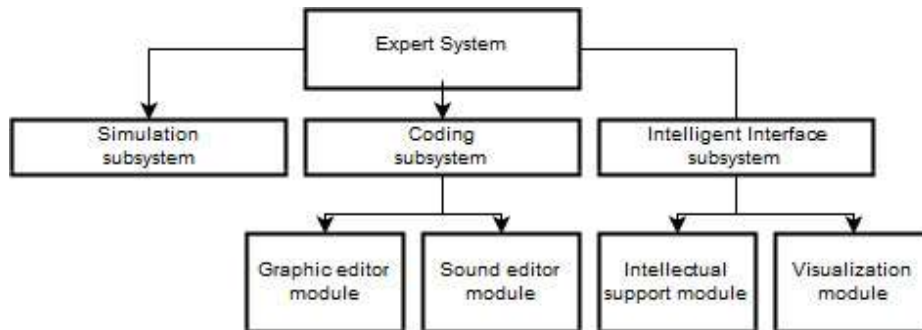


Fig. 3. The general scheme of the Expert System

The "Simulation" subsystem provides interactive extraction of expert knowledge about TP, OOC and their parameters, algorithms of the operator's activity. The goal of TP may be to decompose into other sub-levels (tasks, procedures, functions and operations). Knowledge about the sub-levels could be formalized in the form semantic graph. The nodes of this graph are the names of the functional units F_i of different scale, and the arcs show the relation R_{FiFj} - "to implement the functional unit F_i (goal, task), it is necessary to perform the functional unit F_j (task, procedure, function, operation)". The root of the tree is the concept of "Technological process". The description of the TP, given by experts, is saved in the tptree.xml file.

The description of TP objects, given by experts, is saved in the objects.xml file. Here is a fragment of this file that describes the "Operative" state of the Compressor-01 object of TP. This object has one quantitative parameter "Pressure" and one qualitative parameter "Noise". The parameter "Pressure" should be in the range from 1,0 to 2,1 Pa in "Operative" state. The operator must observe (tag "observed") and control (tag "controlled") the parameter "Pressure". The "Noise" parameter is only observed and can be "true" or "false":

```
<objects>
  <object name="Compressor-01" uuid="{550e8400-e29b-41d4-a716-446655440000}">
    <state name="Operative" uuid="s001">
      <parameters>
        <qnParameter name="Pressure" uuid="p001-p001" union="Pa"
          controlled="true" observed="true">
          <minValue>1.0</minValue>
          <maxValue>2.1</maxValue>
        </qnParameter>
        <qParameter name="Noise" uuid="q001" union="" controlled="false"
          observed="true">
          <values>
            <value> true </value>
            <value> false </value>
          </values>
        </qParameter>
      </parameters>
    </state>
  </object>
</objects>
```

The use of simulation at the stage prior to the development of the design specification for HMI allows: determine the probabilities of various MMS states; estimate the ability of the operator to process the incoming information necessary for decision making and draw up a list of those situations that he cannot solve due to his psychophysiological limitations. The results obtained will allow the designer to reasonably accomplish the task allocation between the operator and the machine. In the next stages of HMI design, this subsystem, if there are several variants of "paper" HMI prototypes, ensures the selection of the optimal variant. The choice is based on such quantitative indicators of the operator's activity, such as the probability of correct execution and execution time of the control algorithms obtained as a result of a simulation experiment.

The extracted expert knowledge about TP, OOC and their parameters are input to the Control System HMI design. The HMI design process is performed with the participation of the "Coding" subsystem and "Intelligent Support" module of the "Intelligent Interface" subsystem. The scenario of the HMI designer interaction with "Coding" subsystem can be described by the use case diagram shown in Fig. 4. The "Intelligent Interface" subsystem is shown as an external entity.

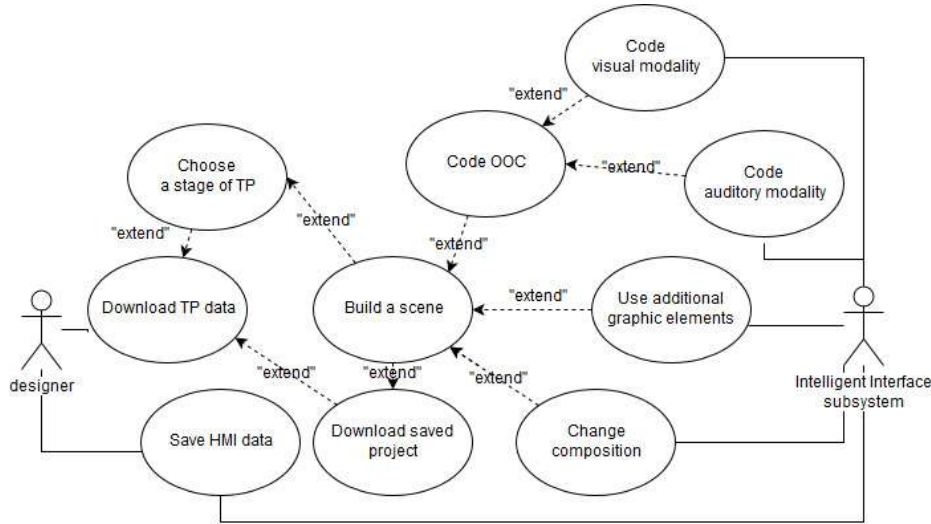


Fig. 4. The use case "HMI design" diagram

For each stage of TP a scene (video frame) is built, representing the TP mode. The video frame can be divided into zones with various formats of information presentation and filled with graphic elements. Thus, HMI consists of a set of scenes.

The process of building a scene consists in coding OOC, their states and parameters, as well as developing a scene composition with the allocation of operational space zones using various data presentation formats. The result of the expert's interaction with the intellectual support module is filling the KB with the following knowledge: requirements for the information about the OOC (see position 2 in Fig. 5); about the specifics of the operator's tasks and environment (see position 3 in Fig. 5). In accordance with the results of interaction, the HMI designer receives the ES recommendations on the choice of alphabet types for coding (see position 1 in Fig. 5). All subsequent steps of building the scene are also supported by recommendations for the designer. They are generated by the intellectual support module.

The graphic and sound editors modules are intended, respectively, for coding OOC and basic interface elements using alphabets coding visual modality (symbols, color, flicker frequency, brightness, etc.) and alphabets of auditory modality (verbal and sound signals) in accordance with ergonomic requirements.

Declarative knowledge is represented in the ES in the form of predicates (j); $P(x_1, x_2)$, where j is a name of fact (the ordinal number of the fact in the KB); P is a name of predicate; $x_i, i=1, 2$ is a subject constants.

Procedural knowledge is presented in the ES in the form of production rules (k); $R_i: A_i \rightarrow K_i$, where k is a name of the production (the ordinal number of the rule in the KB); R is a condition for the applicability of the production core; $A_i \rightarrow K_i$ is a production core.

After adding an element to the scene, HMI designer can apply coding to it in accordance with the received recommendations. This is done by changing the corre-

sponding properties of the element. An example of the correspondence between the properties of the Button control and the categories of codes (alphabets) that can be applied by using these properties are listed in Table 2.

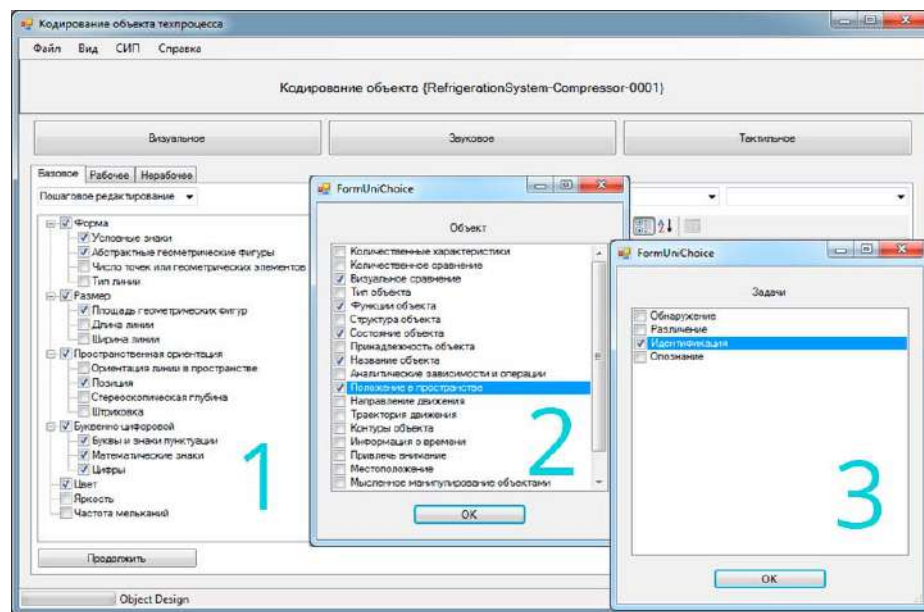


Fig. 5. An example of a expert survey and the generation of recommendations on the choice of the alphabet

Table 2. Table of correspondence of the Button control properties and the code category

Property	Code
FlatStyle	Stereo Depth
Font	Alphanumeric
FontHeight	Size
ForeColor	Color
Height	Size
Image	Geometric figures, Conventional symbols, Type of line, Number of points or geometric elements, Orientation in space, Alphanumeric
Size	Area of figure
Text	Alphanumeric
Width	Area of figure

These properties are changed and set using the element control properties panel.

The developed HMI project is saved in the .xml file format. Data from this file is input to the "Intelligent Interface" subsystem for the visualization of HMI. An example of the design results is presented in Fig. 6.

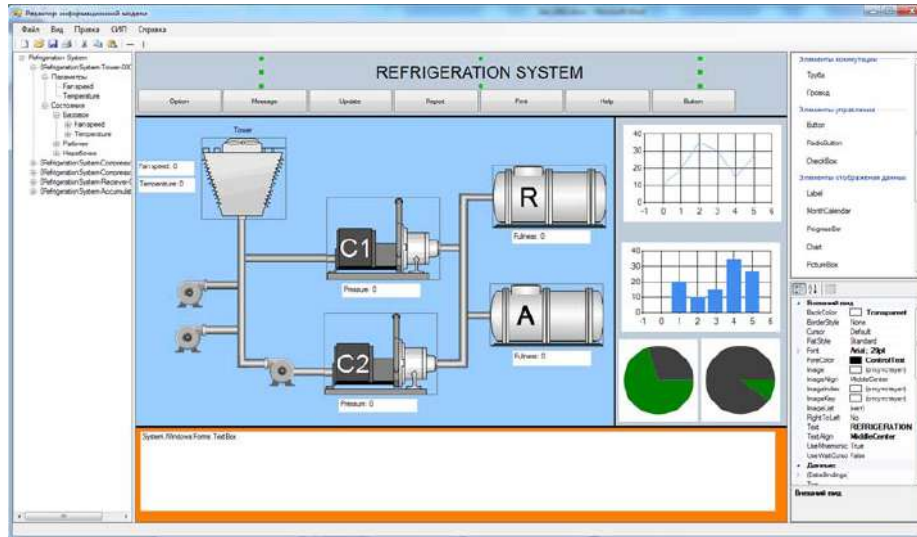


Fig. 6. An example of the HMI design results

The ES is implemented using Microsoft .Net Framework in C#. The Asp.Net, Ajax, CSS, HTML and SQL technologies were used to implement the visualization module. The WaveGenerator.dll library was used to work with the Wav sound format.

5 Conclusion

The urgent problem of information and mathematical support development is solved to automate the HMI designer activities.

The method of intellectual support for Control System HMI ergonomic design is proposed for the first time. Information technology that provides a methodology covering both design and pre-design stages of HMI design has been developed. At the pre-design stages of the HMI development, it is planned to carry out simulation modeling of TP and control processes (operator activities). The simulation results will allow more precise specify of the requirements for the operator activities and HMI at the terms of reference creation stage and to increase the validity of the task allocations between operator and machine in the early stages of HMI developing. If there are several “paper” prototypes or real HMI, the simulation will allow justifying the choice of the optimal variant based on quantitative indicators.

The practical significance of the obtained results is that the software realizing the proposed method is developed and can be used to solve practical problems of HMI design.

The experimental results of the Automated Control System operating analysis and modeling for aerated concrete production [26] have shown that the proposed method can be used to solve practical problems of HMI designing, as well as optimizing the structure and methods of TP controlling.

Acknowledgment

The work was performed as part of the project “Development of methods, models and information technology for managing complex objects under conditions of stochastic uncertainty” (number of 04918) of Zaporizhzhia National Technical University.

References

1. Five contributing factors of human error. <https://wonderwarepaocwest.com/news/operatr-five-contributing-factors-human-error/>
2. Muratov, O., Tihonov, M. The human factor at nuclear object. <http://www.proatom.ru/modules.php?name=News&file=print&sid=1555> (In Russian)
3. 2018 Self-driving safety report. General Motors. <https://www.gm.com/content/dam/company/docs/us/en/gmcom/gmsafetyreport.pdf>
4. Cochran, E., Bullemer, P.: Abnormal Conditions Management: Not by New Technology Alone. Proceeding of the AIChE Safety Conference, Houston, TX (1996)
5. Lorenzo, D.K.: A manager's guide to reducing human errors: Improving human performance in the chemical industry. Arlington, VA: Chemical Manufacturer's Association (1990)
6. EU-OSHA: The human-machine interface as an emerging risk. Luxembourg: Office for Official Publications of the European Communities (2009). doi: 10.2802/21813
7. The concept of the SCADA-system TRACE MODE functions development based on the expert systems of decision-making and execution technology. <http://isup.ru/articles/2/241/> (In Russian)
8. Be an Expert Builder of Expert System / Gensym. <http://www.gensym.com/product>
9. Lange, T.: Intelligent SCADA Systems. In: Engineer IT. Automation and Technical Control April 2007, pp. 26-30 (2007)
10. Munipov, V.M.: Ergonomics: human-oriented design of equipment, software and environments: Textbook. – M.: Logos, 2001. – 356 p. (In Russian) (2001)
11. Kaminska, Zh. K.: Analiz of problems of automation of process ergonomic design of information models of technological processes. In: Bulletin of Zhytomyr State Technological University. Technical science. Issue 1 (52), Zhitomir, – 2010, pp.103-108 (In Ukrainian) (2010). doi: [https://doi.org/10.26642/tn-2010-1\(52\)-103-108](https://doi.org/10.26642/tn-2010-1(52)-103-108)
12. Sukhov, S. A.: Development of a system for monitoring operational data based on WEB-technologies: Phd Theses: 05.13.11. Institute of High Energy Physics. – Protvino. – 87 p. (In Russian) (2007)
13. Poletykin, A. G.: Research and development of methods and software for information control systems of the upper block level: Doctoral Theses: 05.13.11. Institute of Control Sciences. – Moscow. – 160 p. (In Russian) (2009)
14. Sukhov, A. A.: Software systems for improving the quality of the interface of computer training systems: Phd Theses: 05.13.11. Mosk. state un-t instrument and informatics. – Moscow. – 132 p. (In Russian) (2010)
15. GOST 21480-76. Man-machine system. Mnemo-schemes. General ergonomic requirements. Introduction term 01.01.1977. – M.: Standards publishing house.– 6 p. (1976)
16. GOST 28392-89 Aircraft mnemo-schemes. General ergonomic requirements. Introduction term 01.01.1991. – M.: Standards publishing house. – 15 p. (1990)

17. ANSI/ISA-101.01-2015, Human Machine Interfaces for Process Automation Systems. July 2015, 64 p. <https://www.isa.org/store/ansi/isa-10101-2015,-human-machine-interfaces-for-process-automation-systems/41921133> (2015)
18. Burns, C.M., Haidukiewicz, J.R.: Ecological interface design. – Boca Raton: CRC Press (2004)
19. Upton, C., Doherty, G.: Extending ecological interface design principles: a manufacturing case study. In: International Journal of Human Computer Studies. 2008. 66 (4), pp. 271–286 (2008). doi: <https://doi.org/10.1016/j.ijhcs.2007.10.007>
20. Innovating for people: Handbook of human-centered design methods. Pittsburgh, PA: LUMA Institute, LLC (2012)
21. Fiset, Jean-Yves: Human-machine Interface Design for Process Control Applications. ISA. USA. 171 p. (2009)
22. O'Brien, L.: ISA-101 Standard Addresses the HMI Lifecycle. <https://www.arcweb.com/blog/isa-101-standard-addresses-hmi-lifecycle>
23. Nazarenko, N.A., Kaminska, Zh.K.: Method of human-machine interface designing of SCADA-system. In: News SPbGETU "LETI" № 7. 2017, pp. 31–38. (In Russian) (2017)
24. Information-Control Man-Machine Systems: Research, design, test: Handbook / Under the total. ed. Gubinskogo A.I. and Evgrafova V.G. – M.: Mashinostroenie, 1993. – 528 p. (in Russian) (1993)
25. Ergonomics: principles and recommendations. Methodical manual. – M.: VNIITE. – 184 p. (in Russian) (1983)
26. Kaminska, Zh. K., Kulynych, E. M., Serdiuk, S. N.: The Automated Control System operating analysis and modeling for aerated concrete production. In: Radio Electronics, Computer Science, Control. 2018. № 4, pp. 214–225. (In Russian) (2018). doi: [10.15588/1607-3274-2018-4-21](https://doi.org/10.15588/1607-3274-2018-4-21)

Models and Methods of Information Technologies of Spatial Configurations Synthesis

Sergiy Yakovlev¹[0000-0003-1707-843X], Oleksii Kartashov²[0000-0002-6282-553X],
Kyryl Korobchynskyi³[0000-0002-3676-6070]

¹ National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
svsyak7@gmail.com

² National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
alexeykartashov@gmail.com

³ National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
kirill.korobchynskiy@gmail.com

Abstract. The paper discusses the models and methods of information technology for the synthesis of optimal configurations of spatial objects. Models are based on the concept of geometric information and the formation of the configuration space of geometric objects. Information technologies carry out the transformation of geometric information in the process of synthesis of optimal configurations, the formation of an appropriate database structure, its consolidation and the organization of automatic exchange of information between the original software components and specialized software. An information-analytical model for the synthesis of optimal configurations has been built. The problem of forming a data structure and creating a data storage is considered.

Keywords: Geometric Information, Spatial Configuration, Information-analytical Model, Optimization.

1 Introduction

Currently, most publications are devoted to problems of mathematical modeling of systems with objects of a given spatial form and the development of special nonlinear optimization methods. The issues of information support, analysis of the data structure and methods of transformation of geometric information in the optimization process are not fully covered.

The purpose of this paper is to present the concept of creating information technology for the synthesis of spatial configurations of objects of a given shape and arbitrary metric characteristics. The use of such information technology allows to solve the following tasks:

- organize automatic exchange of information between the original software components and specialized software systems used to obtain local solutions and display the results of the solution in a convenient and multifunctional form;
- parallelize the search for local minima at the global search stage, which significantly reduces the time spent on solving the problem;
- visualize both final and intermediate solution results

- carry out an interactive search for solutions, which allows you to get better results in a shorter time.

The present paper continues the research [1] related to the use of information technologies for the synthesis of spatial configurations of various classes of geometric objects.

2 Object - Oriented Model and Configuration Space of Geometric Objects

The papers [2,3] introduced the configuration space of geometric objects (GO), which is based on the formalization of the concept of geometric information. Geometric information $\mathbf{g} = (\{s\}, \{m\}, \{p\})$ about the object S includes a spatial form $\{s\}$ as an equivalence class of point sets in space $R^3(R^2)$; metric parameters $\{m\} = (m_1, \dots, m_k)$ of the form, specifying the dimensions of the object; placement parameters $\{p\} = (p_1, \dots, p_l)$, that determine the position of an object in space. On a plurality of geometric information, the linear space of canonical informations and the general information space are constructed. In accordance with the general concept of constructing such spaces, we define their structure in the following way.

To specify the components $\{s\}$ and $\{m\}$ of the geometric information $\mathbf{g}$ of the object S we use the equation of its boundary $f(\xi, \mathbf{m}) = 0$, where $\xi = (x, y)$, if $S \subset R^2$, and $\xi = (x, y, z)$, if $S \subset R^3$. Let the variables $\mathbf{m} = (m_1, \dots, m_k)$ have an admissible domain of values $D \subseteq R^k$, and the function $f(\xi, \mathbf{m})$ is such that for any fixed $\mathbf{m} \in D$

$$f(\xi, \mathbf{m}) = 0, \text{ if } \xi \in fr S(\mathbf{m});$$

$$f(\xi, \mathbf{m}) > 0, \text{ if } \xi \in int S(\mathbf{m});$$

$$f(\xi, \mathbf{m}) < 0, \text{ if } \xi \in c(cl S(\mathbf{m})),$$

where c, fr, int, cl are operators of topological complement, frontier, interior and closure respectively.

We put the equation of the boundary of an object in a space of corresponding dimension in the basis of the creation of its object-oriented model (see Fig. 1).

In the papers [4,5], classes of so-called basic 2D and 3D objects were identified. Circles, ellipses, rectangles, convex polygons, as well as their closed additions to the whole space are offered as basic 2D-objects. The main three-dimensional objects are spheres, rectangular parallelepipeds, straight circular cylinders, circular cones, convex polyhedra and closures of additions of these objects in space R^3 .

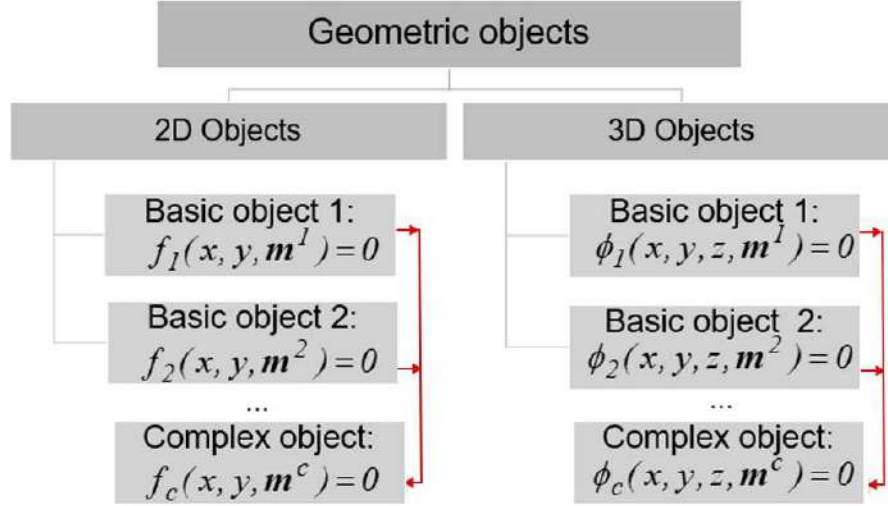


Fig. 1. Formation of geometric objects.

We propose to generalize the indicated results and form classical basic objects using the equations $f^c(x, y, \mathbf{m}) = 0$ and $\phi^c(x, y, z, \mathbf{m}) = 0$ of their boundaries in the corresponding spaces $R^2(R^3)$ (see Fig. 1). To construct an equation of the complex object boundaries, can be used the theory of R-functions [6].

Let us choose a system of coordinates in space $R^3(R^2)$, which we will call stationary, and we will associate with an object S their own (moving) coordinate system, the beginning of which is called a pole. The mutual position of the specified coordinate systems characterizes the placement parameters $\mathbf{p} = (p_1, \dots, p_\beta) = (\mathbf{v}, \boldsymbol{\theta})$, where $\mathbf{v}$ - the coordinate vector of the pole of the object in a fixed coordinate system, and $\boldsymbol{\theta}$ - the vector of the angle parameters that determine the mutual position of the axes of the eigen and fixed coordinate systems. For $S \subset R^3$ we have $\mathbf{p} = (p_1, \dots, p_\beta) = (\mathbf{v}, \boldsymbol{\theta})$. In the general case $\beta = 6$, $\mathbf{v} = (v_1, v_2, v_3)$ $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. For $S \subset R^2$ we put $\mathbf{p} = (\mathbf{v}, \boldsymbol{\theta}) = (v_1, v_2, \theta)$.

A position of the GO with respect to a stationary coordinate system may be given by the equation of its general position

$$F(\xi, \mathbf{m}, \mathbf{p}) = f[\mathbf{A}(\xi - \mathbf{u}), \mathbf{p}] = 0, \quad (1)$$

where $\mathbf{A}$ - the orthogonal operator, expressed through angular parameters $\mathbf{v}$.

The equation (1) is form the basis of the configuration space of the GO. A configuration space specifies a set of values of geometric variables called generalized coordinates, and defines the location in the space of the system of objects and their

parts both relative to one another and with respect to a given fixed reference system. We will formulate for a GO S its configuration space $\Xi(S)$, choosing as generalized variables $\mathbf{g} = (\mathbf{m}, \mathbf{p})$.

Let us consider a set of objects $\Omega = \{S_1, \dots, S_n\}$. Denote $\Xi(S_i)$ the configuration space of an object S_i with generalized variables $\mathbf{g}^i = (\mathbf{m}^i, \mathbf{p}^i)$, $i \in J_n$, where we use the notation $J_n = \{1, \dots, n\}$. Each point $\mathbf{g}^i \in \Xi(S_i)$ corresponds to a parameterized object $S_i(\mathbf{g}^i) \subset R^3(R^2)$. We will form the configuration space of a set of geometric basic objects $\Xi(\Omega) = \Xi(S_1) \times \dots \times \Xi(S_n)$ with generalized variables $\mathbf{g} = (\mathbf{g}^1, \dots, \mathbf{g}^n)$.

Definition. The mapping $\xi: \Omega \rightarrow \Xi(\Omega)$ of a set $\Omega = \{S_1, \dots, S_n\}$ in the configuration space $\Xi(\Omega) = \Xi(S_1) \times \dots \times \Xi(S_n)$, that satisfies a given set of constraints Λ , specifies the spatial configuration of objects.

Thus, the spatial configuration defines a set of parameterized GOs $S_i(\mathbf{g}^i)$, $i \in J_n$, which together totality form a complex object of a particular structure. Let us form a complex GO

$$S_B = B(S_1, \dots, S_n),$$

where the operator B defines the structure S_i , $i \in J_n$ of the object system. A complicated object S_B in the configuration space $\Xi(\Omega)$ corresponds to a parameterized GO

$$S_B(\mathbf{g}^1, \dots, \mathbf{g}^n) = B(S_1(\mathbf{g}^1), \dots, S_n(\mathbf{g}^n)), \quad (2)$$

and the point $\mathbf{g} = (\mathbf{g}^1, \dots, \mathbf{g}^n) \in \Xi(\Omega)$ specifies the spatial configuration of the objects S_i , $i \in J_n$ of the given structure. For fixed generalized variables $\mathbf{g}^i = \hat{\mathbf{g}}^i$, $i \in J_n$ the point $\hat{\mathbf{g}} = (\hat{\mathbf{g}}^1, \dots, \hat{\mathbf{g}}^n) \in \Xi(\Omega)$ defines an image of a complex object $S_B(\hat{\mathbf{g}}^1, \dots, \hat{\mathbf{g}}^n) = B(S_1(\hat{\mathbf{g}}^1), \dots, S_n(\hat{\mathbf{g}}^n))$.

On the basis of the above typology, taking into account equation (1) an object-oriented model of the GO is proposed, the fragment of which is shown in Fig.2. This polymorphic model contains an abstract class GeometryObjectBase that describes a set of virtual methods that implement operations common to all objects. Such operations may include receiving information from a file, saving a file, visualizing on the screen, and so on. The descendants of the GeometryObjectBase class have two implementations in accordance with the dimension of the GO space. Each of these classes contains fields and virtual methods that provide affine transformations of motion in the corresponding space.

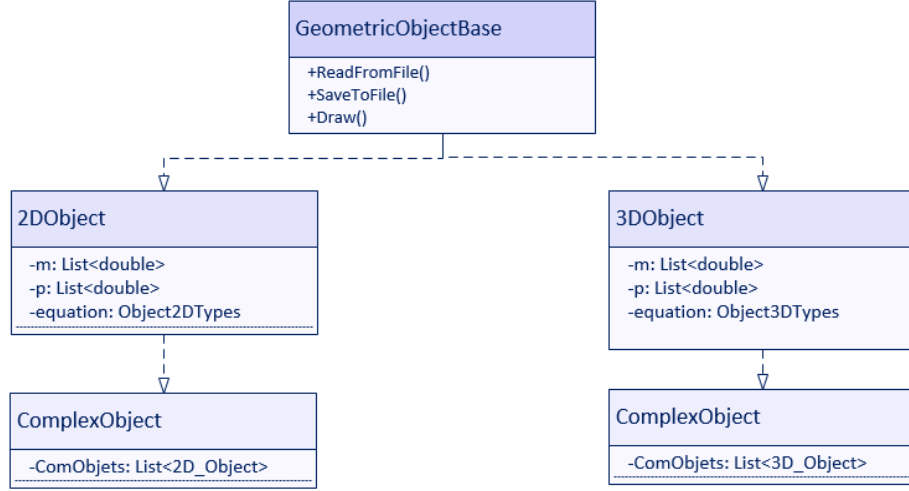


Fig. 2. Hierarchical object-oriented model of geometric object.

For example, for a 3D object, such fields will have three coordinates of the translation vector and three turning angles. The descendants, implementing complex objects form the next level. Each of them contains information about the metric parameters of this type of object and implements the necessary methods. The complex object will contain information about the set of objects of the corresponding space and the implementation of the methods that will be applied to each element of this set.

In general, the spatial configuration of a GO must satisfy a system of constraints that allow us to distinguish the corresponding class of spatial configurations. Such restrictions are due to the fact that the objects that make up the spatial configuration are different. Methods of formalizing such relationships depend on the choice of generalized variables of the configuration space, on the relative position restrictions of the GO and their physic mechanical properties (see Fig. 3).

Usually an additional object S_0 is used, called a container. In this case, all objects $S_i, i \in J_n$ must belong to the container S_0 . We introduce on the set of objects the binary relation of inclusion $\{\circ\}$. We will assume that $S' \circ S''$, if $int S' \subset S''$.

Let the object S_0 in the configuration space $\Xi(S_0)$ have generalized variables $\mathbf{g}^0$. We form the configuration space $\Xi(S_0) \times \Xi(\Omega)$. Then the set of generalized variables $(\mathbf{g}^0, \mathbf{g}^1, \dots, \mathbf{g}^n) \in \Xi(S_0) \times \Xi(\Omega)$ specifies the packing configuration, if $S_j(\mathbf{g}^j) \circ S_0(\mathbf{g}^0)$, $S_i(\mathbf{g}^i) * S_j(\mathbf{g}^j)$ for any $i, j \in J_n, i < j$. Note that in this case, the pole and its own coordinate system of the object S_0 coincides with the beginning and the axes of the fixed coordinate system, and the placement parameters of the container are $\mathbf{p}^0 = (0, \dots, 0)$.

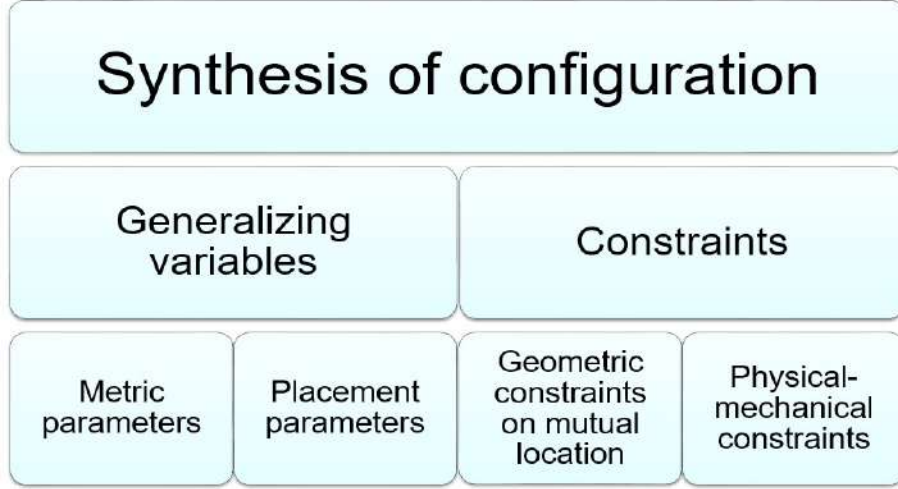


Fig. 3. Synthesis of spatial configurations.

Generalized configuration variables $\mathbf{g}^0, \mathbf{g}^1, \dots, \mathbf{g}^n$ of the configuration space $\Xi(S_0) \times \Xi(\Omega)$ may be subject to additional constraints that create special classes of packing configurations, including layout configuration, in the case of restrictions on the minimum and the maximum admissible distances between objects. In the case of objects representing solids $S_i, i \in \mathbf{J}_n$, of given masses $q_i, i \in \mathbf{J}_n$, a balanced system of such bodies defines a configuration of the balanced packing configuration. If the poles of the objects $S_i, i \in \mathbf{J}_n$ coincide with the centers of their masses, balanced packing takes place under condition

$$\sum_{i=1}^n x_i q_i = 0, \sum_{i=1}^n y_i q_i = 0, \sum_{i=1}^n z_i q_i = 0.$$

In accordance with (2) consider a complex object

$$S_B(\mathbf{g}^1, \dots, \mathbf{g}^n) = \bigcup_{i=1}^n S_i(\mathbf{g}^i)$$

in configuration space $\Xi(\Omega)$. Then the set of generalized variables $(\mathbf{g}^0, \mathbf{g}^1, \dots, \mathbf{g}^n) \in \Xi(S_0) \times \Xi(\Omega)$ specifies the covering configuration, if $S_0(\mathbf{g}^0) \supset S_B(\mathbf{g}^1, \dots, \mathbf{g}^n)$. Objects $S_0, S_i, i \in \mathbf{J}_n$ are called the coverage domain and covering objects, respectively.

We note some papers devoted to the study of spatial configurations of packing [4, 5, 7–11], layout [12–16], balanced packing [17, 18] and covering [19–21]. In the case when both metrical and placement parameters of GO take discrete values, we have a

class of combinatorial configurations. Since, in this case, combinatorial objects are characterized by the vector of their generalized variables, when they are mapped into Euclidean space, so-called Euclidean combinatorial configurations are formed, the properties of which are described in [22-24].

To formalize constraints in packing and layout problems with fixed sizes and shapes of objects, Yu.G. Stoyan developed the theory of Φ -functions. The study of spatial configurations and the corresponding configuration spaces allows generalization of the Φ -function concept to the case of variable metric parameters GO. The analysis of existing methods for constructing the Φ -function for basic 2D and 3D objects allowed us to naturally transfer known results to the class of packing and layout problems with variable metric parameters of objects. For an analytical description of the coverage conditions, a special class of ω -functions was proposed in [25].

3 Informational and Analytical Models Synthesis of Optimal Configurations

Taking into account the object-oriented model of GO and methods of their relationships modeling, an information-analytical model of the process of spatial configurations synthesis is presented (see Fig. 4).

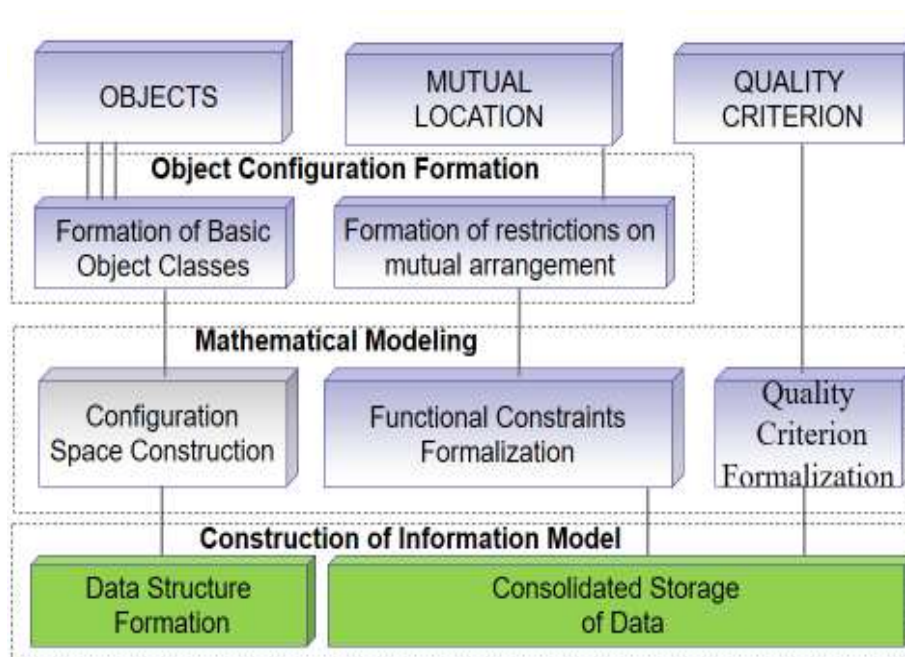


Fig. 4. Information-analytical model of optimal spatial configurations synthesis process.

Depending on the quality criteria and the choice of specific constraints that form the set of admissible spatial configurations, the task of finding optimal configurations can be attributed to the corresponding class of mathematical programming problems. The corresponding process of synthesis of optimal configurations is presented in Fig. 5.

On the one hand, methods of solving such problems can be implemented independently of the subject area. From the mathematical point of view, it is enough to formalize the target function, functional constraints, Jacobi and Hesse matrices to find local extremum. In this case, the development of modern information technology for the synthesis of optimal configurations of complex systems requires the construction of their information-analytical models in the automatic mode. On the other hand, the considered problems belong to the class of NP-complete, with a high dimension are multyextremes. Thus, the synthesis of optimal configurations is a creative intellectual process that makes it possible for the decision maker to be involved.

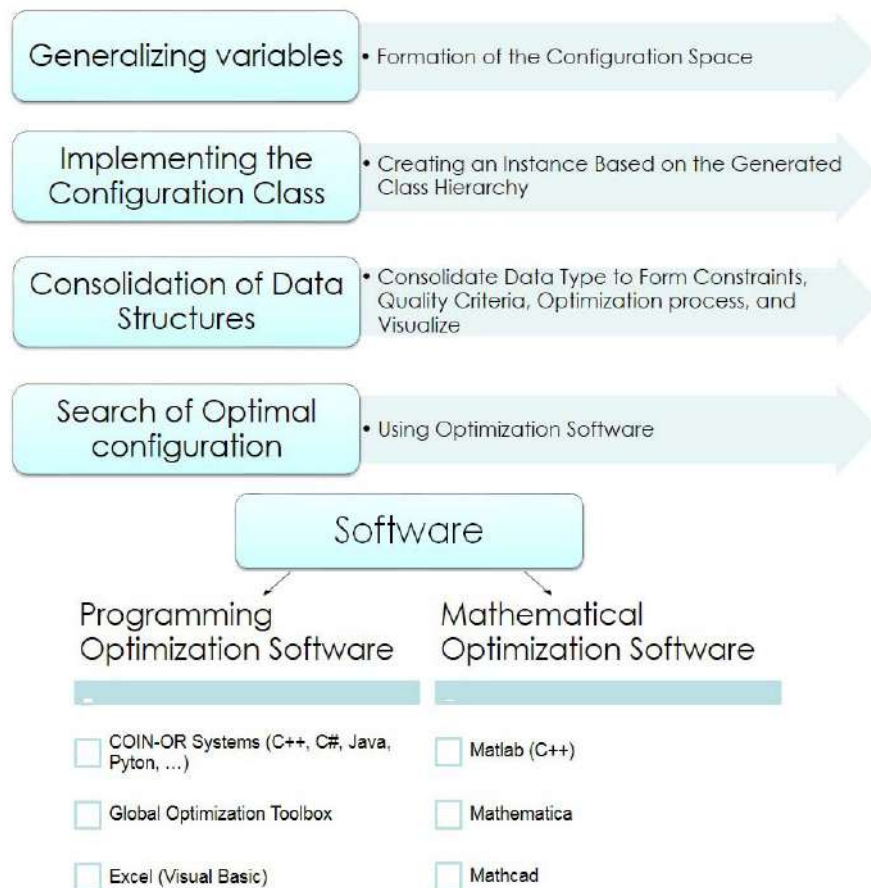


Fig. 5. Synthesis of optimal configurations.

The practice of using existing software packages to solve optimization tasks requires the construction of information-analytical models in a certain format with the help of external methods developed by the user. The available software leaves stage of the model construction the non-automated, which makes it urgent to search for new computer technologies for modeling problems in various technical fields and to develop information technologies for the transformation and visualization of geometric information in the process of synthesizing optimal configurations of complex systems.

3D visualizing of the solution process on each iteration is an important and rather difficult task, so it makes sense to use special software packages for this. For each iteration, excluding display intermediate solutions, the user according to the given parameters can influence the optimization process itself by changing the generalized variables of the model. Such changes may make the spatial configuration inadmissible. However, in accordance with the proposed technology, a locally optimal configuration is automatically synthesized that satisfies all the requirements. Data exchange in the implementation of this technology using a special consolidated data warehouse is shown in Fig. 6.

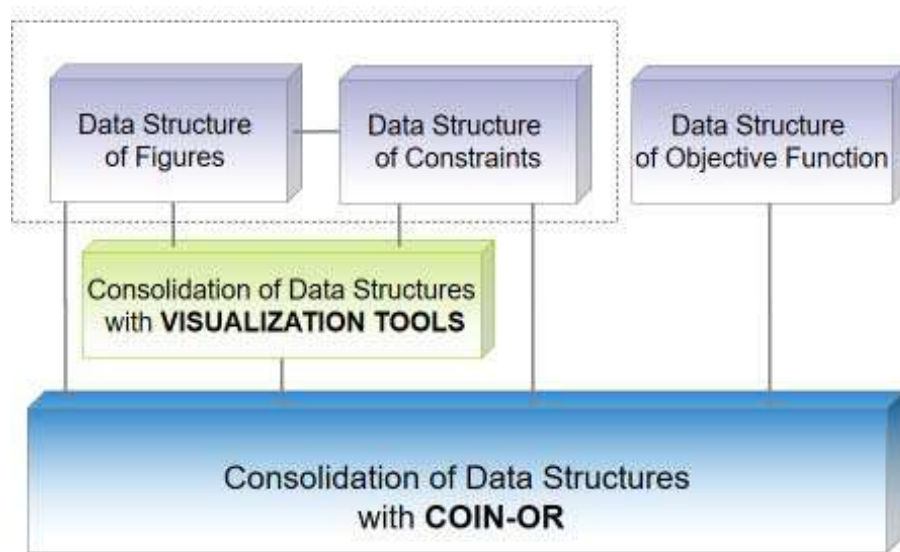


Fig. 6. Structure of the consolidated data storage.

The process of synthesis of spatial configurations is accompanied by multiple transformations of data of GO. In the analysis of the obtained results, detection, processing and transformation of suitable for use of the formed GO, formalized constraints and object function are performed. When using traditional tools, redundant data volumes arise and it need significant resources to perform their transformation.

4 Conclusions

The spatial configurations synthesis problem is so complex that it cannot be solved automatically and requires integration of different information technologies, including those involving the decision maker. This combination of different technologies is impossible without the use of information technologies related to the storage and transformation of data in different formats. To successful use of information technologies the paper formalizes the concept of geometric information and obtained its structure.

The configuration space of geometrical object was built generated by the parameters of geometric information. Depending on the type of geometric objects, the type of constraints and the function of the target, various spatial configurations were identified and analyzed. Their features were taken into account when implementing the method of solving a problem based on the use of information technologies. A general solution scheme was proposed and a consolidated database was developed, which allows combining various existing software packages and original elements in the process of solution. It is shown information technology using for exchanging of information between the original software components and specialized software systems, visualization both final and intermediate solution results, carrying out an interactive search for solutions.

References

1. Yakovlev, S.V., Kartashov, O., Korobchynskiy, K.: The informational analytical technologies of synthesis of optimal spatial configuration. In: IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, pp. 374-377 (2018).
2. Stoyan, Y.G., Yakovlev, S.V.: Configuration space of geometric objects. *Cybernetics and Systems Analysis*, vol. 54, no.5, pp. 716-726 (2018).
3. Yakovlev, S.V.: On some classes of spatial configurations of geometric objects and their formalization. *Journal of Automation and Information Sciences*, vol. 50, no. 9, pp. 38-50 (2018).
4. Stoyan, Y., Romanova, T.: Mathematical models of placement optimization: two- and three-dimensional problems and applications. *Modeling and Optimization in Space Engineering*, vol. 73, pp. 363-388 (2013).
5. Stoyan, Y., Pankratov, A., Romanova, T.: Placement problems for irregular objects: mathematical modeling, optimization and applications. *Optimization Methods and Applications*, vol. 130, pp. 521-559 (2017).
6. Rvachov, V.L.: *Theory R-function and its applications*. Kiev: Nauk. Dumka (1982).
7. Fasano, G.: A modeling-based approach for non-standard packing problems. *Optimized Packings with Applications*, vol. 105, pp. 67-85 (2015).
8. Sriramya, P., Parvatha, B.V.: A state-of-the-art review of bin packing techniques. *European Journal of Scientific Research*, vol. 86(3), pp. 360-364 (2012).
9. Fadel, G.M., Wiecek, M.M. Packing Optimization of free-form objects in engineering design. *Optimized Packings with Applications*, vol. 105, pp. 37-66 (2015).
10. Bortfeldt, A., Wascher, G.: Constraints in container loading: a state-of-the-art review. *European Journal of Operational Research*, vol. 229, no.1, pp. 1-20 (2013).

11. Yakovlev, S.V.: The method of artificial space dilation in problems of optimal packing of geometric objects. *Cybernetics and Systems Analysis*, vol. 53 no. 5, pp. 725-732 (2017).
12. Sun, Z.-G., Teng, H.-F.: Optimal layout design of a satellite module. *Engineering Optimization*, vol. 35(5), pp. 513-529 (2003).
13. Drira, A., Pierreval, H., Hajri-Gabouj, S.: Facility layout problems: a survey. *Annual Reviews in Control*, vol. 31(2), pp. 255-267 (2007).
14. Coggan J., Shimada K., Yin S.: A survey of computational approaches to three-dimensional layout problems. *CAD Computer Aided Design*, vol. 34(8), pp. 597-611 (2002).
15. Tian T., Zhu W., Lim A., Wei L.: The multiple container loading problem with preference. *European Journal of Operational Research*, vol. 248 (1), pp. 84-94 (2016).
16. Stoyan, Yu.G., Semkin, V.V., Chugay, A.M.: Optimization of 3D objects layout into a multiply connected domain with account for shortest distances. *Cybernetics and Systems Analysis*, vol. 50(3), pp. 374-385 (2014).
17. Kovalenko, A.A., et al.: Balance layout problem for 3D-objects: mathematical model and solution methods. *Cybernetics and Systems Analysis*, vol. 51(4), pp. 556-565 (2015).
18. Stoyan, Yu.G., et al.: Method of balancing rotating discretely distributed masses. *Energomashinostroenie*, no. 2, pp. 4-5 (1982).
19. Stoyan, Yu.G., Patsuk, V.M.: Covering a convex 3D polytope by a minimal number of congruent spheres. *International Journal of Computer Mathematics*, vol. 91(9), pp. 2010-2020 (2014).
20. Shekhovtsov, S.B., Yakovlev, S.V.: Formalization and solution of one class of covering problem in design of control and monitoring systems. *Avtomatika i Telemekhanika*, no. 5, pp. 160-168 (1989).
21. Yakovlev, S.V.: On a class of problems on covering of a bounded set. *Acta Mathematica Hungarica*, vol. 53(3), pp. 253-262 (1999).
22. Yakovlev, S.: Convex extensions in combinatorial optimization and their applications. *Springer Optimization and its Applications*, vol. 130, pp. 567-584 (2017).
23. Yakovlev, S.V., Pichugina, O.S.: Properties of combinatorial optimization problems over polyhedral-spherical sets. *Cybernetics and Systems Analysis*, vol. 54(1), pp. 99-109 (2018).
24. Yakovlev, S.V., Pichugina, O.S., Yarovaya, O.V.: Polyhedral spherical configuration in discrete optimization. *Journal of Automation and Information Sciences* 51(1), 38-50 (2019).
25. Yakovlev, S.V.: Formalization of spatial configuration optimization problems with a special function class. *Cybernetics and Systems Analysis*, vol. 55 (4), pp. 512-523 (2019).

Research and Development of the API for Personal Health Record

Anzhelika Parkhomenko¹ [0000-0002-6008-1610], Ivan Tyshchenko²
¹ Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine
parhom@zntu.edu.ua
² Vanderslab Ltd, Zaporizhzhia, Ukraine
johan.willfred@gmail.com

Abstract. The problems of electronic personal health record development and application are considered in this paper. The features of popular on the market software for working with patients' medical information have been analyzed. It is shown that the main restrictions on the usage of the considered solutions in Ukraine are: the orientation towards integration with the state medical systems of another countries, the absence of localizations, the limited functions of searching the necessary data in previously downloaded information, closed API and the lack of documentation. Therefore, this work is devoted to the research and development of public API for Personal health record as well as its documentation based on the OpenAPI specification.

Keywords: informatization of medicine, personal health record, application programming interface, web-application.

1 Introduction

Thanks to modern information technologies, medicine acquires completely new features today. The widespread introduction of computer technology and specialized software leads to significant changes in medical theory and practice, as well as in the training of healthcare professionals. [1-3].

Today, a significant number of medical institutions of various forms of ownership provides Ukrainians with a wide range of services, especially in large cities. Due to the fact that the majority of private medical institutions are relatively small, they are usually specialized in certain healthcare sectors, that generally leads to the usage of several unrelated clinics by citizens. As a result, patients have medical documents (advisory opinions, research results, prescriptions, etc.) in various formats and forms. Therefore, the patient's medical information is kept distributed and most often inaccessible to the person himself. This makes difficulty in changing the place of medical services receiving and increases the number of examinations that have to be repeated. The lack of a holistic "history" of the patient increases the risk of incorrect diagnoses and reduces the effectiveness of the treatment as a whole.

All this factors contribute to the direct interest of people in the availability of their own medical information, regardless of the medical institution. Especially the collection and processing of medical information is complicated for people who travel

frequently or change their place of residence, which is quite common in the modern world. Thus, a personalized electronic medical record can provide the patients with the possibility to more closely monitor their health status and have access to the necessary health information at anytime, anywhere, from any platform. Moreover, it can be not only a medical history or medical examination data, but various information related to human health (physiological parameters and features of a healthy person's development, dietology issues, healthy lifestyle, medical bills, etc.) [4-7].

Today, a significant number of software systems are offered on the market for creating a personalized electronic medical record and working with it. Each solution has its own characteristics of implementation and usage. The selecting of an appropriate software is a complex task, because it requires to take into account the peculiarities of the process of informatization of medicine in Ukraine, the issues of personal data protection and reliable information storage [8].

Therefore, this work is devoted to the study of methods and ways of organizing personalized electronic medical records that will ensure the effective storage and retrieval of information about patient's health.

2 State-of-the-Art

As studies have shown, in [9] several important definitions are given related to the electronic information about the patients' health:

- Electronic Medical Record (EMR) is the electronic information related to the health of the patient, which is created, stored, maintained and used by certified medical professionals and staff in one medical institution.
- Electronic Health Record (EHR) is the electronic information related to the health of the patient, complying with national interoperability standards, which is created, maintained and used by certified medical professionals and staff of more than one medical institution.
- Personal Health Record (PHR) is the electronic information related to the health of the patient, complying with national interoperability standards, obtained from various sources, which the patient self-manages, controls and provides access to it [9].

Thus, EMR, EHR and PHR are the sets of information about the health of a particular person, differing from each other in terms of the collection of information and the way it is managed. These three concepts indirectly determine the most important areas of health care system informatization:

- Informatization of specific medical organizations.
- Creation of integration projects for the information exchange between medical organizations, at the national and international levels, that means the creation of a single information space for physicians (professionals).
- Creation of a common information space for non-professionals (patients and even those who do not consider themselves to be patients yet), that means the creation of services and facilities for maintaining medical information and electronic interaction between doctors and patients.

Informatization of medicine in Ukraine is an important part of the health care system reforming [10,11]. But, as studies have shown, today in Ukraine there are only attempts to create systems for working with EMR and EHR, with a significant lag of the third component - PHR. Considering the experience of the United States, Great Britain and other leading countries in which PHRs are actively used, their development and implementation is also a key step in the informatization of medicine in Ukraine.

The analysis of existing on the market software solutions for the PHR allowed to identify the most popular: Microsoft HealthVault, Zweena, Health Companion, Healthspek.

Microsoft HealthVault is a popular environment for working with PHR, which allows to collect, store, use and monitor information about the health of a person and his family, including medical images. This solution can be called one of the best on the market due to the well-known name of the developer - Microsoft company and a fairly large number of platforms on which different versions of HealthVault can be installed [12].

The studies have shown that the disadvantages of HealthVault are:

- Glut of functionality, for example, the program combines task manager and calendar functionality that it is usually more convenient for users to have them as separate standard applications for their platforms (for example, Google Calendar for Android smartphones).
- Ultimate approach to interface organization.
- Commercial overtones: the application of this software leads to the usage of the services of certain companies and promotes certain products that are often provided by Microsoft partners.
- Restrictions by regional principle: not all countries are available when creating a profile (for example, Ukraine is not available for selection).
- Lack of speech recognition and optical character recognition (OCR).

Zweena is an environment for working with PHR that supports import from HealthVault and requires an account and pre-registration to log in. This product has less functionality than HealthVault, and it can be used both for free and with a monthly payment [13]. Zweena is only available as a website and does not have its own mobile applications for working with it. The shortcomings that were identified during analysis of this software are:

- Completely non-adaptive interface, the web site is in no way adapted for use on a device with the screen size less than 960px.
- Lack of mobile applications.
- Closed API (information about it can be obtained only after contacting the company on the phone).
- Restrictions of free account (1000 entries maximum).
- Chargeable recognition of scanned documents.

Health Companion is a complete and long-term system for working with PHR that can integrate personal health information from various sources, track medical financial bills and make preventive health and wellness recommendations based on personal risk factors [14].

Health Companion has an adaptive web-site. Android and iOS apps are available for usage free of charge. The interface languages are English or Spanish, so the applications are oriented for usage in Americas. The system has a lot of integration with polyclinics and other medical services, and it also meets the requirements of the USA law.

The shortcomings of this system revealed during the analysis are:

- Orientation to receive information primarily through integration with other systems.
- Mobile applications have some issues with compatibility with the latest operating systems of the platforms which they operate on.
- Closed API.
- Limited search capabilities.

Healthspek is a free system for managing all personal and family health information in one secure location accessible from multiple devices anywhere in the world. The system has an adaptive website, Android and iOS applications that are available for usage for free. Interface language is only English [15]. The shortcomings identified during the analysis of this system are:

- Closed API.
- Adaptation problems for countries without insurance medicine.
- Limited information retrieval capabilities.
- Poor support for people with disabilities.

Summing up the results of analysis of the functionality and features of popular all over the world software systems for working with PHR, it can be concluded that most of them are aimed at integration with certain information providers, as well as the usage for certain countries. Another important limitation of the considered PHRs is the lack of documents' recognition as well as full search tools (only in HealthVault there is a relatively comprehensive search, but it also does not allow finding audio recordings by content). One of the major drawbacks of all considered software systems is the closed API. Nevertheless, API is an interface that allows developers to use ready-made blocks to build a program code and, thus, to speed up software development and make it more attractive economically [16-18].

Therefore, the task of development of an open API, which will enable the users to maintain their own electronic PHR, that will combine information from different sources and provide convenient tools for working with personal data, is relevant.

3 Architectural Approach and the Methods of Organizing the API for PHR

The REST implementation was chosen as an architectural style for the API-oriented development of PHR, which has several advantages over the SOAP approach [19-21].

The developed public API for implementing client applications (mobile, desktop and web) will provide basic PHR functionality: storing, viewing, and searching of the health information. In this case, the following possibilities are given to user:

- Creation of profiles for other people (family members).
- Creation of events.

- Storing of information in the form of photos, text and audio records.
- Search among stored information.
- Deletion of the records.
- Export of documents as PDF files.

Based on the completed assessment of the level of technologies and platforms development, it was decided to develop the system using the JavaScript programming language, which provides sufficiently high speed, convenience and performance of writing code. Node.js was used for the software interface on the server side development. During the analysis of the development toolkit, it was decided to use a number of libraries, such as Elasticsearch, Express.js, Mongoose, etc. These solutions make it possible to reduce the cost of program developing while increasing its quality parameters, and therefore to increase its competitiveness.

The Express was chosen as the Node.js framework, which provides easy maintenance, high speed, relatively cheap hosting and a wide range of hosting selection. Node.js support is the minimum for most hosting sites, and therefore there is no need to pay for “exoticism”. The flexibility of the structure and the minimum amount of time spent on performing routine actions are also typical for it.

WebStorm from JetBrains was chosen as the development environment for the website. This environment is characterized by convenience and wide functionality. Its usage reduces time costs and increases programmer productivity. Developed API can be run on any modern Linux based VPS (Virtual Dedicated Server) server.

The created project can be divided into the following parts:

- ElasticSearch is the search engine, which provides high search performance among users' artifacts in the database.
- API documentation that provides developers with information on incoming and outgoing data for the route.
- API-core is the part of the system that is responsible for user requests routing and processing.

The developed API is the application that installs on an ordinary web server with Node.js 6 (LTS Boron), MongoDB 3.4. *, Elasticsearch 5.6. *. It is available for the usage by clients via the http / https protocol. JSON data exchange format is used.

It is necessary to perform the following steps for developed API functions usage.

Step 1. To form the HTTP request that includes the following data: server address, the API version (/ api / 1), the controller name (/ users), the action (/ profile). Parameters can be transmitted both via the URL "? Param = 123" and in the body of the request.

Step 2. To send the generated request to the server (at the specified address) via the GET / POST / PUT / DELETE method: POST https://server-url.com/api/profiles

Step 3. In response, to get the result of the command execution by the server in the form of JSON:

```
{
  "fullName": "string",
  "phone": "string",
  "birthday": "string",
  "relationship": "string",
```

```

    "gender": "string",
    "height": {
        "ft": 0,
        "in": 0
    },
    "weight": "string",
    "emergencyContactName": "string"
}

```

The created list of requests to API is as following.

POST / auth / sign-up – the user registration.

POST / auth / sign-in – the user authorization.

POST / auth / forgot – the password recovery.

POST / profiles / - to create a profile.

PUT / profiles / {id} – to edit a profile.

DELETE / profiles / {id} – to delete a profile.

POST / profiles / {profile} / events – to create a medical event.

GET / profiles / {profile} / events – to get a list of profile medical events.

DELETE / profiles / {profile} / events / {id} – to delete the event.

GET / profiles / {profile} / events / {id} – to get information about the event.

PUT / profiles / {profile} / events / {id} – to edit the event.

POST / profiles / {profileId} / events / {eventId} / recordings – to add an audio record;

DELETE / profiles / {profileId} / events / {eventId} / recordings / {recordId} – to delete an audio record;

GET / aprofiles / {profileId} events / search – the search of information.

GET / profiles / {profileId} / events / {eventId} / notes - the list of notes.

PUT / profiles / {profileId} / events / {eventId} / notes – to add the notes.

DELETE / profiles / {profileId} / events / {eventId} / notes – to remove the notes.

POST / profiles / {profileId} / events / {eventId} / photos – to upload photos.

DELETE / profiles / {profileId} / events / {eventId} / photos / {id} – to delete photos.

The convenience of working with any API substantially depends on how its documentation is written and executed.

There are different specialized services and tools (API BluePrint, Swagger, API Designer, RESTUnited, MireDot, Kittn API, apiDOC, apiGen and others) for API documentation development. All generators have their own advantages and disadvantages. Some of them are not for free and support limited number of programming languages. The process of documents generation is based on the usage of standardized format (Markdown, JSON, YAML, RAML). For example, the API BluePrint uses the Markdown format which is intended primarily for text formatting and not for the usage as a basis of documentation generation. The Swagger supports YAML and JSON formats. The YAML is much more convenient than the JSON, but it does not allow to describe the repeating elements easily, which are often present in the API descriptions. So, it is very difficult to adjust it under API documenting. API Designer is the platform for testing as well as the interactive editor, based on RAML

format for creating documents online. The simplicity and consistency are its advantages, but the lack of tools developed by the community is its open issue [18].

So, the Swagger was chosen for API documentation organization and management after the provided analysis [22]. Swagger User Interface is a small collection of scripts for creating interactive documentation for APIs of web applications with the REST architecture.

The main Swagger online documentation page is shown in Fig. 1. There are two text fields on the top panel. The first field contains the path to the JSON documentation file. The developed API methods are described in this format. The second field is the access token that will be used for requests to the API creation with the usage of Swagger. During the page opening, the value of the token is taken from the configuration file and the test user is added to the database using migration. If it is necessary, the access token can be changed with using the documentation. For example, for several users work simulation in the application, it is necessary to find the access token of the desired user in the database. Then enter this token in the field and click the Explore button. The token will be valid until the user exits this page. The token from the configuration file will be loaded after the page re-opening.

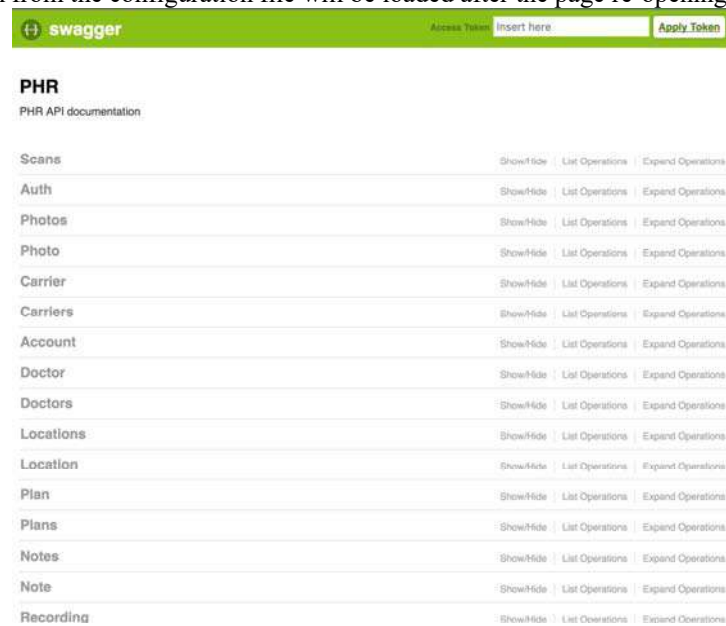


Fig. 1. API documentation page.

The documentation page contains the accordion interface element with addresses which are contained in the API. A brief information about the API (the base URL that is attached to the server address and the current version of the API) is shown after the list of display controllers. The list of groups' routes is opened with clicking on its name (Fig. 2).

It is possible to find out from the list next data: the method which you can call this route with (POST / GET / ...), the required URL that is attached to the base and a brief description of the assignment. Routes are sorted by call method and alphabetical order. If the method is selected, detailed information about it opens: a text description, a list of parameters and responses from the server, as well as a “Try it out” button for creating a request to the server (Fig. 3). The result of calling the API method is displayed after clicking on the button (Fig. 4).

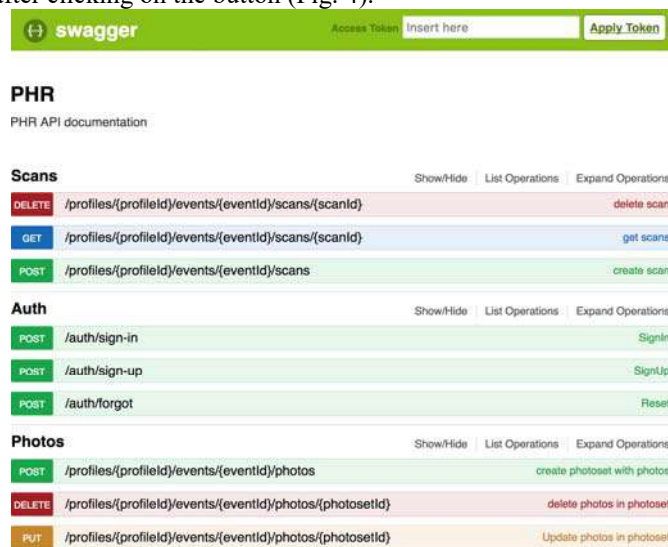


Fig. 2. Grouped API resources

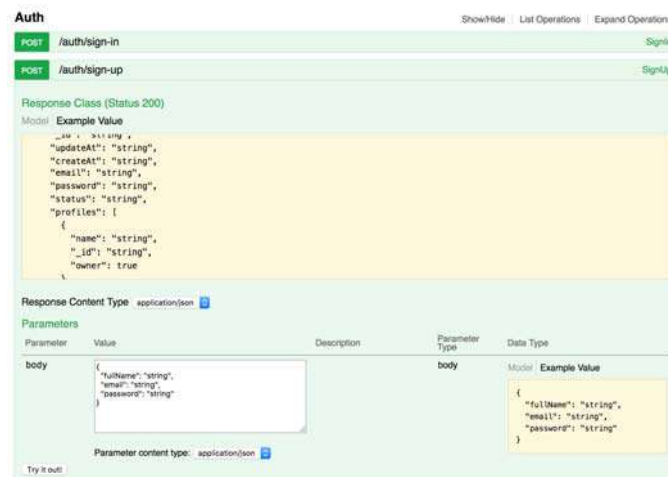
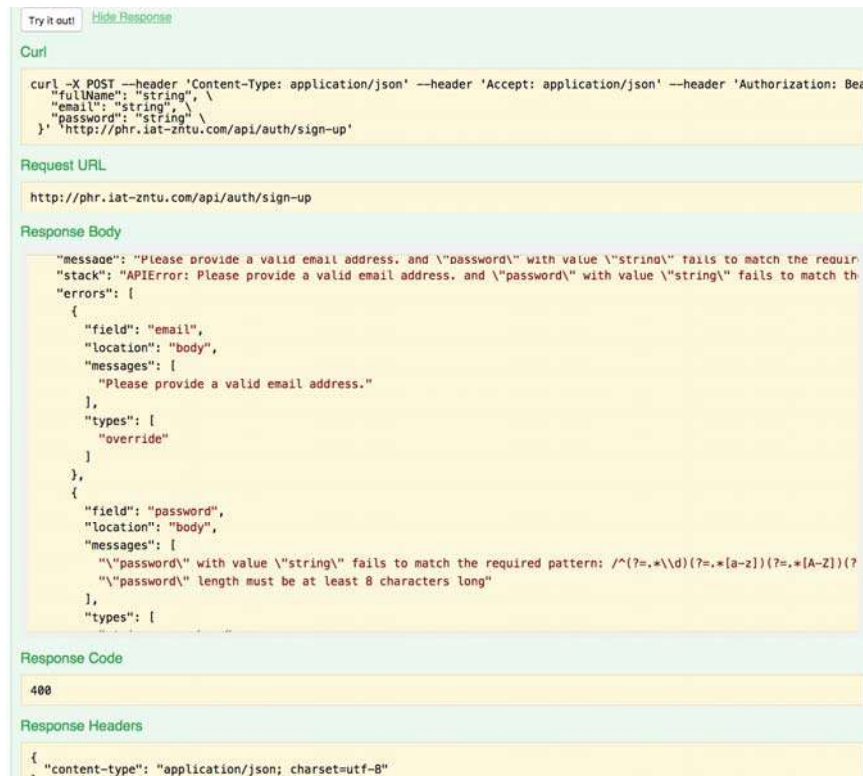


Fig. 3. Description of the method



```
Try it out! Hide Response

Curl
curl -X POST --header 'Content-Type: application/json' --header 'Accept: application/json' --header 'Authorization: Bearer '
  "fullName": "string", \
  "email": "string", \
  "password": "string" \
} 'http://phr.iat-zntu.com/api/auth/sign-up'

Request URL
http://phr.iat-zntu.com/api/auth/sign-up

Response Body
{
  "message": "Please provide a valid email address, and \"password\" with value \"string\" fails to match the required pattern: /^(?=.*\\d){?=[a-z]}{?=[A-Z]}{?=[0-9A-Z]{8,}$",
  "stack": "APIError: Please provide a valid email address, and \"password\" with value \"string\" fails to match the required pattern: /^(?=.*\\d){?=[a-z]}{?=[A-Z]}{?=[0-9A-Z]{8,}$",
  "errors": [
    {
      "field": "email",
      "location": "body",
      "messages": [
        "Please provide a valid email address."
      ],
      "types": [
        "override"
      ]
    },
    {
      "field": "password",
      "location": "body",
      "messages": [
        "\"password\" with value \"string\" fails to match the required pattern: /^(?=.*\\d){?=[a-z]}{?=[A-Z]}{?=[0-9A-Z]{8,}$",
        "\"password\" length must be at least 8 characters long"
      ],
      "types": [
        "override"
      ]
    }
  ]
}

Response Code
400

Response Headers
{
  "content-type": "application/json; charset=utf-8"
}
```

Fig. 4. Answer from the server

4 Conclusion

Nowadays Ukraine is only at the beginning of informatization process of medicine. Considering the experience of the United States, Great Britain and other advanced countries, the implementation of PHR is a relevant task and it can be considered as an important step in the informatization of medicine, together with the introduction of EMR and EHR. Therefore, the problems considered in this paper are relevant, and the proposed solution is seemed to be promising. As a work result, an open API was created, which simplifies the development of mobile and web applications for the necessary platform for working with PHR, and it also provides visual documentation according to the OpenAPI specifications.

Acknowledgements. This work is carried out partly with the support of Erasmus + KA2 project BIOART «Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees» (586114 – EPP – 1 – 2017 – 1 – ES - EPPKA2 – CBHE – JP)

References

1. Kachmar, V.O.: Medical information systems - the state of development in Ukraine. *Ukrainian Journal of Telemedicine and Medical Telematics*, **8** (1), 12-17 (2010)
2. Churpiy, I.K., Churpiy, N.V., Skripko, V.D. Current state of informatization in medicine. *Bukovinsky medical bulletin*, **15**(1), 171-173 (2011)
3. Nazarenko, G.I., Guliyev, Ya.I., Ermakov, D.E.: Medical information systems: theory and practice, general editor G.I. Nazarenko, 320 p. FIZMAT-LIT, Moscow (2005)
4. Bates, D.W.: Physicians and ambulatory electronic health records *Health Aff*, Millwood, **24** (5), 1180-1189 (2005)
5. Menachemi, N., Perkins, R.M., Durme, D.J.: Examining the adoption of electronic health records and personal digital assistants by family physicians in Florida. *Inform. Prim. Care*, **14**(1), 1-9 (2006)
6. Miller, R.H., Sim, I.: Physicians' use of electronic medical records: barriers and solutions. *Health Aff*, Millwood, **23**(2), 116-126 (2004)
7. Mark, A., Hall, J.D., Kevin, A., Schulman, M.D.: Ownership of Medical Information, *JAMA*, **301**(12) 1283 (2009)
8. Gravanova, Yu.: Features of the storage of medical information. <http://www.cnews.ru/reviews/free/national2006/articles/safe/>
9. Defining Key Health Information Technology Terms. <http://www.hitechanswers.net/wp-content/uploads/2013/05/NAHIT-Definitions2008.pdf>
10. Decree of the President of Ukraine No. 187/2012 The National plan of actions for 2012 on the implementation of the program of economic reforms for 2010-2014. "Prosperous Society, Competitive Economy, Efficient State". <https://zakon.rada.gov.ua/laws/show/187/2012>
11. Draft Order of the Ministry of Health of Ukraine "On approval of the concept of informatization of the healthcare of Ukraine on 2013-2018 years" http://uacm.kharkov.ua/download/2013_10/148-154_Konzeptiya_10_sc_P.pdf
12. Healthvault. <https://international.healthvault.com>
13. Zweena Health. <https://www.pinterest.com/zweenahealth/>
14. Health Companion. <https://www.healthcompanion.com>
15. Healthspek. <https://www.healthspek.com>
16. Anuff, Ed.: APIs are Different then Integration. <https://pages.apigee.com/rs/apigee/images/APIs-not-integration-ebook-05-2014.pdf>
17. Poliakov, M., Larionova, T., Tabunshchyk, G., Parkhomenko, A., Henke, K.: Remote laboratory for teaching of control systems design as an integrated system. In: The XIII international conference on Remote engineering and virtual instrumentation, Madrid, Spain, pp. 333-340 (2016)
18. Parkhomenko, A., Gladkova, O., Sokolyanskii, A., Shepelenko, V., Zalyubovskiy, Y.: Implementation of reusable solutions for remote laboratory development. *International Journal of Online Engineering*. **12**(7), 24-29 (2016)
19. Why Should We Choose REST (Client-Server) Model to Develop Web Apps? <https://medium.com/@audira98/why-should-we-choose-rest-client-server-model-to-develop-web-apps-c3bb2451b13a>
20. Types of HTTP requests and REST philosophy. <http://habrahabr.ru/post/50147/>
21. W3C. SOAP Specification. <https://www.w3.org/TR/soap/>
22. Maltseva, D. Why use Swagger for creating and documenting APIs. <https://dev.to/dianamaltseva8/why-use-swagger-for-creating-and-documenting-apis-1151>

Detection of Social Network Toxic Comments with Usage of Syntactic Dependencies in the Sentences

Serhiy Shtovba^[0000-0003-1302-4899], Olena Shtovba^[0000-0003-1418-4907],
Mykola Petrychko^[0000-0001-6836-7843]

Vinnytsia National Technical University, Khmelnytske Shose, 95, Vinnytsia, 21021, Ukraine
shtovba@vntu.edu.ua
olena.shtovba@yahoo.com
petrychko.myckola@gmail.com

Abstract. Social networks sometimes become a medium for threats, insults and other components of cyberbullying. A huge number of people are involved in online social networks. Hence, a protection of network users from anti-social behavior is an important activity. One of the major tasks of such activity is automated detecting the toxic comments with threats, insults, obscene etc. The bag of words statistics and bag of symbols statistics are typical features for the toxic comments detection. The effect of syntactic dependencies in sentences on the quality of detection of the social network toxic comments is studied in the article for the first time. Syntactic dependences are relationships with proper nouns, personal pronouns, possessive pronouns, etc. Twenty syntactic features of sentences have been verified in the total. The paper shows that 3 additional specific features significantly improve the quality of toxic comments detection. These three features are: the number of dependences with proper nouns in the singular, the number of dependences that contain bad words, and the number of dependences between personal pronouns and bad words. The experiments are based on data from kaggle competition "Toxic Comment Classification Challenge". For our experiments, the original dataset with 159751 comments was reduced to 106590 comments due to problems with human-free extraction of the syntactic features. We use mean of the error rates for each types of misclassification as the metric of quality due to unbalanced dataset. A decision tree is used as a classifier. The decision trees were synthesized for two splitting rules: Gini index and deviance criterion.

Keywords: natural language processing, syntactic dependencies, toxic comments, social network, machine learning, features selection, balanced accuracy, decision tree.

1 Introduction

Social networks sometimes become a place for threats, insults and other components of cyberbullying. A huge number of people are involved in online social networks. Hence, a protection of network users from anti-social behavior is an important activity. One of the major tasks of such activity is automated detecting the toxic com-

ments. Toxic comments are textual comments with threats, insults, obscene, racism etc.

The various techniques are used for human-free detecting the toxic comments. Bag of words statistics and bag of symbols statistics are typical source information for the toxic comments detection. Usually the following statistics-based features are used: length of the comment, number of capital letters, number of exclamation marks, number of question marks, number of spelling errors, number of tokens with non-alphabet symbols, number of abusive, aggressive, and threatening words in the comment, etc. [1]. High count of bad words in the comment increases a chance to classify it as toxic. However, there are some difficulties with usage of the bad words statistics. Some out-of-vocabulary words are produced by typos and by spelling errors. Often authors of toxic comments distort their bad words purposely. They convert the bad words to phonetically identical forms by replacing letter combinations *oo* to *u*, *for* to *4*, *too* to *2* etc. Another variant is to distort to visual similar forms, for example, *5h1t*, *b!tch*, *b1tch*. Scientists develop special technologies for detecting the masked bad words [2, 3], but vandals have a reserve in time and in persons. In addition to analyzing the separated keywords, some methods take into account the order of the words in sentences. For example, authors of [4, 5] used n-grams-based approach, but such modeling does not reflect the whole relations in sentences.

The aim of the paper is to study an effect of syntactic dependencies in sentences on the quality of detecting the social network toxic comments. Syntactic dependences are relationships with proper nouns, personal pronouns, possessive pronouns, etc. Opposite to n-gram method and naive Bayesian approach, the model based on the syntactic dependencies does not directly tie with the training set vocabulary. All the various proper names, personal pronouns, possessive pronouns are allocated into separate groups. It allows to use the vocabulary-free generalized features in the model. Another instance from this group in the test set will not affect the simulation negatively. We use the information technology from [6] for extraction the syntactic features from the data set. We compare the results of toxic comments detection on two sets of features. The first set is typical features that based on bag of words statistics and bag of symbols statistics. The second one is extended set that contains typical features together with syntactic features. The experiments are performed on the “Toxic Comment Classification Challenge” data set.

2 Data sets and preprocessing

Data set “Toxic Comment Classification Challenge” is collected by Conversation AI team, a research initiative founded by Jigsaw and Google, both a part of Alphabet. The data set is used in kaggle-competition [7]. The data set consists of 159751 Wikipedia comments which have been labeled by human raters for toxic behavior. Most of the comments are English [8].

Each comment is manually categorized with 6 binary labels: toxic, severe toxic, obscene, threat, insult, and identity hate. Some comments have toxic multiplicity. In this case a comment belongs to 2, 3, and even 6 toxic categories simultaneously (Fig-

ure 1). Also a comment may be neutral, i.e. it does not belong to any toxic category. For example, the following comment “Your vandalism to the Matt Shirvington article has been reverted. Please don’t do it again, or you will be banned.” is neutral. Comment “Hi! I am back again! Last warning! Stop undoing my edits or die!” is toxic and threatened, and comment “Would you both shut up, you don’t run Wikipedia, especially a stupid kid.” is toxic and insult.

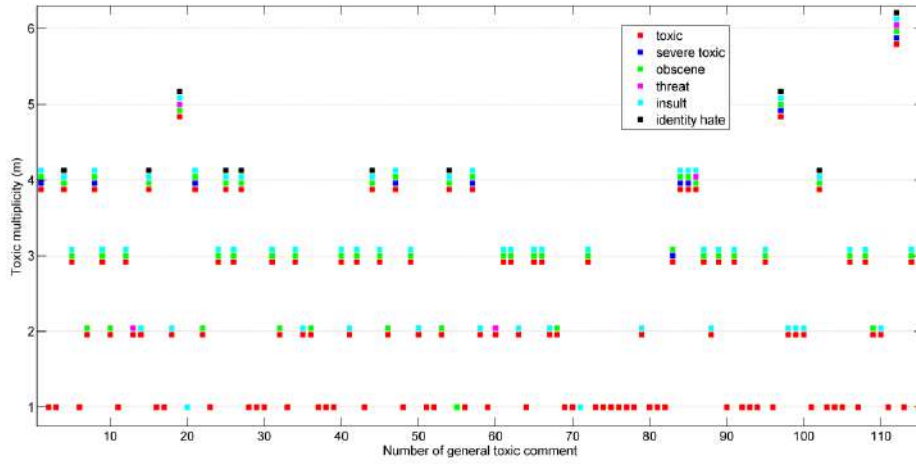


Fig. 1. Categories of the first 115 non-neutral comments

16225 comments have the toxic labels. The rest of the comments are neutral. A distribution of the comments on toxic multiplicities is presented on Figure 2. It shows that only comments with high toxicity multiplicity are rarely encountered. Most of toxic comments (60.8%) belong to several toxic categories ($m > 1$).

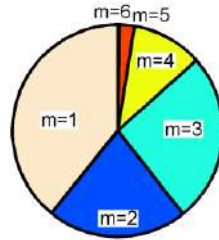


Fig. 2. Distribution of multiplicity (m) of toxic comments

Figure 3 shows the combinations of toxic categories in one comment. The set of toxic comments with the same category is represented by a color square. Each toxic category represents the corresponding color. The area of the square equals to the number of comments with the same toxic category. The intersection of squares reflects the number of comments that belong to two relevant toxic categories simultaneously. Figure 3 shows that all the severe toxic comments also belong to toxic cate-

gory – the blue square is completely inside the red square. Also, almost all the severe toxic comments are obscene and insult. There are 3 very low intersecting categories: severe toxic, threat, and identity hate. Few comments belong simultaneously to two out of these three categories. Figure 3 also shows the degree of similarity for two finite sets in form of Jaccard index (k_j). It is calculated as the cardinality of the intersection of the sets divided by the cardinality of the union of the sets. For our case Jaccard index corresponds to the ratio of the area of intersection of two squares over the area of the union of two squares.

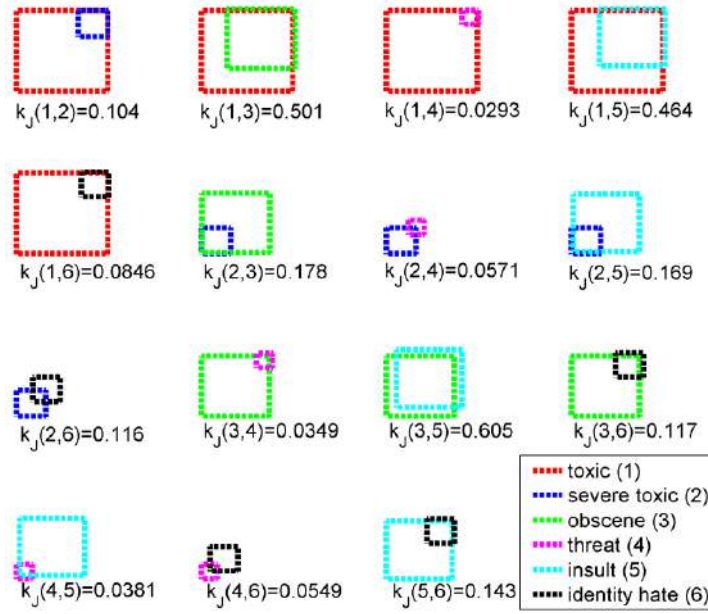


Fig. 3. Jaccard similarity indexes for various toxic categories

We propose to add several specific features to the typical feature set that based on statistics of a bag of the words and statistics of a bag of the symbols. The specific features are taking into account some syntax dependencies between words in comment. The specific features extraction was done using the technology from [6]. The specific features were extracted automatically for 106590 comments. Features extraction for some comments was unsuccessful due to non-English text and out-of-vocabulary words. As a result, the modified data set consists 66.8% of the source data set. Neutral comments compose 87.2% of the modified data set. It is slightly less than in the source data set where the neutral ratio is 89.8%. Distributions of the comments on toxic categories are almost equal for two data sets (Table 1).

Table 1. Source data sets and modified data sets

Category	Comments in source data set	Comments in modified data set	Share of source data set, %
Toxic	15294	12948	84.7
Severe toxic	1595	1492	93.5
Obscene	8449	7303	86.4
Threat	478	442	92.5
Insult	7877	6943	88.1
Identity hate	1405	1251	89

3 Features and quality metric

The following features are used for formalized description of each comment:

- x_1 is a number of words;
- x_2 is a number of unique words;
- x_3 is a ration of unique words;
- x_4 is a number of tokens without the stop-words;
- x_5 is a number of spelling errors;
- x_6 is a number of all-caps words;
- x_7 is a ratio of all-caps words;
- x_8 is a length of the comment;
- x_9 is a number of capital letters;
- x_{10} is a number of explanation marks;
- x_{11} is a number of question marks;
- x_{12} is a number of punctuation marks;
- x_{13} is a number of masking symbols (*, &, \$, %);
- x_{14} is a number of happy smiles;
- x_{15} is a ratio of explanation marks;
- x_{16} is a ratio of question marks;
- x_{17} is a ratio of spaces;
- x_{18} is a ratio of capital letters;
- x_{19} is a ratio of lowercase letters;
- x_{20} is a number of the comment's words that included into the bad word list at <https://www.cs.cmu.edu/~biglou/resources/bad-words.txt>;
- x_{21} is a number of the comment's words that included into the swear word list at <http://www.bannedwordlist.com>;

x_{22} is a number of the comment's words that included into facebook black list at <https://www.frontgatemedia.com/a-list-of-723-bad-words-to-blacklist-and-how-to-use-facebooks-moderation-tool/>;

x_{23} is a number of the comment's words that included into google blacklist at <https://www.freewebheaders.com/full-list-of-bad-words-banned-by-google/>;

x_{24} is a number of the comment's words that included into the naughty word list at <https://gist.github.com/ryanlewis/a37739d710ccdb4b406d>;

x_{25} is a number of the comment's words that included into 5 mentioned lists;

x_{26} is a number of dependencies with proper nouns in the singular;

x_{27} is a number of dependencies with proper nouns in the plural;

x_{28} is a number of dependencies with personal pronouns;

x_{29} is a number of dependencies with possessive pronouns;

x_{30} is a number of dependencies with denial (with words never or not);

x_{31} is a number of dependencies with denial that contain proper nouns in the singular;

x_{32} is a number of dependencies with denial that contain proper nouns in the plural;

x_{33} is a number of dependencies with denial that contain personal pronouns;

x_{34} is a number of dependencies with denial that contain possessive pronouns;

x_{35} is a number of dependencies between proper nouns in the singular and the words from dependencies with denial;

x_{36} is a number of dependencies between proper nouns in the plural and the words from dependencies with denial;

x_{37} is a number of dependencies between personal pronouns and the words from dependencies with denial;

x_{38} is a number of dependencies between possessive pronouns and the words from dependencies with denial;

x_{39} is a number of dependencies that contain the bad words;

x_{40} is a number of dependencies with denial that contain the bad words;

x_{41} is a number of dependencies between proper nouns in the singular and the bad words;

x_{42} is a number of dependencies between proper nouns in the plural and the bad words;

x_{43} is a number of dependencies between personal pronouns and the bad words;

x_{44} is a number of dependencies between possessive pronouns and the bad words;

x_{45} is a number of dependencies between pronouns and the bad words.

Twenty specific features $x_{26} - x_{45}$ are examined for toxic comments detection for the first time. Let us modify the original kaggle-task of categorizing the toxic comments to the classification one with two alternatives: a neutral comment and a general

toxic comment. It allows to easy checkup the informative levels of the proposed syntactic features.

The data set is unbalanced with class proportion about 9 to 1. Hence, misclassification rate is not suitable metric for quality of the classifier. According to [9] we use balanced accuracy approach. The metric of quality of the classifier is as follows:

$$Q_{aver} = \frac{P_{nt} + P_{tn}}{2},$$

where P_{nt} denotes probability of $n \rightarrow t$ type classifying errors, when a neutral comment is recognized as a general toxic comment; P_{tn} denotes probability of $t \rightarrow n$ type classifying errors, when a general toxic comment is recognized as a neutral comment. Q_{aver} is mean of probabilities of each type misclassification. It is simple and interpretable metric for examination a classifier on unbalanced data set.

4 Computational experiments

A decision tree is used as a classifier. We choose this kind of classifier taking into account the following reasons: 1) a synthesis of the decision tree is a fast procedure even for large training set, hence, it is possible to carry out several experiments; 2) features selection is carried during the decision tree synthesis; it is easy to check the informative levels of the proposed syntactic features. We divide the data set on training data and test data. The test set consists of every sixth comment. The rest comments are in the training set. Thus, the test set contains 17765 comments and training set contains 88825 comments. We use the training data for decision tree synthesis. After this, the decision tree is pruned for minimization Q_{aver} on the test set. We check up two sets of the features: typical set – $x_1 - x_{25}$ and extended set - $x_1 - x_{44}$.

Rebalancing the class distribution is yielded by a sampling in way of increasing the weight of minor class objects. We suppose that correct classification of the comment with high toxic multiplicity is more important than the comment with low toxic multiplicity. Weight w of toxic comment C is defined by the following heuristic formula:

$$w(C) = b + \sqrt{m(C)},$$

where b denotes a bias of toxic comment weight; $m(C) \in \{1, 2, \dots, 6\}$ denotes toxic multiplicity of comment C .

Figure 4 shows the dependences of the classifier quality under the bias of toxic comment weight. The decision trees were synthesized with two splitting rules: Gini index-based rule and deviance criterion-based rule. The experiments show that Gini index-based rule provides better decision trees. Q_{aver} is low, when the bias of toxic comment weight belongs to $[4.5, 5.8]$. Minimal value of $Q_{aver} = 0.118$ is obtained for

$b \in [5.2, 5.5]$. Figure 4 shows that the extended set of features significantly improves the classifier quality.

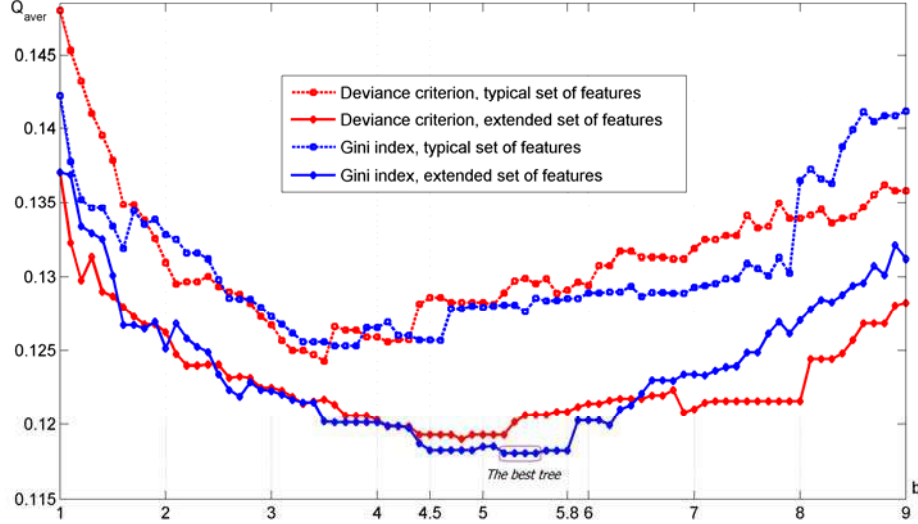


Fig. 4. Experimental dependencies of toxic comments classifier quality

The best model is a decision tree with minimal value of Q_{aver} . The best decision tree is presented on Figure 5. Misclassification rate for the best decision tree is $Q=0.0987$. The other metrics for the best decision tree are as follows: $Q_{aver}=0.118$, $P_{nt}=0.0919$, and $P_{in}=0.1442$. The best tree correctly detects almost all comments with high and average toxic multiplicities (Figure 6). The best tree correctly detects almost all the toxic comments with labels severe toxic, obscene, and identity hate (Figure 7).

Let us analyze 5 best trees. All the trees use the following features: $x_3 - x_9$, x_{15} , $x_{17} - x_{19}$, x_{22} , $x_{24} - x_{26}$, x_{39} , and x_{43} . 4 out of 5 trees use feature x_1 additionally. Among their most important features are 3 new syntactic ones: a number of dependencies with proper nouns in the singular (x_{26}); a number of dependencies that contain the bad words (x_{39}) and a number of dependencies between personal pronouns and the bad words (x_{43}).

We also point to 4 following slightly less important features. Typical features x_2 , x_{10} , and x_{12} are in 2 out of 5 the best trees. Syntactic feature x_{28} is selected for 1 out of 5 the best trees. The mentioned 4 extra features may be used for more complicated models for toxic comment detection.

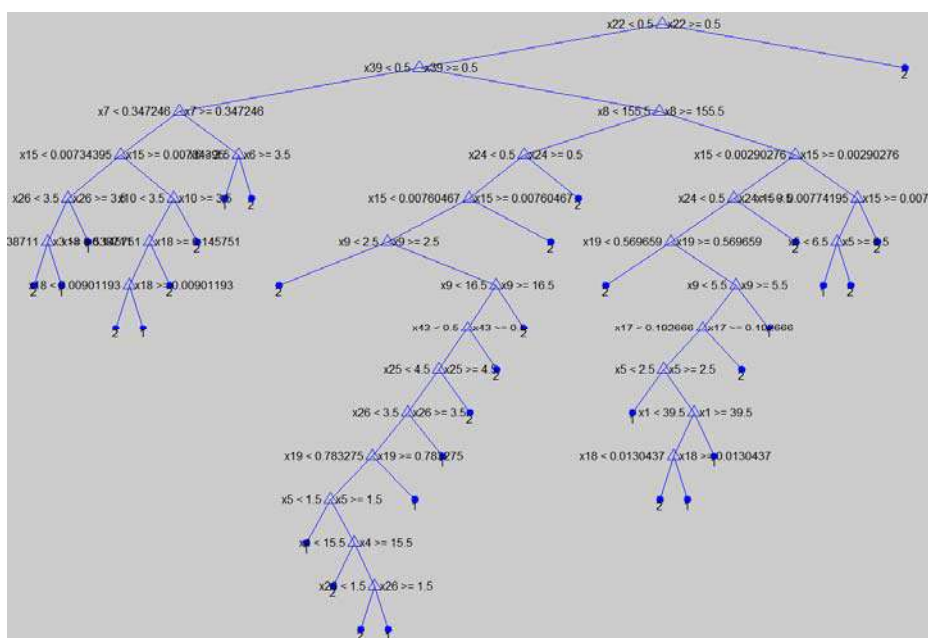


Fig. 5. The best decision tree (1 – neutral comment, 2 – general toxic comment)

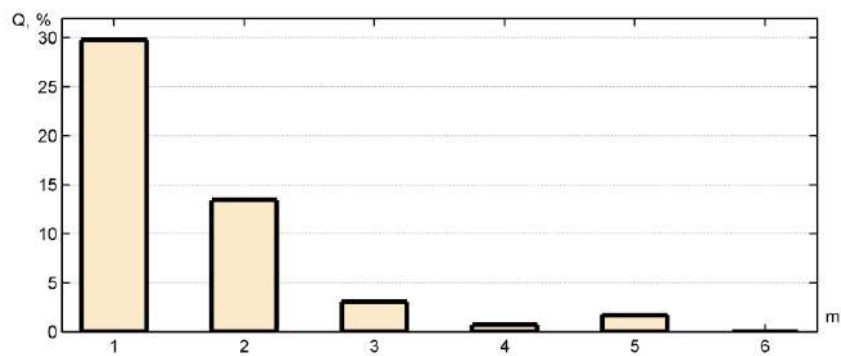


Fig. 6. Distribution of misclassification rate (Q) on the comments with various toxic multiplicity (m)

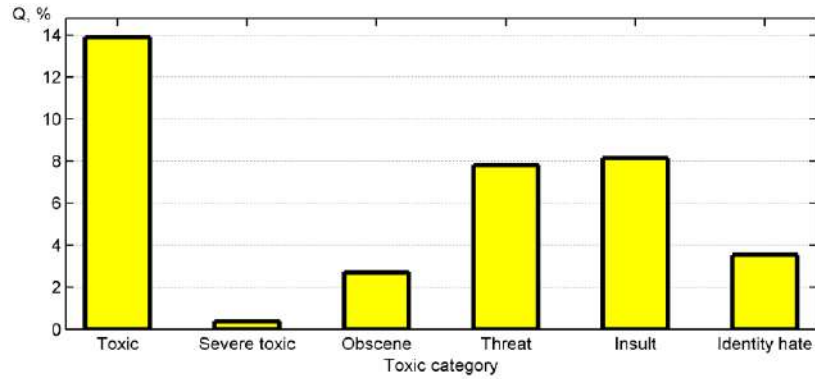


Fig. 7. Distribution of misclassification rates (Q) on various toxic categories

5 Conclusion

The problem of detecting the toxic comments in social networks was considered. For our experiments we used kaggle data set "Toxic Comment Classification Challenge". The bag of words statistics and bag of symbols statistics are typical features for detecting the toxic comments. The effect of syntactic dependencies in sentences on the quality of the social network toxic comments detection was studied in the article. Syntactic dependences are relationships with proper nouns, personal pronouns, possessive pronouns, etc. In total 20 syntactic features of sentences had been checked.

A novelty of the research consists of the experimental confirmation that 3 additional specific features significantly improve the quality of toxic comments detection. Those three features are: the number of dependences with proper nouns in the singular, the number of dependences that contain bad words, and the number of dependences between personal pronouns and bad words. The selection of 3 specific features allows to significantly reduces the computational complexity of text comment pre-processing, since the calculation of all 20 specific features requires a lot of resources. Accordingly, with 3 specific features added to the typical set, the identification of the toxic comments can be done in real time with good quality.

Acknowledgements. Authors thank Olexandr Yahimovych for extraction the syntactic features from the data set of toxic comments. This research is supported by government scientific project 46-G-388 «Fuzzy logic and computational linguistics based the identification of hidden dependencies in online social networks».

References

1. Salminen, J., et al.: Anatomy of online hate: developing a taxonomy and machine learning models for identifying and classifying hate in online news media. In: Proceeding of the

Twelfth International AAAI Conference on Web and Social Media, 2018, pp. 330-339 (2018).

2. Srivastava, S., Khurana, P., Tewari, V.: Identifying Aggression and Toxicity in Comments using Capsule Network. In: Proceeding of the First Workshop on Trolling, Aggression and Cyberbullying (TRAC-2018), pp. 98-105 (2018).
3. Sood, S.O., Antin, J., Churchill, E.F.: Using Crowdsourcing to Improve Profanity Detection. In: Proceeding of Association for the Advancement of Artificial Intelligence. Spring Symposium: Wisdom of the Crowd, 2012, pp. 69–74 (2012).
4. Mohammad, F.: Is preprocessing of text really worth your time for toxic comment classification? In: Proceeding of International Conference on Artificial Intelligence, 2018, pp. 447-453 (2018).
5. Warner, W., Hirschberg, J.: Detecting hate speech on the world wide web. In Proceedings of the Second Workshop on Language in Social Media. Association for Computational Linguistics, 2012, pp. 19–26 (2012).
6. Bisikalo, O., Yahimovich, A., Yahimovich, Y.: Development of the method for filtering verbal noise while search keywords for the English text. Technology Audit and Production Reserves. 6(2): 33–41 (2018).
7. Toxic Comment Classification Challenge. Available: <https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge>.
8. Elnaggar, A. et al.: Stop Illegal Comments: A Multi-Task Deep Learning Approach. arXiv preprint arXiv:1810.06665 (2018).
9. Brodersen, K.H., Ong, C.S., Stephan, K.E., Buhmann, J.M.: The balanced accuracy and its posterior distribution. In Proceedings of the 20th IEEE International Conference on Pattern Recognition, 2010, pp. 3121-3124 (2010).

Computer-Integrated Design of Cast Parts on the Criterion of Performance on the Example of the Turbine Wheel Turbocharger

Oleg Akimov¹[0000-0001-7583-9976], Vadim Soloshenko²[0000-0002-5444-4382],
Kateryna Kostyk³[0000-0003-4139-9970]

¹National Technical University «Kharkiv Polytechnic Institute», Kyrpychova str., 2, Kharkiv,
61002 Ukraine
olak@kpi.kharkov.ua

²General Electric Company Polska Sp. z o.o. Al. Krakowska 110/114, Warsaw, 02-256 Poland
soloshenko.v@gmail.com

³National Technical University «Kharkiv Polytechnic Institute», Kyrpychova str., 2, Kharkiv,
61002 Ukraine
eklitus@gmail.com

Abstract. In this work the method of computer modeling solved the problem of reducing the efficiency of the turbocharger associated with the mismatch of the geometry of the resulting casting petal turbine technology geometry, calculated gas-dynamic calculations. Computer methods was calculated crystallization of the casting and hardening of the stencil. Computer simulation showed that there is a significant change in the geometry of the turbine blade, which adversely affected the geometry of the resulting casting. Was proposed and solved by the method of computer simulation of "reverse" task of the task of such a geometry of the mold, the solidification of the stencil in water and its deformation led to a (given) geometry of the resulting casting of the turbine of the turbocharger. The solution of this problem by the method of computer modeling allowed to reduce the production defect on the discrepancy of the turbine blade geometry of the turbocharger by 90 %.

Keywords: computer-integrated design, optimization, turbocharger, turbine wheel, engine

1 Introduction

It is widely used turbochargers to force the internal combustion engine by supercharging, to utilize the residual energy of the gases of the exhaust in the cylinders of internal combustion engines. A large number of publications are devoted to the problems of design of turbocharger wheel structures and their optimization by gas-dynamic, strength and dynamic criteria [1–10]. Designs of the Ukrainian turbochargers of the TV series (TKR-7.5 TV, TKR-8.5 TV and their modifications) for diesels of agricultural and transport cars on the design characteristics don't concede to the best foreign

analog. In accordance with the criterion of performance characteristics of cast parts of the engine (details determining the performance of power, fuel consumption, efficiency, etc.) in computer-integrated resource design is necessary to conduct the following studies:

- Investigation of the influence of the geometry of cast parts of the internal combustion engine on their performance;
- Identification of possible deviations of the geometry of the cast parts on the stages of the technological process of their production;
- Simulation of phase transitions in the manufacture of cast parts;
- Investigation of VAT and its influence on the geometry of cast parts.

However, the quality of the real turbochargers in addition to the design stage is largely determined by the quality of the manufacturing stage, which is due to the presence of a certain level of technology and quality management (control) system. At the same time, experienced and commercially produced samples can actually have different quality indicators. Among the production operations that determine the consumer level of the turbocharger as a product, are the achievement of the required roughness of gas and air channels, exact compliance with the design of the real dimensions of the sliding bearing, gaps, passage sections of diffusers and snails, high-quality performance of friction welding operations of the turbocharger and shaft, rotor balancing and others. But, as in the design, are particularly significant technological errors in the manufacture of turbochargers' impellers. Turbine and compressor impellers, which are the most functional and loaded elements of the turbocharger, actually determine the quality of the entire structure. The experience of control studies within the framework of the author's supervision of the quality of manufacture of turbochargers, in particular casting of impellers of JSC "AVTRAMAT" (Kharkiv) and JSC "Kupyansk foundry" (Kupyansk Kharkiv region) show the presence of potential opportunities to improve the quality of mass-produced products.

It is known that "quality" is a set of properties and characteristics of products or services that give them the ability to meet conditioned or anticipated needs. For a turbocharger, these are the efficiency of the working process (estimated efficiency of the turbocharger or a decrease in the specific effective fuel consumption of engines in the entire field of operating modes), and the level of reliability. Manufacturing errors, respectively, can both worsen the working (gas-dynamic) characteristics of the turbocharger, and reduce the margin of safety, leading to an increase in the fuel consumption of the engine, the absence of the planned power level, the breakdown of the impellers. In connection with the absence of the results of studies in the literature on the sensitivity of the design of the impellers to technological errors in this article, an attempt is made to analyze the level of errors in the shaping of the impellers of turbines and to identify ways for their further study and minimization [10–12].

The technology of manufacturing of impellers, in the world of engineering relates to so-called "high technologies", which underlines their complexity and knowledge intensity. Turbine impellers are made by casting on melted models. With a relatively low cost and high productivity characteristic of this method, there are a number of technological problems that make it difficult to obtain impellers with the planned

geometry due to the inevitable errors that occur at different stages of production. Such stages, considering them coinciding with the number of reflections of the profile surfaces, for the turbine wheel five [13–14].

Initially, when developing the technology of mass production of turbine wheels, the main attention was paid to the stability of the thickness field for blades having a complex spatial shape. However, in the transition from analog design methods (followed by experimental refinement) to computer-aided modeling and optimal design methods (for gas-dynamic and strength parameters), it was found that the deviations of the turbine wheel blade profile from the nominal, arising in the manufacturing process and difficult to control, significantly affect the main performance characteristics (efficiency of the turbine and turbocharger as a whole).

2 The influence of technological errors of morphogenesis of turbine wheels on gas-dynamic characteristics of turbochargers

The aim of this work was to solve the problem of reducing the efficiency of the turbocharger, associated with the discrepancy between the geometry of the resulting casting turbine blade technological geometry. The basis for the study in this direction was not quite satisfactory results of gas-dynamic tests of turbine impellers TKR-7.5 TV. It was turned out after the study that it was due to technological deviations of the blade profile from the nominal. Sensitivity of gas-dynamic characteristics (boost level, efficiency), determined from the analysis of its sensitivity to the output angle of the blade β (Fig. 1) proved to be so high for optimal design that 20...40 % of the effect (1...2 % wheel efficiency) of optimization was "eaten" by manufacturing inaccuracy (Fig. 2). It is shown in Fig.2 the optimal design (β_2) is more sensitive to manufacturing errors than the previous parameters (β_1), and the average value of the real value of the angle β (β_3) is usually less than optimal.

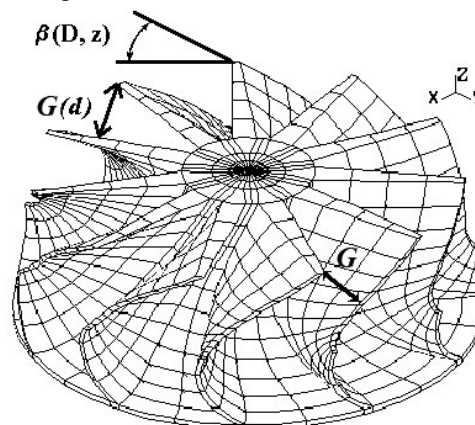


Fig. 1. Controlled parameters of the geometry of the "stencil" turbine TKR-7.5 TV: $G(d)$ - "throat section" of the channel; (D, z) - the angle of the turbine blade at the output

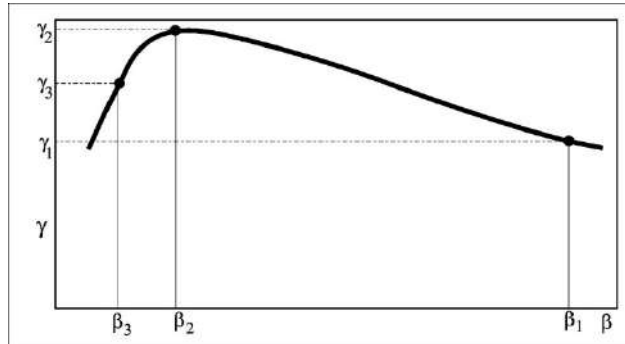


Fig. 2. The dependence efficiency (γ) turbine wheel TKR-7.5 TV from the angle of the scapula (β) output channel: β_1 - before optimization; β_2 - after numerical optimization; β_3 - when making the optimal design

To assess the possibility of changing the geometry of the inter-blade channel due to deformation of the blade profile during manufacture, it was made selective measurements of the "throat section" of the channels G of the turbine wheels (the smallest distance from the surface of the trough at the outlet at a given diameter of one blade to the surface of the back of the adjacent blade, Fig. 1).

For the turbine wheel TKR-7.5 TV measurement "throat section" on the peripheral diameter D_{2n} (71 mm) five stencils gave a stable size of 8.8 mm with a maximum deviation of 0.2 mm (Table 1). The outer diameter of these stencils $D_1 = 79.1$ mm with a nominal size on the mold $D_1 = 80$ mm, $D_{2n} = 72$ mm. if you insert the displacer into the stencil channel, the gap is about 1 mm between the stencil and the form of the displacer. The planned size of the "throat section" (G) at $d = 71$ mm was 9.6 mm. Reducing the size of the "throat section" (G) to 7.1...8.8 mm corresponds to a decrease in the angle of inclination of the guide profile β by $d = 71$ mm at $4.5^\circ \pm 1^\circ$.

Table 1. Measurements of "throat sections" G (d=71mm), mm blades five wax models turbine wheel TKR-7.5 TV

shoulder number										
1	2	3	4	5	6	7	8	9	10	11
8.85	8.85	8.85	8.85	8.7	8.9	8.8	8.8	8.95	8.7	8.9
8.8	8.7	8.9	8.75	8.95	8.85	8.7	8.75	8.7	8.8	8.8
8.75	8.85	8.85	8.75	8.95	8.75	8.75	8.75	8.8	8.8	8.8
8.7	8.75	8.8	8.5	8.9	8.7	8.6	9.0	8.7	8.8	8.8
8.7	8.7	8.65	8.9	8.8	8.7	8.7	8.7	8.65	8.75	8.7

In the study of casting turbine wheel TKR-8.5 found the following. The plastic displacer did not reach the end of the casting channel by about 4 mm. (for different castings: 4.2, 4.5, 4.3, 3.9, 4.4, 4.6, 3.9 mm). It is noted that the penetration of the displacer into the channel (which must completely repeat the shape of the displacer) is prevented by premature contact of the surfaces at the turbine outlet, which character-

izes the decrease in the angle of inclination of the guide profile of the metal casting. A similar operation carried out for the stencil showed that the displacer goes deeper into the stencil (gap -0.7... 1.5 mm). Despite the fact that the stencil seems to be closer to the planned geometry, it is in it that the main part of the final error is laid.

Control measurements of G on the nominal diameter $D_{2n} = 81$ mm in 8 impellers of turbines TKR-8.5 TV, passed the working tests, showed instead of the nominal size of 10.6 mm average 10.0 mm with a deviation of 0.2 mm. One of the wheels had a change in the size of the "throat section" in the range from 9.0 to 11.0 mm, including 10.5 mm. It showed the worst performance. The other wheels had a lower efficiency of 1...2 % due to errors in wheel geometry.

3 Analysis of forming errors by technological stages

For the serial production technology of impellers, the analysis of technological stages was carried out for the appearance of the error of shaping (in particular, the size of the "throat section"). Thus, tasks were set:

- Statistical evaluation of errors introduced at each stage of manufacturing with the allocation of random and natural components;
- Research of the factors determining natural changes of the geometrical dimensions, with the purpose of forecasting, stabilization, and total technological change in the preparation of source tables for CNC machines in the manufacture of the original masterform for displacers;
- Research on determinants of random scatter geometry for a technologically justified to minimize it;
- Development of a rational system of control operations for the technological process of manufacturing wheels.

Step 1. Production of the master form (mold) for molding of the displacer by milling of three surfaces: a back, a trough and the closing surface of a trough on axisymmetric preparations on the CNC machine. Milling is carried out by successive passes along the cutter forming in the end plane of the wheel with a step along the length of the master.001 m. Then surface polished. The error of this stage is small and is mostly random.

Step 2. Production of a mold is made in two operations in the following order: displacers are cast, and then collected in a mold.

In the manufacture of plastic displacers is poured into the master form. The errors of this operations are associated with insufficient accuracy of positioning of the channel displacer and its platform - a conical prism with an angle equal to $360/N$. Manufactured displacers after manual processing are assembled into a mold. This is one of the most important operations. It mates the surface of the trough on the molded displacers N times. Even with a gap between its of about 0.001 m (see *Step 1*), a $0.1 \cdot N$ mm error is accumulated on the last displacer, which is eliminated by manual fitting. This operation (*Step 2*) is defined as the main source of random errors.

At the *Step 3*, wax models of the wheel (stencils) are made. The introduced errors are associated with the shrinkage of the wax model and the spatial deformation of the wax of the impellers due to uneven heat removal from the blades and the hub. The key is the low level of thermal conductivity of the wax model in comparison with the metal casting, technologically determined high cooling rate of the wax (within 5 minutes by immersion in water), low modulus of elasticity of the wax.

This *Step* is estimated as the main source of systematic errors affecting the reduction of linear dimensions and the occurrence of angular deformations (decrease in the angle of flow out of the wheel).

At the *Step 4*, the gating ceramic form of the turbine wheel is made by the method of multiple coating of the wax model (10-12 times) with a special suspension with its subsequent drying. Insertion error associated with the thermal expansion of ceramic forms by pouring of the melt.

During the last *Step*, the heat-resistant alloy is cast into a mold. The introduced errors at this stage can be considered insignificant and compensating errors of the previous stage (due to the expansion of the ceramic form).

Due to the presence of sources of errors of systematic and random nature in the production of the task, first, to stabilize the errors that are natural (by maintaining the mode of casting stencils), calculate and take them into account in the preparation of data for control programs for CNC machines. Secondly, to minimize the errors of random nature, arising mainly at the stage of mold Assembly by increasing the accuracy of the previous stages. Thirdly, without a certain level of production culture, which includes, in addition to high technological discipline, also the control of all components of quality, it is impossible to improve the quality of turbines as a whole. In this regard, it was proposed to make in the technological process of manufacturing castings of the wheels of the operation to control the value of "throat section" of the channel.

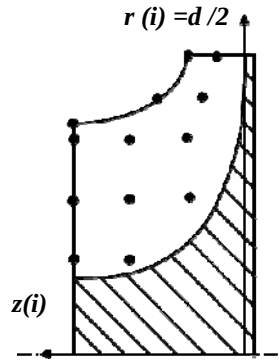
4 Stability of the field thickness of turbine blades of turbochargers

The attention to the stability of the blade thickness field was paid from the very beginning of the development of the technology of serial production of turbine impellers. In order to increase the stability of the blade thickness field, the technology of mold manufacturing was improved, in particular, the separate production of mold displacers by milling was replaced by casting into the master mold. Table 2 presents the results of statistical processing of thickness measurements $T(i)$ at 13 points of 36 turbine blades TKR-11, obtained by successive grinding of the impeller - the average thickness $M[T(i)]$ and the standard deviation $\sigma[T(i)]$. The location of the points for which thickness measurements were carried out is shown on the cylindrical projection of the blade (Fig. 3). As can be seen from the table. 2 the average thickness at some points reaches 19 %.

Table 2. Statistical characteristics of thickness $T(i)$, mm TKR-11

i node	$z(i)$, mm	$d(i)$, mm	$M[T(i)]$ mm	$\sigma[T(i)]$	(σ/M) 100 %
1	12.5	110	1.24	0.139	11
2	6	110	1.70	0.089	5.2
3	16.6	94	1.64	0.211	13
4	8	94	2.19	0.103	4.7
5	26	85	0.86	0.128	15
6	26	78	1.05	0.200	19
7	18	78	1.77	0.159	9
8	8	78	2.56	0.148	5.8
9	26	62	1.40	0.120	8.6
10	18	62	2.24	0.128	5.7
11	8	62	2.97	0.117	4.0
12	26	46	1.72	0.168	9.8
13	18	46	2.42	0.166	7.0

Due to the complexity of the manufacturing technology of molds for smelted models, the level of heterogeneity at similar points of the blades was reduced. So, on control measurements of products of JSC AVTRAMAT this value did not exceed $0.15 \cdot 10^{-3}$ that corresponds to requirements of specifications.

**Fig. 3.** Cylindrical projection TKR-11

5 The influence of the "variable thickness" of the blades on the static strength of impellers of turbines

The characteristic distributions of the stress intensity σ_i of turbine wheels before optimization of the blade profile geometry according to the joint criteria of gas dynamics and strength are given on the Fig.4, *a* and after (Fig.4, *b*) in the field of centrifugal forces. Usually, after the optimization stage, the average stress level increases, and the

resulting project is more sensitive to the technological errors of geometry. The static strength margin of the optimized turbo wheels is about 2.0.

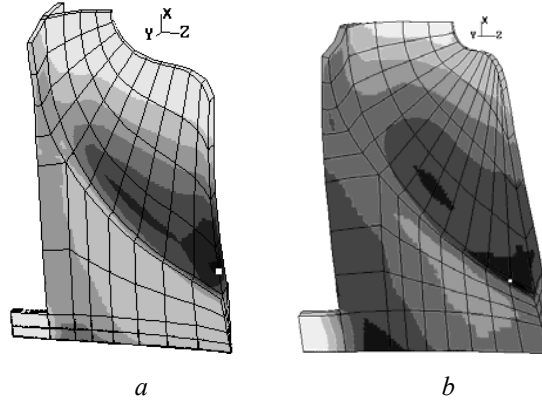


Fig. 4. The characteristic shape of the intensity distribution of the stresses σ_i wheels of turbines in the field of centrifugal forces: *a*-before optimization (TKR-11); *b* - after optimization by gas-dynamic characteristics (TKR-8.5 TV).

For finding the ultimate and the statistical estimates of scatter σ_i^m turbine wheel used the theory of sensitivity analysis applied to the finite element model of the wheel on the basis of three-dimensional ISO-parametric 20-node finite elements (FE). On the Fig. 5 it is given the pattern of sensitivity analysis maximum stress intensity to the normal increments of the nodes of the grid FE $\partial(\sigma_i^{max})/\partial(\vec{n})$ for the turbine wheel TKR-8.5 TV.

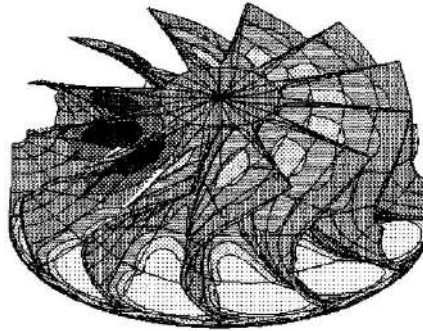


Fig. 5. $\partial(\sigma_i^{max})/\partial(\vec{n})$ for turbine wheel TKR-8.5 TV

The picture of sensitivity analysis σ_i^{max} to the thickness of the turbine TKR-8.5, and the dispersion is acceptable and the actual thickness variation across the surface of the blades and hub is calculated standard deviation σ_i^{max} , equal to 1.6 % of $\sigma_i^{max}_{nom}$, which indicates a relatively low effect of different thicknesses on the maximum stresses in the field of centrifugal forces. Qualitatively from Fig. 5 it can be seen that

the area at the root of the blade at the outlet of the flow for the back and trough has a different sign sensitivity, that is, the greatest influence on the σ_i^{max} is the movement of the middle surface of the blade, and not the change in thickness. For the blade area along the outer meridional circumference opposite - σ_i^{max} grows with increasing thickness. The reduction in the thickness of the interscapular region of the disk leads to a decrease in σ_i^{max} .

Taking into account the errors of the turbine wheel shaping practically does not change the magnitude and nature of the temperature effect from the exhaust gases in the cylinders.

6 Influence of technological errors on the characteristics of dynamic strength of turbine impellers

Real "variable thickness" of the blades also results in frequency detuning of the blades is cyclically symmetric structures (CSS) of the wheel. This affects the stratification of the earlier times the natural frequency (NF) oscillation of the blade crown, the earlier the harmonic distortion in the circumferential direction (for similar points) of their own forms (OF) of the oscillation, causes a redistribution of energy in forced oscillations between the shoulder blades and, ultimately, leads to the spread of the resonance voltages, the increase of dynamic loads of the blade row in general. Given the known nature of the fatigue strength curve, a 50 % increase in dynamic stress (σ_i^{max}) can reduce the life of the product by 3 to 5 times.

Wheel continuity and high NF cause relatively low levels of damping, which increases the risk of resonant stresses. In addition, the structure of the midrange spectrum is (Table. 3) that at the bottom is a package of "blade" oscillation frequencies generated by the first OF vibrations of the isolated blade. The relative density of this package is commensurate with the magnitude of the frequency disorder of the isolated blade, which characterizes the most dangerous, in the sense of the possibility of maximum overload, the region of the ratio of the parameters.

Two statements of a problem - extreme and statistical deserve attention. The first can be formulated as follows: "As far as possible to increase the resonant voltage (this value is called "overload" β) in the worst location of the different frequency of the blades, and feasible if the values for characteristic errors of morphogenesis serial manufacturing technology? Regardless of the nature of the relationship of the blades as subsystems of the CSS (elastic, inertial, aeroelastic, etc.) calculated maximum overload value ($\max\beta$) 11-blade wheel for resonance at multiples of earlier NF $\max\beta^m = 1.67$ (for non-multiple NF - $\max\beta^{nm} = 2.16$).

Based on the specifics of the spectrum of OF wheels (bottom package OF wheels from the first OF of the scapula, areas of high sensitivity fluctuations of the package located on the blades - Fig. 6, the density factor of the OF packet is more important than the exact frequency matching) next, a discrete model of the blade crown was used, (Fig. 7) with an elastic coupling of subsystems corresponding to the wheels of turbines TKR-8.5 TV in value and width of the "blade" midrange package. In this paper, for such a model, the feasibility assessment is carried out (Table. 4)worst fre-

quency detuning. As shown below, all frequency detuning values are in the range of possible frequency deviations of the serial technology used.

Table 3. Eigenfrequencies ω_i of blades, disk and turbocharger TKR-8.5 TV, Hz

	Insulated impeller blade			Wheel hub (without blades)		
i	1	2	3	1	2	3
ω_i	7242	15032	19292	9950*	1005	10875*

No.	The blade wheel of the turbine (blade package OF), * - multiple frequencies					
i	1, 2	3, 4	5, 6	7, 8	9, 10	11
p_i	6075*	6126*	6135*	6140*	6145*	6365
m	1	2	3	4	5	0

	Disk OF		Blade package of SF on the basis of the 2nd SF of the scapula					
i	12	13, 14	15, 16	17, 18	19, 20	21, 22	23, 24	25
p_i	10490	10690	12000	12190	12285	12333	12355	12990
m	0	1	2	3	4	5	6	0

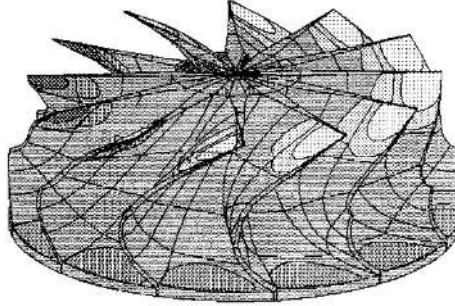


Fig. 6. The sensitivity analysis of common-mode NF $\partial(p_{13})/\partial(\vec{n})$ of the blade NF package (turbine TKR-8.5 TV)

The frequency separation of the blades for the entire wheel casting technology is random and does not allow selective Assembly (unlike axial turbomachines). In the statistical formulation of the problem on the statistical characteristics of the deviations of the geometry (or different frequencies) of the blades (usually the normal distribution hypothesis is used), the density of the probability of overload at resonant oscillations is investigated. To answer the question whether the increase in the accuracy of

wheel shaping (for example, due to the development of new technologies) will lead to a decrease in the overload in the statistical formulation, it is necessary to build a dependence of the stochastic characteristics of the overload on the initial characteristics (dispersion of the different frequencies of the blades).

Table 4. The worst (even) detune $PR_i = (\Delta p_i^2 / p_{nom}^2) / R^2$, causing maximum overload β^{max} at resonance k-th harmonic excitation ($p_{nom} = 7.5$ KHz, viscous friction $n = 0.1$, $R^2 = 0.01$, $N = 11$)

K	PR_i (i - shoulder number)						β
	$i=1$	$i=2, 11$	$i=3, 10$	$i=4, 9$	$i=5, 8$	$i=6, 7$	
$k=0$	-1,042	3,7994	0,5781	0,3419	0,6591	0,4330	2,1573
$k=1$	-,3683	3,6298	0,7651	0,6905	0,8534	0,8005	1,6725
$k=2$	-,0217	5,9146	0,5371	0,5371	0,6255	0,5366	1,6723
$k=3$	-,0551	-12,363	-0,2113	-0,2464	-0,1550	-0,2646	1,6715
$k=4$	-,0018	-3,8313	0,2001	-0,9743	-0,8505	-0,7502	1,6699
$k=5$	1,7835	-2,0662	0,4598	0,4121	0,5366	-0,3232	1,6689

Experimentally (using a frequency analyzer of the firm "Bruel and Kjaer" for isolated blades) found that the magnitude scatter of the fundamental tone of real blades (6700 Hz) is found for the investigated wheels turbine TKR-8.5 TV in the range up to 300-500 Hz c dispersion $\sigma[\Delta p_i] \cong 100$ Hz. We will use the "detuning parameter" (PR) [15–16], defined for the wheel as the ratio of the square of the maximum relative frequency detuning of the blade to the density of the simulated eigenfrequency package ("generated" by the first eigenform of the isolated blade oscillations):

$$PR = \max_{i=1, N} \{ (p_i^2 - p_{nom}^2) / p_{nom}^2 \} / R^2 \quad (1)$$

where $R^2 = [(f_N - f_1) / f_1] / 4$ is a quarter of the relative width of the eigenvalue package (approximately the square of half the relative width of the NF package), determined in the model by the value of weak elastic bonds. From table. 3 it can be seen that the blade package $R^2 \cong 2.44$ % of 2.44 %. Thus, the applied serial technology is characterized by the level of the detuning parameter $\sigma[PR] \cong 1.2$ ($PR_{max} \cong 5$).

The "analytical solution" of the problem in the statistical formulation within the framework of linear relations of the sensitivity analysis overload to the frequency disorder of the blades and the application of statistical theorems on linear operations over normally distributed random variables (methods of statistical dynamics) is limited by the scope of applicability of the first approximations of the classical (system with strong coupling) or modified (system with very weak coupling) perturbation methods. However, for the studied levels of parameter mismatch ($PR \cong 0.5-5$) is characterized by essentially nonlinear dependence of the system loading on the magnitude of the frequency detuning of blades. This does not allow the use of linear (or quadratic) relations in the description of the dynamic characteristics of the upset systems. Therefore, to find the probabilistic characteristics of the overload is justified by the use of methods of mathematical statistics, where as a sample of observations of a random variable is a set of overloads of the system, calculated by numerical genera-

tion within the normal distribution law. In this paper, the method of numerical experiment was used for the discrete model of Fig. 8. Different frequency of blades was considered to be non-randomized and distributed according to the normal law with the dispersion $\sigma(PR_i)$.

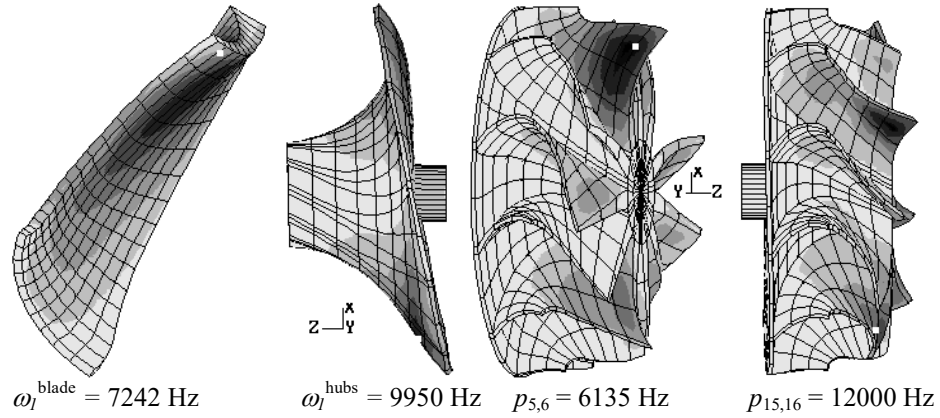


Fig. 7. The natural frequencies and mode shapes of the turbine TKR-8.5 TV

As the random misalignment with a standard subroutine for generating random numbers with a constant law of distribution in the open area $]0,1[$ with the initialization at the current time, which according to the known dependences were converted to normal distribution of the different incidence of the blade with the given mathematical expectation (ME) and standard deviation (SD, σ). The sample of a random variable of different frequencies for each individual blade was previously checked for compliance with the normal distribution law according to the criterion χ^2 by the significance level $\alpha = 75\%$. Then the problem of forced oscillations was solved for 1000 types of frustrated crowns obtained in this way, statistical characteristics $\beta[\sigma(\alpha f_i)]$ or $\beta[\sigma(PR)]$ were calculated. The dependence of $ME\{\beta[\sigma(PR)]\}$ was constructed similarly by 25 values of $\sigma^*(\alpha f_i)$ for a given excitation harmonic (Fig. 8).

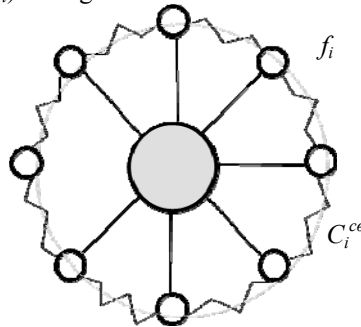


Fig. 8. The discrete model of a quasi-periodic structure consisting of a set of elastically coupled subsystems

According to this method, the 11-blade turbine wheel TKR-8.5 TV ($N = 11$) was studied. For Fig. 9 shows the dependence of ME overload ($ME\{\beta[\sigma(PR)]\}$) from the value of SD "parameter mismatch" ($\sigma(PR)$) at resonance with the 5 (6) harmonic excitation, and the interval $ME(\beta) + \sigma(\beta)^{left}$, $ME(\beta) + 3 \cdot \sigma(\beta)^{left}$, $ME(\beta) + \sigma(\beta)^{right}$, $ME(\beta) + 3 \cdot \sigma(\beta)^{right}$, the limit of rating overload $\beta^{max} = 1.67$, $\beta^{min} = 1.0$. For $\sigma(PR)^{serial} \cong 1.2$ at resonance on the excitation harmonic $k = 5$ or $k = 6$: $ME(\beta) = 1.44$, $\sigma(\beta) = 0.1$, the coefficient of asymmetry – 0.85, coefficient of kurtosis 0.4.

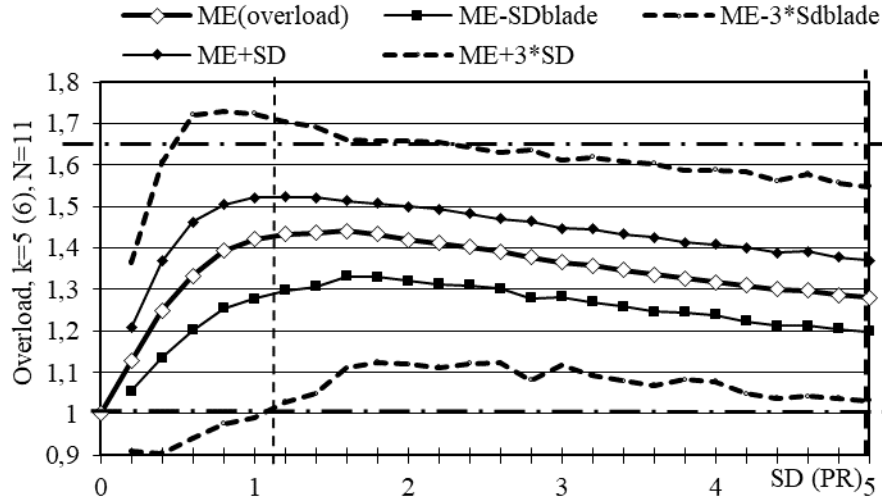


Fig. 9. Dependence of ME overload of $ME\{\beta[\sigma(PR)]\}$ on the value of the SD of the "detuning parameter" at resonance with 5 (6) excitation harmonic, $N = 11$. Dotted vertical line - $SD(PR) = 1.2$ a serial manufacturing technology turbowheels, dashed "bold" - max level ($PR_{max} \cong 5$), the dot-dash horizontal limit of rating overload ($\beta_{max} = 1.67$)

7 Conclusion

It is obvious that the accuracy of the real optimal resource design of the structure should correspond to the accuracy of its manufacture. Casting technology of manufacturing turbine wheels and compressors, having undoubted advantages (low cost and high performance), requires consideration and modeling of the errors of forming a natural character at the stage of formation (phase transition and cooling) wax model (stencil). In comparison with steel casting, the wax has more significant temperature gradients after the phase transition (due to the significantly lower thermal conductivity of the wax mixture compared to steel) and is much more prone to deformation when thermal stresses occur due to the lower (compared to steel) modulus of elasticity. With a certain degree of error, it is possible to analyze the movements of the stencil by determining the temperature field after the phase transition (assuming at this point in time the wax shape of the wheel is perfectly accurate and unstrained) and

solving the thermoelastic cooling problem. This solution will also allow to establish zones of possible residual stresses.

To stabilize and reduce these effects, it is necessary to improve the technological mode of the phase transition of the stencil (fixing and increasing the time), improve the wax composition (increasing the modulus of elasticity), pre-registration of the deformation of the stencil in the manufacture of the mold of the displacer or the development of technology that uses not wax models, but, for example, Quick-Cast and 3D-Prototyping technologies.

It is shown that the most dangerous consequence of technological deviations of the casting geometry of turbine impellers is the possibility of growth of resonant dynamic stresses. For serial manufacturing technology, estimates (for a discrete system with the corresponding partial frequency of the subsystem and the spectrum density) of the maximum realized overload ($\max \beta^{kr} = 1.67$) and statistical characteristics of the overload ($ME(\beta) = 1.3$, $\sigma(\beta) = 0.125$). The conclusion is made about the rationality of further reduction of different thicknesses in order to reduce possible overloads at resonant oscillations ($M^{serial}(\beta)$ are positioned relative to $M^{max}(\beta) [\sigma(\alpha fi)]$) due to improvement of technology of manufacture of castings turbochargers' wheels.

Thus, computer simulation showed that there is a significant change in the geometry of the turbine blade, which adversely affected the geometry of the resulting casting. Was proposed and solved by the method of computer simulation of "reverse" task of the task of such a geometry of the mold, the solidification of the stencil in water and its deformation led to a (given) geometry of the resulting casting of the turbine of the turbocharger. The solution of this problem by the method of computer modeling allowed to reduce the production defect on the discrepancy of the turbine blade geometry of the turbocharger by 90 %.

References

1. Wang, Z., Chen, H.: Combustion System Development of GDI-T Engine for China Stage III Fuel Consumption Regulation. Society of Automotive Engineers (SAE)-China Congress, pp. 13-33. Springer, Singapore (2017). doi: 10.1007/978-981-10-8506-2_2
2. Salnikov, A., Danilov M.: A centrifugal compressor impeller: A multidisciplinary optimization to improve its mass, strength, and gas-dynamic characteristics. ASME Turbo Expo 2017: Turbomachinery Technical Conference and Exposition, 26–30 June 2017. American Society of Mechanical Engineers, pp. V02CT47A013-V02CT47A013 (2017) doi: 10.1115/GT2017-64123
3. Chan, S., Liu, H.: Mass-based design and optimization of wave rotors for gas turbine engine enhancement. Shock Waves 27.2, pp. 313-324 (2017). doi: 10.1007/s00193-016-0646-z
4. Shouren, W., Peiquan G., Liying Y.: Centrifugal precision cast TiAl turbocharger wheel using ceramic mold. Journal of Materials Processing Technology, vol. 204. (1-3), pp. 492-497 (2008). doi: 10.1016/j.jmatprotec.2008.01.062

5. Kler, A.M., Zakharov, Y.B.: Optimizing parameters of GTU cycle and design values of air-gas channel in a gas turbine with cooled nozzle and rotor blades. *Thermophys. Aeromech.*, vol. 19 (3), pp. 403-413 (2012). doi: 10.1134/S0869864312030067
6. Kler, A., Zakharov, Y.: Joint optimization of power plant cycle parameters and gas turbine flow path parameters with blade airfoils represented by cubic splines. *Energy*, vol. 137, pp. 183-192 (2017). doi: 10.1016/j.energy.2017.07.020
7. Kolmakova, D., Popov, G., Shklovets, A., Ermakov, A.: Rational Design of Gas Turbine Engine Compressor to Provide the Required Dynamic Strength Level of Rotor Blades. *ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition*, June 2016. American Society of Mechanical Engineers, pp. V07AT32A015-V07AT32A015 (2016). doi: 10.1115/GT2016-56991
8. Korn, A., Huurdeman, B., Bühl, H.: Reducing fuel consumption and CO₂ emissions by optimizing active intake systems for naturally aspirated and turbocharged engines. *Internationaler Motorenkongress 2014*, pp. 169-191. Springer Vieweg, Wiesbaden (2014). doi.org/10.1007/978-3-658-05016-0_11
9. Hiereth, H., Prenninger, P.: Exhaust gas turbocharging. *Charging the Internal Combustion Engine*, pp. 60-104 (2007). doi: 10.1007/978-3-211-47113-5_5
10. Weber, C.: Modelling products and product development based on characteristics and properties. *An Anthology of Theories and Models of Design*, pp. 327-352. Springer, London (2014). doi: 10.1007/978-1-4471-6338-1_16
11. Zhuang, J., Qiao, X., Bai, J., Hu, Z.: Effect of diesel from direct coal liquefaction–biodiesel blends on combustion, performance and emission characteristics of a turbocharged di diesel engine. *Fuel Processing Technology*, vol. 123, pp. 82-91 (2014). doi: 10.1016/j.fuproc.2014.01.029
12. Murthy, D. N. P., Blischke, W. R.: Products and Product Quality. *Warranty Management and Product Manufacture*, pp. 15-34 (2006). doi: 10.1007/1-84628-258-6_2
13. Wang, D., Dong, A., Zhu, G., Shu, D., Sun, J., Li, F., Sun, B.: Rapid casting of complex impeller based on 3D printing wax pattern and simulation optimization. *The International Journal of Advanced Manufacturing Technology*, pp. 1-7. (2018). doi: 10.1007/s00170-018-2736-9
14. Yamamoto, T., Suzuki, A., Komarov, S. V., Ishiwata, Y.: Investigation of impeller design and flow structures in mechanical stirring of molten aluminum. *Journal of Materials Processing Technology*, vol. 261, pp. 164-172 (2018). doi.org/10.1016/j.jmatprotec.2018.06.012
15. Grigor'ev, V. A., Kalabukhov, D. S., Rad'ko V. M.: Application of neural network approximation methods in the generalization and presentation of the aircraft gas turbine engine turbomachinery characteristics. *Russian Aeronautics (Iz VUZ)*, vol. 58 (1), pp. 48-53 (2015). doi: 10.3103/S1068799815010080
16. Xu, J., Wang, Y., Xu, L.: PHM-oriented sensor optimization selection based on multiobjective model for aircraft engines. *IEEE Sensors Journal*, vol. 15 (9), pp. 4836-4844 (2015). doi: 10.1109/JSEN.2015.2430361

Interrelation Between the Class of Bent-Sequences and the Class of Perfect Binary Arrays

Artem Sokolov^[0000-0003-0283-7229]

Odesa National Polytechnic University, Shevchenko ave. 1, Odesa, 65044, Ukraine
radiosquid@gmail.com

Abstract. The paper is devoted to the research of the interrelation between classes of such perfect algebraic constructions as perfect binary arrays and bent-sequences. The algebraic normal form of bent-sequences of length $n=16$ that generate perfect binary arrays of order $N=4$, are presented. The exact number of perfect binary arrays in the full set of bent-sequences of length $n=64$ is found. The lower bound of cardinality of the full class of perfect binary arrays of order $N=8$ is improved.

Keywords: perfect binary array, 2DPACF, bent-sequence, Walsh-Hadamard transform.

1 Introduction and problem statement

Perfect binary arrays (PBA) are an important class of algebraic constructions that have found numerous applications in the tasks of cryptographic information protection [1], the synthesis of error correction codes [2], construction of orthogonal and biorthogonal signal systems, antenna aperture synthesis, as well as in many other applications of science and technology [3].

Nevertheless, despite the numerous applications and a large number of publications devoted to the problems of the synthesis of PBA, in the general case, there are no methods for constructing their full classes for the derived value of the PBA order N . Moreover, today there is not even an accurate estimation of the cardinality of the full class of PBA of practically valuable orders $N > 4$, in particular order $N = 8$. Significant progress in solving the problem of synthesizing the full class of PBA of order $N = 8$ was obtained in [4], in particular, it was found that the cardinality of the PBA class of order $N = 8$ is not less than $J_{8 \times 8} \geq 688128$, while 688128 PBA were constructed using the original constructive method.

Another major class of perfect algebraic constructions is the class of the bent-sequences (the truth tables of bent-functions), which was introduced in [5] and also found their numerous applications in cryptography and coding theory [6]. Methods for the synthesis of a full class of bent-sequences of length $n=16$ (the truth tables of bent-functions of four variables) are described in [7], while constructive methods for

the synthesis of a full class of bent-sequences of length $n = 64$ are proposed in [8, 9]. Recent researches of the PBA class carried out in [7] made it possible to establish that the full class of bent-sequences of length $n = 16$ and cardinality $J_{bent} = 896$ includes the full class of PBA of order $N = 4$ and cardinality $J_{PBA} = 384$.

However, the characteristics of the interrelation between the classes of bent-sequences of length $n = 16$ and PBA of order $N = 4$ remains unspecified. Researches on the interrelation between the class of bent-sequences of length $n = 64$ and PBA of order $N = 8$ are absent in the literature.

The purpose of this paper is to determine the interrelation between the class of bent-sequences of practically significant lengths $n = 16; 64$ and PBA of orders $N = 4; 8$.

2 Basic definitions

We introduce the basic definitions:

Definition 1 [10]. A perfect binary array is a two-dimensional sequence (matrix)

$$H(N) = \|h_{i,j}\|, \quad i, j = 0, 1, \dots, N-1, \quad h_{i,j} \in \{-1, 1\}, \quad (1)$$

having an ideal two-dimensional periodic autocorrelation function (2DPACF), whose elements

$$R(m, \tau) = PACF(m, \tau) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} h_{i,j} h_{i+m, j+\tau} = \begin{cases} N^2, & \text{for } m = \tau = 0; \\ 0, & \text{for any other } m \text{ and } \tau, \end{cases} \quad (2)$$

where $m, \tau = 0, 1, \dots, N-1$, and all indices of elements $h_{i+m, j+\tau}$ are reduced modulo N .

Let us give as an example a perfect binary array of order $N = 4$ as well as its two-dimensional periodic autocorrelation function

$$H = \begin{bmatrix} + & + & + & - \\ + & + & + & - \\ + & + & + & - \\ - & - & - & + \end{bmatrix}, \quad R = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3)$$

where the symbol “+” denotes +1, and the symbol “-” denotes -1, respectively.

Definition 2 [11]. The Walsh-Hadamard transform (WHT) of a vector $F(n)$ in matrix form is defined as

$$W_f(w) = A(n)F(n), \quad (4)$$

where $F(n)$ is the binary sequence of length n , $A(n)$ is the Hadamard matrix of order n , which is constructed in accordance with the following recurrence relation

$$A(n) = \begin{bmatrix} A(n/2) & A(n/2) \\ A(n/2) & -A(n/2) \end{bmatrix}, \quad A(1) = [+]. \quad (5)$$

Definition 3 [7]. A binary sequence $B = [b_0, b_1, \dots, b_i, \dots, b_{n-1}]$ of length n , where $b_i \in \{\pm 1\}$ are the coefficients, $i = 0, 1, \dots, n-1$, $n = 2^k$, $k = 2, 4, 6, 8, \dots$, is called a bent-sequence, if it has a uniform Walsh-Hadamard spectrum $W_B(\omega)$.

3 Interrelation Between the Class of Bent-Sequences of Length $n = 16$ and the Class of Perfect Binary Arrays of Order $N = 4$

Let us consider the currently known methods of classification of PBA and bent-sequences in order to establish the interrelation between these classes of perfect algebraic constructions. The modern approach to the classification of PBA involves the use of the following proposition:

Proposition 1 [3]. Each PBA of order N generates a $E(N)$ -class of equivalent PBA matrices by using the cyclic rows and columns shift and inversion, with the cardinality of the equivalent matrices class

$$J_{E(N)} = 2N^2. \quad (6)$$

Thus, in accordance with **Proposition 1**, for the order of the PBA $N = 4$, the cardinality of each equivalent class is $J_{E(8)} = 2 \cdot 4^2 = 32$, and accordingly, the full set of PBA of cardinality $J_{PBA} = 384$ can be divided into $384/32 = 12$ nonequivalent classes. Representatives of these non-equivalent classes are given in [3].

In [7] it was shown that the full class of bent-sequences of length $n = 16$ includes the full class of PBA of order $N = 4$ in the case of their representation as vectors by successive concatenation of the rows (columns) of the corresponding PBA.

For example, we concatenate the rows of PBA (3), as a result of which we obtain the following sequence and its Walsh-Hadamard transform coefficients in accordance with (4)

$$\begin{aligned} B_1 &= [++++-++-++-+-+], \\ W_{B_1}(w) &= A(16)B_1 = \\ &= [4 \ 4 \ 4 \ -4 \ 4 \ 4 \ 4 \ -4 \ 4 \ 4 \ 4 \ -4 \ -4 \ -4 \ -4 \ 4]. \end{aligned} \quad (7)$$

It is easy to see that the sequence (7) really satisfies the condition of **Definition 3** and is a bent-sequence of length $n = 16$.

A modern approach to the classification of bent-sequences is based on the consideration of affine-equivalent classes. This classification is easiest to make on the basis of the representation of bent-sequences in algebraic normal form.

Definition 4 [11]. The algebraic normal form (ANF) $\varphi(x_1, x_2, \dots, x_k)$ of a sequence T is a polynomial of $k \leq \log_2 n$ variables with coefficients $a_i \in \{0, 1\}$, where the AND operation is used as the multiplication, and the XOR operation is used as the addition operation

$$\varphi(x_1, x_2, \dots, x_k) = \bigoplus_{i=0}^{N-1} a_i X_i^s, \quad (8)$$

where X_i^s are the terms of the ANF polynomial of degree $s = wt\{X\}$; wt is the Hamming's weight.

The coefficients $a_i = \{a_0, a_1, \dots, a_{N-1}\}$ can be found by performing the Reed-Muller transform [11], i.e. by multiplying the original sequence by the Reed-Muller matrix RM_v

$$\{a_i\} = T \cdot RM_v, \quad T = \{a_i\} \cdot RM_v, \quad (9)$$

where the original sequence T is represented above the alphabet $\{0, 1\}$ using a bijective mapping $+1 \leftrightarrow 0, -1 \leftrightarrow 1$, and the Reed-Muller matrix RM_v is determined using the following recurrent rule

$$RM_0 = [1], \quad RM_v = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \otimes RM_{v-1} = \begin{bmatrix} RM_{v-1} & 0 \\ RM_{v-1} & RM_{v-1} \end{bmatrix}, \quad (10)$$

where $\otimes$ is the Kronecker product.

Definition 5 [11]. Terms of ANF of the degree $s = wt\{X\} \leq 1$ are called as affine.

For example, for sequence length $n = 16$ there are the following possible affine terms: $1, x_0, x_1, x_2, x_3$ on the basis of which corresponding affine codewords can be formed.

For example, we can represent as the ANF obtained from the PBA bent-sequence (7)

$$B_1 = x_3 x_4 + x_1 x_2. \quad (11)$$

It is known [8] that the sum of a bent-sequence with an affine function (which is equivalent to adding one or several affine terms to the ANF coefficients sequence) leads to the formation of other bent-sequences. Thus, the full set of bent-sequences of cardinality $J = 896$ can be classified into $896 / 32 = 28$ affine non-equivalent classes,

in each of which it is possible to distinguish a bent-sequence that does not have affine terms.

In this paper, through numerous experiments, the following statement was established:

Proposition 2. Let H_0 to be PBA of order $N = 4$, and T_0 to be the sequence obtained by concatenating its rows (columns). Then the sequences $T_0, T_1, \dots, T_{2^{k+1}-1}$ obtained by adding to the ANF sequences one or several affine terms construct, by line-by-line (column) filling, a set of matrices $H_0, H_1, \dots, H_{2^{k+1}-1}$ that are also PBA.

Note that the **Proposition 2** is valid only for PBA of order $N = 4$, and fully corresponds to **Proposition 1**, in terms of structure. However, **Proposition 2** makes it easy to establish the interconnection between the generating PBA represented as their ANF (which contain affine terms) and the corresponding generating bent-sequences. We present all 28 generating bent-sequences, among which 12 (in bold font) generates affine non-equivalent classes of PBA of cardinality 32 PBA in each one, corresponding to **Proposition 2**

$$\begin{aligned}
b_1 &= x_2 x_3 + x_1 x_4; & b_{15} &= x_3 x_4 + x_1 x_4 + x_1 x_2; \\
b_2 &= x_2 x_3 + x_1 x_4 + x_1 x_2; & b_{16} &= x_3 x_4 + x_1 x_4 + x_1 x_3 + x_1 x_2; \\
b_3 &= x_2 x_3 + x_1 x_4 + x_1 x_3; & b_{17} &= x_3 x_4 + x_2 x_3 + x_1 x_2; \\
b_4 &= x_2 x_3 + x_1 x_4 + x_1 x_3 + x_1 x_2; & b_{18} &= x_3 x_4 + x_2 x_3 + x_1 x_3 + x_1 x_2; \\
b_5 &= x_2 x_4 + x_1 x_3; & b_{19} &= x_3 x_4 + x_2 x_3 + x_1 x_4; \\
b_6 &= x_2 x_4 + x_1 x_3 + x_1 x_2; & b_{20} &= x_3 x_4 + x_2 x_3 + x_1 x_4 + x_1 x_3; \\
b_7 &= x_2 x_4 + x_1 x_4 + x_1 x_3; & b_{21} &= x_3 x_4 + x_2 x_4 + x_1 x_2; \\
b_8 &= x_2 x_4 + x_1 x_4 + x_1 x_3 + x_1 x_2; & b_{22} &= x_3 x_4 + x_2 x_4 + x_1 x_3; \\
b_9 &= x_2 x_4 + x_2 x_3 + x_1 x_3; & b_{23} &= x_3 x_4 + x_2 x_4 + x_1 x_4 + x_1 x_2; \\
b_{10} &= x_2 x_4 + x_2 x_3 + x_1 x_3 + x_1 x_2; & b_{24} &= x_3 x_4 + x_2 x_4 + x_1 x_4 + x_1 x_3; \\
b_{11} &= x_2 x_4 + x_2 x_3 + x_1 x_4; & b_{25} &= x_3 x_4 + x_2 x_4 + x_2 x_3 + x_1 x_2; \\
b_{12} &= x_2 x_4 + x_2 x_3 + x_1 x_4 + x_1 x_2; & b_{26} &= x_3 x_4 + x_2 x_4 + x_2 x_3 + x_1 x_3; \\
b_{13} &= x_3 x_4 + x_1 x_2; & b_{27} &= x_3 x_4 + x_2 x_4 + x_2 x_3 + x_1 x_4; \\
b_{14} &= x_3 x_4 + x_1 x_3 + x_1 x_2; & b_{28} &= x_3 x_4 + x_2 x_4 + x_2 x_3 + x_1 x_4 + x_1 x_3 + x_1 x_2.
\end{aligned} \tag{12}$$

4 Interrelation Between the Class of Bent-Sequences of Length $n = 64$ and the Class of Perfect Binary Arrays of Order $N = 8$

In the general case, the problem of synthesizing a complete class of PBA of order $N = 8$ is computationally complex and has not been solved yet. Significant progress in the construction of PBA classes was made in [4], where a method for synthesizing the PBA class based on the classes of thinned matrices was proposed and the constructions for their reproduction and superposition were found.

The results of [4] are based on the following proposition:

Proposition 3 [3]. The PBA $H_0(N)$ of the arbitrary order N can always be represented as an interleaving ($\cup$) of its thinned matrices

$$\begin{aligned} H_0(N) &= \|h_{i,j}\| = \|a_{i,j}\| \cup \|b_{i,j}\| \cup \|c_{i,j}\| \cup \|d_{i,j}\| = \\ &= A_0(N/2) \cup B_0(N/2) \cup C_0(N/2) \cup D_0(N/2), \end{aligned} \quad (13)$$

where $\|a_{i,j}\| = h_{2i,2j}$, $\|b_{i,j}\| = h_{2i,2j+1}$, $\|c_{i,j}\| = h_{2i+1,2j}$, $\|d_{i,j}\| = h_{2i+1,2j+1}$ are the corresponding thinned matrices, $i, j = 0, 1, \dots, N/2 - 1$, and the indices $h_{i,j}$ vary within the limits $i, j = 0, 1, \dots, N - 1$.

Each PBA can be represented as (13). In the general case, the set of various structures of thinned matrices obtained by thinning the full class of PBA, we denote as

$$\left\{ \{A_i(N/2)\}; \quad \{B_j(N/2)\}; \quad \{C_\mu(N/2)\}; \quad \{D_\nu(N/2)\}; \right\}, \quad (14)$$

$$i = 1, 2, \dots, \Psi_A; \quad j = 1, 2, \dots, \Psi_B; \quad \mu = 1, 2, \dots, \Psi_C; \quad \nu = 1, 2, \dots, \Psi_D$$

where the parameters $\Psi_A, \Psi_B, \Psi_C, \Psi_D$ are the number of different structures (degrees of freedom) of the corresponding thinned matrices A, B, C, D of order $N/2$.

Different matrix structures from (14) can be obtained by using the cyclic shift operations in rows and columns, inversion, transposition, and mirroring of the set of generating matrices.

In [4], such a set of thinned matrices was obtained, the structures of which are presented in Table 1.

Table 1. The known set of thinned matrices of order $N/2 = 8/2$

Thinned matrix	2DPACF	Thinned matrix	2DPACF
$A = \begin{bmatrix} - & + & - & + \\ + & + & + & + \\ - & + & - & + \\ + & + & + & + \end{bmatrix}$	$R_A = \begin{bmatrix} 16 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \\ 16 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$B = \begin{bmatrix} + & + & - & - \\ - & + & + & - \\ + & + & - & - \\ - & + & + & - \end{bmatrix}$	$R_B = \begin{bmatrix} 16 & 0 & -16 & 0 \\ 0 & 0 & 0 & 0 \\ 16 & 0 & -16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
$C = \begin{bmatrix} - & + & - & + \\ + & + & + & + \\ + & - & + & - \\ - & - & - & - \end{bmatrix}$	$R_C = \begin{bmatrix} 16 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \\ -16 & 0 & -16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$D = \begin{bmatrix} + & + & - & - \\ + & + & - & - \\ - & - & + & + \\ - & - & + & + \end{bmatrix}$	$R_D = \begin{bmatrix} 16 & 0 & -16 & 0 \\ 0 & 0 & 0 & 0 \\ -16 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
$A_0 = \begin{bmatrix} - & - & - & - \\ + & + & + & + \\ + & + & + & + \\ + & + & + & + \end{bmatrix}$	$R_{A_0} = \begin{bmatrix} 16 & 16 & 16 & 16 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$E_0 = \begin{bmatrix} - & + & - & + \\ - & + & - & + \\ - & + & - & + \\ + & - & + & - \end{bmatrix}$	$R_{E_0} = \begin{bmatrix} 16 & -16 & 16 & -16 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
$A_1 = \begin{bmatrix} - & + & + & + \\ - & + & + & + \\ - & + & + & + \\ - & + & + & + \end{bmatrix}$	$R_{A_1} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 16 & 0 & 0 & 0 \\ 16 & 0 & 0 & 0 \\ 16 & 0 & 0 & 0 \end{bmatrix}$	$E_1 = \begin{bmatrix} - & - & - & + \\ + & + & + & - \\ - & - & - & + \\ + & + & + & - \end{bmatrix}$	$R_{E_1} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ -16 & 0 & 0 & 0 \\ 16 & 0 & 0 & 0 \\ -16 & 0 & 0 & 0 \end{bmatrix}$
$A_2 = \begin{bmatrix} - & + & + & + \\ + & - & + & + \\ + & + & - & + \\ + & + & + & - \end{bmatrix}$	$R_{A_2} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{bmatrix}$	$E_2 = \begin{bmatrix} - & - & - & + \\ - & + & + & - \\ - & + & + & - \\ + & + & + & - \end{bmatrix}$	$R_{E_2} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & -16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & -16 \end{bmatrix}$

$A_3 = \begin{bmatrix} + & + & + & - \\ + & + & - & + \\ + & - & + & + \\ - & + & + & + \end{bmatrix}$	$R_{A_3} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 0 & 0 & 16 \\ 0 & 0 & 16 & 0 \\ 0 & 16 & 0 & 0 \end{bmatrix}$	$E_3 = \begin{bmatrix} + & + & - & + \\ - & + & - & - \\ - & + & + & + \\ - & - & - & + \end{bmatrix}$	$R_{E_3} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 0 & 0 & -16 \\ 0 & 0 & 16 & 0 \\ 0 & -16 & 0 & 0 \end{bmatrix}$
$A_4 = \begin{bmatrix} - & + & - & + \\ + & + & + & + \\ + & - & + & - \\ + & + & + & + \end{bmatrix}$	$R_{A_4} = \begin{bmatrix} 16 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 16 & 0 & 16 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$E_4 = \begin{bmatrix} - & - & - & - \\ - & + & - & + \\ + & + & + & + \\ - & + & - & + \end{bmatrix}$	$R_{E_4} = \begin{bmatrix} 16 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -16 & 0 & -16 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
$A_5 = \begin{bmatrix} - & + & + & + \\ + & + & - & + \\ - & + & + & + \\ + & + & - & + \end{bmatrix}$	$R_{A_5} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 16 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \end{bmatrix}$	$E_5 = \begin{bmatrix} - & - & - & + \\ + & - & + & + \\ - & - & - & + \\ + & - & + & + \end{bmatrix}$	$R_{E_5} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 0 & -16 & 0 \\ 16 & 0 & 0 & 0 \\ 0 & 0 & -16 & 0 \end{bmatrix}$

As a result, the PBA class of cardinality $J_{PBA8 \times 8} = 688\,128$ was built, which increased the lower bound estimation of the cardinality of PBA class of order $N = 8$ in factor of 7 compared to the previous result obtained in [12].

It is known, that all the representatives of PBA class of the order $N = 4$, when concatenating the PBA rows (columns) generates bent-sequences. The research carried out in this paper allowed us to establish that there are just $J'_{PBA8 \times 8, bent} = 98\,304$ representatives in the PBA class of order $N = 8$ (synthesized in [4]) that forms the bent-sequences of length $n = 64$ when concatenating their rows (columns).

At the same time, the other $688\,128 - 98\,304 = 589\,824$ PBA (when concatenating rows or columns) have non-uniform Walsh-Hadamard transform coefficients (absolute values). In order to classify these spectral coefficients, it is most convenient to use the definition of the elementary structure of the Walsh-Hadamard transform coefficients [13].

Definition 6 [13]. The elementary structure of the vector $W(\omega)$ of Walsh-Hadamard transform coefficients is the set of absolute values of its spectral components.

It was established experimentally that all remaining $589\,824$ PBA, on the basis of which it is impossible to form bent-sequences by applying the operation of concatenation of their rows (columns), according to **Definition 6**, have an elementary structure $\{0(12), 8(48), 16(4)\}$.

This notation of the elementary structure should be understood as follows: the number in front of the parentheses characterizes the absolute value of the Walsh-Hadamard transform coefficient, whereas the number in parentheses indicates how many times it occurs in the vector of the Walsh-Hadamard transform coefficients.

Let us consider, for example, one of these PBA, as well as its 2DPACF

[illegible]

Applying the concatenation of the rows (columns) of PBA (15), we obtain a binary sequence and, according to **Definition 2**, corresponding to it vector of Walsh-Hadamard transform coefficients, which has an elementary structure $\{0(12), 8(48), 16(4)\}$

$$\begin{aligned}
T &= \{-\text{---}+ - + + \text{---} - + \text{---} - + + + \text{---} + + + \text{---} + + + \text{---} + \\
&\quad - \text{---} - + - + + + + + - + - + - + - + - + - + - + - \}; \\
W_T &= \begin{Bmatrix} -8 & -8 & -8 & -8 & 8 & 8 & -8 & 8 & -8 & -8 & -8 & -8 & 8 & -8 & 8 \\
& -8 & 8 & 8 & 8 & -8 & 8 & 8 & 0 & -8 & 8 & 8 & 8 & 8 & 8 & 8 \\
& & -8 & 8 & -8 & 8 & 8 & 0 & 0 & 0 & 8 & -8 & 8 & -8 & 0 & 0 \\
& & & -8 & 8 & 8 & -8 & 0 & 0 & 16 & 16 & 8 & -8 & -8 & 8 & 0 & 0 & -16 & -16 \end{Bmatrix}.
\end{aligned} \tag{16}$$

The research performed in this paper shows that establishing the interrelation between the class of PBA of order $N = 8$ and the class of bent-sequences of length $n = 64$ can significantly increase the lower bound of the number of PBA due to their new structures, that exist in the class of bent-sequences.

Note that, in the general case, the problem of synthesizing bent-sequences of length $n = 64$ is a complex computational problem, coupled with the enumeration of a set of $J = 2^{64} = 18446744073709551616$ elements. Nevertheless, the theory of bent-squares that was proposed in [14], with the help of which it was possible to synthesize the full set of bent-sequences of length $n = 64$ which have the cardinality $J_{bent64} = 5\,425\,430\,528$.

We established that within this set there is a set of PBA of cardinality $J_{sdr-bent64} = 2\,326\,528$ in which, of course, we found as a subset of $J'_{PBA8 \times 8, bent} = 98\,304$ PBA, that was constructed in [4].

As an example, we present one of the PBA and it's 2DPACF, that was found in the bent-sequences of the length $n = 64$ class and is not member of set of PBA synthesized in [4]

[illegible]

We concatenate the PBA (17) rows, as a result of which we obtain the following sequence and its Walsh-Hadamard transform coefficients in accordance with (4)

$$\begin{aligned}
B = & [+++++-----+-+--+--+--+--+--+ \\
& +---+-+---+---+---+---+---+---+]; \\
W_B = & \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & -8 & 8 & 8 & 8 & 8 & -8 & -8 & -8 \\
8 & 8 & -8 & -8 & 8 & 8 & -8 & 8 & 8 & 8 & -8 & -8 & 8 & -8 & -8 \\
8 & -8 & 8 & -8 & 8 & 8 & -8 & 8 & 8 & -8 & -8 & -8 & 8 & -8 & 8 \\
8 & -8 & -8 & 8 & -8 & -8 & -8 & 8 & 8 & -8 & -8 & -8 & 8 & 8 & 8 \end{bmatrix},
\end{aligned} \tag{18}$$

which proves that (18) is indeed a bent-sequence.

Note that PBA (17) consists of thinned matrices presented in Table. 2. These structures of thinned matrices differ from the matrices presented in Table 1. This fact shows that there exist other structures of thinned matrices that differ from those found in [4].

Table 2. Thinned matrices of PBA (17)

Thinned matrix	2DPACF	Thinned matrix	2DPACF
$A' = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & - & + & - \\ + & + & + & + \end{bmatrix}$	$R_{A'} = \begin{bmatrix} 16 & 0 & 16 & 0 \\ 8 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 \\ 8 & 0 & 8 & 0 \end{bmatrix}$	$B' = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ - & + & - & + \\ - & - & - & - \end{bmatrix}$	$R_{B'} = \begin{bmatrix} 16 & 0 & 16 & 0 \\ -8 & 0 & -8 & 0 \\ 0 & 0 & 0 & 0 \\ -8 & 0 & -8 & 0 \end{bmatrix}$
$C' = \begin{bmatrix} + & + & - & - \\ + & + & - & - \\ + & - & - & + \\ - & + & + & - \end{bmatrix}$	$R_{C'} = \begin{bmatrix} 16 & 0 & -16 & 0 \\ 0 & -8 & 0 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & -8 \end{bmatrix}$	$D' = \begin{bmatrix} + & + & - & - \\ - & - & + & + \\ + & - & - & + \\ + & - & - & + \end{bmatrix}$	$R_{D'} = \begin{bmatrix} 16 & 0 & -16 & 0 \\ 0 & 8 & 0 & -8 \\ 0 & 0 & 0 & 0 \\ 0 & -8 & 0 & 8 \end{bmatrix}$

Thus, the discovering of the interrelation between the class of PBA of order $N = 8$ and bent-sequences of length $n = 64$ allows us to improve the estimation of the lower bound of the cardinality of PBA class of order $N = 8$. Summarizing, it was established that in [4], the PBA class that produces the bent-sequences has cardinality $J'_{PBA8 \times 8, bent} = 98\,304$, as well as PBA class that does not produce bent-sequences by concatenating rows (columns) has cardinality $J'_{PBA8 \times 8, nonbent} = 589\,824$. In this paper, it is clarified that the full class of bent-sequences of length $n = 64$ includes the PBA class of cardinality $J_{PBA8 \times 8, bent} = 2\,326\,528$. Thereby, the cardinality of class of all the known PBA is

$$J_{PBA8 \times 8} \geq 589\,824 + 2\,326\,528 = 2\,916\,352, \tag{19}$$

which is larger by a factor of ~ 4.2 compared to the estimation in [4].

5 Conclusion

We note the main results obtained in the paper:

1. The lower bound estimation for the cardinality of the class of the PBA of order $N = 8$ is improved. In particular, it has been established that the cardinality of the

PBA class of order $N = 8$ is not less than $J_{PBA8 \times 8} \geq 589\,824 + 2\,326\,528 = 2\,916\,352$ that is a factor of ~ 4.2 greater than the known estimation.

2. The total cardinality of the PBA class of order $N = 8$, which are producing the bent-sequences of length $n = 64$ by concatenating the rows (columns), is established and equal to $J_{PBA8 \times 8, bent} = 2\,326\,528$. It is shown that the existence of new structures of thinned matrices, which differ from the previously known ones, are possible.
3. The interrelation between the PBA class of order $N = 4$ and bent-sequences of length $n = 16$ is established. In particular, 12 ANF polynomials of bent-sequences that produce the full PBA class of order $N = 4$, are presented.

It should be noted that the search for new structures of thinned matrices, as well as the rules for their interleaving for formal enumeration of the PBA full class that generates bent-sequences, is an actual direction for further research. The number of PBA, which generate sequences with other (different from the bent-sequences) elementary structures of the Walsh-Hadamard transform vectors also remains unknown and can be the actual direction for further research.

References

1. Mazurkov M.I., Chechelnytskyi, V.Y., Murr P. Information security method based on perfect binary arrays. *Radioelectronics and Communications Systems*, vol. 51, no. 11, pp. 612-614. doi:10.3103/s0735272708110095 (2008)
2. Baranov P.Y., Mazurkov M.I., Chechelnytskyi V.Y., Yakovenko A.A. Family of two-dimensional correcting codes on a basis of perfect binary array. *Radioelectronics and Communications Systems*, vol. 52, no. 9, pp. 501-506. doi:10.3103/s0735272709090088 (2009)
3. Mazurkov M.I. *Broadband radio communication systems*. Science and Technology, p. 340. (2010)
4. Chechelnytskyi, V. Y. A method for generation of the full class of perfect binary arrays of the order $N = 8 \times 8$. *Radioelectronics and Communications Systems* vol. 48, no. 11 pp. 48-52. doi:10.3103/S0735272705110087 (2005)
5. Rothaus, O.S. On "bent" functions. *J. Comb. Theory Ser. A.*, USA: Academic Press Inc, no. 20(3), pp. 300—305. doi:10.1016/0097-3165(76)90024-8 (1976)
6. Tokareva N. *Bent Functions: Results and Applications to Cryptography*. Academic Press, p. 220. (2015)
7. Mazurkov M.I., Sokolov A.V. The regular rules of constructing the complete class of bent-sequences of length 16. *Proceedings of ONPU*, no. 2(41), pp. 231-237 (2013)
8. Qingshu Meng, Min Yang, Huanguo Zhang, Jingsong Cui. A novel algorithm enumerating bent functions. *Discrete Mathematics*, vol. 308, issue 23, pp. 5576–5584. doi:10.1016/j.disc.2007.10.024 (2008)
9. Sokolov, A.V. Synthesis method of a complete class of bent-functions of six variables. *Problems of physics, mathematics and technics*, no. 4(29), pp. 94-102 (2016)
10. Jedwab J., Mitchell C. Constructing new perfect binary arrays. *Electronics Letters*, vol. 24, no. 11, pp. 650-652. doi:10.1049/el:19880440 (1988)
11. Logachev O.A., Salnikov A.A., Yashchenko V.V. *Boolean Functions in Coding Theory and Cryptography*, Amer Mathematical Society, p. 334 (2012)

12. Mazurkov M. I., Chechelitsky V.Ya. Classes of equivalent and generating perfect binary arrays for CDMA technologies. Proceedings of the universities. Radioelectronics, vol. 46, no. 5, pp. 54–63 (2003)
13. Sokolov A.V., Barabanov N.A. Algorithm for removing the spectral equivalence of component Boolean functions of Nyberg-design S-boxes. Radioelectronics and Communications Systems, vol. 58, no. 5, pp. 220-227. doi:10.3103/s0735272715050040 (2015)
14. Agievich S.V. On the representation of bent functions by bent rectangles. — Probabilistic Methods in Discrete Mathematics: Proceedings of the Fifth International Petrozavodsk Conference (Petrozavodsk, June 1–6, 2000). Utrecht, Boston: VSP, pp. 121—135 (2002)

Development the Heat Leak Detection Method for Hidden Thermal Objects by Means the Information-Measuring Computer System

Iryna Zhuravska¹[0000-0002-8102-9854], Maksym Musyienko²[0000-0001-9228-2233],
Oleksii Tohoiev³[0000-0003-3465-7767]

¹ Petro Mohyla Black Sea National University, 68 Desantnykiv str., 10, Mykolaiv, 54003, Ukraine

iryna.zhuravska@chmnu.edu.ua

² Petro Mohyla Black Sea National University, 68 Desantnykiv str., 10, Mykolaiv, 54003, Ukraine

musienko2001@ukr.net

³ Petro Mohyla Black Sea National University, 68 Desantnykiv str., 10, Mykolaiv, 54003, Ukraine

oleksiitohoiev@gmail.com

Abstract. The purpose of this paper is to develop a thermographic method for detecting the places of heat energy leak on the underground thermal pipes using the computer system. Innovativeness of this work would be to develop an analytical model for constructing a heat map with an account of seasonal changes on separate temperature scales. It's proposed to upload the data that was received from the information-measuring computer system into the cloud storage. Due to the heat map laid on Google Maps, the repair equipment could be precisely allocated in place of the heat leakage displayed on the heat map. The implemented process using a Testo 805i infrared thermometer, which data is transmitting to the mobile devices (smartphone, tablet, etc.) with iOS version 10 (or later) and Android 4.3 (or later). The experiments confirm the correctness of the theoretical assumptions and sufficient accuracy for practical usage while elimination the heat pipes leaks.

Keywords: thermal pipeline, leak detection, pyrometer, GPS coordinates, information-measuring computer system, Firebase Authentication.

1 Introduction

The urgency of the searching problem and the reduction of heat loss via the engineering construction elements is due to the need for a rigorous energy saving. This problem is exacerbated by the work of heat power engineering in many countries by reason of limited natural energy resources. Therefore, big attention is paid to reducing operating costs, and timely detection and elimination of emergency situations on the

heat mains. The most problematic this is the detection of heat energy losses on hidden thermal pipes.

New methods the detection and repairing the founded defects in transmission pipelines have been described by many researchers and scientists. All of the world common attention is paid for non-destructive techniques and technologies like acoustic leak detection, tracer gas, radiotracers, the combination of visual and thermographic inspection, etc. [1–5].

It should be noted that manual acoustic and pulse measuring devices are harmful to the health of the operator of such equipment [4, 5].

It's available to detect places of the local heat leak on the underground thermal pipelines using the computer system. Such a system should use infrared thermography devices which is able to transmit data through wireless ports.

The proposed solution would allow pinpoint precision thermal energy loss location to eliminate unnecessary digging, destruction and expense.

2 Formal problem statement

The detection and recovery time of the carrying capacity of pipeline sections with local damage can be significantly reduced through the measuring and numerical analytical procedures with using of the information-measuring computer systems (IMCS).

Infrared is a proven noninvasive technology, which locates leaks in the systems of heat pipes without any destruction owing to the digging of highways, under which heating mains in cities are usually laid.

The purpose of the work is to develop a thermographic method for detection the places of heat energy leak on the underground thermal pipe by means computer system.

The object of the work is the data handling process for detection heat loss using a combination of terrain maps and results of heat measurements.

The subject of the work is an IMCS for creating heat maps of energy leak on the thermal pipelines according to the data of pyrometric non-contact measurements.

The tasks of the study include:

- analysis of existing algorithms and methods implementing heat map construction on the basis of collected data;
- analysis of the existing base of devices for research and analysis the current temperature of objects;
- development of an analytical model for forming a heat map with an account of seasonal changes on separate temperature scales;
- development of a software application for detection of heat energy leaks on thermal pipelines and visualizing the results of measurements on heat maps of the area.

The practical value of the work is:

- detection and localization of the breakthroughs of the heating main for the planning of repairing works, etc.;

- the possibility of identifying objects that emit heat.

3 Literature review

A hot-water pipeline heats the surface of the ground and, therefore, can be detected by a thermographic method. In Fig. 1 shows how using a thermal camera located on the drone, areas can be detected where the heat main is broken [6]. The positive is that the movement of such IMCS can be set strictly within the spatial corridor [7], that is, the street. Also, it can be a study at once of the entire specified area, such as city block, sub-district. etc. [8].

The leak imaging system based on rotor-wing and fixed-wing UAVs [8, 9]. One of the best modern hardware solutions for studying energy losses on the thermal line is the Thermal imaging cameras Zenmuse XT developed by FLIR technology (that costs about \$7,000 USD) on board at Drone DJI Matrice 210 (that costs about \$12,000 USD), are shown in Fig 2.

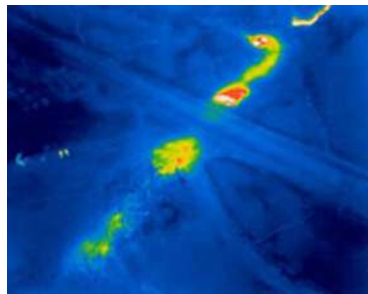


Fig. 1. Infrared detection using a drone the placements of heat energy loss [6].



Fig. 2. Drone DJI Matrice 210 with Thermal Imaging Cameras Zenmuse XT and video camera on board [10].

Usually, such systems don't determine accurate GPS coordinates of the accident and are therefore useless for the repair team. At present, the leak detection team should wait for repairmen at the identified accident site or create written, scheme and photos documentation [2].

Due to the lack of methods for identifying technological objects, is done a manual inspection with accordance to the route was visualized on the map, after which a heat map is created [11].

B Czech Republic has developed the system Workswell WIRIS, which can be fitted with a GPS sensor for storing location information on the position of the drone when producing a record about detection of hot water leaks [6]. This system is designed to be mounted on a drone (UAV) too.

But there are two problems for the implementation of this method of study of energy losses. First, the above version of the thermal measurements is too expensive. Second, on June 11, 2018, the State Aviation Service of Ukraine banned flights of

remotely manned aircraft (drones or UAVs) over roads of state importance, streets of cities and villages, etc. [12]. Similar prohibitions on the use of drones on the territory of settlements exist in other countries – Italy, Austria, Czech Republic, Great Britain, etc. [13, 14]. Thus, the problem under consideration is relevant not only for Ukraine.

In this article, we proposed to collect thermal data using an on-duty car with a manual pyrometer. One of the best solutions is the Testo 805i Bluetooth pyrometer (Fig. 3).



Fig. 3. Infrared thermometer Testo 805i using wireless transmitting data to a smartphone.

Unlike the above examples, this device (that costs about \$100 USD) has several advantages such as:

- measuring the temperature of the object with the help of non-contact infrared beam;
- the ability to transmit data through Bluetooth port 4.0 to computer equipment;
- laser sight in the form of a circle with a point in the middle;
- it is compact and comfortable when carried [15].

4 The modified thermographic method to detect heat loss place improved by the addition of leak geo coordinates

4.1 A Development of the IMCS Structure and Technological Process

Thus, due to the inability to use drones for the inspection of heating mains in urban areas, it is proposed to inspect the thermal pipe using a non-contact pyrometer. For example, the model Testo 805i could be used, which is equipped with a diagnostic team moving the on-duty car.

Bluetooth infrared thermometer Testo 805i has the means to transmit the measurement data to the mobile device (smartphone, tablet, etc.) via Bluetooth port 4.0 at a distance of 15–20 m. It requires iOS 8.3 (or later) / Android 4.3 (or later).

Unfortunately, the Testo 805i doesn't measure automatically, so it is necessary to press the button on the device to receive data. If it is to do every 10 seconds, then at a speed of 40 km/h (10 m/s) the temperature will be measured every 100 m.

The measured temperature is recorded into the database of the mobile device and can be uploaded to the cloud storage. Then measured data can be plotted on the terrain map as values or as a heat map.

The structural scheme of the technological process is shown in Fig. 4.

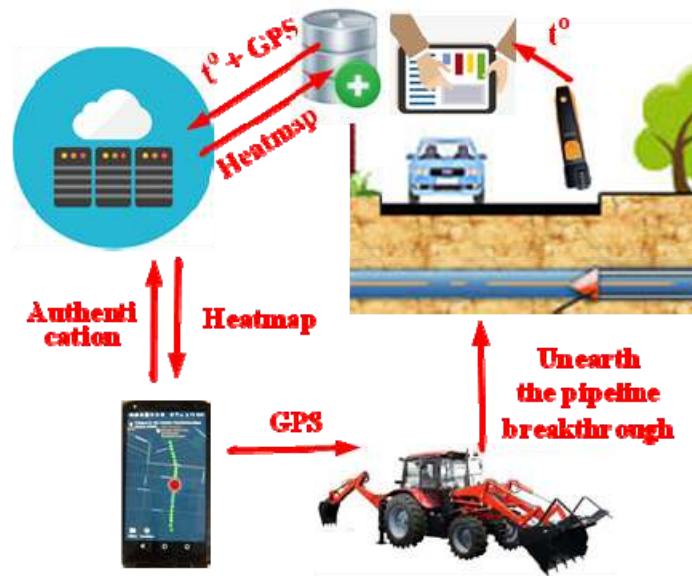


Fig. 4. Scheme of the technological process.

In Fig. 5 shows a block diagram of the formation of a heat map.

Initially, temperature data are measured with a pyrometer, and if these data have measured – they are recorded in the database, if not – the heat map will not be built, further measurements are needed.

Further, when the temperature data to the phone (or tablet) received there is a process of tracking the coordinates, namely latitude and longitude, using a GPS module of mobile phone or tablet. If the data have received – they are also entered in the database and in the case of their absence – the heat map will not be built. In this case, the on-duty car with measuring equipment moves further, to a point with new geolocation, where subsequent measurements would be made.

The final stage is the very construction of a heat map. Then it adds to Google Maps [16]. To do this, it needs to use the library *com.google.maps.android:android-maps-utils.Subsequent*.

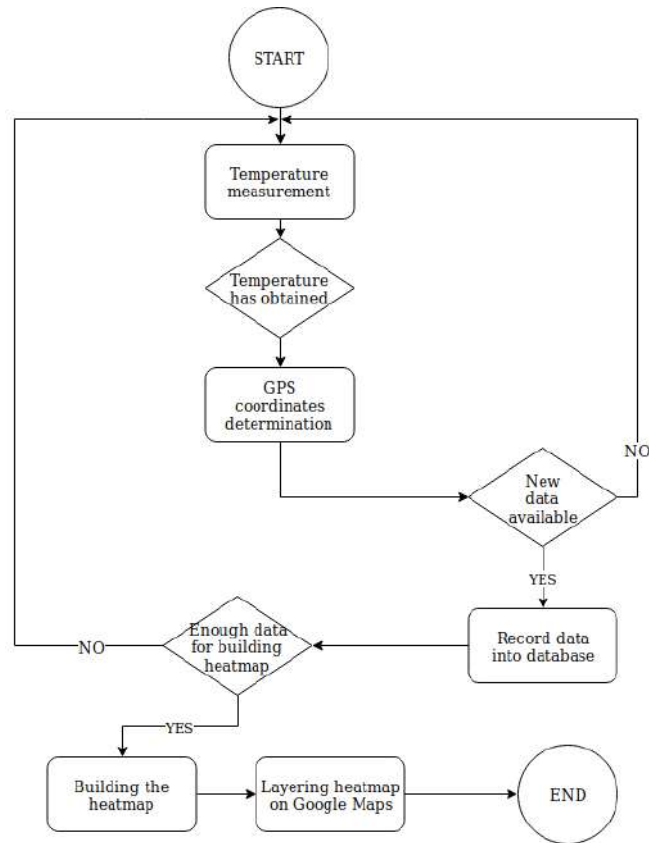


Fig. 5. Block scheme of the heatmap building algorithm.

The further development of the proposed method is the application of cloud technologies when the obtained data (temperature and GPS coordinates) are uploaded to Google Firebase [17, 18]. In this case, the complete technological process is shown in Fig. 4.

4.2 Results of measurements and input data for research

Since Mykolaiv is one of the large industrial cities of Ukraine with a significant number of heating mains under urban highways, it is possible on its example to study the measurements of the temperature above the heating main were carried out in Mykolaiv (Ukraine).

As a result of the temperature measurements, a breakthrough of the heat pipe at the intersection of the streets was detected as shown in Table 1. The temperature data from the pyrometer to the tablet are transmitted in tabular form or as a graph (Fig. 2) and are supplemented in the database combined with geocoordinates from the GPS module of the tablet and from Google Maps at the time of temperature measurement (Table 1).

Table 1. The fragment of a training sample for temperature measurements on the thermal pipe in preparation for the heating season (in October)

Data (DD.MM)	Time (HH:MM:SS)	Temperature (°C)	North latitude (degree)	East longitude (degree)	Street name	Conclusion on the state of the thermal pipe
18.10	16:24:10	+20.1	46.970312	32.011009	Velyka Morska Str.	holistic
18.10	...	...	...	...	...	...
18.10	16:25:50	+20.1	46.967368	32.012510	Kolodyazna Str.	holistic
18.10	...	...	...	...	...	...
18.10	16:27:02	+21.1	46.965597	32.013154	Central Avenue	holistic
18.10	...	...	...	...	...	...
18.10	16:27:16	+28.1	46.962030	32.013060	Chkalova Str.	break-through
18.10	16:27:24	+28.2	46.961526	32.012981	Chkalova Str.	break-through
18.10	...	...	...	...	...	...
18.10	16:29:05	+16.4	46.958510	32.012688	Pogranychna Str.	holistic
18.10	...	...	...	...	...	...
18.10	16:30:30	+16.4	46.953942	32.011994	Kuznetska Str.	holistic
18.10	...	...	...	...	...	...
18.10	16:33:15	+15.1	46.949474	32.012019	Zavodska Str.	holistic

As can be seen from Table 1, the temperature above the heat main on the map segment with Chkalova Str. differs significantly from others.

After carrying out repairs, the second temperature measurements on the heating main were made at the same points during the heating season (in January).

The environment temperature was lower, and therefore the average temperature of the thermal pipelines fell compared with the previous measurements made in October.

But the temperature doesn't differ in a large measure from each other along the heat pipe. This is why we conclude that the breakthrough in the pipeline has been eliminated.

4.3 An analytical model for the formation of a heat map

It should be noted that tabular data is very inconvenient for perception and rapid analysis. Therefore, it was decided to build on the basis of the measurements of the

heat map of the process of finding emergency sections on the heating main. For this purpose, an analytical model for the formation of a heat map was developed.

From the data analysis in Table 1 it can be seen that the measured temperature of the road surface over the heating main was above the ambient temperature at the date of research [19] on 2°C, and at the place of the leakage was recorded at an additional exceeding of 1...4°C. Even the direction of the flow of water through the heating main channel is revealed –the temperature of the road along the pipeline in this direction gradually drops to the ambient temperature.

From this, it follows that in the analytical model it is advisable to take into account the different levels of surface temperature above the heating main with the help of a color gradient.

So, two color models – RGB and HSV – are discussed [20]. The RGB model describes the color space based on Red, Green and Blue colors (Fig. 6, a). The HSV model operates with parameters Hue, Saturation and Value (Fig. 6, b).

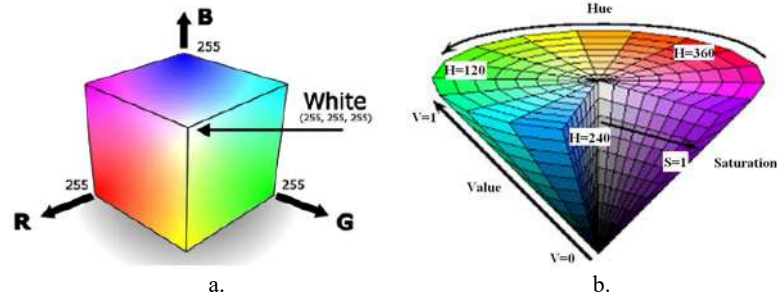


Fig. 6. RGB (a) and HSV (b) color models with fixed value of parameters.

However, if the colors are supposed to be interpreted using the human visual system, it is advisable to integrate the RGB space and the HSV space [21]. In addition, diagnostic and repair work on heating mains can be carried out at different times of the day, and the HSV color space is more resistant to changes in lighting [22]. Therefore, it is advisable to formulate the mathematical equivalent of the projection of the HSV model vectors in the RGB color space.

In this article, only parameters Hue and Value of the HSV color model are used to construct the color gradient in the RGB color model.

The pyrometer readings, corresponding to the seasonal temperature value of the section of the pipeline, are successively added to the array of temperatures $\{t_{i,j}\}$, where $i = \|N_{Jan} N_{Feb} \dots N_{Dec}\|$ – the numbers of all N_{day} rows of maximal temperature values in the frame of month, $j = \overline{0, X}$ – the numbers of columns with each of the X measured temperature value in the matrix of the input file data for heat map.

The maximal temperature values N_{day} of each scale are calculated from the maximal temperature $T_{day}^{\max}$ in researching area from the weather web resource such as [19] in the date (for example, on October 18). The parameter N_{day} takes into account the possible rise of the surface temperature of the road due to existing of pipeline T_{heat} and due to a leak in heat main T_{leak} on values described above on October 18:

$$N_{day} = 19^{\circ}C + 2^{\circ}C + 4^{\circ}C = 25^{\circ}C.$$

It should be noted that the numbers of rows i correspond to the number of temperature scales and can be less than the number of months. It would be possible, since the months in different seasons may have equal temperature scales. This approach saves the computational resources needed to build a heat map.

To recount the parameters of the HSV color model in the value of each spectrum of the RGB color model, the temperature coefficient k of the HSV model should be calculated. This will allow scaling the temperature values to the color of a certain tone. To switch to the parameters of the color model, it is also necessary to calculate the basic statistical parameters, namely the maximum and minimum values of temperature (2) – (3) from the set of temperatures (4) – (5):

$$k = \frac{(sp_{\max} - sp_{\min})}{(t_{\max} - t_{\min})}, \quad (1)$$

$$t_{\max} = \max \{t_{i,j}\}, \quad (2)$$

$$t_{\min} = \min \{t_{i,j}\}, \quad (3)$$

$$i = \|\overline{N_{Jan}}, \overline{N_{Feb}}, \dots, \overline{N_{Dec}}\|, \quad j = \overline{0}, \overline{X}, \quad (4)$$

$$N_{day} = T_{day}^{\max} + T_{heat} + T_{leak}. \quad (5)$$

Therefore, the next step in forming a thermal image in the proposed analytical model is the conversion of temperature data into a color RGB model based on the normed values of parameters of the HSV model.

The values of the red, green, and blue components of the RGB spectrum are labeled $R_{i,j}$, $G_{i,j}$ and $B_{i,j}$, respectively and calculated based on the set of normalized $H_{i,j}$ values as follows:

$$R_{i,j} = \left\{ \begin{array}{l} V_{\min}, H_{i,j} = 0.or.H_{i,j} = 1; \\ |V_{\min} - ad_{i,j}|, H_{i,j} = 2; \\ V, H_{i,j} = 3; \\ |V_{\min} + ad_{i,j}|, H_{i,j} = 4 \end{array} \right\}, \quad (6)$$

$$G_{i,j} = \left\{ \begin{array}{l} V_{\min}, H_{i,j} = 0.or.H_{i,j} = 1; \\ |V_{\min} - ad_{i,j}|, H_{i,j} = 2; \\ V, H_{i,j} = 3; \\ |V_{\min} + ad_{i,j}|, H_{i,j} = 4 \end{array} \right\}, \quad (7)$$

$$B_{i,j} = \left\{ \begin{array}{l} V_{\min}, H_{i,j} = 0.or.H_{i,j} = 1; \\ |V_{\min} - ad_{i,j}|, H_{i,j} = 2; \\ V, H_{i,j} = 3; \\ |V_{\min} + ad_{i,j}|, H_{i,j} = 4 \end{array} \right\}, \quad (8)$$

$$ad_{i,j} = (V - V_{\min}) \cdot H_{i,j}. \quad (9)$$

In this case, the higher the temperature the more the color is close to the red, and vice versa – the colder the temperature the more the color is closer to the blue. As shown in Fig. 7, a, in the described model, the temperature scale is implemented in the temperature range from 0 to +25°C, which corresponds to the season when the heating pipes are cooked for the heating season.

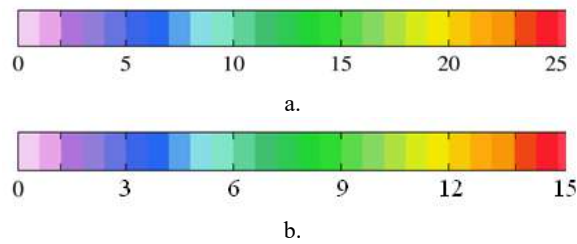


Fig. 7. Temperature Scale-25 for the pre-heating (a) and Scale-15 for the heating (b) season

It should be noted that thermal measurements can be carried out in different seasons when the temperature of the environment varies considerably over the seasons. For

example, in the region under consideration, seasonal temperature fluctuations are from 0 to +25°C in the autumn and to +15° C in the winter. In this case, the visual temperature scale should have been replaced by the scale shown in Fig. 7, b.

When constructing temperature scales, it should be noted that the operating temperature range for the pyrometer is from minus 10°C to +50°C [15]. But with negative values of the ambient temperature, the measurement error is significantly increased. Therefore, further temperature scales were built only for ranges in which the pyrometer measurement error did not exceed 1.5 %.

4.4 Software development

The software was developed in Android Studio 3.3.1, an application is created for the Android platform with JDK ver. 11.0 [23, 24].

Consider the life cycle of the application. On the initial phase of the application, the task should be sent to diagnostic teams for inspection of the area.

Using the Firebase Cloud Messaging (FCM), the system server sends messages to clients about current tasks. There is an ability to send notifications (up to 2K bytes) and data messages (up to 4K bytes).

Upon receiving the assignment, the diagnostic teams on a duty car are headed to places of information gathering. After their arrival, the system reads the GPS coordinates and sends a message to the server about the diagnostic team location.

The diagnostic team specialists, using pyrometers, measures the temperature values. Data via Bluetooth 4.0 technology are transmitted to a tablet/notebook storing a local database in SQLite. When the tablet is connected to the Internet, information is sent to the cloud database.

After completing the inspections by the diagnostic team, other services (mobile clients) will receive a message through the FCM for further works on the territory.

Using the web version of the application helps to create reports based on the data collected. The *JasperReports* library is used to generate reports on the server.

5 Experiments and results

The proposed methods and indicators characterizing their quality were implemented as software and experimentally investigated in solving the problem to detect the heat leak in underground thermal mains on Chkalova Street of Mykolaiv city (Ukraine) during preparation to the heating season in October.

Temperature values along this street were obtained as the table and as the graphic (Fig. 8) from pyrometer Testo 805i to the tablet of developed IMSC mounted to the on-duty car. Based on the graph of the ambient temperature from the weather site (Fig. 9), using the algorithm discussed above, a heat map was built for the thermal main under the road on Chkalova str. and was layered on Google Map of the specified area of the Mykolaiv city.

Based on Fig. 8 it is also possible to conclude that on pipe segment after the detected leakage point, the temperature over the surface of the heating main has been reduced to the ambient temperature for the specified date of Oct. 18, 2018 (Fig. 9).

In Fig. 10 shows screenshots of the interface of the developed Android application. You can see the map and two buttons on the smartphone of IMCS. The button "Location" is responsible for the current location of your mobile device (or tablet). The button "Open" is responsible for opening a file from the database needed to build a heat map.

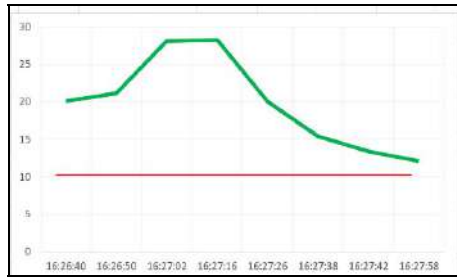


Fig. 8. Temperature graphic from pyrometer to a smartphone on October

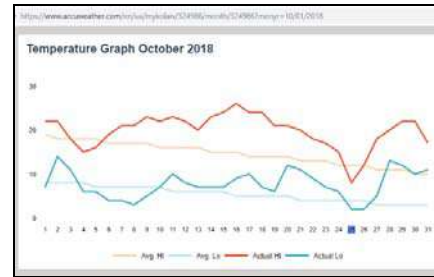
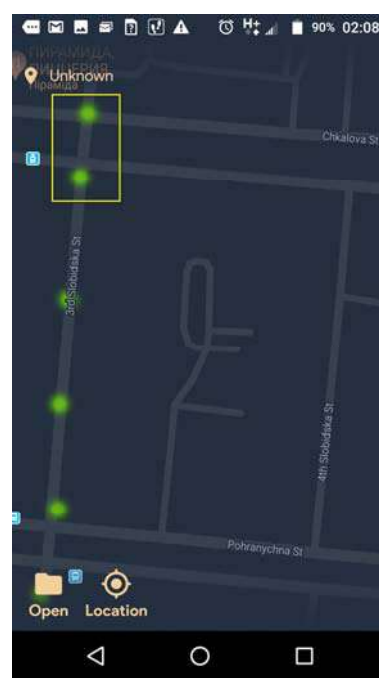


Fig. 9. The temperature of the environment on the date of measurement [19]



a.



b.

Fig. 10. The heatmap with reflected the accident (a) and its eliminated (b)

From the analysis of the heat map in Fig. 10, a can be seen that on the corner of the 3rd Slobidska Str. and Chkalova Str. the heat energy leak from the underground pipeline.

The advantages of the proposed method are that the repair team can operate independently of the work schedule and the location of the diagnostic team specialists. When the repair team arrives at the place of heat energy leakage, the brigadier from any mobile device (smartphone, tablet, etc.) is authorized on Firebase. In this case, his profiles in any social networks can be used – Google, Facebook, Twitter, etc. (Fig. 11).

After the computer diagnostics were carried out, the received data was transferred to the corresponding repair company. The unearthing of the heating pipeline by the found coordinates of the breakthrough has confirmed the fact of the accident exactly on the corner of the 3rd Slobidska Str. and Chkalova Str. in the detected place (Fig. 12).



Fig. 11. Integrate Firebase Authentication with popular services [25].



Fig. 12. The detected places of the local heat leak on the underground thermal pipe.

When the access to the heat map of the thermal pipeline has given, brigadier of the repair team can compare the own GPS coordinates with the coordinates received from the cloud service. Then he is able to accurately to combine the repair equipment (excavator, pump, etc.) place with detected places of the heat energy leak from heat map together.

In Fig. 12 shows the same breakthrough of the heating main that was detected as a result of temperature measurements and reflected in Fig. 10, a. The red color in the center of the spot corresponds to the raised temperature. And in Fig. 10, b shows the same place, but already with the eliminated breakthrough (see the yellow rectangle).

6 Conclusion

Analysis of leak detection methods in thermal pipelines suggested that the thermographic method in combination with the determination of the geographic coordinates of the measurement point should increase the accuracy of determining the location of the breakthrough of the heating main.

Innovativeness of this work would be to develop an analytical model for constructing a heat map with an account of seasonal changes on separate temperature scales. The proposed reduction of the HSV color model basis due to the fixing value of the "Saturation" parameter as constant significantly reduces the number of computations when converting the temperature values measured by the digital pyrometer to the image of the color RGB model using the parameters of the color model HSV.

A software application for building heat maps based on measured temperature data for the Android platform with JDK 11.0 version was developed. Introduced Firebase Authentication, integrated with popular social services.

The software implementing the proposed methods has been developed and studied for the detection places of the heat energy leak on the underground thermal pipelines. Due to the heat map laid on Google Maps, the repair team gained access to detected places of the heat energy leak regardless of the work schedule and the location of the diagnostic team is provided. The conducted experiments have confirmed the performance of the developed software and allow recommending it for use in practice.

The prospects for further research are to a wider of applied developed IMCS while solution to the problems associated with various heat leak, not only thermal pipelines. Also, It would be useful to study the accuracy (error) of developed methods for the different temperature scales.

References

1. Non-destructive Testing and Repair of Pipelines. Barkanov, E. N., Dumitrescu, A., Parinov, I. A. (eds). Springer, Cham (2018). doi: 10.1007/978-3-319-56579-8
2. Fahmy, M., Moselhi, O.: Detecting and locating leaks in Underground Water Mains Using Thermography. In: 26th Int. Symp. on Automation and Robotics in Construction, pp. 61–67 (2009), <https://www.irbnet.de/daten/iconda/CIB14819.pdf>, last accessed 2019/03/16
3. Timashev, S., Bushinskaya, A.: Diagnostics and Reliability of Pipeline Systems. Springer International Publishing (2016). doi: 10.1007/978-3-319-25307-7
4. Leak Detection in Heat Exchangers and Underground Pipelines Using Radiotracers. material for education and on-the-job training for practitioners of radiotracer technology. International Atomic Energy Agency, Vienna (2009), https://www-pub.iaea.org/MTCD/Publications/PDF/TCS-38_web.pdf, last accessed 2019/03/16
5. Practical Acoustic Emission Testing. The Japanese Society for Non-Destructive Inspection (ed.). Springer, Japan (2016). doi: 10.1007/978-4-431-55072-3
6. Pipeline inspection with thermal diagnostics [The Workswell WIRIS System]. Workswell s.r.o., Prague, Czech Republic (2018), <https://www.drone-thermal-camera.com/wp-content/uploads/2018/05/pipeline.pdf>, last accessed 2019/03/16
7. Musiyenko, M. P., Zhuravska, I. M., Kulakovska, I. V., Kulakovska, A. V.: Simulation the behavior of robot sub-swarm in spatial corridors. In: Proceeding of the IEEE 36th Int. Conf. on Electronics and Nanotechnology (ELNANO–2016), Kyiv, Ukraine, pp. 382–387 (2016). doi: 10.1109/ELNANO.2016.7493090
8. Zhuravska, I., Kulakovska, I., Musiyenko M.: Development of a method for determining the area of operation of unmanned vehicles formation by using the graph theory. Eastern-European Journal of Enterprise Technologies. 2, 3 (92): 4–12 (2018). doi: 10.15587/1729-4061.2018.128745
9. Stockton, G. R.: Leak Detection – Aerial Thermal Mapping Reveals Leaks in All Types of Systems. In: CampusEnergy2017 [Int. District Energy Association]. Miami, FL. (2017),

- <https://www.districtenergy.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=5ae4b471-ee47-9f2f-4443-94a8fe505273>, last accessed 2019/03/16
10. Thermal Mapping – Inspection of District Heating. In: Case Study for Drone Systems – with Thermal Capture Recorder 2.0 640. TeAx Technology GmbH, Copenhagen, Denmark (2018), <https://thermalcapture.com/thermal-mapping-inspection-of-district-heating/>, last accessed 2019/03/16
 11. Atkinson, D.: DJI: Zenmuse XT.vs.Zenmuse XT2. Publ. April 20th 2018, <https://www.heliguy.com/blog/2018/04/20/zenmuse-xt-vs-zenmuse-xt2/>, last accessed 2019/03/16
 12. Cornelis, E., Meinke-Hubeny, F.: The STRATEGO Project No: IEE/13/650 (Multi-level actions for enhanced Heating & Cooling plans), May 29 2015. The European Union (2015), http://stratego-project.eu/wp-content/uploads/2014/09/D3_7a-STRATEGO-WP3-Mapping-methodology-and-data-sources.pdf, last accessed 2019/03/16
 13. Tymchasovyi poriadok vykorystannia povitrianoho prostoru Ukrainy [Temporary rules of use of airspace of Ukraine]. Approved May 31 2018. The State Aviation Administration of Ukraine (2018), https://avia.gov.ua/wp-content/uploads/2018/05/TYMCHASOVYJ-poryadok-VPPU_31.05.18-2.doc, last accessed 2019/03/16 [in Ukrainian]
 14. Regulation of the European Parliament and of the Council on common rules in the field of civil aviation, December 7 2015. Brussels (2015), http://ec.europa.eu/smart-regulation/impact/ia_carried_out/docs/ia_2015/swd_2015_0262_en.pdf, last accessed 2019/03/16
 15. Airspace restrictions for unmanned aircraft and drones [Civil Aviation Authority – CAA in the United Kingdom], <https://www.caa.co.uk/Consumers/Unmanned-aircraft/Our-role/Airspace-restrictions-for-unmanned-aircraft-and-drones/>, last accessed 2019/03/16.
 16. Testo – Smart Probes: instruction manual. Publ. November 11 (2015), http://www.testequipmentdepot.com/testo/pdf/smartprobes_manual.pdf, last accessed 2019/03/16
 17. Price, E.: Create a heat map of your Google location history with this tool. Publ. Apr. 11 (2018), <https://lifehacker.com/create-a-heat-map-of-your-google-location-history-with-1830215410>, last accessed 2019/03/16
 18. Smyth, N.: Firebase Essentials – Android Edition. Payload Media (2017)
 19. Ashok, K. S.: Mastering Firebase for Android Development: Build real-time, scalable, and cloud-enabled Android apps with Firebase. Packt Publishing Ltd. (2018)
 20. Temperature Graph October 2018 [web-site of AccuWeather, Inc.], <https://www.accuweather.com/en/ua/mykolaiv/324986/month/324986?monyr=10/18/2018>, last accessed 2019/03/16
 21. Kolkur, S., Kalbande, D., Shimpi, P., Bapat, C., Jatakia, J.: Human Skin Detection Using RGB, HSV and YCbCr Color Models. ICCASP/ICMMD-2016. In: Advances in Intelligent Systems Research, vol. 137, pp. 324–332. Atlantis Press (2017)
 22. Loesdau, M., Chabrier, S., Gabillon, A.: Hue and Saturation in the RGB Color Space. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mamass, D. (eds) Image and Signal Processing. ICISP [Lecture Notes in Computer Science], vol. 8509, pp. 203–212. Springer, Cham (2014). doi: 10.1007/978-3-319-07998-1_23
 23. Nursabillilah, M. A.: Performance Comparison between RGB and HSV Color Segmentations for Road Signs Detection. Applied Mechanics and Materials. 393: 550–555 (Sept. 2013). doi: 10.4028/www.scientific.net/AMM.393.550.
 24. Mainkar, P.: Expert Android Programming: Master skills to build enterprise grade Android applications. Packt Publishing Ltd (2017)
 25. How to add Firebase Authentication to your Android Studio, <http://blog.androidrich.com/2016/09/how-to-add-firebase-authentication.html>, last accessed 2019/03/16

Techniques of DNA Microarray Data Pre-processing Based on the Complex Use of Bioconductor Tools and Shannon Entropy

Sergii Babichev^{1,2}[0000-0001-6797-1467], Bohdan Durnyak²[0000-0003-1526-9005], Valeriy Zhydetskyi²[0000-0002-2880-9616], Iryna Pikh²[0000-0002-9909-8444], Vsevolod Senkivskyi²[0000-0002-4510-540X]

¹Jan Evangelista Purkyně University in Ústí nad Labem, České mládeže, 8, Ústí nad Labem, 40096, Czech Republic
sergii.babichev@ujep.cz

²Ukrainian Academy of Printing, Pid Holoskom, 19, Lviv, 79000, Ukraine
durnyak@uad.lviv.ua, v_uad@ukr.net, pikhirena@gmail.com, senk.vm@gmail.com

Abstract. The paper presents the comparison analysis of various techniques of DNA microarrays data pre-processing in order to choose the optimal combination of the methods in terms of the minimum value of Shannon entropy criterium. The Bioconductor package tools of R software were used during the simulation process. The DNA microarray data of patients, which were investigated on lung cancer from database Array Express, were used as the experimental data. The algorithm of step-by-step procedure of the data processing for purpose of determination of the optimal combination of the methods has been proposed as the results of the research. The results of the simulation have shown that the optimal combination of the methods for the investigated data is the following one: *rma* method background correction, *invariant set* method normalization and *mas* method PM correction and summarization. This combination of the methods corresponds to the minimum value of the Shannon entropy criterion.

Keywords: DNA microarray, gene expression profiles, background correction, normalization, PM correction, summarization, Shannon entropy

1 Introduction

One of the current directions of modern bioinformatics is reconstruction and simulation of gene regulatory networks based on the data from DNA-microchip experiments [1]. Implementation of this process involves the preliminary experimental data pre-processing in order to form the array of gene expression profiles. DNA microchip data are presented as a matrix of light intensities, the values of which are proportional to the expression of the appropriate genes. The genes expression value determines the amount of appropriate type of protein which will be generated by this gene. Each of the DNA microchips includes the results of the

experiment for appropriate investigated object or for appropriate conditions of the experiment performing. In this case the quantity of the DNA microchips corresponds to the quantity of the investigated objects. It should be noted, that conditions of the experiments performing are differed in the most cases. In this case very important is the stage of the obtained data pre-processing. This procedure involves the use of four steps: background correction, normalization, PM correction and summarization [2–5]. Each of the steps can be implemented using various methods.

The technique to determine the optimal combination of methods to form the gene expression profiles array objectively is absent nowadays. In this paper we propose the technique of gene expression array formation based on Shannon entropy criterion which is calculated with the use of James-Stein shrinkage estimator method [6]. The optimal combination of the methods is determined based on the minimum value of the Shannon entropy criterion.

2 Formal problem statement

The data, which are obtained during the DNA microchip experiments, are presented as a matrix of light intensities. A block chart of procedure of DNA microchip light intensities matrix formation during the experiment performing is presented in fig. 1.

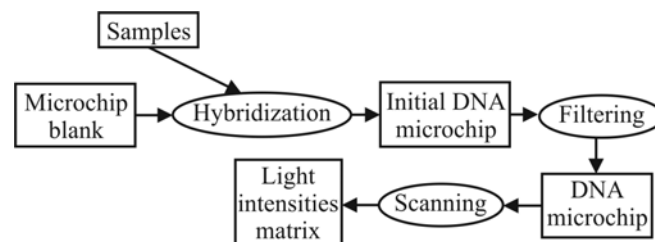


Fig. 1. A block chart of the procedure of DNA microchip light intensities matrix formation

Joining of complementary single-chain nucleotides with fluorescent labels to a single molecule is performed during the hybridization process. It is obvious, that the level of light intensities in appropriate point of the microchip is proportional to quantity of the hybridized RNA molecules, which correspond to appropriate type of the protein. The following stages of the DNA microchip processing are filtering in order to remove unhybridized samples and scanning for purpose of the matrix of light intensities formation. Fig. 2 presents the step-by-step procedure to transform the light intensities values to the expression of the corresponding genes. As it can be seen from fig. 2, each of the steps assumes the use of various methods and choice of the combination of these methods influences directly to the quality of the obtained genes expression data. Thus, the main problem consists in determination of the optimal combination of methods to process the DNA microchip data in order to increase the informativity level of the obtained gene expression data.

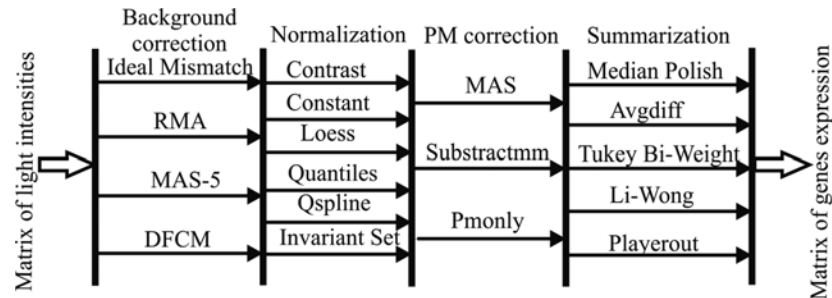


Fig. 2. A chart of step-by-step procedure to transform the light intensities matrix to the matrix of genes expression

3 Literature review

The issues concerning the DNA microarray processing are presented in [7–9]. The authors considered in detail the stages of DNA microarrays creation and the peculiarities of their processing. In [10] the author considered the possibility of the neuro-fuzzy modeling implementation for purpose of the microarrays experiments data processing. In [11] the authors presented the results of research concerning analysis of genes expression for the purpose of the objects classification using Bayesian network. However, it should be noted, that the hereinbefore works do not contain the investigations concerning choice of the appropriate combination of the methods to process the DNA microchip data based on quantitative criteria.

Classification and detail description of the background correction methods are presented in [2–5]. Ideal Mismatch method was proposed by Affymetrix company [3]. This method involves the complex use of both the Perfect Match (PM) nucleotide samples which fully correspond to the investigated genes and the Miss Match (MM) samples, in which the mean nucleotide is changed to complementary one. Robust Multichip Average (RMA) background correction method involves the use only PM samples [4]. This fact decreases the costs to the microchip preparing due to absence of the MM samples. The values of light intensities in this case are presented as the sum of the useful signal, which is distributed exponentially, and the normally distributed noise component. Distribution Free Convolution Model (DFCM) background correction method [5] also assumes that values of light intensities are presented as the combination of both the useful signal and the noise component. But in this case do not any assumes about the character of the components distribution. This method involves the use of both the PM and MM samples. The main idea and the detail description of the Affymetrix Micro Array Suite 5.0 (MAS 5.0) technique of background correction are presented in [2,3].

The techniques of DNA microchip data normalization are presented in [2,12–16]. The necessity of this stage is determined by low correlation of the data which were determined when different conditions of the experiment performing. The aim of the normalization process is the reduction of the microchip empirical data to the same distribution. This step allows minimizing the technological differences between the

parameters of different genes and, as the result, to carry out the comparison of the expression values of the corresponding genes obtained under different conditions of the experiment performing. The results of the research concerning comparison analysis of various methods of PM corrections and summarization of the DNA microchip data are presented in [2,14,17,18]. PM correction stage is performed in order to reduce the nonspecific hybridization effect by correction of the PM samples light intensities taking into account the light intensities of the corresponding MM samples. The summarization process assumes the calculation of gene expressions values from light intensities of the samples for investigated genes.

However, it should be noted that in spite to the achievements in this subject area the effective technique to choose the optimal combination of methods of DNA microchip data processing is absent nowadays. This problem can be solved based on modern techniques of the complex data processing which are used in different field of scientific research nowadays [19–23].

The aim of the paper is the improvement of technique of DNA microarray data processing based on the complex use of Bioconductor tools and Shannon entropy for purpose of gene expression array formation.

4 Materials and methods

The Shannon entropy criterion, which is calculated based on James-Stein shrinkage estimator [6], was used as the main criterion to estimate the gene expression informativity during the simulation process. This technique is based on the complex use of the two different models: a high-dimensional model with low bias and high variance, and a lower dimensional model with larger bias but smaller variance. Evaluation of the probability of values distribution in cells in accordance with James-Stein shrinkage technique is calculated by the formula:

$$p_i^{Srink} = \lambda p_i + (1 - \lambda) p_i^{ML} \quad (1)$$

where p_i^{ML} is the probability of the gene expression values distribution in the i -th cell which is calculated by maximum likelihood method, $p_i = 1/n_i$ is the shrinkage target or probability in the i -th cell in the case of uniform distribution of the gene expression values, n_i is quantity of the features in the i -th cell. It is obvious, that p_i^{ML} corresponds to the high-dimensional model with low bias and high variance and p_i corresponds to the models with higher bias and lower variance of the features distribution. Parameter of the intensity λ in this case is calculated as follows:

$$\lambda = \frac{1 - \sum_{i=1}^k (p_i^{ML})^2}{(n-1) \sum_{i=1}^k (p_i - p_i^{ML})^2} \quad (2)$$

where n is the quantity of the features in the investigated vector. The value of the Shannon entropy is calculated with the use of standard formula taking into account the method of the probability estimation:

$$H^{Shrink} = - \sum_{i=1}^k p_i^{Shrink} \log_2 p_i^{Shrink} \quad (3)$$

Less value of the criterion (3) corresponds to the higher level of the investigated vector informativity.

A structural block chart of the algorithm which was used to determine the optimal combination of the methods of DNA microarray data processing is shown in fig. 3. Implementation of this algorithm involves the following steps:

1. Loading of the DNA microarray data.
2. Setup of the stage of data processing (background correction, normalization, PM correction, summarization). Fixation of the methods, which do not correspond to this stage randomly.
3. Choice of the first method for current stage.
4. DNA microarray data processing by selected methods.
5. Calculation of the Shannon entropy by formulas (1)–(3) for each of the microchips. Calculation of average value of the Shannon entropy for all DNA microarrays.
6. If the number of the method is less than the maximum quantity of the methods at this stage, then choice the next method and go to the step 4 of this procedure. Otherwise fixation of the method which correspond to the minimum value of the Shannon entropy.
7. If the number of the stage is less than maximum quantity of the stages, then go to the next stage and go to the step 3 of this procedure. Otherwise, DNA microarray data processing with the use of determined combination of the methods.

5 Experiments

Simulation process of DNA microchip data pre-processing was performed based on R software [24] using functions of *Bioconductor* package [25]. The lung cancer patients' gene expression profiles E-GEOD-68571 [26] from database ArrayExpress [27] were used as the experimental data during the simulation process. These data includes 96 of DNA microchips of patients which were investigated on lung cancer.

Each of the DNA microchips includes 7129 of genes. 10 patients were identified as healthy and 86 sick patients were divided by the state of their health into three groups.

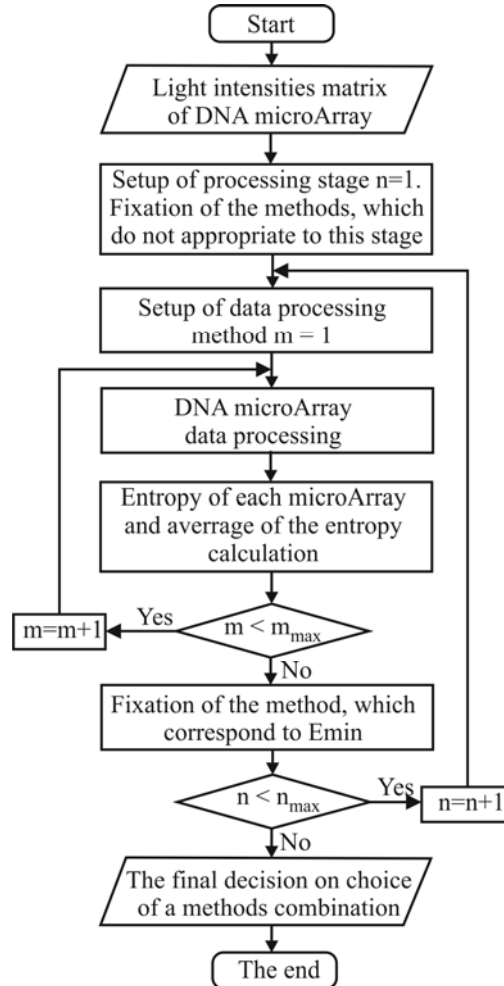


Fig. 3. A chart of the step-by-step procedure to transform the light intensities matrix to the matrix of genes expression

Fig. 4 shows the images of nine of the investigated DNA microchips which were obtained by scanning of the appropriate objects. The following step of the data processing is the transform of the light intensity matrixes into array of genes expressions using the hereinbefore presented methods.

6 Results and discussions

The character of light intensities values distribution at the selected DNA microarray is presented in fig. 5. Fig. 6 shows the MA charts for all pairs of five selected DNA microchips. The *MA* chart presents the difference of logarithms of the PM (Perfect Math) samples values (*M*) versus the average of logarithms of the PM samples values (*A*). Parameters *M* and *A* for *i*-th gene and samples *k* and *n* are calculated in the following way:

$$M = \log_2 \left(\frac{x_{ki}}{x_{ni}} \right), \quad A = \frac{1.2}{2} \log_2 (x_{ki} \cdot x_{ni}) \quad (4)$$

The chart is created for PM values for all possible pairs of the investigated samples. In the case of the highest quality of the data processing, the data should be distributed in a rather narrow range, and the points at *MA* diagram should be located along the axis of *M* = 0 with the lowest averages.

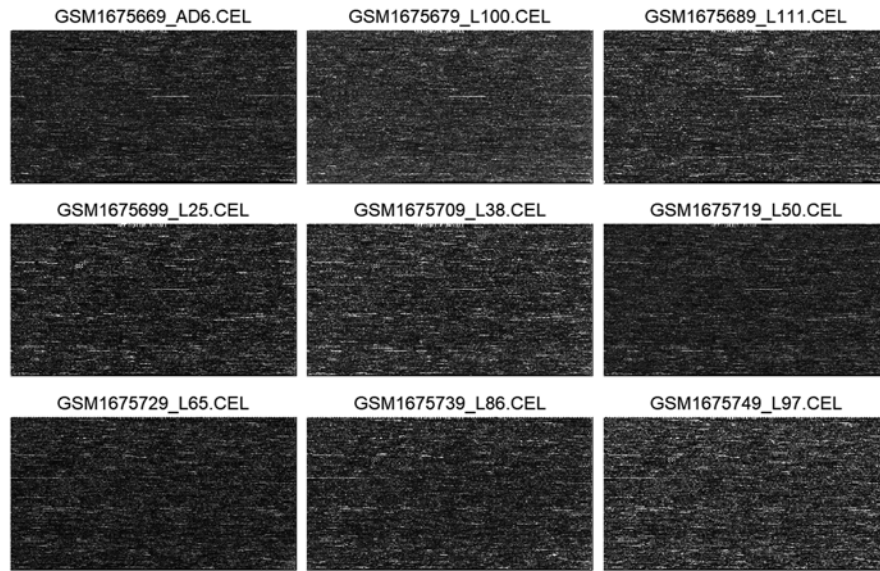


Fig. 4. Scanning images of nine of the investigated DNA microchips

The analysis of the received diagrams confirms the assumption concerning the necessity of the initial data preprocessing. The character of the data distribution for various microchips is differed significantly (Fig. 5a). The kernel density plots which are shown in fig. 5b are distributed along axis of the light intensities logarithm randomly too. Finally, the corresponding points on the MA diagrams (Fig. 6) have different distributions too. These facts do not allow us to compare the investigated gene expression profiles objectively. Fig. 7 and Fig. 8 present the results of the research concerning background correction of the DNA microarrays by methods:

“rma”, “mas” and “DFCM”. The “Ideal Mismatch” method has not used due to lower quality of its operation [28]. The analysis of the obtained charts allows us to conclude that the background correction increases the image quality. The processed data are distributed more uniformly to compare with unprocessed data. However, it should be noted that visual analysis of the diagrams does not allow us to compare the quality of the used methods objectively in order to choose the best one. Fig. 9 presents the results of the research concerning determination of the optimal combination of the methods to process the DNA microarray data based on the minimum value of the Shannon entropy in accordance with hereinbefore technique.

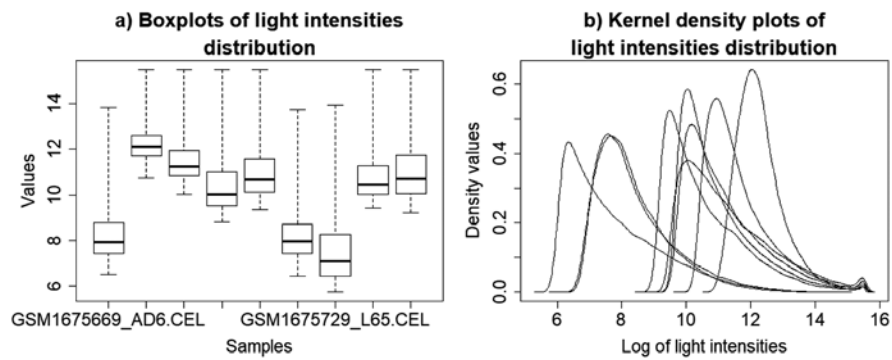


Fig. 5. Estimation of light intensities distribution at the selected DNA microchips

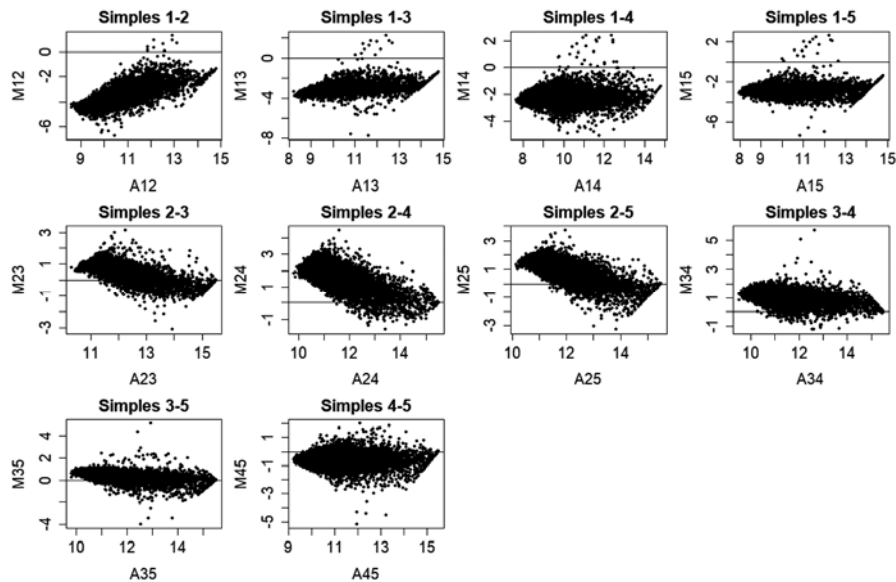


Fig. 6. MA charts of light intensities distribution for PM samples

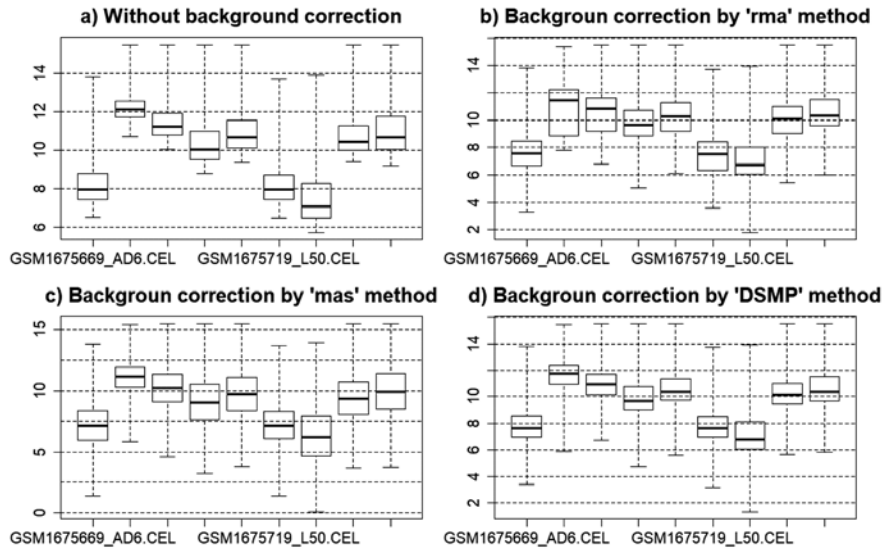


Fig. 7. Boxplot charts of unprocessed and processed data when the various background correction methods are used

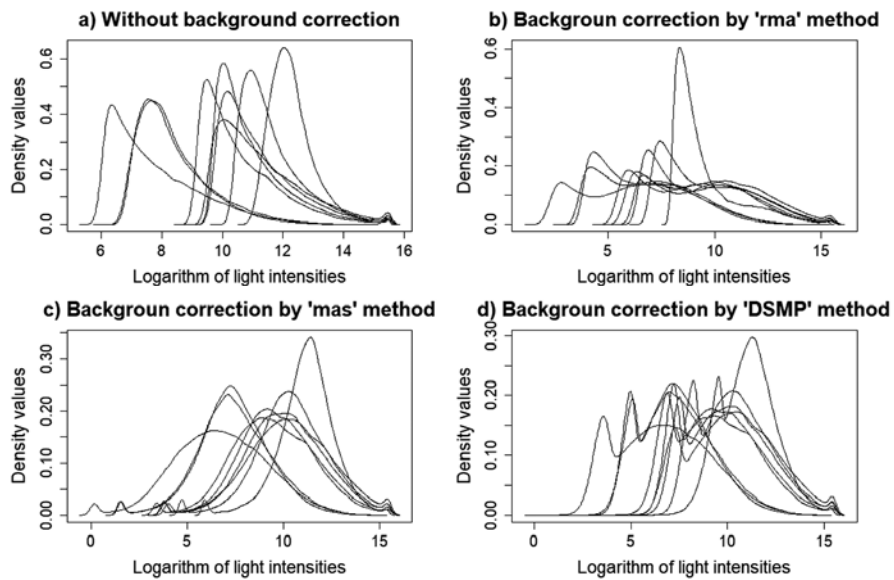


Fig. 8. Kernel density plots of unprocessed and processed data when the various background correction methods are used

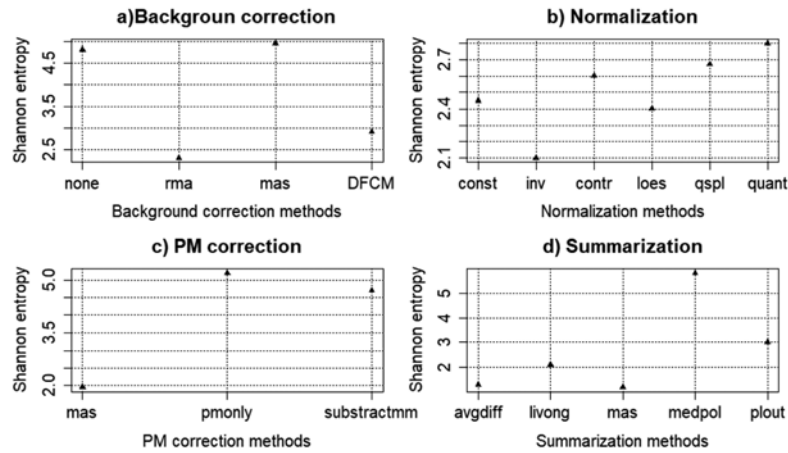


Fig. 9. Charts of the Shannon entropy distribution versus the methods of the data processing at the stages: a) background correction; b) normalization; c) PM correction; d) summarization

The analysis of the obtained charts allows us to conclude that the optimal methods in terms of the minimum value of Shannon entropy criterion are the following ones: “*rma*” background correction method; “*invariant set*” normalization method; “*mas*” methods PM correction and summarization. This combination of the methods was used to process the investigated DNA microarrays. Fig. 10 presents the boxplots of genes expression vectors for the investigated samples of both the non-processed (fig. 10a) and processed (fig. 10b) data.

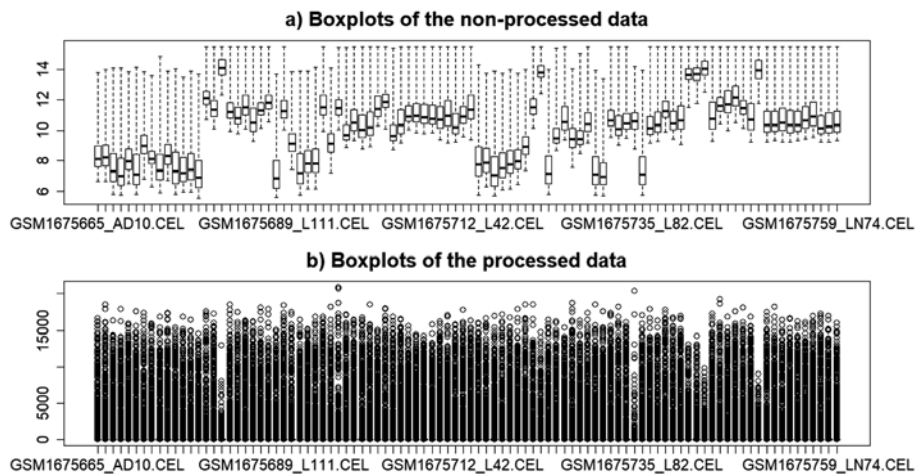


Fig. 10. Results of the DNA microchips processing

As it can be seen from fig. 10b, the values of genes expression are distributed in the same range. The change of this range can be explained in the following way. The expression values of the largest quantity of genes are low. But some of the genes have significantly higher values of expression. It means that these genes determine some important processes in the investigated objects. The expression values of these genes determine the variation range of another genes expression. The analysis of the boxplots allows us also to conclude that the values of the largest quantity of gene expressions for various objects lie in a very narrow range. This can mean that these genes are responsible for the functions that are inherent for all investigated objects. However, each of the investigated samples contains genes, the expression of which goes beyond the inter-quartile range. These genes are very important for the following research since they allow us to distinguish the investigated objects by their particularities.

7 Conclusions

In this paper we have proposed the technique of gene expression array formation which were obtained based on DNA microarray experiments. The initial data is presented as a set of DNA microchips, each of which contains the matrix of light intensities, the values of which are proportional the expression values of the appropriate genes. Four stages have been performed during the simulation process: background correction, normalization, PM correction and summarization. Each of the stage assumed the use of different methods. The Shannon entropy criterion which is calculated based on James-Stein shrinkage estimator has been used as the main criterion to estimate the genes expression informativity.

The simulation process has been performed based on R software with the use of *Bioconductor* package functions. The lung cancer patients' gene expression profiles E-GEOD-68571 from database ArrayExpress have been used as the experimental data during the simulation process. The results of the simulation have shown that the optimal combination of the methods in terms of the minimum value of the Shannon entropy is the following one: “*rma*” background correction method, “*invariant set*” normalization method and “*mas*” methods PM correction and summarization. This combination of the methods has been used to process the investigated DNA microchips.

The boxplots of both the non-processed and processed data have been created as the simulation results. The analysis of the obtained results has shown that the values of the largest quantity of gene expressions for various objects lie in a very narrow range. It means that these genes are responsible for the functions that are inherent for all investigated objects. However, each of the investigated samples contains genes, the expression of which goes beyond the inter-quartile range. This fact can mean that these genes are very important for the following research since they allow us to distinguish the investigated objects by their particularities.

References

1. Zak, D.E., Vadigepalli, R., Gonye, G.E., Doyle, F.J., Schwaber, J.S., Ogunnaike, B.A.: Unconventional systems analysis problems in molecular biology: A case study in gene regulatory network modeling. *Computers and Chemical Engineering*, 29 (3), pp. 547-563 (2005) doi: 10.1016/j.compchemeng.2004.08.016
2. Bolstad, B.M., Irizarry, R.A., Åstrand, M., Speed, T.P.: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19 (2), pp. 185-193 (2003) doi: 10.1093/bioinformatics/19.2.185
3. Affymetrix. Statistical Algorithms Description Document. Affymetrix, Inc., Santa Clara, CA, pp. 1-27 (2002)
4. Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., Speed, T.P.: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Selected Works of Terry Speed*, pp. 601-616 (2012) doi: 10.1007/978-1-4614-1347-9_15
5. Chen, Z., McGee, M., Liu, Q., Kong, M., Deng, Y., Scheuermann, R.H.: A distribution-free convolution model for background correction of oligonucleotide microarray data. *BMC Genomics*, 10 (SUPPL. 1), art. no. S19 (2009) doi: 10.1186/1471-2164-10-S1-S19
6. Hausser, J., Strimmer, K.: Entropy inference and the james-stein estimator, with application to nonlinear gene association networks. *Journal of Machine Learning Research*, 10, pp. 1469-1484 (2009)
7. Kohane, I.S., Kho, A.T., Butte, A.J.: *Microarrays for an integrative genomics*. Cambridge, Massachusetts, England: A Bradford book, the MIT press, 236 p. (2003)
8. Ivakhno, S.S., Kornelyuk, A.I.: Microarrays: Technologies overview and data analysis. *Ukrain'skyi Biokhimichnyi Zhurnal*, 76 (2), pp. 5-19 (2004)
9. Babichev, S.A., Kornelyuk, A.I., Lytvynenko, V.I., Osypenko, V.V.: Computational analysis of microarray gene expression profiles of lung cancer. *Biopolymers and Cell*, 32 (1), pp. 70-79 (2016) doi: 10.7124/bc.00090F
10. Wang, Z.: *Neuro-Fuzzy Modeling for Microarray Cancer Gene Expression Data*. Oxford University Computing Laboratory, 107 p. (2005)
11. Loren van Themaat, E.V.: *On the Use of Learning Bayesian Networks to Analyze Gene Expression Data: Classification and Gene Network Reconstruction*. University of Amsterdam, 73 p. (2005)
12. Gentleman, R., Carey, V., Huber, W., Irizarry, R., Dudoit, S.: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, 473 p. (2005)
13. Park, T., Yi, S.-G., Kang, S.-H., Lee, S.Y., Lee, Y.-S., Simon, R.: Evaluation of normalization methods for microarray data. *BMC Bioinformatics*, 4, art. no. 33, 13 p. (2003) doi: 10.1186/1471-2105-4-33
14. Raddatz, B.B., Spitzbarth, I., Matheis, K.A., Kalkuhl, A., Deschl, U., Baumgärtner, W., Ulrich, R.: Microarray-Based Gene Expression Analysis for Veterinary Pathologists: A Review. *Veterinary Pathology*, 54(5), pp. 734-755 (2017) doi: 10.1177/0300985817709887
15. Åstrand, M.: Contrast normalization of oligonucleotide arrays. *Journal of Computational Biology*, 10 (1), pp. 95-102 (2003) doi: 10.1089/106652703763255697
16. Chen, Y.-J., Kodell, R., Sistare, F., Thompson, K.L., Morris, S., Chen, J.J.: Normalization methods for analysis of microarray gene-expression data. *Journal of Biopharmaceutical Statistics*, 13 (1), pp. 57-74 (2003) doi: 10.1081/BIP-120017726
17. Barbará, D., Wu, X.: *An Approximate Median Polish Algorithm for Large Multidimensional Data Sets*. Springer-Verlag London Ltd. Knowledge and Information Systems, vol. 5, pp. 416-438 (2003)

18. Lazaridis, E.N., Sinibaldi, D., Bloom, G., Mane, S., Jove, R.: A simple method to improve probe set estimates from oligonucleotide arrays. *Mathematical Biosciences*, 176 (1), pp. 53-58 (2002) doi: 10.1016/S0025-5564(01)00100-6
19. Babichev, S., Lytvynenko, V., Osypenko, V.: Implementation of the objective clustering inductive technology based on DBSCAN clustering algorithm. *Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2017*, 1, art. no. 8098832, pp. 479-484 (2017) doi: 10.1109/STC-CSIT.2017.8098832
20. Babichev, S., Korobchynskiy, M., Lahodynskiy, O., Korchomnyi, O., Basanets, V., Borynskyi, V.: Development of a technique for the reconstruction and validation of gene network models based on gene expression profiles. *Eastern-European Journal of Enterprise Technologies*, 1 (4-91), pp. 19-32 (2018) doi: 10.15587/1729-4061.2018.123634
21. Babichev, S., Krejci, J., Bicanek, J., Lytvynenko, V.: Gene expression sequences clustering based on the internal and external clustering quality criteria. *Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2017*, 1, art. no. 8098744, pp. 91-94 (2017) doi: 10.1109/STC-CSIT.2017.8098744
22. Tkachenko, R., Doroshenko, A., Izonin, I., Tsymbal, Y., Havrysh, B.: Imbalance data classification via neural-like structures of geometric transformations model: Local and global approaches. *Advances in Intelligent Systems and Computing*, 754, pp. 112-122 (2019) doi: 10.1007/978-3-319-91008-6_12
23. Peleshko, D., Ivanov, Y., Sharov, B., Izonin, I., Borzov, Y.: Design and implementation of visitors queue density analysis and registration method for retail videosurveillance purposes. *Proceedings of the 2016 IEEE 1st International Conference on Data Stream Mining and Processing, DSMP 2016*, art. no. 7583531, pp. 159-162 (2016) doi: 10.1109/DSMP.2016.7583531
24. Ihaka, R., Gentleman, R.: R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5 (3), pp. 299-314 (1996) doi: 10.1080/10618600.1996.10474713
25. El. Resource: <https://www.bioconductor.org/about/>
26. Beer, D.G., Kardia, S.L.R., Huang, C.-C., Giordano, T.J., Levin, A.M., Misek, D.E., Lin, L., Chen, G., Gharib, T.G., Thomas, D.G., Lizyness, M.L., Kuick, R., Hayasaka, S., Taylor, J.M.G., Iannettoni, M.D., Orringer, M.B., Hanash, S.: Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nature Medicine*, 8 (8), pp. 816-824 (2002) doi: 10.1038/nm733
27. El. Resource: <https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-6857/>
28. Baldi, P., Hatfield, G.W.: DNA Microarrays and gene expression: From experiments to data analysis modeling. Cambridge University Press, pp. 22-23 (2002)

Automated information system for the rehabilitation of post-stroke patients in residual period

Alexander Azarkhov¹[000-0003-0062-0616], Irina Fedosova²[0000-0003-3923-8270],
Tetiana Levytska³[0000-0003-3359-1313], Vasily Efremenko⁴[000-0002-4537-6939],
Oleksandr Cheiliakh⁵[0000-0002-5794-9036]

¹Priazovsky State Technical University, Universitetskaya str., 7, Mariupol, 87555, Ukraine
alexazarhov@gmail.com

²Priazovsky State Technical University, Universitetskaya str., 7, Mariupol, 87555, Ukraine
irivasilevna@gmail.com

³Priazovsky State Technical University, Universitetskaya str., 7, Mariupol, 87555, Ukraine
tlevitiisys@gmail.com

⁴Priazovsky State Technical University, Universitetskaya str., 7, Mariupol, 87555, Ukraine
vgefremenko@gmail.com

⁵Priazovsky State Technical University, Universitetskaya str., 7, Mariupol, 87555, Ukraine
cheylyakh_o_p@pstu.edu

Abstract. The article outlines the theoretical and practical aspects of the development of new approaches for improving the effectiveness of health care based on information computational technologies. A medical information system for the process of rehabilitation and rehabilitation of post-stroke patients was designed and implemented, based on structuring each of the levels according by relevant criteria, which allows to take into account the effect of endogenous and exogenous risks of patients' therapy in residual period.

Keywords: medical information system, post-stroke patient, automated computer system, rehabilitation process, program complex, infological scheme

1 Introduction

Medical information systems (MISs) and technologies provide intellectual support for the physician's activities, speeding up the decision making while improving the quality and reliability of the treatment tactics. This is particularly important for the area of treatment of patients with disorders of cerebral circulation, the number of which is increasing worldwide every year. Among the causes of primary disability of the adult population, stroke ranks the second position. In the structure of total mortality stroke accounts 11.1 %, sharing fifth position with diseases of the nervous system, which is one of the highest indexes in the world [1].

According to the data of the World Health Organization, 4.6 million people die of a stroke every year, which is 9-12 % of all causes of death. Among adults, 25 % of cases of disability are due to stroke, and only 10–20 % of stroke patients return to work. After a stroke the cognitive impairment, hemiparesis and speech disorders are

observed in 33 %, 30 % and 27 % cases, accordingly.

Neuro-rehabilitation can be classified according to the duration of the recovery period as follows: a) an early recovery period - up to 6 months after the acute phase of the disease; b) late recovery period - from 6 months to 1 year after the acute phase of the disease; c) residual period - after 1 year [2]. The post-stroke treatment of patients is complicated task which requires the development of novel approaches including information computational technologies.

2 Literature review

The systemic study of the residual period of post-stroke rehabilitation was not conducted in Ukraine previously. In particular, the risks in the rehabilitation period were not classified. The concept of monitoring the health status of post-stroke patients was not developed while the principles of the continuity of rehabilitation activities were not worked out. Notably, there was no specialized medical information systems focused of post-stroke patients' rehabilitation.

The development of medicine, focused on rehabilitation and maximum restoration of human functional capabilities, includes new approaches for building up-to-date information technologies for managing sanatorium-resort activities [3-5]. Among the current problems there is the creation of integrated informational systems aimed to optimizing information flows for bringing up the medical care to a new level [6, 7].

Despite the achievements in this field [8-11], the high efficiency of information support in the rehabilitation process during the residual period was not reached yet. This could be possible on the basement of integrating the medical treatment/rehabilitation process with computer models, algorithms and other hardware/software tools. This link allows to record, analyze and process the information in order to find appropriate strategy for patients in the residual period.

Recent studies showed that in the residual period there is a fairly large number of ambiguities. This is due to the fact that in the period from 1-2 years to 3-5 years after the acute phase of the disease, its clinical picture may differ significantly for different patients [1]. In this connection, the concepts of "early residual period" and "late residual period" were proposed for implementing into clinical studies and practice [12, 13]. It is emphasized that the main task of neuro-rehabilitation (especially in the later period) is to maintain and stabilize stroke patients for improving the quality of their health [14]. The resolving of the problem of informatization of post-stroke patients' rehabilitation is possible only on the basement of up-to-date information systems and technologies.

There are a number of known information systems used for managing the health care institutions [15, 16]. The "Institute of Medical Rehabilitation Problems" has developed an automated information system for supporting rehabilitation treatment for patients with disorders of the musculoskeletal system. This system covers the process from optimizing the patient's entrance examination to formulating the recommendations on appropriate rehabilitation procedures [17]. The software core of the information system is adapted to the specifics of the treatment/rehabilitation process,

focusing on setting up the workplaces with a wide range of new functionalities.

An automated computer system was developed in “Research Institute of Medical and Social Expertise and Rehabilitation” to predict the results of rehabilitation for cerebral stroke and for traumatic brain injury [12]. This automated system allows to consistently solve the tasks as: (a) the prediction of the results of rehabilitation at different periods of the disease; (b) the probabilistic prediction of outcome of rehabilitation using differentiated diagnostic indicators; (c) the selection of the optimal rehabilitation measures depending on the period of the disease and severity of functional disorders; (d) the developing patients databases with the possibility of statistical processing the results of the examination and rehabilitation; (e) formulating and printing the forecast and recommended rehabilitation measures. The mentioned system offers the physician the tactic of medical rehabilitation taking into account the period of the disease, the functional class of disorders/disability and the clinical and rehabilitation prognosis. The use of this system allows to adequately predicts the outcome of rehabilitation for cerebral stroke (traumatic brain injury) in the early stages. This leads to formulating an optimal set of rehabilitation measures aimed at the maximum reduction in the patient’s disability.

However, the above mentioned medical information systems do not take into account the risks and specifics of the residual period of post-stroke patients. The aim of the paper is to develop an adequate medical information system for the rehabilitation of post-stroke patients in the residual period, focusing on effectiveness of the quality of the rehabilitation and recovery process of post-stroke patients.

3 The structural organization of MIS for the management of rehabilitation treatment of post-stroke patients

The development of a medical information system is based on a specific set of the design principles such as:

- the subject area of MIS functioning is determined by the triad “Patient - Disease – Diagnostic/ Therapeutic Process”. A feature of the subject area is the ability to change the relevant fragments of the system as well as its functionality;
- focusing on medical personnel, which has a fairly high user level relative to computing technology and complex biomedical equipment.

The principle of “New tasks” and the principle of “First leader” were implemented as basic principles for MIS designing [18].

The principle of “New tasks” assumes that the system developers should be sufficiently acquainted with the subject area and the peculiarities of the functioning of information system [4]. The principle of “First leader” requires the participation of the head of the institution in the process of development and implementation of an information system. This is connected with the need of understanding the problems which could be caused by implementation of computer system with further managing decisions made by the first manager.

MIS should comply the following design principles: comfort for user, modularity,

adaptability, functional completeness, openness of the system. The block-diagram of developed MIS for sanatorium-resort rehabilitation of post-stroke patients is shown in Figure 1.

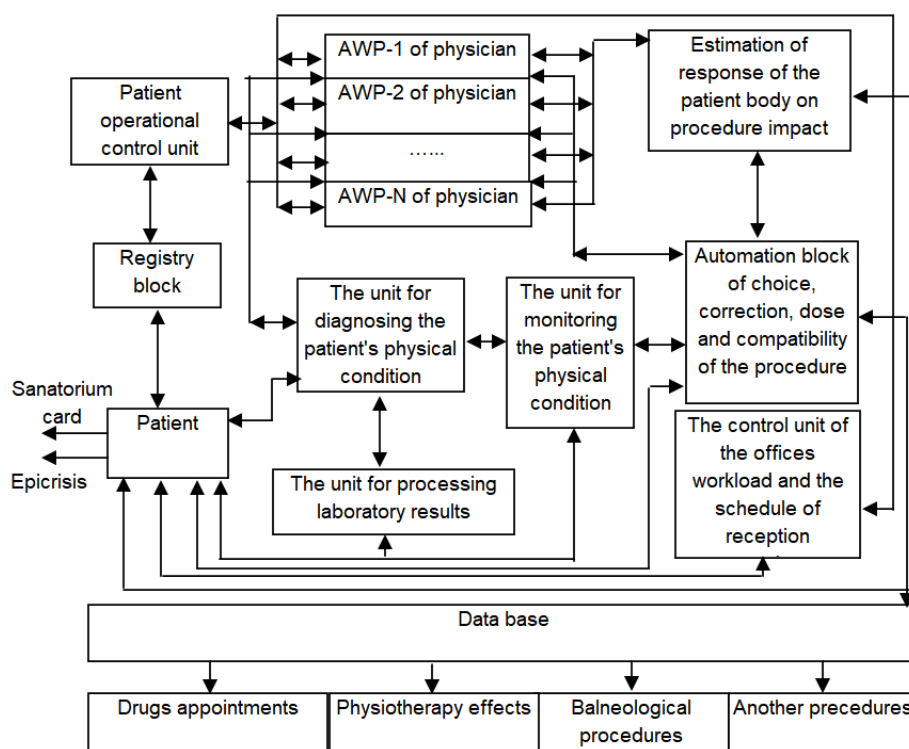


Fig. 1. Structural diagram of the operational dispatch management of post-stroke patients' rehabilitation

The main feature of the health care system in sanatorium-resort institutions is the close intersection of management tasks and medical care technology, which are based on the use of resort factors.

4 Functional organization of MIS

The complex of functional capabilities of MIS is formed on the basis of the requirement to reduce the risks predicted for each of the components of the triad "Patient - Disease – Diagnostic/Therapeutic Process" [19]. The features and capabilities of the created MIS are:

- MIS ensures the formation and management of patient flows having disorders of the musculoskeletal system, which require specific approach for rehabilitation tactic;
- MIS contains the complex of information support for modern telemedicine technologies and formation of medical electronic passport of a patient;

- MIS is able to analyze and forecast medical and financial activities of a sanatorium ensuring optimal management in order to reduce the medical and financial risks.

Application of designed MIS provides post-stroke patients with the following options:

- focusing on quantitative and qualitative assessment of the functional and adaptive capabilities of the human body for the health restoration/correction in order to elevate patient's social status;
- a significant improvement in the quality of health care services through the information interaction between doctors and specialists taking part in the treatment of the patients;
- continuous monitoring and using the patients' database through a personal medical electronic passport;
- formation of individual or personalized programs of rehabilitation and medical prevention, both primary and secondary;
- the use of common criteria for assessing the effectiveness and adequacy of therapeutic and rehabilitation/recreational activities in the clinic, sanatorium and at home.

Ultimately, this leads to an increase in health quality, to prolongation of professional activity, to decreasing premature mortality and disability, to increasing the duration and quality of life.

In turn, the authority of the sanatorium is provided by MIS with the following:

- operational information about the health and functional level of both patients and staff; about the course of the treatment and rehabilitation process; about condition and dynamics of medical place capacity;
- the results of the analysis of the causes of non-compliance with medical prescriptions with further correction of individual treatment and rehabilitation schemes in compliance with medical standards;
- the ability to control and manage the medical, financial and other activities of the sanatorium with the aim of rational targeted expense of financial resources;
- the protection of all medical, financial and other information from unauthorized access;
- the possibility of saving and the subsequent redistribution of financial resources by optimizing the structure of the sanatorium, medical-diagnostic and rehabilitation processes, eliminating duplication of expensive research and appointments.

MIS supplies the medical staff of the sanatorium with the opportunity to ensure the high quality of treatment and rehabilitation of post-stroke patients due to:

- prompt receipt of the necessary information in a standardized form about medical appointments and procedures, results of diagnostic, clinical and other studies, etc.;
- information support of the processes of treatment and decision-making to determine the strategy of treatment of each patient individually;
- informational support for choosing and controlling the therapeutic process with providing the drugs, physiotherapeutic and balneological procedures, etc.;
- information support for the distribution of medical and rehabilitation procedures between the patients taking into account work schedules and workload of the respective medical rooms;
- elimination of duplication of diagnostic studies, procedures and medical prescrip-

tions;

- controlling the performance of staff functional duties by the management;
- controlling the supply and using the medical drugs.

5 User interface

MIS is designed as a computer program “Registry of the automated polyclinic” for automate maintenance of the registry in the clinic. It is aimed at increasing the reliability and efficiency of receiving and controlling the patients flow. Some of the program’s windows are shown in Fig. 2 and Fig. 3.

Регістратура

Картка - Оліф

Дані про пацієнта

Прізвище: Костин

Ім'я: Сергей

По-батькові: Володимирович

Дата народження: 09.07.1993

Стать: Ч Ж

ID номер пацієнта: 1

Адреса

Область: Вінницька

Район: м. Вінниця

Населений пункт: Келецька

Вулиця: Келецька

Будинок: 126

Квартира: 8

Телефон: 432292

Прізвище	Ім'я
Костин	Сергей
Куба	Юлія
Матв	Віра
Романенко	Татьяна

Медичні відомості

Збереження відомостей

1/4

Fig. 2. The window of the module “Registry of the automated polyclinic”

Бібліотека тестових методик

Файл Тест

Анкета

Прізвище: Костин

Ім'я: Сергей

По-батькові: Володимирович

Дата народж.: 09.07.1987

Список тестів

Тести самоопитування	Тести опитування лікарем	Тести фізичні
☒ Самоопитник депресії	☐ Шкала Гамільтона	☐ Швидкість ходи
☐ Тест Кренеліна	☐ Шкала NIHSS	☐ Динамометрія
☐ Тест Бенгтона	☒ Тест Річі	☒ Проба Озерського
☐ Тест ЛІ	☐ Загальна шкала порушень	
☐ Проба Шульце	☐ Тест Еартела	

Розрахунок тестування (Ctrl+H)

Fig. 3. The window of the testing methods' kit

The program “Registry of the automated polyclinic” allows to enter information about patients into the database providing the doctor with necessary information about the patient. It also provides an effective control of the workload of the clinic.

The program allows a convenient widescreen search on such criteria as the family name, date of birth, the region and the area of the residence etc. These parameters can be used when searching both individually and in a complex.

The fields used in the program are the patient biography, address, contact information and medical information about his visits to clinic, information about examinations, diagnoses, previous Botkin's disease, the fluorography, the severity of diabetes and other necessary information.

After these data are entered and stored in a database, they become accessible to any user from any workstation under appropriate access right. Under the first program run the connection to server database is needed. On subsequent runs the connection will occur automatically.

The program complex involves thirteen test methods, five of which are classical, which the patient must undergo, and five more are the test methods that the doctor must consider, basing on the survey and examination of the patient. Three more methods are the physical tests to indicate the patient health state.

The structure of the physical test is different, since it is envisaged that the patient will pass it separately. Therefore the only thing that interests the doctor is the result of the test and the date of its passage.

6 Automated workplace (AWP) of the neurologist

The hierarchy of every medical system includes several levels such as regional, medical institution and local (the human body). Each of them is characterized by its own functioning laws and the range of tasks [20].

At the local level, automated systems are widely used to improve the efficiency and quality of medical care due to computer collecting, processing, storing, presenting and using medical information necessary to adequately address diagnostic and treatment tasks.

As a rule, for each patient, all stages of the therapeutic and diagnostic process are to be reflected in chronological order in certain medical records. The doctor and other medical staff involved, make new records reflecting the nature of their activity and its specific results. A typical infological scheme of the doctor's automated workplace and informational interconnections between the individual components are presented in Figure 4.

The diagnosis formulated for various diseases is formed in a single format which reflects: the etiology of the disease; clinical (clinical and morphological) version of the disease; phase of the disease (remission, exacerbation, etc.); flow stage (initial, unfolded, etc.); some of the most pronounced syndromes; complications.

An important step in the development and operation of the doctor's AWP is certification of the AWP, which is carried out at three levels: technical, organizational, and in terms of working and safety conditions.

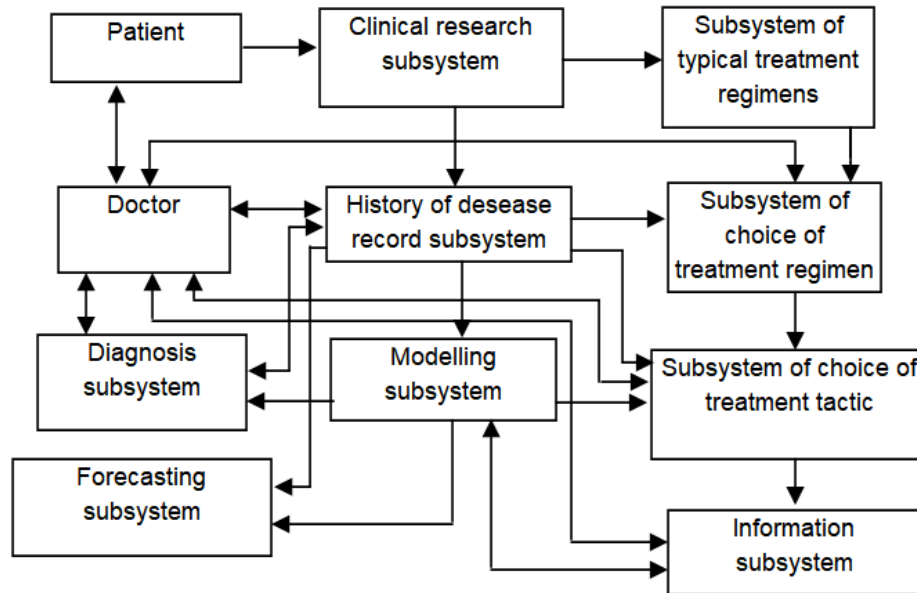


Fig. 4. Infological scheme of the automated workplace of the neurologist

In the process of certification, AWP receives a comprehensive assessment, defined as the arithmetic average of three group coefficients namely: assessment at technical level (K_1), assessment at organizational level (K_2), and assessment at the level of working conditions and safety (K_3). Each of these levels, in turn, is also defined as the arithmetic average of the estimates for each of the elements (criteria) that make up this level. According to the results of AWP certification, each workplace should be assigned to one of three categories: certified, conditionally certified, not certified. Schematically, the work on optimization and certification of automated doctor's workplaces can be represented by the scheme shown in Fig. 5.

The generalized results of certification allow to obtain systematic information about using doctors' AWP in any clinical or sanatorium structure in order to optimize their structure, functional and operational capabilities.

7 Experiments and results

Achieving an optimal level of health care is possible basing on an integral information technology for managing the treatment and diagnostic process, which includes the planning, motivation and control, diagnosis, rehabilitation, and developmental sequence according to present-day medical standards.

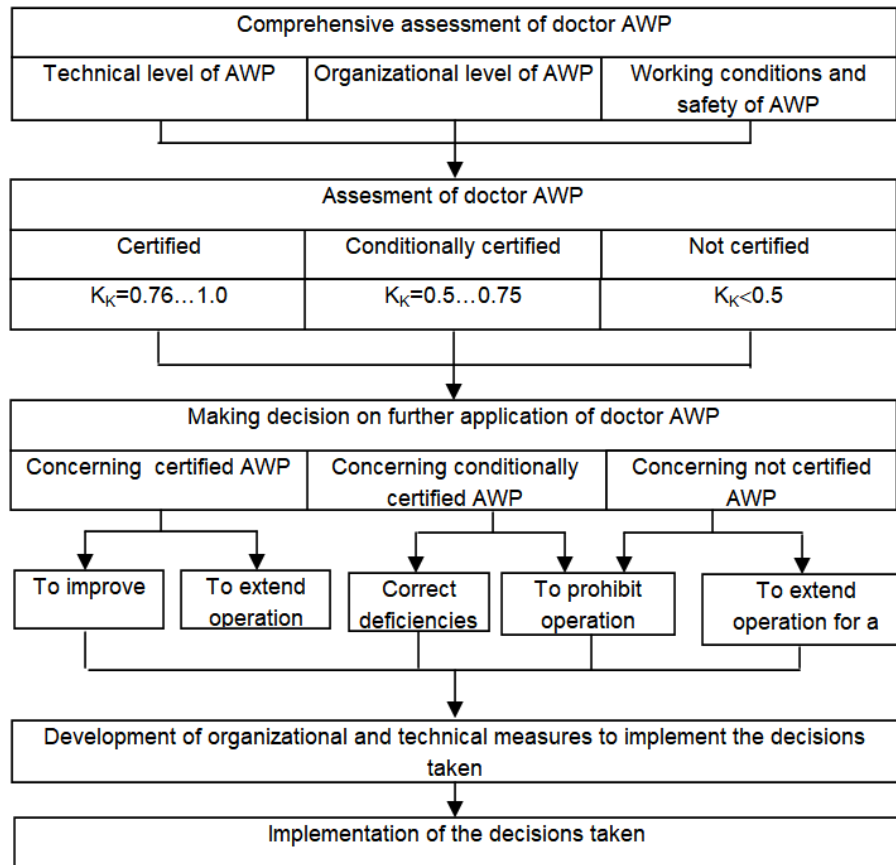


Fig. 5. Scheme of optimization and certification of doctor's automated workplace

One approach to rise an effectiveness of treatment and rehabilitation procedure (TRP) is applying the method of sequential comparison using standard estimates. According to this method, the efficiency is determined by expertly calculating the values of individual indicators according to groups of criteria and measuring their relative significance using an interval scale. At the first stage, the organizational system of TRP is identified basing on the goals, functions and resources of this system. The second stage is the formation of performance criteria. The final procedure for the formation of a system of criteria is their ranking by the degree of influence on the efficiency of TRP [21]. The ranking is carried out on a quantitative scale in the range from 0 to 10. The maximum rating is assigned to the indicator having the greatest advantage. If the expert's mark of the i -th attribute is designated as a_{ij} , then the relative weight of the indicator (V_i) can be calculated using the formula:

$$V_i = \frac{\sum_i a_{ij}}{\sum_i \sum_j a_{ij}} \quad (1)$$

The indicator having the greatest weight gets rank 1.0. For each of the most important criteria, a utility scale is developed with the range from 0 to 1.0. The main purpose of the scale is to convert dissimilar gauges to their equivalent points. An example of constructing such a scale is presented in Figure 6.

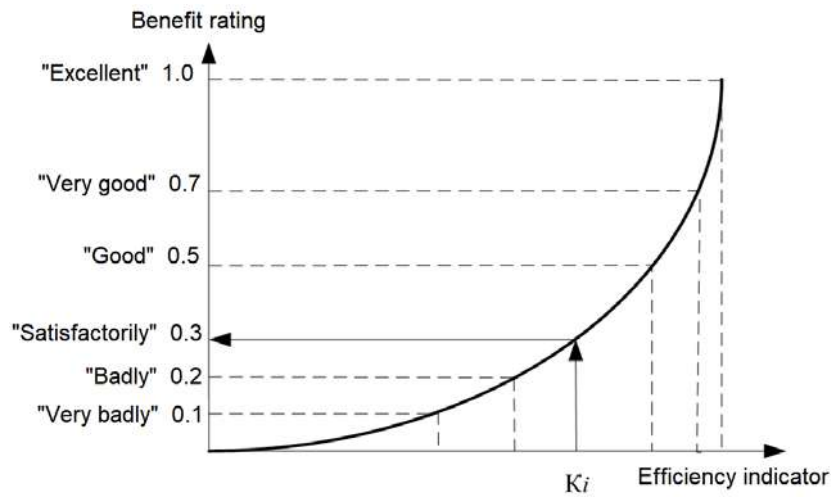


Fig. 6. The scale of utility for assessing the effectiveness of TRP

The third stage includes the calculation of the actual values of the indicators and the conversion of the obtained numerical estimates into points using the utility scale.

The fourth stage is the analysis of the effectiveness of TRP organization. This stage includes the calculation of reserves of efficiency and the relative assessment of the significance of the criteria for improving the effectiveness of the system. Standardized indicators and normalized assessments are used to assess the effectiveness (E) of the treatment or rehabilitation as:

$$E = \frac{A_{br} - A_{ar}}{A_{br}} \cdot 100 \% \quad (2)$$

where A_{br} and A_{ar} are the assessment of the state before and after rehabilitation, respectively.

MIS was probated upon pilot project conducted in sanatorium “Metallurgist” in city of Mariupol. The project allowed to compare the sanatorium activity indicators as a result of MIS application. The effectiveness of medical information system was studied through an anonymous survey of the users (47 people in total). Twenty nine persons of them were the doctors, the rest were employees and nursing staff. The

persons who had previous computer experience were: doctors (72 %), employees (29 %), nursing staff (38 %). The duration of work with the system ranged from 10 to 18 months, while almost half of the users worked with MIS throughout the entire period.

Analysis showed (Fig. 7) that MIS implementation led to significant improvement of all the considered indicators.

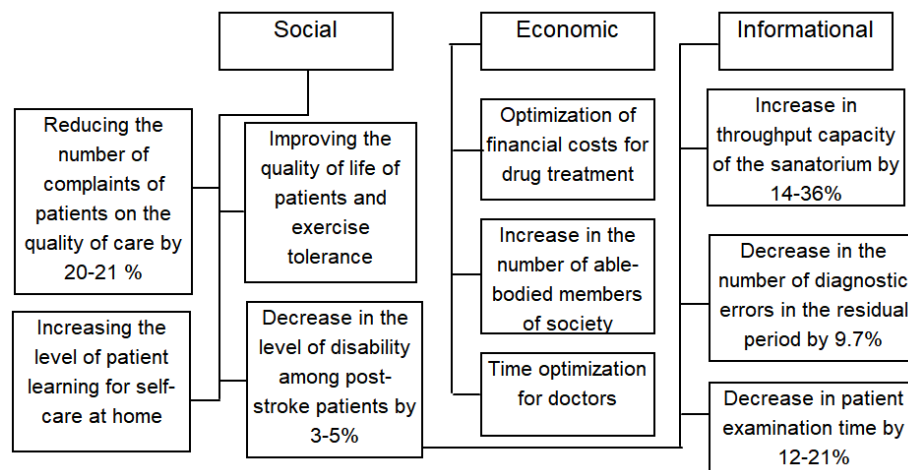


Fig. 7. The results of MIS usage in practice

Specifically, the doctors' time decreased: for the doctor core work - by 25 %, for paperwork - by 33 %, for patient examination - by 12-21 %. The number of patients served during the same time increased by 14-36 %. The number of complaints for the quality of medical care decreased by 20-21 %; the number of medical errors decreased by 9.7 %. The time spent by the doctors for professional development in relation to the total working time increased by 15 %.

If the overall assessment of the rehabilitation treatment is rated as: 1 - excellent, 2 - good, 3 - poor, then the analysis showed the following. The doctors evaluated rehabilitation treatment without MIS as 1.9 and with MIS as 1.0. The patients evaluated rehabilitation treatment without MIS as 1.7 and with MIS as 1.1. This reflects the improvement of rehabilitation strategy and treatment measures.

The presented results show the high effectiveness of developed medical information system.

8 Conclusion

The analysis of the existing approaches of information support for the rehabilitation treatment of post-stroke patients allowed to set the measures aimed at minimizing the risks of complications of the central and peripheral nervous systems. It also resulted

in development of criteria and principles for more effective information support. Some of these principles and criteria allowed to substantiate the strategy of managing the residual period of rehabilitation of post-stroke patients. The concept of monitoring the state of post-stroke patients and principles for ensuring the continuity of rehabilitation measures using medical information system were formulated.

A medical information system for the rehabilitation of post-stroke patients was developed in this work. The system is based on structuring each of the levels according to relevant criteria, which allows to take into account the effect of endogenous and exogenous risks.

Two-step control of medical appointments was implemented based on information model of the rehabilitation system, providing an automated control of appointments and procedures for compliance with the diagnosis with further adjusting the rehabilitation procedures. This resulted in decreasing the number of prescribing drugs and therapeutic procedures which were inadequate to the state of the patients.

References

1. Kuklina, E.V., Tong, X., George, M.G., Bansil, P.: Epidemiology and prevention of stroke: a worldwide perspective. *Expert Review of Neurotherapeutics*. 12(2), 199–208 (2012). doi: 10.1586/ern.11.99
2. Lee, K.B., Lim, S.H., Kim, K.H., Kim, K.J., Kim, Y.R., Chang, W.N., Yeom, J.W., Kim, Y.D., Hwang, B.Y.: Six-month functional recovery of stroke patients: a multi-time-point study. *International Journal of Rehabilitation Research*. 38(2), 173–180 (2015). doi: 10.1097/MRR.000000000000108
3. Afandi, R.R., Radman, A., Bahari, M., Zakaria, L.Q., Mustapha, M., Ismail, W.: Ontology development in patients information system for stroke rehabilitation. In: 8th International Conference on Biomedical Ontology, ICBO 2017. pp 1-5. *CEUR Workshop Proceedings* 2137, Newcastle (2017)
4. Berg, M.: The search for synergy: interrelating medical work and patient care information systems. *Methods of Information in Medicine*. 42(4), 337–44 (2003). doi: 10.1267/METH03040337
5. Mitrovic, A., Mathews, M., Ohlsson, S., Holland, J., McKinlay, A.: Computer-based post-stroke rehabilitation of prospective memory. *Journal of Applied Research in Memory and Cognition*. 5(2), 204–214 (2016). doi: 10.1016/j.jarmac.2016.03.006
6. Du, H.-S., Ma, J.-J., Li, M.: High-quality health information provision for stroke patients. *Chinese Medical Journal*. 129(17), 2115–2122 (2016). doi: 10.4103/0366-6999.189065
7. Nam, H.S., Park, E., Heo, J.H.: Facilitating stroke management using modern information technology. *Journal of Stroke*. 15(3), 135–143 (2013). doi: 10.5853/jos.2013.15.3.135
8. Bohil, C.J., Alicea, B., Biocca, F.A.: Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*. 12(12), 752–762 (2011). doi: 10.1038/nrn3122
9. Lin, W.-Y., Chen, C.-H., Tseng, Y.-J., Tsai, Y.-T., Chen, C.-K.: Predicting post-stroke activities of daily living through a machine learning-based approach on initiating rehabilitation. *International Journal of Medical Informatics*. 111, 159–164 (2018). doi: 10.1016/j.ijmedinf.2018.01.002

10. Burilich, I.N., Korenevskiy, N.A., Shtotland, I.M.: Complex diagnostics of functional states using psychological and physiological experiments data. *Harold of New Medical Technologies*. 10(3), 44-46 (2003)
11. Nam, H.S., Cha, M.J., Kim, Y.D., Kim, E.H., Park, E., Lee, H.S.: Use of a handheld, computerized device as a decision support tool for stroke classification. *European Journal of Neurology*. 19, 426-430 (2012). doi: 10.1111/j.1468-1331.2011.03530.x
12. Belousova, O.V., Azarkhov, O.Y., Zlepko, S.M., Shtofel, D.Ch., Kostishyn, S.V.: Model of prognostication of complications and rehabilitation process control of post-apoplectic patients. *Bulletin of the National Technical University "KhPI". A series of "Information and Modeling"*. 36, 23-28 (2011) (in Russian)
13. Zlepko, S., Sierhieieva, V., Azarkhov, O., Makogon, V.: Biomedical system for emotional stress evaluation. In: 11th International Conference of Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET'2012. pp. 230 (2012) (in Russian)
14. Prokopenko, S.V., Mozheyko, E.Y., Petrova, M.M., Koryagina, T.D., Bezdenezhniy, A.F.: Correction of post-stroke cognitive impairments using computer programs. *Journal of the Neurological Sciences*. 325(1-2), 148-153 (2013)
15. Park, E., Nam, H.S.: A service-oriented medical framework for fast and adaptive information delivery in mobile environment. *IEEE Transactions on Information Technology in Biomedicine*. 13, 1049-1056 (2009). doi: 10.1109/TITB.2009.2031495
16. Roy, P.M., Durieux, P., Gillaizeau, F., Legall, C., Armand-Perroux, A., Martino, L.: A computerized handheld decision-support system to improve pulmonary embolism diagnosis: a randomized trial. *Annals of Internal Medicine*. 151, 677-686 (2009). doi: 10.7326/0003-4819-151-10-200911170-00003
17. Koziavkin, V.I.: Information security system for the Institute for problems of medical rehabilitation. *Ukrains'kyi Visyk Psykhonevrolohi*. 8(2), 10-12 (2000) (in Russian)
18. Ay, H., Benner, T., Arsava, E.M., Furie, K.L., Singhal, A.B., Jensen, M.B.: A computerized algorithm for etiologic classification of ischemic stroke: the Causative classification of stroke system. *Stroke*. 38, 2979-2984 (2007). doi: 10.1161/STROKEAHA.107.490896
19. Kozelkin, A.A., Revenko, A.V., Medvedkova, S.A., Subbotovskaya, L.V., Tolstikova, E.D.: Modern aspects of neurorehabilitation of post-stroke patients. *International Neurological Journal*. 8, 20-24 (2010) (in Russian)
20. Kahraman, A., Jones, F.: Factors influencing and shaping the lived experience after stroke: A systematic review of qualitative studies. *JBHI Library Systematic Review*. 7(8 Suppl), 1-12 (2009)
21. Bogale, M.A., Yu, H., Sarkodie-Gyan, T., Abdelgawad, A.: Characterization and quantification of gait deficits within gait phases using fuzzy-granular computing. *J. Biomedical Science and Engineering*. 5, 720-728 (2012). doi: 10.4236/jbise.2012.512090

Inversion of Dynamic Systems for Certain Classes of Signals

Alexander Kutsenko¹[0000-0001-6059-3694], Sergey Kovalenko¹[0000-0001-6770-6778], Vladimir Tovazhnyansky¹[0000-0002-0713-3514],

¹National Technical University "Kharkiv Polytechnic Institute", 2, Kyrpychova Str., 61002, Kharkiv, Ukraine
kuzenko@kpi.kharkov.ua, adbc@ukr.net, vtovazhnianskyi@gmail.com

Abstract. Methods of inversion of dynamic systems are widely used for solving problems of control of mechanical and electrical systems. Solving the inversion problems raises a number of difficulties related to the high sensitivity of the results with respect to the accuracy of setting the parameters of a mathematical model of object parameters, instability in controlling non-minimum phase objects, and violation of the conditions of physical realizability.

In this paper, an approximate method of solving the inversion problem for linear stationary dynamic systems is proposed which is largely free from those disadvantages. The method is based on the representation of the input and output signals by their approximations in the linear space of specially selected D -functions of time. The feature of the proposed method of inversion of dynamic systems is the representation of multidimensional polynomials approximating the input and output signals as a product of rectangular matrices and a vector of powers of time.

Mathematical models of linear dynamic systems in the form of differential equations in the state space and in the equivalent input-output form, as well as SISO and MIMO dynamical systems are considered in the paper.

Keywords: dynamical systems, polynomial signals, quasi-harmonic functions, matrix equations.

1 Introduction

The inversion problem of dynamic systems has an extensive bibliography and a long history. The papers [1, 2] can be considered as the fundamental work in this direction, where the criteria and methods for constructing inverse operators are justified. A significant contribution to the development of the theory and practice of inversion of dynamic systems was made in the works [3–5]. In them new criteria for the invertibility of linear dynamic systems are proposed and specific ways to solve the inversion problem are given. A number of practical results of solving inversion problems with regard to electrical and mechanical systems are given in [6, 7].

The inversion problem become of particular importance due to the solution of the problem of the synthesis of combined automatic control systems. Various aspects of the inversion problem for combined control systems in the most general statement are presented in [4]. Despite significant progress in solving the inversion problem of dynamic systems, in practice there are a number of difficulties associated with the high sensitivity of the results to the accuracy of the parameters of the mathematical model of a controlled object, the instability in controlling non-minimally phase objects, and violation of conditions of physical realizability of inverse operators. Generally, the listed problems do not allow us to find a practically realizable solution of the problem of finding the inverse operator in the control problem. Nevertheless, for solving a number of practical problems, it seems natural to consider approximate mathematical models of a controlled object and signals at its inputs and outputs, for which the inversion problem has the correct solution. In combined control systems, these assumptions are compensated for by the deviation control loop.

Thus, the purpose of this work is to develop an approximate, practically realizable numerical method for solving the inversion problem for linear dynamic systems.

2 Statement of the research problem

We will consider linear stationary dynamical systems whose mathematical models are presented either in the form of equations of state

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx,\end{aligned}\tag{1}$$

where $x \in R^n$ is the state vector, $u \in R^m$ is the control vector; $y \in R^s$ the output vector; A, B, C are matrices of the corresponding dimensions, or in the equivalent form “input-output”:

$$A_0 y^{(p)} + A_1 y^{(p-1)} + \dots + A_p y = B_0 u^{(q)} + B_1 u^{(q-1)} + \dots + B_q u, \tag{2}$$

where $A_0, A_1, \dots, A_p$ are $(s \times s)$ matrices, $B_0, B_1, \dots, B_q$ are $(s \times m)$ matrices. In the following, we will assume that the controlled system under consideration is asymptotically stable, and the dimensions of the control vector and the output vector coincide, i.e. $s = m$.

Consider the linear space of continuous Φ of continuous differentiable vector functions $\varphi(t)$, which satisfy the condition

$$\frac{d\varphi(t)}{dt} \in \Phi. \tag{3}$$

These D -functions include the class of functions of the form

$$\varphi(t) = \sum_{k=1}^N e^{\alpha_k t} (R_k(t) \sin \omega_k t + Q_k(t) \cos \omega_k t), \quad (4)$$

where α_k , ω_k are some constants, but $R_k(t)$ and $Q_k(t)$ are vector polynomials of the degree not higher than l . It is not difficult to verify that a function of the form (4) satisfies condition (3). For different values of the parameters α_k , ω_k , $R_k(t)$, $Q_k(t)$ different classes of D -functions can be obtained: polynomials, trigonometric polynomials, quasi-harmonic functions.

Prove the following statement: if the input action $u(t)$ is a D -function of the Φ class, then the forced response of the dynamic system (1) is also a function of the Φ class.

To prove the statement, we will seek a solution of equation (1) in the form of an infinite series

$$x(t) = \sum_{k=1}^{\infty} C_k u^{(k-1)}(t), \quad (5)$$

where C_k are some $m \times n$ matrices to be determined.

After substitution (5) in (1) we get

$$\sum_{k=1}^{\infty} C_k u^{(k)} = \sum_{k=1}^{\infty} A C_k u^{(k-1)} + B u. \quad (6)$$

By equating matrix coefficients of derivatives $u^{(k)}$ of the same order in the left and right sides of (6), we obtain a system of recurrence relations for calculating matrices C_k :

$$A C_1 + B = 0, \quad A C_{k+1} = C_k, \quad k = \overline{1, \infty}$$

of which directly follows

$$C_1 = -A^{-1}B, \quad C_2 = -A^{-2}B, \dots, C_k = -A^{-k}B, \dots \quad (7)$$

Thus, the output response of the dynamic system (1) at zero initial conditions will take the form

$$y(t) = -C \sum_{k=1}^{\infty} A^{-k} B u^{(k-1)}(t). \quad (8)$$

Since all $u^{(k-1)}(t)$ in (8) belong to the class Φ , their linear transformation $y(t)$ is also an element of Φ , i.e. the statement is proved.

Within these assumptions about the structures of the dynamic system and the signals at the inputs and outputs, the statement of the inversion problem can be formulated as follows: find an input action $u(t)$ on an interval $[t_0, t_1]$ belonging to a certain

class Φ of D -functions, under which the output $y(t)$ of the system (1), (2) will be a specified D -function of the same class Φ at zero initial conditions.

3 Inverting dynamic systems in a class of polynomials

Let the set Φ of input and output signals be a set of vector polynomials of degree not higher than l

$$\varphi(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_l t^l, \quad (9)$$

where $a_0, a_1, \dots, a_l$ are n -dimensional vectors of coefficients.

Instead of polynomials of the form (9), we will consider equivalent polynomials

$$\varphi(t) = b_0 + b_1 t + b_2 \frac{t^2}{2!} + b_3 \frac{t^3}{3!} + \dots + b_l \frac{t^l}{l!}, \quad (10)$$

where vectors a_k and b_k are connected by the ratio $b_k = a_k k!$, $k = \overline{0, l}$.

Further, the vector polynomial (10) will be represented in a vector-matrix form.

$$\varphi(t) = BT, \quad (11)$$

where B is an $n \times (l+1)$ -dimensional matrix, whose columns correspond to the vector coefficients of the polynomial (10), and $T = \left(1, t, \frac{t^2}{2!}, \dots, \frac{t^l}{l!}\right)^T$ is the $(l+1)$ -dimensional column vector.

Representation of multidimensional polynomials in the vector-matrix form (11) will allow in the future to effectively apply matrix methods to solving specific inversion problems of dynamic systems.

Consider some elementary operations with polynomials and their analogues in vector-matrix form:

- addition: $\varphi_1(t) + \varphi_2(t) = (B_1 + B_2)T$;
- multiplication by number: $\alpha\varphi(t) = (\alpha B)T$;
- multiplication by the matrix C on the left: $C\varphi(t) = CBT$;
- differentiation:

$$\frac{d\varphi(t)}{dt} = B\Lambda T, \quad (12)$$

where the elements of the $(l+1) \times (l+1)$ dimension matrix Λ are the form

$$\lambda_{ij} = \delta_{i, j+1}, \quad i, j = \overline{1, l+1}, \quad (13)$$

where δ_{ij} is the Kronecker delta.

Subsequent derivatives are found in accordance with the formula

$$\frac{d^k \varphi(t)}{dt^k} = B \Lambda^k T, \quad (14)$$

where the elements of the matrix Λ^k are found by the formula which is similar to (13)

$$\lambda_{ij}^k = \delta_{i,j+k}, \quad i, j = \overline{1, l+1}. \quad (15)$$

Consider first the SISO system defined in the form of "input-output" (2)

$$a_0 y^{(p)} + a_1 y^{(p-1)} + \dots + a_p y = b_0 u^{(q)} + b_1 u^{(q-1)} + \dots + b_q u, \quad (16)$$

where $a_0, a_1, \dots, a_p, b_0, b_1, \dots, b_q$ are constant coefficients, $p \geq q$.

The polynomial signals at the input and at the output will be represented in a vector-matrix form

$$y(t) = YT, \quad u(t) = UT, \quad (17)$$

where Y and U are $(l+1)$ -dimensional row vector composed of the coefficients of the polynomials $y(t)$ and $u(t)$.

After substitution (17) into (16) we get

$$a_0 Y \Lambda^p T + a_1 Y \Lambda^{p-1} T + \dots + a_p Y T = b_0 U \Lambda^q T + b_1 U \Lambda^{q-1} T + \dots + b_q U T. \quad (18)$$

From (18) directly follows

$$Y(a_0 \Lambda^p + a_1 \Lambda^{p-1} + \dots + a_p E) = U(b_0 \Lambda^q + b_1 \Lambda^{q-1} + \dots + b_q E). \quad (19)$$

Introducing the corresponding notations, we rewrite (19) in the form

$$\overline{Y} \overline{A} = \overline{U} \overline{B}, \quad (20)$$

where $\overline{A}$ and $\overline{B}$ are the square $(l+1) \times (l+1)$ lower triangular matrices, which are formed in accordance with the following rules:

$$a_{ij} = \begin{cases} 0 & \text{if } j+p < i \vee j < j, \\ a_{j-i+p} & \text{if } j+p \geq i \vee i \geq j, \end{cases} \quad i, j = \overline{1, l+1}.$$

$$b_{ij} = \begin{cases} 0 & \text{if } j+q < i \vee i < j, \\ b_{j-i+q} & \text{if } j+q \geq i \vee i \geq j, \end{cases}$$

and have the form

Table 1. View of the lower triangular matrix $\bar{A}$

$\bar{A} =$		1	2	3	4	...	$l+1$
	1	a_p	0	0	0	...	0
	2	a_{p-1}	a_p	0	0	...	0
	3	a_{p-2}	a_{p-1}	a_p	0	...	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
	$p+1$	a_0	a_1	a_2	a_3	...	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
	$l+1$	0	0	0	0	...	a_p

Table 2. View of the lower triangular matrix $\bar{B}$

$\bar{B} =$		1	2	3	4	...	$l+1$
	1	b_q	0	0	0	...	0
	2	b_{q-1}	b_q	0	0	...	0
	3	b_{q-2}	b_{q-1}	b_q	0	...	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
	$q+1$	b_0	b_1	b_2	b_3	...	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
	$l+1$	0	0	0	0	...	b_q

Let us analyze the relation (20). It is easy to see that (20) is a linear mapping between vectors Y and U , written in a symmetric form. Since $\bar{A}$ and $\bar{B}$ are lower triangular matrices with diagonal elements $a_{ii} = a_p$ and $b_{ii} = b_q (i = \overline{1, l+1})$, then $\bar{A}$ and $\bar{B}$, in the general case, are non-degenerate and system (20) has a unique solution.

If a solution of the direct control problem is sought, then $Y = U\bar{B}\bar{A}^{-1}$.

In the case that is of interest to us, the solution of the inversion problem, the desired vector of coefficients U is found from the relation $U = \bar{Y}\bar{A}\bar{B}^{-1}$.

The matrix $\bar{B}^{-1}$ can be calculated as follows.

$$\bar{B}^{-1} = \begin{pmatrix} \frac{1}{b_q} & 0 & 0 & \dots \\ -\frac{\Delta_1}{b_q^2} & \frac{1}{b_q} & 0 & \dots \\ \frac{\Delta_2}{b_q^3} & -\frac{\Delta_1}{b_q^2} & \frac{1}{b_q} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $\Delta_1, \Delta_2, \dots, \Delta_l$ is the sequence of minors of matrix $\bar{B}$, in which each successive one is formed as a result of the bordering of the previous minor, starting from $\Delta_1 = b_{q-1}$ – the element of the matrix b_{21} .

An effective method for calculating minors Δ_k is a method based on the consistent application of the Schur and Frobenius formulas [8] for calculating the determinant and inversion of block matrices. From the computational point of view, the solution (20) with respect to the vector U is easier to find by sequential calculation of the components U , starting from u_l . In this case, due to the triangular structure of the matrix $\bar{B}$, a linear equation with one unknown is formed at each step of the iterative process. Thus, the solution of the inversion problem is found for $l+1$ steps.

Consider the solution of the inversion problem for MIMO systems in the polynomial signals environment. Let the mathematical model of the system under consideration be of the form (1). Then the solution of the direct control problem in the general form is represented in the form (8). Substitute in (8) vector-matrix expressions for polynomials $y(t)$ and $u(t)$, taking into account that the powers of the matrix Λ satisfy the condition $\Lambda^k = 0, \forall k > l$.

As a result, we obtain the analytical relationship between the matrices Y and U :

$$Y = -\sum_{k=1}^{l+1} CA^{-k}BU\Lambda^{k-1}. \quad (21)$$

The result (21) for solving the direct control problem in the polynomial signal environment makes it easy to solve the inverse problem by solving the matrix equation (21) with respect to the matrix U under a specified matrix Y .

The solution of the matrix equation (21) can be obtained by vectoring matrices Y and U and constructions based on the Kronecker product of matrices [9], which leads to a linear system of algebraic equations of $m \times (l+1)$ dimension:

$$\sum_{k=1}^{l+1} CA^{-k}B \otimes (\Lambda^{k-1})^T \text{vec}U + \text{vec}Y = 0, \quad (22)$$

where the column vectors $\text{vec}U$ and $\text{vec}Y$ are composed of transposed rows of matrices Y and U matched in ascending order of the row number.

4 Inverting dynamic systems in the class of quasiharmonic functions

Let the signals at the input and output of a dynamic system have the form

$$\varphi(t) = \sum_{k=1}^N (R_k(t) \sin \omega_k t + Q_k(t) \cos \omega_k t), \quad (23)$$

where $R_k(t)$ and $Q_k(t)$ are vector polynomials of the degree not higher than l , ω_k are some positive numbers.

Using the matrix-vector representation of polynomials, the relation (23) can be written as

$$\varphi(t) = \sum_{k=1}^N (R_k \sin \omega_k t + Q_k \cos \omega_k t) T, \quad (24)$$

where R_k and Q_k are matrices whose rows correspond to the coefficients of the components of vector polynomials in (23).

To find the derivatives of the function (24), we first consider the single-frequency function of the form (24)

$$\varphi(t) = (R \sin \omega t + Q \cos \omega t) T. \quad (25)$$

The sequence of derivative functions of the form (25) can be written as

$$\begin{aligned} \frac{d\varphi}{dt} &= [(R\Lambda - Q\omega) \sin \omega t + (R\omega + Q\Lambda) \cos \omega t] T, \\ \frac{d^2\varphi}{dt^2} &= [(R\Lambda^2 - 2Q\Lambda\omega - R\omega^2) \sin \omega t + (Q\Lambda^2 + 2R\Lambda\omega - Q\omega^2) \cos \omega t] T, \\ &\dots\dots\dots \\ \frac{d^k\varphi}{dt^k} &= (R^k \sin \omega t + Q^k \cos \omega t) T; \end{aligned}$$

where R^k and Q^k are calculated by the recurrence formula

$$(R^k \mid Q^k) = (R^{k-1} \mid Q^{k-1}) \Omega, \quad (26)$$

where Ω is a $2(l+1) \times 2(l+1)$ block matrix

$$\Omega = \begin{pmatrix} \Lambda & \omega E \\ -\omega E & \Lambda \end{pmatrix}, \quad (27)$$

each block of which has a dimension of $(l+1) \times (l+1)$.

Based on the relations (26) and (27), the sequence of matrices of polynomial coefficients of derivatives of a single-frequency function (25) can be written as

$$(R^k \mid Q^k) = (R \mid Q) \Omega^k, \quad (28)$$

where the matrix Ω^k is calculated by successive raising to the power of the matrix (27).

Consider the most general case when an exponential multiplier is present at a specified output of the system, i. e.

$$\varphi(t) = \sum_{k=1}^N e^{\lambda_k t} (R_k(t) \sin \omega_k t + Q_k(t) \cos \omega_k t). \quad (29)$$

Spreading the method for finding the derivatives of a single-frequency quasi-harmonic signal, discussed earlier, we will write the sequence of derivatives of single-frequency functions (29) in the form

$$\frac{d^k \varphi}{dt^k} = e^{\lambda t} [R^k \sin \omega t + Q^k \cos \omega t] T,$$

where

$$\begin{aligned} R^k &= R^{k-1} \Lambda + R^{k-1} \lambda - Q^{k-1} \omega, \\ Q^k &= Q^{k-1} \Lambda + Q^{k-1} \lambda + R^{k-1} \omega, \quad k = \overline{1, l} \\ R^0 &= R, \quad Q^0 = Q. \end{aligned}$$

or in a matrix form

$$\begin{pmatrix} R^k \\ Q^k \end{pmatrix} = \begin{pmatrix} R^{k-1} \\ Q^{k-1} \end{pmatrix} \begin{pmatrix} \Lambda + \lambda E & | & -\omega E \\ \hline -\omega E & | & \Lambda + \lambda E \end{pmatrix}.$$

Introducing the notation

$$\overline{\Omega} = \begin{pmatrix} \Lambda + \lambda E & | & -\omega E \\ \hline -\omega E & | & \Lambda + \lambda E \end{pmatrix},$$

we obtain the final ratio to calculate the coefficients R^k and Q^k of derivatives $\varphi^{(k)}(t)$:

$$\begin{pmatrix} R^k \\ Q^k \end{pmatrix} = \begin{pmatrix} R \\ Q \end{pmatrix} \overline{\Omega}^k, \quad (30)$$

where the power of the matrix $\overline{\Omega}^k$ are calculated by successive raising $\overline{\Omega}$ to a power.

Let the matrix representation of the input and output signals be in the form

$$\begin{aligned} u(t) &= e^{\lambda t} (R_u \sin \omega t + Q_u \cos \omega t) T, \\ y(t) &= e^{\lambda t} (R_y \sin \omega t + Q_y \cos \omega t) T, \end{aligned} \quad (31)$$

where R_u , Q_u , R_y , Q_y are the $q \times (l+1)$ -dimensional matrix of coefficients of polynomial multiplier.

Then the derivatives of these signals in accordance with (30) can be represented as

$$\begin{aligned} \frac{d^k u}{dt^k} &= e^{\lambda t} (R_u^k \sin \omega t + Q_u^k \cos \omega t) T, \\ \frac{d^k y}{dt^k} &= e^{\lambda t} (R_y^k \sin \omega t + Q_y^k \cos \omega t) T, \end{aligned} \quad (32)$$

where the matrix coefficients R_u^k , Q_u^k , R_y^k , Q_y^k are found in accordance with the formula (30).

After substituting the (31) and (32) into (2) and reducing by $e^{\lambda t}$ and T we get

$$\sum_{k=0}^p A_k Y \bar{\Omega}^{p-k} = \sum_{k=0}^q B_k U \bar{\Omega}^{q-k}, \quad (33)$$

where $Y = (R_y \mid Q_y)$ and $U = (R_u \mid Q_u)$ are matrix representations of polynomial factors in front of functions $\sin \omega t$ and $\cos \omega t$ for output and input signals.

The obtained matrix mapping (33) is symmetrical and allows us to find solutions to both direct and inverse control problems by vectoring the desired control of matrix Y and constructions based on the Kronecker product of matrices [9] by analogy with (22).

The obtained result is easily distributed to the SISO system. In this case, the matrices A_k and B_k are scalars a_k and b_k , accordingly, and the matrix equation (33) takes the form

$$Y \sum_{k=0}^p a_k \bar{\Omega}^{p-k} = U \sum_{k=0}^q b_k \bar{\Omega}^{q-k},$$

whose solution with respect to vectors Y or U (for the problem of inversion) is connected to the procedure of inversion of matrices $\sum_{k=0}^p a_k \bar{\Omega}^{p-k}$ or $\sum_{k=0}^q b_k \bar{\Omega}^{q-k}$ of $2(l+1) \times 2(l+1)$ dimension.

5 Software package for solving inversion problems

The software is fully developed in relation to polynomial models of input and output signals and contains the following basic structural blocks:

1. The block of input the initial information i.e. matrices A , B , C and their dimensions for the case of specifying the controlled object in the form (1), and coefficients $a_0, a_1, \dots, a_p$, $b_0, b_1, \dots, b_q$ for SISO systems specified in the form of "input-output" (16).

2. A task formation block that includes a random or fixed-step selection of N values of components of the output vector $y^*(t)$ at a fixed time interval, as well as setting the degree l of approximating polynomials $y^*(t)$ and calculating their coefficients using the least squares method. The approximation of signals by cubic splines is provided.

3. The block of formation of matrices $\bar{A}$ and $\bar{B}$ as well as the matrix of the system of linear equations (22) and the calculation of the condition number $cond(\cdot)$ of systems of linear equations based on the Euclidean norm. If $cond(\cdot) \leq 100$, then the solu-

tion of systems (20) or (22) is followed. Otherwise, the degree of approximating polynomials l is incremented by one.

4. The correctness of the inversion problem solution is controlled by numerical integration of the initial systems of differential equations with zero initial conditions and comparison of the result of integration $y(t)$ to the corresponding values of the initial output function $y^*(t)$.

6 Conclusion

Simplified mathematical models of signals based on polynomials and harmonic functions with polynomial varying amplitude are proposed in the paper.

A matrix representation of polynomial signals was proposed and substantiated, which made it possible to represent controlled dynamic processes as static linear transformations in the space of rectangular matrices. Rather simple and effective algorithms for the numerical solution of the inversion problem, as well as a method for estimating the degree of robustness of the results, are obtained.

References

1. Sain, M. K., Massey, J. L.: Invertibility of linear time-invariant dynamical systems // IEEE Trans. Automatic Control, Vol. 14(2), pp. 141–149 (1969)
2. Silverman, L.M.: Inversion of multivariable linear systems // IEEE Trans. Automatic Control, Vol. 14(3), pp. 270–276 (1969)
3. Ilin, A. V., Korovin, S. K., Fomichev, V. V.: Robust methods of dynamic systems inversion. FIZMATLIT, Moscow (2009)
4. Kostenko, Yu. T., Lyubchik, L. M.: Control systems with dynamic models. Basis, Kharkov (1996)
5. Borukhov, V. T.: Criteria of invertibility of linear stationary multidimensional systems // Avtomatika i Telemekhanika, Vol. 11, pp. 5–11 (1978)
6. Pukhov, G. E., Zhuk, K. D.: Synthesis of multiply connected control systems by the method of inverse operators. Naukova dumka, Kyiv (1966)
7. Krutko, P. D.: Inverse problems of dynamic control systems. Linear models. Nauka, Moscow (1987)
8. Sigorsky, V. P.: Engineering mathematics. Tekhnika, Kyiv (1975)
9. Lancaster, P.: Theory of matrices. FIZMATLIT, Moscow (1978)

Practical Implementation Effectiveness of the Speed Increasing Method of Group Matrix Cryptographic Transformation

Svitlana Sysoienko¹[0000-0002-0009-337X], Iryna Myronets²[0000-0003-2007-9943],
Vira Babenko³[0000-0003-2039-2841],
Cherkasy State Technological University, Shevchenko str., 460, Cherkasy, 18006, Ukraine
¹s.sysoienko@gmail.com, ²irenmir30@gmail.com,
³verababenko84@gmail.com

Abstract. This material is devoted to the development of the speed increasing method of implementing a group matrix cryptographic transformation based on a generalized mathematical model of a group matrix cryptographic transformation, by reducing the complexity of building and implementing an reverse transformation, which provided a decrease in mathematical complexity and an increase in the speed of cryptographic transformation. On the basis of the mathematical apparatus of block matrices, checked the correctness of the mathematical model for constructing the group matrix of reverse cryptographic transformation.

Keywords: operation results, group operations, speed encryption, cryptographic transformation, block matrices, mathematical model, reverse transformation, cryptographic protection, logical functions.

1 Introduction

The cyber-attack environment, in which cybernetic influence is possible, is cybernetic space (cyberspace). Cyberspace is an artificial electronic environment for the existence of information objects in digital form, formed as a result of the functioning of cybernetic computer control and information processing systems and provides users with access to computing and information resources of the systems, the development of electronic computing products, the exchange of electronic messages information images in real time to enter into a relationship (to interact) on the sharing of computing and information resources (the provision of information services, conducting e-commerce, etc.) [1].

Ensuring the confidentiality, integrity, significance of information, protection from illegal actions of users is the basis for the functioning of modern computer systems [2-9]. One of the ways to combat cybercrime is to use cryptography.

ments in Ukraine and the world, the increasing number of attacks on computer systems need to solve new security information tasks that are facing the relevant specialists.

Ensuring the protection of confidential information about the social, political, economic, military, scientific and technological condition of the state and personal information of national persons is an extremely important task. In the context of increasing the number of threats, there is a need to develop new and improve existing information security systems. In the field of information security, cryptographic protection is one of the promising directions of scientific research both in our country and abroad.

A significant contribution to the improvement of existing and development of new methods and means of cryptographic protection made such foreign and domestic scientists: C. E. Shannon, B. Schneier, G. Brassard, J. L. Massey, W. Diffie, M. E. Hellman, R. L. Rivest, A. Shamir, N. Koblitz, O. A. Moldovian, M. A. Moldovian, I. D. Gorbenko, V. K. Zadiraka, M. A. Ivanov, A. N. Fionov, V. V. Yashchenko, A. O. Logachev, B. Ya. Riabko, A. M. Oleksiichuk, L. V. Kovalchuk, A. Ya. Biletskyi, O. G. Korchenko and others.

However, due to the development of computer systems, there have always been and will remain unresolved tasks of increasing the level of information security and reducing the time of encryption. The main characteristics of cryptographic systems are the stability, speed and reliability of cryptographic transformations that need to be constantly raised. To date, not all possibilities have been exhausted for improving the stability of cryptographic systems on the basis of the use of logical operations of cryptographic transformation, a significant contribution to the development of which have been made: K. G. Samofalov, V. A. Luzhetskyi, O. V. Dmytryshyn, O. M. Romankevych, R. P. Melnyk and others. Therefore, there is a need for additional research aimed at developing methods for improving the quality of pseudorandom sequences, as well as improving the quality of crypto primitives for streaming and block encryption. An outstanding solution to the problem of increasing the speed and stability of the matrix cryptographic information transformation is the use of group transformations.

4 Evaluating the effectiveness of the speed increasing method of the implementation of group matrix cryptographic transformation

On the basis of the proposed generalized models of group matrix cryptographic transformation [22, 24], which are presented in the following form:

$$G^k = \begin{bmatrix} a_{11}F_1^k(z_1) \oplus a_{12}F_2^k(z_2) \oplus \dots \oplus a_{1n}F_n^k(z_n) \\ a_{21}F_1^k(z_1) \oplus a_{22}F_2^k(z_2) \oplus \dots \oplus a_{2n}F_n^k(z_n) \\ \vdots \\ a_{m1}F_1^k(z_1) \oplus a_{m2}F_2^k(z_2) \oplus \dots \oplus a_{mn}F_n^k(z_n) \end{bmatrix}, \quad (2)$$

where $a_{ij} \in [0,1]$ – the coefficients of the matrix of direct group cryptographic transformation, F_i^k – operations of non-group two operand cryptographic transformations, $\oplus$ – operation of addition modulo 2, Z_i – input data for direct transformation, w_i – input data (results of direct transformation) for reverse transformation, $i \in \{1..n\}$, $j \in \{1..n\}$.

$$G^d = \begin{bmatrix} b_{11}F_1^d(w_1) \oplus b_{12}F_1^d(w_2) \oplus ... \oplus b_{1n}F_1^d(w_n) \\ b_{21}F_2^d(w_1) \oplus b_{22}F_2^d(w_2) \oplus ... \oplus b_{2n}F_2^d(w_n) \\ \\ b_{n1}F_n^d(w_1) \oplus b_{n2}F_n^d(w_2) \oplus ... \oplus b_{nn}F_n^d(w_n) \end{bmatrix}, \quad (3)$$

where $b_{ij} \in [0,1]$ – the matrix coefficients of reverse group cryptographic transformation, F_i^d – operations of reverse non-group two operand cryptographic transformations, $i \in \{1...n\}$, $j \in \{1..n\}$.

For practical implementation, the schemes was constructed for the direct (Fig. 1, which implements the proposed model (2)), and the reverse (Fig. 2, which implements the proposed model (3)), group cryptographic information transformations.

It should be noted that a direct and reverse non-group matrix cryptographic transformation is performed in the first round of encryption and decryption. The direct and reverse group matrix transformations are implemented in other rounds of cryptographic transformation.

Taking into account the general technology of the encryption and decryption process with the key (Fig. 3) [25]:

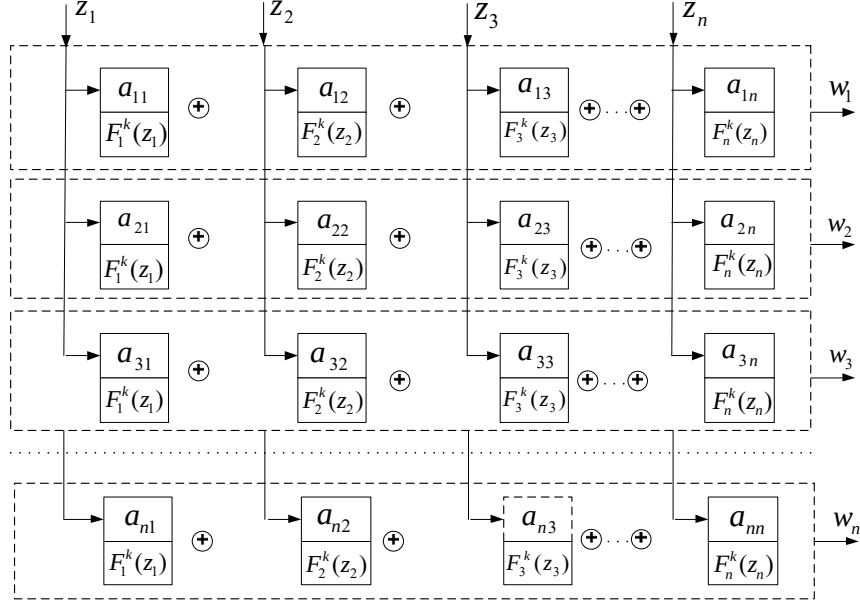


Fig. 1. The block diagram of direct group cryptographic information transformation in the implementation of the model (2)

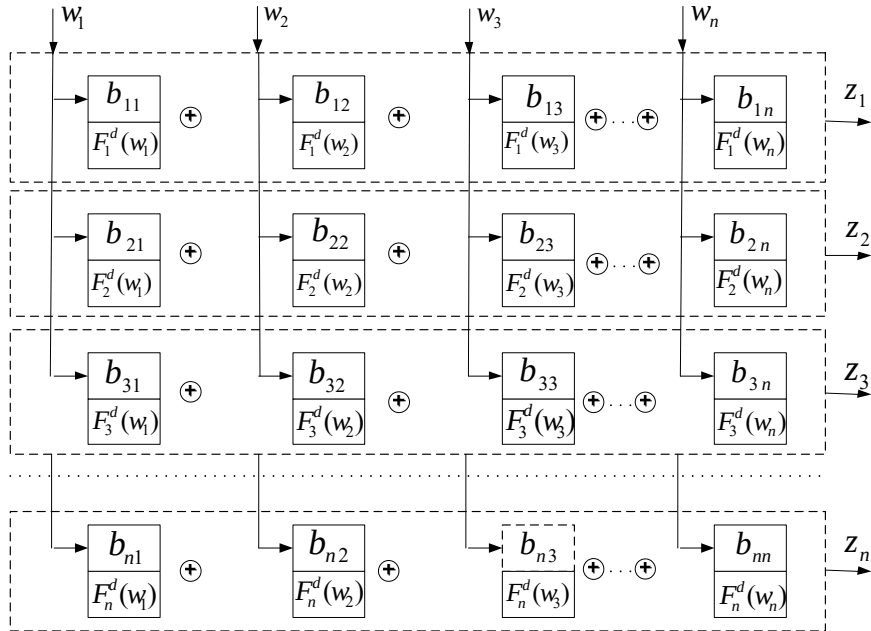


Fig. 2. The block diagram of reverse group cryptographic information transformation for model implementation (3)

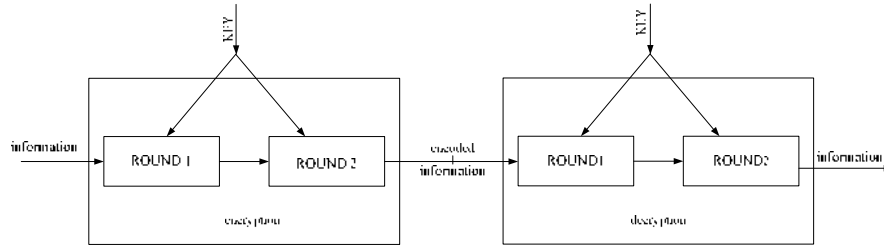


Fig. 3. Encryption and decryption with the key

Consider in more detail two-round hierarchical matrix encryption.

If in a group matrix transformation, n blocks of data are encrypted at the same time for m bits each, then the block diagram of implementing two-round hierarchical matrix encryption can be presented as (Fig. 4):

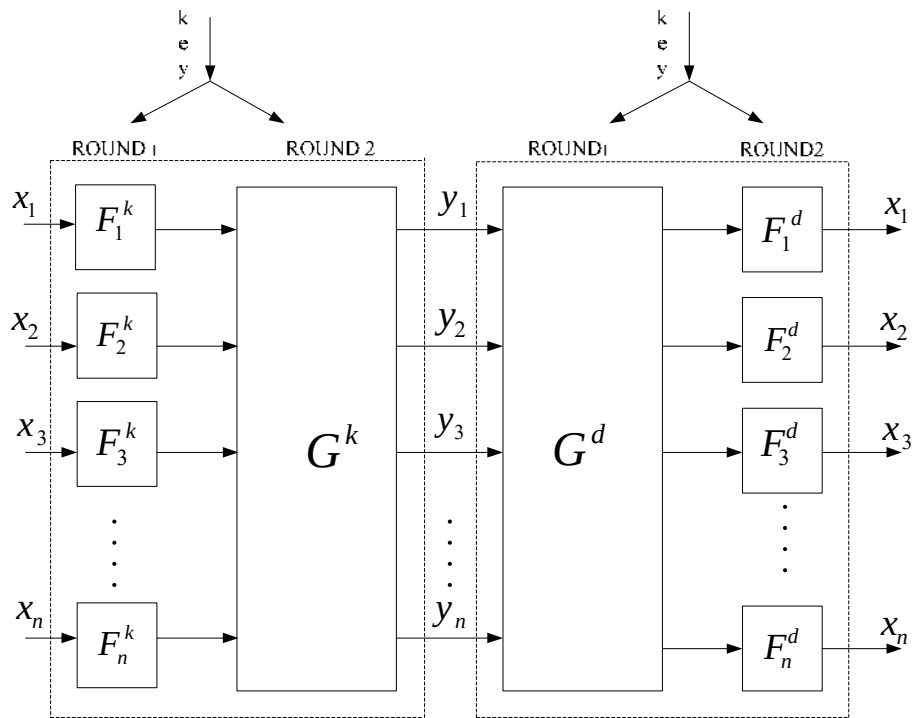


Fig. 4. Block diagram of implementing two-round hierarchical matrix encryption

In accordance with the developed speed increasing method of a group matrix cryptographic transformation, which implements mathematical models (2) and (3), the block diagram of implementing two-round hierarchical matrix encryption can be improved by reducing the elements and blocks of the reverse transform diagram.

The block diagram of implementing the speed increasing method of group matrix cryptographic transformation is presented in Fig. 5.

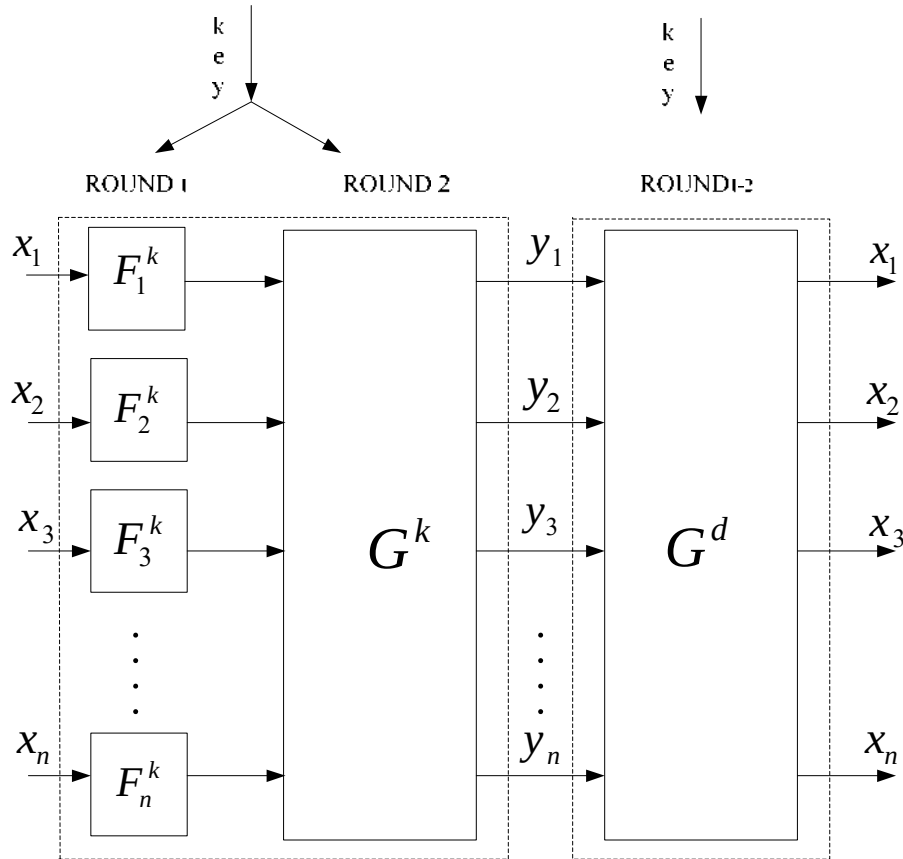


Fig. 5. The block diagram of the speed increasing method of group matrix cryptographic transformation

It should be noted that the block diagram of implementation of the speed increasing method of group matrix cryptographic transformation presented in Fig. 5 provides the implementation of an improved two-round hierarchical matrix encryption. Due to the improvement, which consists of combining two rounds of decryption based on a mathematical model (3), achieved by increasing the speed of the reverse transformation, leading to the reduction of time spent on this transformation.

Let be $t_{F_{\max}}^k$ – the maximal time of non-group direct matrix cryptographic transformation ($t_{F_{\max}}^k = \max\{t_{F_1}^k, t_{F_2}^k, t_{F_3}^k, \dots, t_{F_n}^k\}$), $t_{F_i}^k$ – time of non-group direct ma-

trix cryptographic transformation of Fi operation), t_G^k – time of group direct matrix cryptographic transform, $t_{F_{\max}}^d$ – maximum time of non-group reverse matrix cryptographic transformation ($t_{F_{\max}}^d = \max\{t_{F_1}^d, t_{F_2}^d, t_{F_3}^d, \dots, t_{F_n}^d\}$, $t_{F_l}^d$ – time of non-group reverse matrix cryptographic transformation Fi operation), t_G^k – the time of group reverse matrix cryptographic transformation. It should be noted that in the general case $t_{F_{\max}}^k = t_{F_{\max}}^d$, and $t_G^k = t_G^d$.

Then the speed of implementation of two-round hierarchical matrix encryption (Fig. 4) will be determined:

$$t_{mp} = t_{F_{\max}}^k + t_G^k + t_{F_{\max}}^d + t_G^d. \quad (4)$$

The speed of implementation of the speed increasing method of group matrix cryptographic encryption (Fig. 5) will be determined:

$$t_{mpg} = t_{F_{\max}}^k + t_G^k + t_{\max}^d + t_{\oplus}, \quad (5)$$

where $t_{\max}^d = \max\{t_{F_{\max}}^d, t_G^d\}$, $t_{\oplus}$ – the implementation time of the addition operation by module, $t_{\max}^d > t_{\oplus}$.

Having analyzed expressions (4) and (5), we can state that:

$$t_{mp} > t_{mpg}. \quad (6)$$

Consequently, implementation of the speed increasing method of group matrix cryptographic encryption provides an increase in the total speed of encryption and decryption of information, both by reducing the complexity of finding the reverse transformation and by combining group and non-group matrix transformations when decoding information.

5 Conclusion

As a result of conducted research, an expression was obtained for calculating the complexity of logical determinants, depending on their order:

$$C_n = [(17 + (n-3) \times 6) \times 4 + 3]n + (n-1),$$

where $n > 4$ – the order of the logical determinant.

In this expression, the complexity is determined by the number of inputs of the logical elements of the functional scheme, which implements the construction of the determinant.

It was proved that the model (2) and (3) provide for reducing the complexity of implementing the matrix transformation from 8 to 33 times, depending on the matrix dimensions.

The proposed hierarchical structure of the group transformation made it possible to make certain changes in the matrix encryption and decryption process, that is, it made it possible to significantly reduce the time for the process of passing the cryptographic matrix transform (Fig. 6):

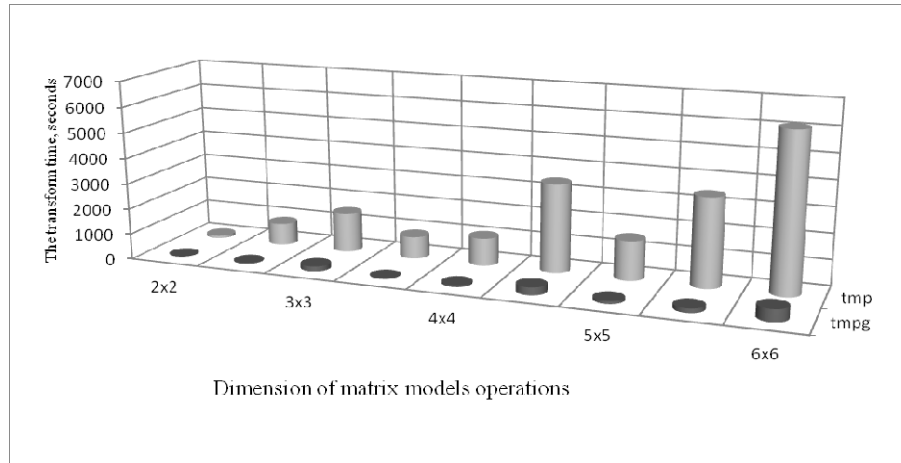


Fig. 6. Time of matrix transformation and group matrix cryptographic transformation

In the classical case, the implementation of group transformation in direct transformation is performed non-group, and then in a group operation, and in the case of inverse transformation, the group is initially performed, and then non-group operations. The spent time goes to all four stages of implementation.

The application of developed models of group matrix cryptographic transformation allows us to combine the stages of reverse group and non-group transformations. According to the results of practical implementation, the implementation of this method provides an increase in speed of 6-8% depending on the matrix dimensions.

References

1. Pogoretsky M. A., Shelomentsev V. P.: The concept of cyberspace as a medium for committing a crime. Information security of a person, a society, a state, vol. 2 (2), pp. 77-81 (2009).
2. Bernet S., Payne S. Cryptography. Official RSA Security Guide. Publishing house «Binom Press», Moscow (2002).
3. Schneier B. Secrets and Lies. Data security in the digital world. Publishing house «Peter», St. Petersburg (2003).
4. Maftyk S. Protection mechanism in computer networks. Publishing house «Myr», Moscow, pp. 126-128 (1993).
5. Zapechnikov S. V., Miloslavskaya N. G., Tolstoy A. I., Ushakov D. V. Threats, vulnerabilities, attacks and approaches to protection. Publishing house «Goryachaya Liniya – Telecom», Moscow, Information security of open systems: in 2 volumes. T. 1. 536 pp. (2006).
6. Shangin V. F. Information security of computer systems and networks. Publishing house «FORUM, INFRA-M», Moscow (2011).
7. Authentication. Theory and practice of ensuring secure access to information resources. Ed. A. A. Shelupanov, S. L. Gruzdeva, Yu. S. Nakhayev. Publishing house «Goryachaya Liniya – Telecom», Moscow, (2009).
8. Faure E.V., Shvydkyi V.V., Shcherba A.I.: Information integrity control based on the factorial number system. Journal of Baku engineering university. Mathematics and computer science, vol. 1, no. 1, pp. 3-13 (2017).
9. Faure E.V., Shcherba A.I., Kharin A.A.: Factorial code with a given number of inversions. Radio electronics, computer science, control, no. 2, pp. 143-153 (2018). DOI 10.15588/1607-3274-2018-2-16.
10. Marushchak A. I. Information law: regulation of information activities. Publishing house «Skif, KNT», Kiev, (2008).
11. Decree of the President of Ukraine “On the Regulations on the Procedure for Implementing Cryptographic Information Protection in Ukraine” No. 505/98 dated 05.22.1998. [Electronic resource]: Official website of the President of Ukraine. URL: <http://www.president.gov.ua/stateauthority/authofstate/prezidlist/Prezidentadmin>.
12. Decree of the President of Ukraine “On Amendments to Paragraph 5 of the Regulations on the Procedure for Implementing Cryptographic Information Protection in Ukraine” No. 333/2008 of April 11, 2008. [Electronic resource]: Official website of the President of Ukraine. URL: <http://zakon5.rada.gov.ua/laws/show/333/2008>.
13. Order of the Department of Special Telecommunication Systems and Information Protection of the Security Service of Ukraine “On Approval of the Regulation on Control over the Operation of the Technical Information Protection System” No. 61 of December 22, 1999 [Electronic resource]: Legislation of Ukraine: Official site of the Verkhovna Rada of Ukraine. URL: <http://www.zakon2.rada.gov.ua>.
14. Rudnytskyi V. M., Myronets I. V., Babenko V. G.: Methodology for increasing the access speed to confidential information resources. Systems of information processing, vol. 5 (86), pp. 15-19 (2010).
15. Moldovian A.A., Moldovian N.A., Sovietov B.Ya. Cryptography: a series of “Textbook for universities. Special literature”. Publishing house «Lan», St. Petersburg, (2000).
16. Chernetskyi I. F. Computer Security Technologies: Monograph. Publishing house MEGU, Rivne (2011).

17. Ivanov M. A.: Cryptographic methods for protecting information in computer systems and networks. Publishing house KUDITS-OBRAZ, Moscow (2001).
18. N. Ferguson, B. Schneier: Practical cryptography. Publishing house «Vyliams» (2005).
19. Myronets I. V.: Algorithm of the method of increasing the access speed to confidential information resources. Methods and means of encoding, protection and compaction of information: Proceedings of the Third International scientific-practic. conf. Publishing house VNTU, Vinnytsya. pp. 308 (2011)
20. Faure E.V., Shcherba A.I., and Rudnytskyi V.M.: The method and criterion for quality assessment of random number sequences. Cybernetics and systems analysis, vol. 52, issue 2, pp. 277-284 (2016). DOI: 10.1007/s10559-016-9824-3.
21. Venbo, Mao: Modern cryptography: theory and practice. Publishing house «Vylyams», Moscow (2005).
22. B. Schneider: Applied cryptography. Protocols, Algorithms, and Source Code in C. Publishing house «TRIUMF», Moscow (2002).
23. Sysoienko S. V., Babenko V. G.: Analysis of the complexity of implementation of operations models of group matrix cryptographic transformation. Naukowy i innowacyjny potencjał prezentacji: kolekcja praca naukowych "ΛόξΟΣ" z materiału międzynarodowej naukowo-praktycznej konferencji, Opole, 18 listopada 2018 r. Równe: «Volynsky Oberegi» Publishing House, t. 7, pp. 50-53 (2018).
24. Sysoienko S. V., Myronets I. V., Babenko V. G.: Construction of a generalized mathematical model of group matrix cryptographic transformation. Modern special technics, vol. 4 (55) (in print).
25. Sysoienko S. V.: Evaluation of the speed of group matrix cryptographic transformation. Systems of control, navigation and communication, vol. 1(47), pp. 141–145 (2018).

Information Technology for Entering Text Based on Tools of the Special Virtual Keyboard Mobile and Auxiliary Devices

Olexander Barmak¹[0000-0003-0739-9678], Ruslan Bahrii¹[0000-0001-5219-1185],
Iurii Krak^{2,3}[0000-0002-8043-0785], Veda Kasianiuk²[0000-0003-3268-303X]

¹National University of Khmelnytsky, 11, Institutes str., 29016, Ukraine
barmakov@khnu.km.ua, bagrijro@khnu.km.ua

²Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska str., 01601, Ukraine
krak@univ.kiev.ua, veda.kasianiuk@gmail.com

³Glushkov Cybernetics Institute, Kyiv, 40 Glushkov avenue, 03187, Ukraine
krak@nas.gov.ua

Abstract. Information technology for the realization of human communication using residual human capabilities, obtained by organizing text entry using mobile and auxiliary devices is proposed. The components of the proposed technology are described in detail: the method for entering text information to realize the possibility of introducing a limited number of controls and the method of predicting words that are most often encountered after words already entered in the sentence. A generalized representation of the process of entering text is described with the aid of an ambiguous virtual keyboard and the representation of control signals for the selection of control elements. The approaches to finding the optimal distribution of the set of alphabet characters for different numbers of control signals are given. The method of word prediction is generalized and improved, the "back-off" statistical language model is used, and the approach to the formation of the training corpus of the spoken Ukrainian language is proposed.

Keywords: virtual keyboard, distribution of symbol, text prediction, corpus of words, statistical language model, N-gram.

1 Introduction

The article presents the results of research and creation of alternative communication technology for people who have temporarily lost their ability to verbal communication. It is proposed to organize communication by entering text and following voice reading it using standard mobile devices. Propose approaches that allow entering text with a limited number of controls, for example, using the four keys of the virtual keyboard.

In modern society communication is a vital necessity for a person, one of its main needs. A large layer of people with speech disorders needs additional means of alter-

native communication for communication. Alternative communication is all communication techniques that supplement or replace ordinary speech for people who do not have the opportunity to communicate as a result of congenital or acquired disease [1].

In the world there are various systems for the implementation of non-verbal communication. People with hearing disabilities use sign language to communicate. The authors proposed approaches for implementing communication in sign languages using virtual reality systems [2]. Many alternative means of communication are limited to devices that can be used only in stationary conditions. Significant constraints on existing systems, what using text entry are the low input speed and, as result, communication that associated with the use of slow methods of selecting controls [3].

The main requirement for an alternative communication system is the speed of text entry considering individual human features and quick adaptation without additional training. The problem is the small amount of available control signals that can be used to generate messages. One possible solution to this problem is the use of virtual ambiguous keyboards, which requires research of optimization methods to effectively solving the problem of ambiguous selection [4]. Another direction is to speeds up the text entry process by reducing additional user actions. This requires conducting research using Natural Language Processing methods (NLP) and Statistical Language Models (SLM) to improve prediction of words [5].

Research objective was to develop information technology for the implementation of communication for people who have temporarily lost their ability to speak. As a result of the realization of this objective: analyzed alternative information channels what suitable for communication and suggested ways to use them, developed mechanisms for rapid text input in Ukrainian with a limited number of control signals, developed system of prediction of the text, developed information technology for the implementation of alternative communication using standard mobile devices.

2 Information technology of alternative communication

Information technology (IT) what based on mobile devices was created for solving set tasks. It implements communication by replacing verbal communication to communication by voice reading of text messages, which entering by a limited number of controls (Fig. 1) [6].

It is proposed to use the characters of verbal communication of the Ukrainian language (letters) for the transmission of text messages. This is due to the fact that people with temporary speech disorders usually prefer to use the language they know, and not learn new communication paradigms.

The input information for the proposed IT is information obtained from alternative human information channels. To provide alternative communication was necessary to intellectualize the process of entering text information for a limited number of control signals. Acceleration of the text input process is possible through the use of redundancy of natural language, which involves using the virtual keyboard, the keys of which contain grouped letters of the alphabet.

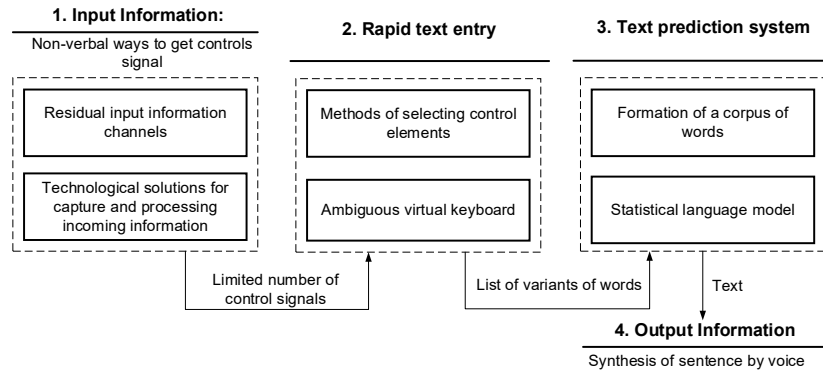


Fig. 1. Diagram of Information Technology

To improve the rapid of entering information and minimize user interaction with the IT device, a prediction system was proposed that automatically suggests the following words that are most commonly encountered after the words already entered in the sentence. For the text prediction system proposed language model and formed training corpus of words adapted to the required type of communication.

Consider the components of information technology in more detail.

3 Method of entering text information

The main difficulty of any means of alternative communication is that a large set of language symbols must be associated with a very limited set of controls. Selection techniques, which are used for the alternative communication, have significant limitations and disadvantages [7], so to speed up the selection process it is proposed to group the letters of the alphabet into control elements.

In general, the method for entering of text information a limited number of control signals is presented as follows (see Fig. 2):

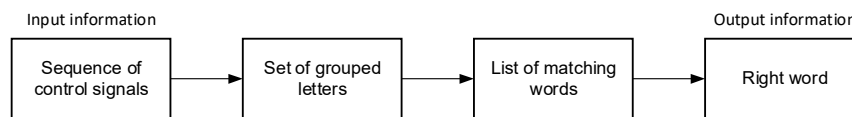


Fig. 2. Method of entering of text information

- 1) *input information*: sequence of control signals to enter the desired word;
- 2) representation of control signals in the form of a sequence of associated sets containing groups of letters;
- 3) list of words-candidates existing in the language (corpus) is possible with this sequence;
- 4) *output information*: choosing the right word.

In addition to non-traditional ways of transferring control signals for a significant number of modern people, it is already typical to manage the process of entering text information using various types of keyboards (both physical and virtual). Without loss of generality, we will consider this method of transferring the control signals in the proposed information technology the main one.

4 Ambiguous virtual keyboard

The virtual keyboard displays the character layout on the device screen. Ambiguous virtual keyboards have several characters per key that make them more productive than normal ones. The purpose of such a keyboard is to reduce the effort at entering text.

In general, the components of the ambiguous keyboard can be represented with 5 parameters:

$$K_{AMB} = \langle P_{sel}, k_{abc}, k_{sec}, A_{ord}, P_{dis} \rangle, \quad (1)$$

where $P_{sel} = \{direct\ selection, scanning, encoding\}$ – selection techniques; $k_{abc} = \{4...8\}$ – number of keys with grouped letters; $k_{sec} = \{1...5\}$ – number of auxiliary keys; $A_{ord} = \{alphabetic, keyboard, statistical\}$ – order of following the letters in the grouped keys; $P_{dis} = \{multi - press, two - key, scanning, T9\}$ – techniques disambiguate selection.

In Figure 3 shows a generalized diagram of the components ambiguous keyboard.

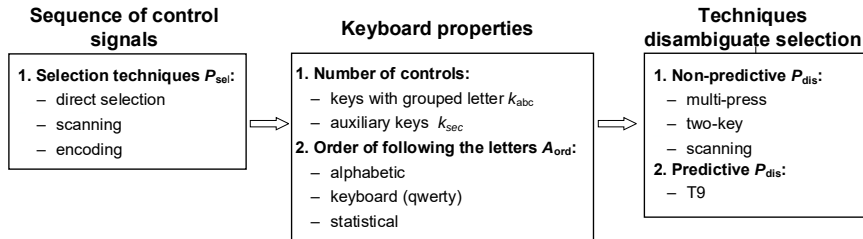


Fig. 3. Components of the ambiguous keyboard

With these components you can describe the known types of entering text information by ambiguous keyboards. For example, according to the above description, the ambiguous keyboard used to enter text on mobile devices with numeric buttons can be represented as: $K_{AMB} = \langle P_{sel} = direct\ selection, k_{abc} = 8, k_{sec} = 1, A_{ord} = alphabetic, P_{dis} = T9 \rangle$. That is, a direct selection method is used, 8 keys with grouped letters and 1 auxiliary, alphabetic order of letters and prediction method T9 with disambiguate.

The improvement of the method for entering text information with a limited number of control signals (Figure 4) within the framework of the above generalized components of the keyboard (1) is given in the following steps:

- 1) representation of the sequence of control signals by any of the indicated selection techniques P_{sel} ;
- 2) reducing the number of controls k_{abc} , k_{sec} and optimizing the distribution of the alphabet set A_{ord} ;
- 3) solving the problem of ambiguous selection P_{dis} by the predictive method for increasing the efficiency of text entry.

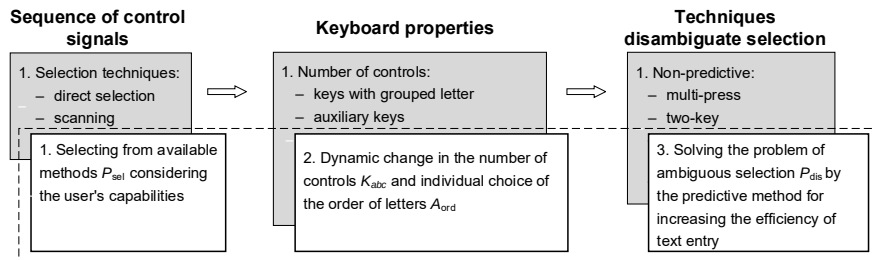


Fig. 4. Diagram the improvement of the method of entering text information with a limited number of control signals

5 Representation of the control signal for selection

In today's development of IT devices, there is a tendency to minimize their size, including the perspective development of chip devices that can be implanted in the human body. As a result, it became possible to use these devices not only under fixed conditions, but also with free movement. The use of such devices by people with disabilities makes it possible to implement alternative communication for them, which will not be tied to a particular place.

Known techniques for selecting symbols can be assigned to three main categories: *direct selection*, *scanning* and *encoding* [8]. The speed of selection directly affects the speed of communication and therefore it should be measured to determine the fastest method, which will be used in the future. In the framework of IT, the representation of a continuous and discrete control signal for the selection is proposed. The result of processing the input signal is to determine its type and the fixation of the methods of selecting control elements.

For a *continuous* control signal, the selection is carried out by a cursor, which can be moved using special devices that capture the movement of the eyes, heads, hands, etc. To make a selection, hold the cursor over the desired control. If this is the keys of the virtual keyboard, then for each time interval holding the cursor over them, the next letter that belongs to the key will be selected. This mode allows you to implement a direct selection of letters and can be used in the absence of words in the dictionary.

For a *discrete* control signal, each of its states is associated with a separate function or control element. If these states are only two, then the scanning selection mode is applied. In this mode, the internal timer cyclically "highlights" the controls and awaits

confirmation of the choice from the user. Of course, the speed of text entry in this mode is quite small, but this mode allows you to use it to people who other ways of communication are not available due to physical constraints.

In the case where the number of states is not sufficient to assign each control element unique control state, a combination of selection modes is used. First of all, it is proposed to bind the keys of the virtual keyboard, as most operations are performed when selecting letters. Other interface features can be selected using encoding. Then the sequence of signal states is used to select them. This allows the maximum use of all available states of the discrete signal coming from a special device.

You can also use the usual physical keyboard and touch type control using the direct selection method to text entry. Touch control allows the use of alternative communication on different mobile devices, which is optimal for people with temporary communication disorders.

6 Research of variants distribution of set of the letters Ukrainian alphabet for different order followings

Decreasing the number of control signals will improve the efficiency of text entry for people with disabilities. In addition, it expands the list of alternative means of communication that can be used to generate control signals. If the scanning selection method is used, decreasing the number of keys should significantly reduce the search time for each character.

Determining the minimum number of keys for ambiguous keypads is an optimization problem, as with a decrease in the number of keys, the number of letters in them increases, which leads to an increase interpretations error of user actions. This requires additional refinement of choice and reduces the efficiency of text entry. So, for a keyboard with 9 keys (T9 standard), the number of letters located on each of them is on average 4th. In this case, the standard algorithm T9 successfully solves the problem of ambiguous selection [9]. Reducing the number of keys to 4 requires placing on each key, on average, about 8 letters, which leads to an increase in the list of possible words. This much complicates the selection of the user's expected words, and requires optimization of the distribution of the letters on the keys and the use of additional prediction algorithms.

Alphabet and keyboard («QWERTY») alphabetic ordering are well-known. For both of these cases, it is not possible to change the order of their sequence, so a study with a uniform distribution of letters in groups and the best variants were determined [10].

The frequency order of the letters is a list of letters of the Ukrainian alphabet in order of decreasing the frequency of their use. To find the optimal letter distributions a set of letters divided into classes according to the frequency of their use. Within these classes, it is possible to change the order of following letters, keeping to the distribution of letters with similar frequencies in different groups. Based on these conditions, an optimization problem has been formed:

- 1) *Constraint 1*: letters with similar frequencies should be placed in different groups;
- 2) *Constraint 2*: letters in groups should be evenly distributed;
- 3) *The criterion of the optimization problem*: the total number of words with the same code should be the smallest.

To solve the optimization problem, an algorithm for actions that considers the distribution constraints is proposed:

- 1) *input information*: a set of letters in a certain order of follow;
- 2) set of letters is divided into classes so that in one class there are letters with similar frequencies;
- 3) for each class, a random letter is selected and placed in a group that does not already have a letter from this class;
- 4) *output information*: formed groups.

For the Ukrainian alphabet, the number of classes is defined in the range from 4 to 8, depending on the number of groups. Thus, each letter that belongs to a particular class belongs to a unique group that enforces the *constraint 1*. Compliance with such an order of distribution of letters ensures the fulfillment of the constraint 2 - the letters are distributed evenly.

7 Solving the problem of ambiguous selection

In order to minimize the interaction between the user and the virtual keyboard, it is proposed to use an algorithm for eliminating ambiguity, similar to T9 (TNKey algorithm), since this method demonstrates better performance than methods that use a several actions to select a letter [11].

To solve the problem of ambiguous selection an algorithm is proposed which allows obtaining a set of words that correspond to a sequence of user actions with a virtual ambiguous keyboard. The algorithm consists of the following steps:

- 1) encoding of all words in the dictionary (corpus) for a given distribution of letters on the keys;
- 2) obtaining the code of the current word corresponding to the sequence of actions (keystrokes) of the user;
- 3) searching the word dictionary corresponding to the code.

In the process of text entry the current word often does not refers to the words of the Ukrainian language, but only reflects the sequence of actions. The code obtained when encoding this sequence is compared to an internal dictionary and generates a list of words corresponding to it.

The use of ambiguous virtual keyboards leads to ambiguous selection. When encoding all dictionary words for a given letter distribution, words that have the same code can be mistakenly offered to the user as an expected word. Such situations interrupt the text entry process and require a refinement of the choice. The quantitative definition of such situations is the frequency of errors, that is, the percentage of words that are misinterpreted.

Such a quantitative estimate of the error rate is determined based on the frequency of word use in the texts. Assume that in a set of words with the same code, the word having the highest frequency of meetings corresponds to the expected one, and other words are not offered correctly. Thus, the quantitative estimation of the error rate is defined as the total number of mistakenly proposed words, normalized to the total number of words, consideration their frequency.

In [10], variants of the distribution of the set of letters of the Ukrainian alphabet into groups in a certain order of succession were investigated. Using the results obtained and the proposed approach, a quantitative estimate of the error rate for a different number of groups and different orders of letters are calculated (Fig. 5).

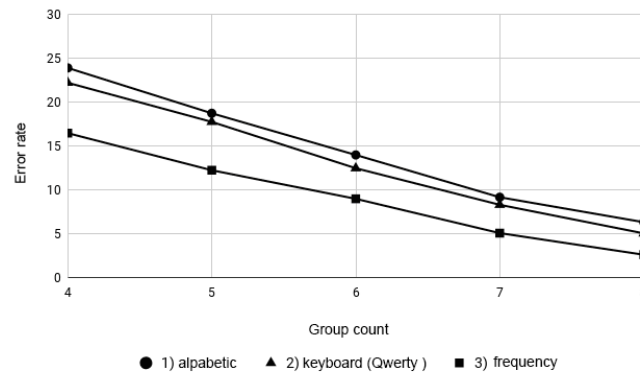


Fig. 5. Error rate for different sequence of letters

Thus, the error rate for 4 groups was 16.43% for the frequency order of letters, which is the best option compared to the keyboard (22.16%) or alphabetic (23.85%) orders. Similar relationships have also been confirmed for another number of groups.

The use of recommended letter distributions for the Ukrainian alphabet provides a more convenient way to enter text for people with different experience with digital devices and ensures the efficiency of the set with decreasing controls.

8 Words prediction method

Word prediction speeds up the text entry process by reducing additional user actions. To do this, it is necessary that as many as possible input words are predicted as the words "by default", that is, they correspond to the word that the user expects.

In general, the method of words prediction, which most often occurs after the words entered in the sentence, can be represented as follows (Figure 6):

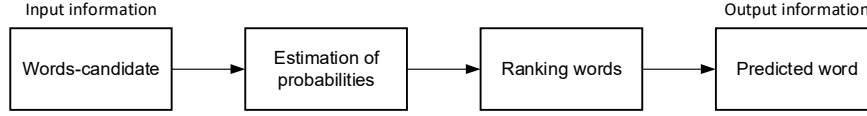


Fig. 6. Method of words prediction

- 1) *input information*: list of words-candidate that correspond to the sequence of user actions;
- 2) the evaluation of the probabilities of words-candidate, considering the previous words of the sentence is presented in the form of different language models;
- 3) the ranking of words for probabilities and the definition of the word "by default";
- 4) *output information*: predicted word.

The purpose of the prediction method is to provide a list of words sorted by probability values. For each candidate word, its probability is estimated considering the language model that uses the statistics of the corpus of the words.

The improvement of the prediction method is submitted by implementing the following two steps:

- 1) prediction of the most probable words using statistical language model, which would allow to realize a problem with acceptable computational complexity (for the possibility of implementation on mobile devices);
- 2) formation of the corpus of words of the spoken Ukrainian language to improve the quality of prediction and reduce user interaction with the device.

Statistical models of the language are used for predictive text. In the field of communication AAC for word prediction is used N-gram model that calculates the probability of the last word, as the probability of a sequence of words in a certain corpus [12].

The maximum likelihood estimation (MLE) method uses to estimate these probabilities. It consists in determining the parameters that maximize the probability of this similarity for given words. Thus, the MLE estimate for the parameters of an N-gram model by getting as normalized counts from a corpus that is statistically representative for language model [13]. For example, we can estimate the probability bigram word w_n , considering the previous word w_{n-1} , calculating the entry (count) of bigrams $C(w_n, w_{n-1})$ and normalize by unigram count for word w_{n-1} :

$$P(w_n | w_{n-1}) = \frac{C(w_{n-1}, w_n)}{C(w_{n-1})} \quad (2)$$

One of the most important problems of N-gram models is the problem of data sparse, which grows rapidly with increasing model order. In fact, the MLE provides zero probability for any sequence of words that is missing from the corpus. To solve the problem of sparse data and improve the overall quality of the prediction without increasing the computational complexity, it is suggested to use a model with a "back-off" with appropriate optimization of the parameters. In the case of a limited corpus of words, the statistical model language, what allows satisfying these requirements is the

Katz's backoff model [14]. The main idea behind Katz's backoff model is to evaluate the conditional probability of a word by a "backoff" to a N-gram of lesser order in the case where N-grams of higher order are not found in the training corpus. So, the model with the most complete information is used to provide the best results.

In particular, for the highest order N, it is proposed to use the trigram model and perform recursively backoff to the bigrams and unigrams:

$$P_{BO}(w_n | w_{n-2}, w_{n-1}) = \begin{cases} P^*(w_n | w_{n-2}, w_{n-1}) & \text{if } C(w_{n-2}, w_{n-1}, w_n) > 0 \\ \alpha(w_{n-2}, w_{n-1}) P_{BO}(w_n | w_{n-1}) & \text{if } C(w_{n-2}, w_{n-1}) > 0 \\ P_{BO}(w_n | w_{n-1}) & \text{else} \end{cases}, \quad (3)$$

where:

$$P_{BO}(w_n | w_{n-1}) = \begin{cases} P^*(w_n | w_{n-1}) & \text{if } C(w_{n-1}, w_n) > 0 \\ \alpha(w_{n-1}) P^*(w_n) & \text{else} \end{cases}, \quad (4)$$

$$P^*(w_n | w_{n-2}, w_{n-1}) = \frac{C^*(w_{n-2}, w_{n-1}, w_n)}{C(w_{n-2}, w_{n-1})}, \quad (5)$$

where P_{BO} – probability calculated by Katz's backoff model, P^* – smoothed probability, estimated by Good-Turing, C – count of N-gram, C^* – smoothed count of N-grams, w_n – predicted word, w_{n-2}, w_{n-1} – previous words, α – backoff coefficients.

In order to simplify the calculation of the coefficient α for the formula (3) offered to be taken equal 0.4, as such, which was heuristically received by Google specialists for the Stupid backoff (SBO) algorithm. It is context-independent and shows approximation to other methods of prediction quality for a limited corpus of words [15].

With each backoff level r the value of the backoff coefficient is calculated according to the formula α^r , where $r = 0$ for trigrams, $r = 1$ for bigrams and $r = 2$ for unigrams. Thus, the probability value calculated by the simplified calculation (S_{Simple}) takes the form:

$$S_{Simple} = \alpha^r P, \quad (6)$$

where:

$$P(w_n | w_{n-2}, w_{n-1}) = \begin{cases} \frac{C(w_{n-2}, w_{n-1}, w_n)}{C(w_{n-2}, w_{n-1})}, & \text{if } C(w_{n-2}, w_{n-1}, w_n) > 0 \\ \frac{C(w_{n-1}, w_n)}{C(w_{n-1})}, & \text{if } C(w_{n-1}, w_n) > 0 \\ \frac{C(w_n)}{N}, & \text{else} \end{cases}, \quad (7)$$

α – backoff coefficient, r – value of backoff level, P – probability calculated by backoff model, C – count of N-gram, w_n – predicted word, w_{n-2}, w_{n-1} – previous words, N – count unigrams.

In this way, word w_n with the highest probability S_{Simple} will be defined as the word "by default".

9 Formation of the corpus of words

The choice of the training corpus is an important stage in the development of any system of predicting the text. To obtain reliable estimates of the probability, statistical language models must be trained in a large set of texts. Also, the more a training corpus is such to this type of communication, the more accurate the probability estimates.

The study modeled the kind of communication that is used in everyday communication. Target users are people with special needs who want to express their thoughts, needs or feelings through alternative communication. Since there are no special corpora for this communication environment, other areas of communication with similar characteristics need to be considered.

To solve this problem, it is suggested to use dialogues on common topics used in vocabulary for studying foreign languages. Such dialogues simulate conversations between people who see each other, cover the most possible common situations and use a limited set of words and phrases.

For the implementation of the communication system of people with disabilities communication was created a limited corpus of spoken Ukrainian on the basis of common topics that simulate conversations between people in similar situations and use a limited set of words and phrases. To create the corpus of Ukrainian language dialogues, a set of texts has been collected consisting of more than 400 dialogues on various subjects, the total volume of which was about 20,000 phrases and 100,000 words. The resulting dialogues, for the further formation of the model, were divided into basic and test sets and experimental studies were conducted to determine the enough filling of the corpus for the task of predicting words and phrases when entering the text [16].

The language model is formed from the text corpus by its partition into N-grams - unigrams, bigrams and trigrams, where each record in N-gram retains the count of words and phrases in a text corpus.

10 Efficiency of information technology

Estimated overall efficiency IT consists of the following properties of alternative means of communication: the speed of text entry a limited number of controls, quantification of the frequency of interpretations error in the process of input and quality predictive value.

The speed of text entry depends on the individual characteristics of the person - his remaining communication capabilities, the experience with digital devices and the

time spent on adaptation. Therefore, for its estimation is used not the indicators of time expenditures, but the productivity of text input, that is, the minimization of the interaction of the device with the user.

The method of entering text information with a limited number of controls, which is an integral part of IT, allows you to implement the possibility of text entry with the help of 4 to 8 keys. The proposed variants of the distribution of letters for the Ukrainian alphabet on the keyboard keys allow you to individually consider the features of the user.

Using the 4 control keys is the fastest way to enter text, because the minimum number of control signals is used, but it requires high IT efficiency to reduce the number of additional actions to refine the choice. In order to compare the effectiveness of the components of the IT research was also conducted using the 6-key controls.

Using an ambiguous input method allows a certain percentage of interpretation errors. The lower the level of interpretation errors, the higher the input performance, since the time spent on correcting an erroneous selection significantly exceeds directly when entering the letters of a word. The distribution of the set of letters of the Ukrainian alphabet for different order of order greatly influences the frequency of interpretation errors.

Word quality prediction is a final component of evaluating the effectiveness of IT and accumulates the impact of all its components. To determine the predictive quality of the any text, a number of experiments were conducted using various statistical language models.

For the 4 keys of the frequency distribution, the prediction quality (Figure 7) for the backoff model was 89.2% for the words known as the N-gram model. Using such a model significantly improves the quality of prediction compared to the use of conventional probability estimates.

For the 6 keys of the keyboard distribution, the prediction quality (Figure 8) exceeded 90% compared to the distribution for 4 keys. This is due to the decrease in the number of candidate words, from which the word "by default".

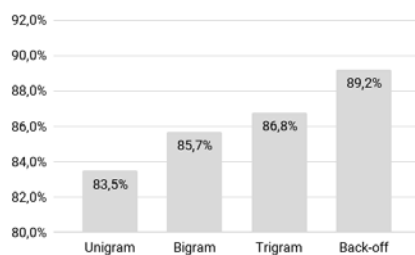


Fig. 7. The quality of prediction of any text on different models for 4 keys of distribution

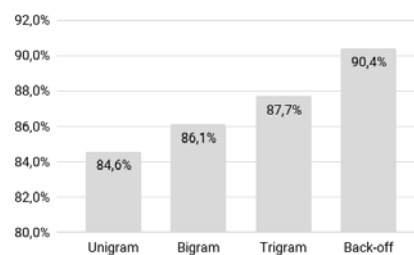


Fig. 8. The quality prediction of any text on different models for the 6-key of distribution

The study of the characteristics of the developed IT made it possible to evaluate the overall effectiveness of the text entry with a limited number of controls and the impact of its components on the final quality of predictive. In the course of the conducted experiments we obtained the performance indicators of the applied methods of input and prediction of text information for different numbers of control elements and the distribution of letters on them (Figure 9). Under the well-known approach is meant a method of entering text information using an ambiguous keyboard with alphabetical order of letters for a specified number of controls.

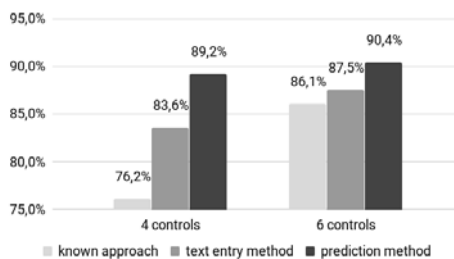


Fig. 9. Effectiveness of the components of information technology

So, the results of an experimental study of the effectiveness of the developed IT alternative communication using methods of input and prediction of text information show that the proposed IT allows people to communicate with the use of residual human capabilities by organizing text input with efficiency exceeding the well-known approaches of 4-13% depending from the number of operating controls.

11 Implementation of information technology alternative communication

A human-computer interaction model was developed for the IT under consideration (Figure 10), which enabled the implementation of software for text entry into digital devices to provide communication for people with disabilities (Figure 11).

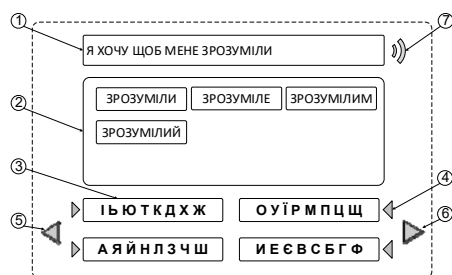


Fig. 10. User interaction model (Ukrainian alphabet)



Fig. 11. Alternative communication system interface (Ukrainian alphabet)

There is also the possibility of implementing information technology of alternative communication for other languages, such as English (Figures 12, 13).

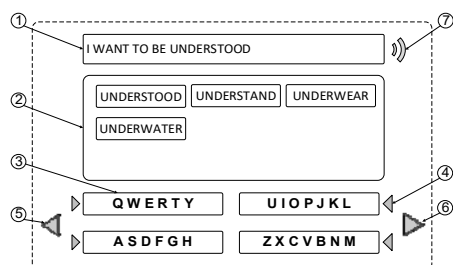


Fig. 12. User interaction model (English alphabet)



Fig. 13. Alternative communication system interface (English alphabet)

The user interaction model with the system for the organization of alternative communication considers individual human constraints (Figure 12, 13). The operating area of the model consists of the following controls: 1 – zone for displaying the entered text; 2 – a zone for displaying suggested words that correspond to the current code of the word entered; 3 – a zone that displays a virtual keyboard with the selected order of letters followed; 4 – control that allows you to select the key containing the required letter; 5 – control that allows you to undo the false choice; 6 – control element that moves to the next word; 7 – a control that activates the function of voice playback of text.

To process the reduced number of states of the input signal, an algorithm for linking them with virtual keyboard keys with ambiguous selection and with separate functional elements of control is proposed.

Object-oriented approach is chosen for information technology design. A general diagram of classes with the description of attributes and methods is developed that allows realizing the information technology on different platforms.

12 Conclusion

The article presents the results of the solution of the actual problem of the implementation of alternative communication for people in whom the channel of verbal communication is temporarily absent. The following basic scientific and applied results are obtained:

- 1) developed method for entering text with a limited number of controls;
- 2) developed a method of prediction the words that most often occur after the words entered in the sentences;
- 3) Predictable text entry system based on mobile devices for implementing communication of people with temporarily speech disorders has been developed.

Using the proposed information system of alternative communication significantly increases the level of socialization of people with special needs, improves the quality

of their lives, develops self-esteem and gives them the opportunity to feel like a full-fledged personality.

References

1. Augmentative and Alternative Communication (AAC) [Internet]. Available from: <http://www.asha.org/public/speech/disorders/AAC/> [Accessed: 2018-09-06]
2. Kryvonos, I.G., Krak, I.V., Barmak, O.V., Kulias, A.I.: Methods to Create Systems for the Analysis and Synthesis of Communicative Information. *Cybernetics and Systems Analysis* 53(6): 847–856 (2017)
3. Koester, H. H., Simpson, R. C.: Method for enhancing text entry rate with single-switch scanning. *Journal of Rehabilitation Research and Development* 51(6): 995–1012 (2014)
4. MacKenzie, I. S., Felzer, T.: SAK: Scanning ambiguous keyboard for efficient one-key text entry. *ACM Transactions on Computer-Human Interaction* 17(11): 1–39 (2010)
5. Higginbotham, J., Lesh, G. W., Moulton, B. J., Roark, B.: The Application of Natural Language Processing to Augmentative and Alternative Communication. *Assistive Technology*, vol. 24, pp. 14–24 (2012)
6. Kryvonos, Iu.G., Krak, Iu.V., Barmak, O.V., Bagriy, R.O.: New tools of alternative communication for persons with verbal communication disorders. *Cybernetics and Systems Analysis* 52(5): 665–673 (2016)
7. Cook, A.M., Cook, P.J.M.: *Hussey's Assistive Technologies: Principles And Practice*. 4th ed. St. Louis, Missouri Mosby, Elsevier (2015)
8. Dowden, P., Cook, A., Reichle, J., Beukelman, D., Light, J.: Choosing effective selection techniques for beginning communicators. Implementing an augmentative communication system: exemplary strategies for beginning communicators, Baltimore, MD: Paul H. Brookes Publishing Co, pp. 395–429 (2002)
9. Silfverberg, M., MacKenzie, I.S., Korhonen, P.: Predicting text entry speed on mobile phones. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '00)*; Hague, 01-06 April 2000, Netherlands: ACM, pp. 9–16 (2000)
10. Krak, Y.V., Barmak, A.V., Bagriy, R.A., Stelya, I.O.: Text entry system for alternative speech communications. *Journal of Automation and Information Sciences* 49(1): 65–75 (2017)
11. Grover, D., King, M., Kuschler, C.: Patent No. US5818437. Reduced keyboard disambiguating computer. Tegic Communications, Inc., Seattle (1998)
12. Ghayoomi, M., Momtazi, S.: An overview on the existing language models for prediction systems as writing assistant tools. In *Proc. IEEE Int. Conf. Syst., Man Cybern.*, San Antonio, TX, USA, pp. 5083–5087 (2009)
13. Jurafsky, D., Martin, J. H.: *Speech and language processing*. 2nd ed. Prentice Hall Inc, New Jersey (2015)
14. Wikipedia: Katz's Back-off Model [Internet]. Available from: https://en.wikipedia.org/wiki/Katz's_back-off_model [Accessed: 2019-01-19]
15. Brants, T., Popat, A. C., Xu, P., Och, F., Dean, J.: Large language models in machine translation. In: *Proceedings Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Prague, Czech Republic, 28–30 June 2007, pp. 858–867 (2007)
16. Kryvonos, Iu.G., Krak, Iu.V., Barmak, O.V., Bagriy, R.O.: Predictive text typing system for the Ukrainian language. *Cybernetics and systems analysis* 53(4): 495–502 (2017)

Information Warfare Model with Migration^{*}

Igor Samoilenko¹[0000-0001-6858-8224], Congzao Dong²[0000-0001-5877-1877],

Bogdan Dovgay³[0000-0003-2840-7884]

¹School of Mathematics and Statistics, Xidian University, 710126 Xi'an, P.R. China
Taras Shevchenko National University of Kyiv, 64 Volodymyrska str., Kyiv, Ukraine
isamoil@i.ua

²School of Mathematics and Statistics, Xidian University, 710126 Xi'an, P.R. China
czdong@xidian.edu.cn

³Taras Shevchenko National University of Kyiv, 64 Volodymyrska str., Kyiv, Ukraine
bogdov@gmail.com

Abstract. We construct and study a discrete time model describing the conflict interaction between two complex systems with non-trivial internal structures. The external conflict interaction is based on the model of alternative interaction between a pair of non-annihilating opponents. The internal conflict dynamics is similar to the one of Lotka-Volterra model, namely information warfare model. We show that the typical trajectory of the complex system converges to an asymptotic attractive cycle. We propose an interpretation of our model in terms of migration processes.

Keywords. Lotka-Volterra equations, information warfare model, conflict interaction, dynamical system, cyclic attractor, limiting distributions, migration

1 Introduction

Since the beginning of 20th century the Lotka-Volterra model of prey-predator interaction is one of the main models for simulation of many processes in population theory, social sciences and economics. As a rule, continuous models where Lotka-Volterra equations have ratio-depended parameters are studied (see, for example [3, 5, 6, 11-13, 17, 19, 21]). Application of this approach to information warfare model was proposed in [16].

Authors regard some social community of quantity N_0 , potentially exposed some information threat (InfT) of two types, that is, for example, the threat of a negative change in its state by transmitting some information relevant to this group by information two different channels. The values $N_1(t)$, $N_2(t)$ – the numbers of “adherents” depending on time t who accepted the new information, ideas, norms, etc. of

^{*} The work was supported by the Chinese Belt and Road Program DL20180077

the type 1 and 2 respectively. These are the main current characteristics of the degree of prevalence of InfT.

The main model assumptions are:

1. Both InfT are distributed among the community through the two information channels:
 - the first one is “external” in relation to the community, for example, advertising media campaign. Its intensity is characterized by the parameters $\alpha_1 > 0$ and $\alpha_2 > 0$ respectively, both are considered to be independent of time;
 - the second, “internal” channel is interpersonal communication between members of the social community (its intensity, that is, the number of equivalent informational contacts, characterized by the parameters $\beta_1 > 0$ and $\beta_2 > 0$ respectively, that are also independent of time). As a result, the adherents of the first idea that have been already “recruited” (their number is equal to $N_1(t)$), make their personal contribution to the recruitment process by affecting non-recruited members (their number is equal to the value of $N_0 - N_1(t) - N_2(t)$). The same is for the adherents of the second idea.
2. The rate of change of the number of adherents $N_1(t)$ and $N_2(t)$ (that is, the number recruited into the unit time) consists of:
 - external recruitment rate (it is proportional to the product of the intensities α_1 and α_2 and on the number of individuals who are not yet recruited $N_0 - N_1(t) - N_2(t)$), that is, $\alpha_1 \cdot (N_0 - N_1(t) - N_2(t))$ and $\alpha_2 \cdot (N_0 - N_1(t) - N_2(t))$ respectively;
 - internal recruitment rate (it is proportional to the product of intensities β_1 and β_2 , on the corresponding number of active adherents $N_1(t)$, $N_2(t)$ and on the number non-recruited $N_0 - N_1(t) - N_2(t)$), that is, $\beta_1 \cdot N_1(t) (N_0 - N_1(t) - N_2(t))$ and $\beta_2 \cdot N_2(t) (N_0 - N_1(t) - N_2(t))$ respectively.

The model is, thus, described by Lotka-Volterra-type equations:

$$\begin{aligned} dN_1(t) / dt &= (\alpha_1 + \beta_1 N_1(t))(N_0 - N_1(t) - N_2(t)), \\ dN_2(t) / dt &= (\alpha_2 + \beta_2 N_2(t))(N_0 - N_1(t) - N_2(t)), \quad t > 0. \end{aligned} \quad (1)$$

The main aim of the work [16] is determination of obvious solution of (1), its stable points, bifurcation points, asymptotic behavior, etc.

In this work we construct a model that describes non-studied variant of information warfare model, i.e., a discrete model with migration. Here individuals migrate not randomly, but according to strategies, discussed in section 4.

We construct the model of the conflict interaction between a pair of complex systems A and B. This means that every system consists of some parameters that interact by some non-trivial law. The system is a finite set of positive numbers: $\mathbf{P} = (P_1, \dots, P_K)$ for A and $\mathbf{R} = (R_1, \dots, R_K)$ for B, where K means the quantity of parameters that characterize the system. We study dynamics in the discrete time. So, the evolution of every system is described by the sequence of vectors with non-negative coordinates $\mathbf{P}^{(n)} = (P_1^{(n)}, \dots, P_K^{(n)})$ for A, and $\mathbf{R}^{(n)} = (R_1^{(n)}, \dots, R_K^{(n)})$ for B, $n = 1, 2, \dots$. The

vectors $\mathbf{P}$ and $\mathbf{R}$ correspond to the moment $n = 0$. Naturally, each system tries to reach the optimal values of its coordinates. In reality, due to the conflict interaction, every coordinate changes in a complicated way. The evolution of all changes is determined by double dependence: by the conflict interaction between systems (which we shortly describe in section 2), and by the mutual “fight” of coordinates (of the information warfare interaction) inside every system.

The law of evolution inside of each (independent) system is described by a discrete variant of equations (1) (here we use the following notations $P := N_1$ and $R := N_2$ to separate the discrete case):

$$\begin{aligned} P^{(n)} &= P^{(n-1)} + (\alpha_1 + \beta_1 P^{(n-1)})(N_0 - P^{(n-1)} - R^{(n-1)}), \\ R^{(n)} &= R^{(n-1)} + (\alpha_2 + \beta_2 R^{(n-1)})(N_0 - P^{(n-1)} - R^{(n-1)}). \end{aligned} \quad (2)$$

Typical behavior of both continuous and discrete information warfare model is shown in Figure 1.

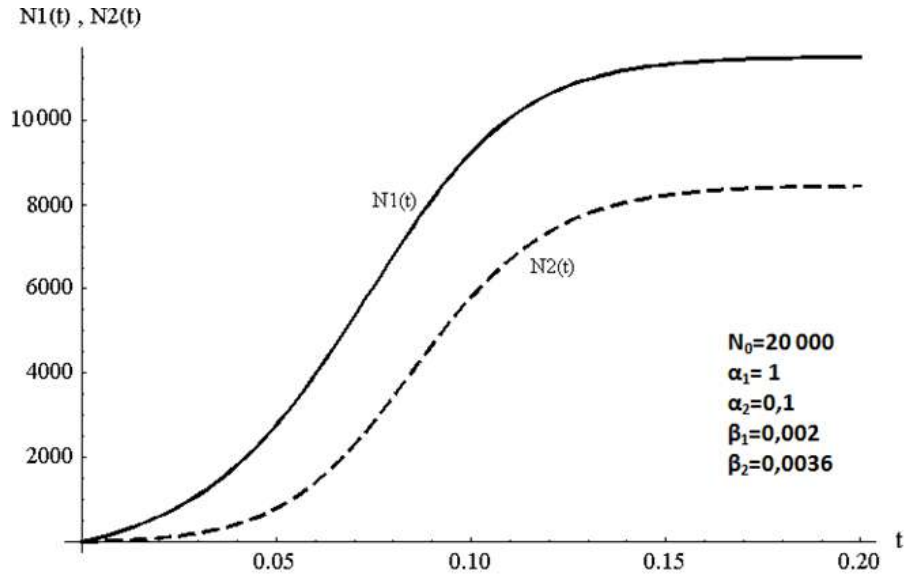


Fig. 1. Typical behavior of continuous information warfare model (Source: [16, p.256])

2 Conflict Interaction Between Non-annihilating Opponents

In this section we shortly remind an alternative approach to describe the redistribution of conflicting positions between two opponents, say A and B, concerning an area of common interests. The main idea of this model is that the influence of every opponent may be redistributed among conflict positions, but no one opponent may be destroyed (that would mean its distribution equals 0 in all the regions). This idea is realized due

to the probabilistic character of opponents' distributions (the sum of the each opponent's presence by all regions should be equal to 1).

We consider the simplest case where the existence space of common interests is a finite set of positions $\mathcal{Q} = \{\omega_1, \dots, \omega_K\}$, $K \geq 2$. Each of the opponents A and B tries to occupy a position ω_i , $i = 1, \dots, K$ with probability $P_A(\omega_i) = p_i \geq 0$ or $P_B(\omega_i) = r_i \geq 0$. The starting distributions of A and B along $\mathcal{Q}$ are arbitrary and normed:

$$\sum_{i=1}^K p_i = 1 = \sum_{i=1}^K r_i.$$

A and B cannot be present simultaneously in a same position ω_i . The interaction between A and B is considered in discrete time. We introduce the noncommutative conflict composition between real-valued stochastic vectors $\mathbf{p}^0 = (p_1, \dots, p_K)$, $\mathbf{r}^0 = (r_1, \dots, r_K)$:

$$\mathbf{p}^1 := \mathbf{p}^0 * \mathbf{r}^0, \quad \mathbf{r}^1 := \mathbf{r}^0 * \mathbf{p}^0,$$

where the coordinates of $\mathbf{p}^l, \mathbf{r}^l$ are defined as follows:

$$p_i^1 := \frac{p_i^0(1 - ar_i^0)}{1 - a \sum_{i=1}^K p_i^0 r_i^0}, \quad r_i^1 := \frac{r_i^0(1 - ap_i^0)}{1 - a \sum_{i=1}^K p_i^0 r_i^0}, \quad (3)$$

where the coefficient a is from the intervals $[-1, 0)$ or $(0, 1]$ and stands for the activity interaction. At the n th step of the conflict dynamics we get two vectors

$$\mathbf{p}^n := \mathbf{p}^{n-1} * \mathbf{r}^{n-1} = \mathbf{p}^0 *^n \mathbf{r}^0, \quad \mathbf{r}^n := \mathbf{r}^{n-1} * \mathbf{p}^{n-1} = \mathbf{r}^0 *^n \mathbf{p}^0,$$

with coordinates

$$p_i^n := \frac{p_i^{n-1}(1 - ar_i^{n-1})}{1 - a \sum_{i=1}^K p_i^{n-1} r_i^{n-1}}, \quad r_i^n := \frac{r_i^{n-1}(1 - ap_i^{n-1})}{1 - a \sum_{i=1}^K p_i^{n-1} r_i^{n-1}}.$$

The behavior of the state $\{\mathbf{p}^n, \mathbf{r}^n\}$ at time $t = n$ for n tending to infinity has been investigated in [1, 4, 7-10]. We shortly describe the results.

Theorem 1. For any pair of non-orthogonal real-valued stochastic vectors $\mathbf{p}, \mathbf{r}$ such that their inner product $(\mathbf{p}, \mathbf{r}) > 0$, and any fixed interaction intensity parameter a not equal to $1/(\mathbf{p}, \mathbf{r})$, the sequence of states $\{\mathbf{p}^n, \mathbf{r}^n\}$ tends to the limit state $\{\mathbf{p}^\infty, \mathbf{r}^\infty\}$. This limit state is invariant with respect to the conflict interaction:

$$\mathbf{p}^\infty := \mathbf{p}^\infty * \mathbf{r}^\infty, \quad \mathbf{r}^\infty := \mathbf{r}^\infty * \mathbf{p}^\infty.$$

Moreover,

$$\begin{cases} (\mathbf{p}^\infty, \mathbf{r}^\infty) = 0, \text{ if } \mathbf{p} \neq \mathbf{r} \text{ and } 0 < a \leq 1, \\ \mathbf{p}^\infty = \mathbf{r}^\infty, \text{ in all other cases.} \end{cases}$$

We emphasize that in the case of a purely repulsive interaction (parameter a belong to the interval $(0,1]$), if the starting distributions are different, then the limiting vectors are orthogonal.

Therefore each of the vectors $\mathbf{p}^\infty, \mathbf{r}^\infty$ contains by necessity some amount of zero coordinates on different positions ω_i . For example the typical limiting picture for $\mathbf{p}^n$ is presented in Figure 2.

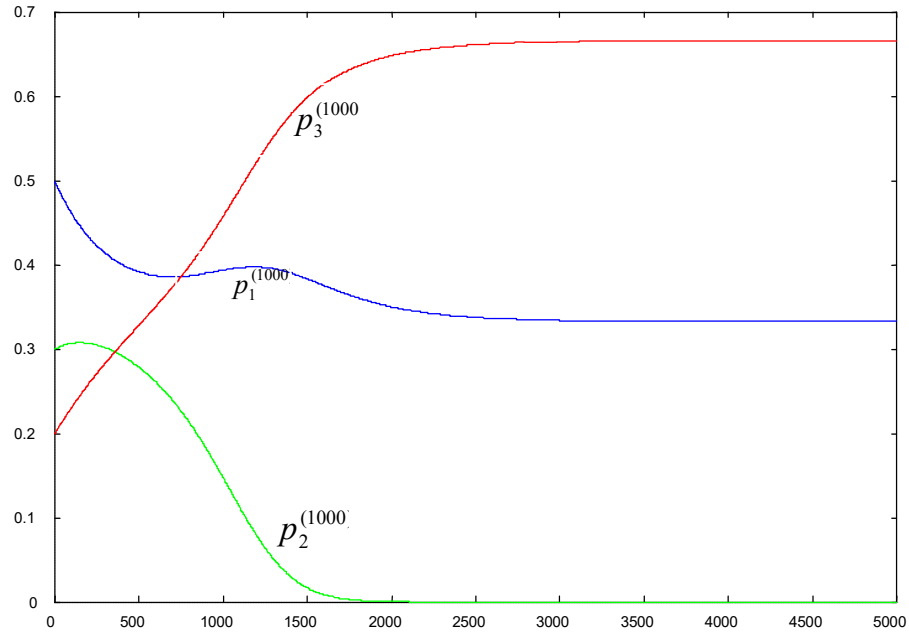


Fig. 2. Typical behavior of a pure conflict model $a=1$, $\mathbf{p}^0=(0.5; 0.3; 0.2)$, $\mathbf{r}^0=(0.48; 0.34; 0.18)$, $\mathbf{p}^\infty=(0.33; 0; 0.67)$, $\mathbf{r}^\infty=(0; 1; 0)$

3 Model of Conflict Interaction Between Complex Systems

In this section we construct a dynamical model of conflict interaction between a pair of complex systems. This again means that every system includes some parameters that interact in non-trivial way described in section 1. But now each of the systems is subject to the inner conflict between their elements. For simplicity, we assume both systems to be similar and described by discrete information warfare models of type (2). We introduce the conflict interaction between these systems using an approach

developed in [1, 2, 4, 7-10]. With such a rather complex situation we may obtain a wide spectrum of evolutions. In this work we study qualitative characteristics of the behavior of corresponding dynamical systems for some choice of parameters $a, \alpha_1, \alpha_2, \beta_1, \beta_2$ (see (2), (3)) and values of initial quantities of adherents P_i, R_i .

The coefficient a , that shows intensity of the interaction between systems, has an important effect. The increasing a from zero to unit causes the appearance of a series of bifurcations. For $a = 0$ we have two copies of independent information warfare models. For small values of a both systems behave like pure information warfare systems, coming them to a stable state.

The role of the coefficients $\alpha_1, \alpha_2, \beta_1, \beta_2$ and initial quantities of adherents P_i, R_i in a pure information warfare model is well-known and described (see, [16]).

The state of our dynamical system is fixed by a pair of vectors $P^n = (P^{(n)}_1, \dots, P^{(n)}_K)$, $R^n = (R^{(n)}_1, \dots, R^{(n)}_K)$ with non-negative coefficients, where $n = 0, 1, \dots$ denotes the discrete time, $K \geq 2$ stands for the number of conflict positions. The complex conflict transformation is denoted by the mapping

$$\begin{pmatrix} P^n \\ R^n \end{pmatrix} \xrightarrow{F} \begin{pmatrix} P^{n+1} \\ R^{n+1} \end{pmatrix},$$

where F is the composition of four operations, the specific mathematical transformations: $F = [N^I * N]U$.

Let us describe them in an explicit form for the first step for the case $K=2$.

The first operation U describes the interaction between elements inside every system separately according to the information warfare model. Corresponding mathematical transformation of vectors (the interaction composition)

$$\{P^0, R^0\} \xrightarrow{U} \{\tilde{P}^0, \tilde{R}^0\}$$

is described by the system of equations of the form (2):

$$\begin{aligned} \tilde{P}_1^{(0)} &= P_1^{(0)} + (\alpha_1 + \beta_1 P_1^{(0)})(N_0 - P_1^{(0)} - P_2^{(0)}), \\ \tilde{P}_2^{(0)} &= P_2^{(0)} + (\alpha_2 + \beta_2 P_2^{(0)})(N_0 - P_1^{(0)} - P_2^{(0)}), \end{aligned}$$

and

$$\begin{aligned} \tilde{R}_1^{(0)} &= R_1^{(0)} + (\alpha_1 + \beta_1 R_1^{(0)})(N_0 - R_1^{(0)} - R_2^{(0)}), \\ \tilde{R}_2^{(0)} &= R_2^{(0)} + (\alpha_2 + \beta_2 R_2^{(0)})(N_0 - R_1^{(0)} - R_2^{(0)}), \end{aligned}$$

where the passage to new values of coordinates is pointed by tilde, but not by changing of upper index, likely to (2).

The following operation involves the interaction $*$ (see (3)) between previous systems according to the theory of the alternative conflict for non-annihilating opponents (see, e.g. [1, 2, 4, 7-10]). To describe this operation we at first have to normalize the vectors $\tilde{P}^0 = (\tilde{P}_1^{(0)}, \tilde{P}_2^{(0)})$, $\tilde{R}^0 = (\tilde{R}_1^{(0)}, \tilde{R}_2^{(0)})$, i.e., to work with stochastic vectors. We use the following notation for normalization: $N\{\tilde{P}^0, \tilde{R}^0\} = \{p^0, r^0\}$, where the coordinates of the stochastic vectors $\{p^0, r^0\}$ are determined by formulae

$$p_1^{(0)} = \frac{\tilde{P}_1^{(0)}}{\tilde{z}_P^{(0)}}, p_2^{(0)} = \frac{\tilde{P}_2^{(0)}}{\tilde{z}_P^{(0)}}, r_1^{(0)} = \frac{\tilde{R}_1^{(0)}}{\tilde{z}_R^{(0)}}, r_2^{(0)} = \frac{\tilde{R}_2^{(0)}}{\tilde{z}_R^{(0)}},$$

where $\tilde{z}_P^{(0)} = \tilde{P}_1^{(0)} + \tilde{P}_2^{(0)}$, $\tilde{z}_R^{(0)} = \tilde{R}_1^{(0)} + \tilde{R}_2^{(0)}$.

The next step exactly corresponds to the conflict interaction between systems. We introduce new stochastic vectors $\{\mathbf{p}^l, \mathbf{r}^l\}$ with coordinates:

$$p_i^{(l)} := \frac{p_i^{(0)}(1 - ar_i^{(0)})}{1 - a \sum_{i=1}^2 p_i^{(0)} r_i^{(0)}}, r_i^{(l)} := \frac{r_i^{(0)}(1 - ap_i^{(0)})}{1 - a \sum_{i=1}^2 p_i^{(0)} r_i^{(0)}}, i = 1, 2.$$

Finally, we have to come back to the non-normalized vectors, which characterize quantitatively populations in both regions after inner and outer conflicts operations. So, at time $n = l$ we have the following vectors $N^l\{\mathbf{p}^l, \mathbf{r}^l\} = \{\mathbf{P}^l, \mathbf{R}^l\}$, where

$$\mathbf{P}^l = (P_1^{(l)}, P_2^{(l)}), \mathbf{R}^l = (R_1^{(l)}, R_2^{(l)}),$$

and $P_i^{(l)} = p_i^{(l)} \tilde{z}_P^{(0)}$, $R_i^{(l)} = r_i^{(l)} \tilde{z}_R^{(0)}$, $i = 1, 2$.

We can repeat this procedure starting from $\{\mathbf{P}^l, \mathbf{R}^l\}$. So we get $\{\mathbf{P}^2, \mathbf{R}^2\}$. And so on for any n th step.

To find the equilibrium points in the case of the complex conflict interaction described above, we have to solve a very complex system of non-linear equations. Thus, pure mathematical approach seems to be not applicable. We use computer simulation methods to see the behavior of the dynamical system under consideration.

Let us consider the case of discrete information warfare model with the conflict interaction between systems. If we take the values of the coefficients $N_0=20000$, $\alpha_1=1$, $\alpha_2=0,1$, $\beta_1=0,002$, $\beta_2=0,0036$ that corresponds to the case of a pure information warfare model presented at Figure 2 and small value of $a=0,00005$, then the influence of the inner conflict is minimized and we have, in fact, two separated information warfare models. In this case the equilibrium points have the coordinates $P_1 = 14.943507$, $P_2 = 35.100629$. The dynamics is constant with these initial data.

In case of larger $a=0,01$, when oscillations appear (see Figure 3), the equilibrium point may also be easily found if we put the initial data in both systems to be equal. In this case the behavior is like in the case of a pure information warfare model, and stabilization occurs.

We should also stress that stable points in the models presented at Figures 1 and 3 are extremely different. This effect is caused by the presence of the outer conflict that initiates oscillations and does not allow exponential growth of the quantity of adherents in all the regions. Thus, if we have some information warfare system and want to change the quantity of adherents of some idea inside this system, we may create an analogous “artificial” system, introduce the conflict interaction and obtain the desired

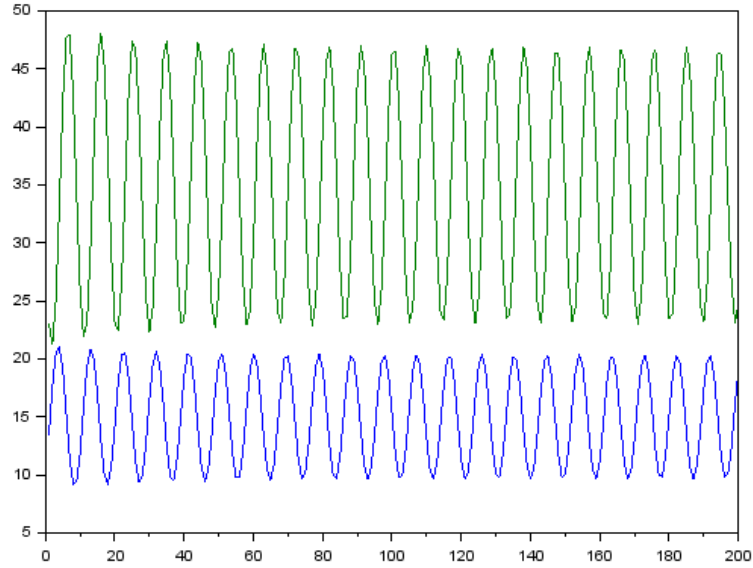


Fig. 3. Oscillations of two adherents' quantities P_1, P_2 inside one region ($0 < n < 200$) $N_0=20000$, $\alpha_1=1$, $\alpha_2=0.1$, $\beta_1=0.002$, $\beta_2=0.0036$, $a=0.01$

shift of the equilibrium point. So, we observed the interesting phenomenon: the equilibrium point of an isolated system is shifted if we come to the case when identical systems are united as an “ensemble”.

However, this equilibrium point is unstable, any perturbation of initial data causes the receding of the system from the equilibrium point.

One of more interesting observations concerns the limit cycles. It is known that no such kind of orbits in discrete information warfare model is possible. But under the effect of the outer conflict, as we see at the pictures, the dynamical system reaches the limit cycle starting both from an inside or outside point with respect to the orbit. Partially, in Figure 4 we present the phase-space picture for (P_1, P_2) in the case of the model presented at Figure 3. As it was pointed above, in case of a pure information warfare model, with the stable initial data there is no dynamics. However, in the case of the model with the outer conflict the process tends to a limit cycle.

Thus, the idea of implementation of outer conflict to the standard information warfare model makes it much more complex and allows observing non-classical effects like oscillations, cyclic attractors, shifting of stable points, etc. We hope to study and describe all these effects in our following research.

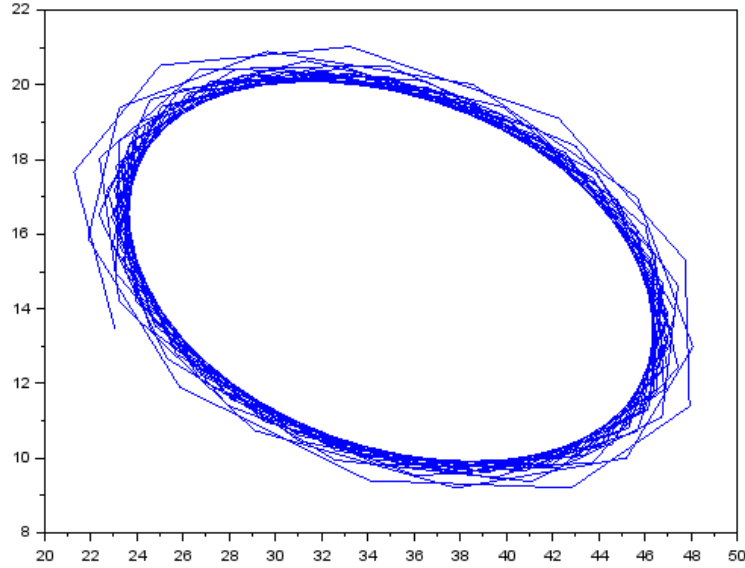


Fig. 4. Phase-space (P_1, P_2) inside one region (corresponding coefficients are presented at Figure 2)

4 Interpretation

In many works on mathematical biology and economics [3, 5, 6, 11-13, 17, 19, 21] the modelling of population dynamics or economical processes is based on Lotka-Volterra type equations. As a rule, continuous, not discrete, models are studied. In some works the migration process is considered. It takes place between different regions, inside which an interaction of the Lotka-Volterra type is present. For example, in [5] the migration rate between regions has some fixed probability.

We study discrete information warfare models with an additional interaction between them. That may be interpreted as some kind of correlation between the inhabitants of different regions. We suppose that discrete models are more natural, partially it is clear that information exchange among individuals happen at some fixed moments of time.

It is well known that in the classical information warfare model [16] a stable point exists. The amount of adherents tends to this point in the phase-space. In this case we observe the following dynamics, after several period of oscillations the populations

stabilize (see Figure 1). Thus, we have an attracting point in phase-space. Such a dynamics exists inside every region when “migration” is absent.

When we introduce an additional interaction between the habitants of different regions a redistribution process appears which we interpret as a migration. In some of our complex models there is no stable point, the amount of adherents in both regions oscillates along fixed orbits. Apparently these orbits in a phase-space are attractors.

We note that explicit formulas of conflict interaction between non-annihilating opponents which describe the redistribution of populations are given by (3). The individuals of a certain kind migrate to the region, where their amount more numerous.

Is the “migration strategy” which is described in our model a natural one? We suppose that in many cases individuals may be right behaving in such a way. If we consider an information warfare model, it is clear that every separated individual is unable to estimate all factors that have an influence on the population dynamics like aggressive information influence, real amount of adherents with his own and alternative position, current population dynamics. In other words, the individual “does not know” the parameters of the information warfare equations and their current influence on the population dynamics.

However the individual has the group reflex and will migrate to the region, where, as he supposes, the information background is best (his population should be concentrated there). He suggests, right there are reliable information resources, possibilities for retranslation of his ideas, better conditions to organize large groups. Formula (3) just describes this tendency.

Similar motivations may be proposed in case of the work migration. Here the unemployed may be regarded as playing the role of “neutrals”, employees and employed workers as playing the role of “information sources”. People, who seek for work and migrate to another country, do not know, as a rule, the real situation in the opposite region. They prefer to migrate to the country where the majority of their friends migrated (group reflex).

So, at the cost of migration accelerates the increasing of one of adherents quantity in one of the regions. But at the same time there is an effect of the inner information warfare “fight” inside every system. As a result, some time later the backward migration starts.

In the Figure 3 we may see the effect of delay, when the amount of adherents of the first idea inside the region decreases, but the adherents of the other idea continue migration to this region, until their amount starts decreasing by following the information warfare model.

We emphasize, that in our model, in comparison with discrete information warfare model, a cyclic oscillations of quantities are observed. Moreover, a cyclic attractor exists in the phase-space, and the adherents trajectory tends to this orbit both from inside or outside point with respect to this cycle. We remark that in our model the normalization was fulfilled by the amount of habitants of the region, so the component of the corresponding vector may be large both at the cost of large population of fixed individuals and at the cost of small whole population of the region. So, a migration to the region with a lot of “free space” is also possible.

We also studied model with the attracting interaction ($a < 0$). In this case we obtained formally a similar dynamics, but now individuals migrate to the region where

they are less numerous. Such a migration strategy might be also natural for some species, e.g. for “missioners” who want to spread their ideas among opponents.

References

1. Alberverio, S., Bodnarchyk, M., Koshmanenko, V.: Dynamics of discrete conflict interactions between non-annihilating opponents. *Methods of Functional Analysis and Topology (MFAT)*. Vol. 11 (4): 309–319 (2005). <http://mfat.imath.kiev.ua/article/?id=330>
2. Alberverio, S., Koshmanenko, V., Pratsiovytyi, M., Torbin, G.: Spectral properties of image measures under infinite conflict interaction. *Positivity*. No. 10: 39–49 (2006). doi: 10.1007/s11117-005-0012-3
3. Alberverio S., Koshmanenko, V.D., Samoilenko, I.V.: The conflict interaction between two complex systems: cyclic migration. *Journal of interdisciplinary mathematics*. V.11, No.2: 163–185 (2008). doi: 10.1007/s11072-006-0054-x
4. Bandyopadhyay, M., Chattopadhyay, J.: Ratio-dependent predator-prey model: effects of environmental fluctuation and stability. *Nonlinearity*. Vol. 18: 913–936 (2005). doi: 10.1088/0951-7715/18/2/022
5. Colato, A., Mizrahi, S. S.: Effects of random migration in population dynamics. *Physical Review E*. Vol. 64: 1–14 (2001). doi: 10.1103/PhysRevE.64.011901
6. Cressman, R., Krivan, V., Garay, J.: Ideal free distributions, evolutionary games, and population dynamics in multiple-species environment. *The American Naturalist*. Vol. 164 (4): 473–489 (2004). doi: 10.1086/423827
7. Koshmanenko, V.: Theorem on conflict for a pair of stochastic vectors. *Ukrainian Math. Journal*. Vol. 55 (4): 555–560 (2003). doi: 10.1023/B:UKMA.0000010167.63115.37
8. Koshmanenko, V.: Theorem of conflicts for a pair of probability measures. *Math. Methods of Operations Research*. Vol. 59 (2): 303–313 (2004). doi: 10.1007/s001860300330
9. Koshmanenko, V. D., Kharchenko, N. V.: Invariant points of a dynamical system of conflict in the space of piecewise-uniformly distributed measures. *Ukrainian Math. Journal*. Vol. 56 (7): 927–938 (2004). doi: 10.1007/s11253-005-0120-z
10. Koshmanenko, V.D., Samoilenko, I.V.: The conflict triad dynamical system. *Communications in nonlinear science and numerical simulation*. No.16: 2917–2935 (2011). doi: 10.1016/j.cnsns.2010.10.028
11. Kuang, Y., Basic properties of mathematical population models. *Journ.Biomath*. Vol. 17: 129–142 (2002). <https://math.la.asu.edu/~kuang/paper/criteria.pdf>
12. Kuang, Y., Beretta, E.: Global qualitative analysis of a ratio-dependent predator-prey system. *J. Math. Biology*. Vol. 36: 389–406 (1998). doi: 10.1007/s002850050105
13. Lonzonn, Y., Solomon, S., Goldenberg, J., Mazarsky, D.: World size global markets lead to economic instability. *Artificial life*.: 357–370 (2003). doi: 10.1162/106454603322694816
14. Lotka, A. J.: Relation between birth rates and death rates. *Science*. Vol. 26: 21–22 (1907). doi: 10.1126/science.26.653.21-a
15. Malthus, T. R. : *An Essay on the Principle of Population*. Reprinted by Macmillan (1894). <http://www.esp.org/books/malthus/population/malthus.pdf>
16. Mikhailov, A.P., Marevtseva, N.A.: Models of information warfare. *Math. Models Comput. Simul.* Volume 4, Issue 3.: 251–259 (2012). doi: 10.1134/S2070048212030076
17. Murray, J. D.: *Mathematical Biology I: An Introduction*. Springer (2002)

18. Takahashi, K. I., Salam, K. Md. M.: Mathematical model of conflict and cooperation with non-annihilating multi-opponent. In: *Unifying Themes in Complex Systems*. Springer, Berlin, Heidelberg, pp 299-306 (2010). doi: 10.1007/978-3-540-85081-6_38
19. Stone, L., Olinky, R.: Phenomena in ecological systems, *Experimental Chaos: 6th Experimental Chaos Conference*, pp. 476-487 (2003)
20. Takahashi, K. I., Salam, K. Md. M.: Mathematical model of conflict with non-annihilating multi-opponent. *J. Interdisciplinary Math.* Vol. 9 (3): 459–473 (2006). doi: 10.1080/09720502.2006.10700457
21. Tufto, J.: Effects of releasing maladapted individuals: a demographic evolutionary model. *The American Naturalist*. Vol. 158 (4): 331–340 (2001). doi: 10.1086/321987
22. Verhulst, P. P.: Notice sur la loi que la population suit dans son accroissement. *Correspondence mathématique et physique publiée par A. Quetelet*. Vol. 10: 113–121 (1838)
23. Volterra, V.: Sui tentativi di applicazione della matematiche alle scienze biologiche e sociali. *Giornale degli Economisti*. Vol. 23: 436–458 (1901). <https://www.jstor.org/stable/23221067>

Development and Software Implementation of the Hot Blast Stove Computer Model

Aleksey Koifman ¹ [0000-0003-2075-7417], Aleksandr Simkin ² [0000-0002-9939-7866]

Pryazovskyi State Technical University, str. Universytets'ka 7, Mariupol, 87555, Ukraine,

¹koifman_o_o@pstu.edu, ²simkin@pstu.edu

Abstract. Based on the evaluation of the existing methods of regenerative heat exchangers modeling, and, in particular, hot blast stoves of blast furnaces, an algorithm for calculating of heat exchange in full cycle “on-gas” and “on-blast” was proposed and implemented as a computer application. The computer model includes adjustment coefficients, the values of which were determined based on the database of technological parameters of an actual hot blast stove block. The modeling results correspond to the actual mill data, it showing the adequacy of the developed model. The computer application will be implemented in the automatic control systems of a hot blast stoves block.

Keywords: model, computer application, regenerator, hot blast stove, HBS, checkerwork, heat exchange, flow-chart

1 Introduction

Several regenerative heat exchangers (hot blast stoves, hereinafter referred to as HBSs) merged in a block are used for the preparation of blast air required for the production of cast iron in a blast furnace. The design of HBSs does not allow to control a number of parameters of their operation, in particular, the temperature of the checkerwork along its height.

There are various approaches to the simulation of the regenerative heat exchange, which happens in HBS.

It is necessary to use computer models to estimate the condition of HBSs, to optimize the operation of the block and stabilize the blast air temperature.

The objective of the research is to develop an HBS computer simulation model, capable of solving technological problems and being used in automation systems.

To achieve the objective it is necessary to solve the following problems:

1. To study the current state of the simulation of the regenerative heat exchange.
2. To develop a mathematical model of the HBS operation in the “on-gas” and “on-blast” periods.
3. To develop an algorithm for the implementation of the above model.
4. To implement the developed algorithm in the form of a computer application.
5. To evaluate the simulation results by comparing them with the actual mill data.

2 Formal problem statement

The HBS checkerwork is a cylinder lined with the chequer bricks in such a way that vertical channels are formed along its height. The checkerwork accumulates the heat during the operation of HBS in the on-blast period. Hot flue gases produced by combustion of blast furnace gas (hereinafter referred as BFG) in the combustion chamber at temperature $t_g = f(\tau)$ enter the checkerwork from the top and descend down through the channels, yielding the heat to the checkerwork, and then enter the flue gas collector. The duration of the checkerwork heating depends on the initial temperature distribution in the checkerwork $t_s = f(y)$, the material of the checkerwork and the parameters of the exchange gas (Fig. 1, a). The on-gas period continues until the flue gases temperature at the outlet of the checkerwork reaches the maximum possible temperature, usually $300 \div 400$ °C, which is due to HBS design features.

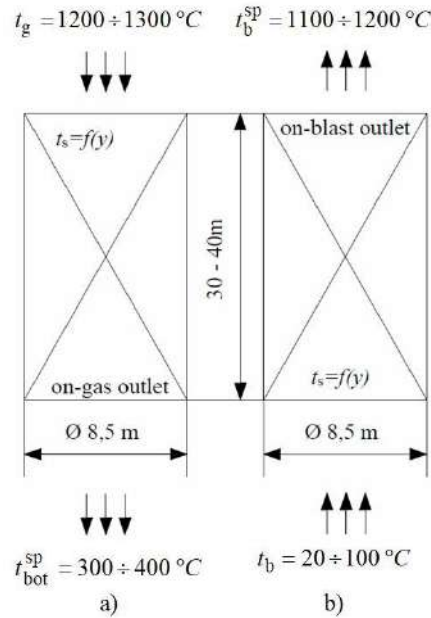


Fig. 1. Physical model of heating (a) and cooling (b) of the checkerwork

The accumulated checkerwork heat Q_{ch}^a is used to heat the blast air. The pressurized blast air at temperature t_b enters the checkerwork from the bottom and passing its channels heats up to the required temperature (Fig. 1, b). The on-blast period lasts until the temperature of the blast air at the checkerwork outlet reaches a set-point temperature t_b^{sp} .

Up to now the methods for the mathematical description of gas dynamics and convective heat exchange have been developed in sufficient detail [1 - 3]. While investi-

gating this problem the authors of the research paid attention to description of these processes in a granular bed. When describing the granular bed the actual structure is replaced with an ideal structure, which is a set of cylindrical vertical channels parallel to each other [4], similar to the checkerwork consisting of shaped elements. Therefore the mathematical methods used for the granular beds can also be fully used for the checkerwork [2, 4].

Now, let us consider the heat exchange in a fixed granular bed. The gas passes in time $d\tau$ at the velocity of open flow u_o through the element of the layer dy at a distance y from the delivery point of the gas to the layer with the porosity ε and a cross section of 1 m^2 . In this case 1 m^3 of the layer accounts for layer porosity $\varepsilon \text{ m}^3$ of gas, and the percentage of particles - $(1 - \varepsilon) \text{ m}^3$ (Fig. 2).

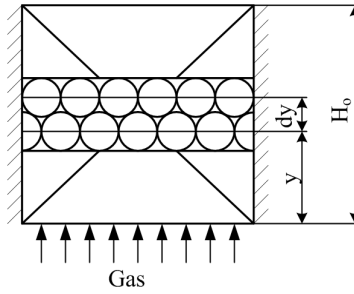


Fig. 2. Fixed granular bed

To describe the heat exchange by simple dependencies, we take a number of assumptions:

- the layer of particles is homogeneous in its fractional composition, and the thermo-physical parameters of the bed and the gas are constant and acquire their average values within the range of operational temperatures, besides, the heat transfer in the gas and in the bed from particle to particle due to thermal conductivity is non-existent;
- the heat flow from the gas to the bed at any point is determined by Newton's law, and the heat exchange coefficient from the gas to the bed is the same along the entire height and section of the bed;
- gas flow rate in time is the same and equally distributed over the cross section of the bed.

In this case the change in gas enthalpy in the elementary layer is determined as follows:

$$d^2 Q = c_g \rho_g \varepsilon \frac{dt_g}{d\tau} dy d\tau = c_g \rho_g \left(\varepsilon \frac{\partial t_g}{\partial \tau} - u_o \frac{\partial t_g}{\partial y} \right) dy d\tau, \quad (1)$$

where c_g – heat capacity, J/kgK ; ρ_g – gas density, kg/m³.

The amount of heat that the gas gives to the bed when goes through it

$$d^2Q = \alpha_V (t_g - t_s) dy d\tau, \quad (2)$$

where α_V – volumetric heat transfer coefficient, W/m^3K .

Considering the heat exchange from the bed material side and bearing in mind that the material will be changed in the elementary layer in time only, given that the material is fixed, we can write:

$$\alpha_V (t_g - t_s) = C_s \rho_s \frac{\partial t_s}{\partial \tau}, \quad (3)$$

where C_s – solid specific heat capacity, J/kgK ; ρ_s – solid density, kg/m^3 .

Let us consider and evaluate different approaches to the description of the heat exchange in regenerative heat exchangers.

3 Literature review

The study of heat exchange in the checkerwork of a regenerative heat exchanger (hereinafter referred to as a regenerator) developed in parallel with the theory of heat exchange in a fixed bed.

Customary mathematical modeling of heat exchange in the bed is based on several approaches: a model of an equivalent recuperator; temperatures of the gases and the material are essentially the same; temperatures of the gases and the bed material are essentially different.

The first approach at the present stage is not relevant, since it does not reflect the principles of regenerative heat exchange. The second approach is not widespread.

The third approach assumes that the heat exchange in the regenerator checkerwork is characterized by the fact that the temperatures of the gases and the bed material differ significantly. Based on the principle, various authors developed a direction in calculating temperature fields in the bed and checkerworks, based on the assumption that the longitudinal thermal conductivity in the bed was non-existent. In this case the equation system describing the heat exchange consists of separate heat exchange equations for the gas and the bed material. These equations were analyzed analytically by T. Schumann [5], H. Hausen [1, 6], A. Willmott [7, 8], etc.

T. Schumann [4] considered the problem of the heat exchange between a liquid and a porous cylinder, while the cylinder material temperature at the start of the heat exchange was uniform (equal along the layer), the gas was supplied at constant temperature and flow. Moreover, the author made a number of other assumptions. The problem does not fully describe the regenerative heat exchange in the regenerator checkerwork. In [9], the solution of T. Schumann's problem was obtained by applying numerical methods, it was used as a model to study the effect of the exchange gas pressure on the intensification of the heating of the regenerator checkerwork.

H. Hausen [6] developed a method for the approximate calculation of temperature fields in a checkerwork at the end of heating (cooling) for a set-point initial temperature distribution, based on the linearity of differential equations. The initial temperature field is approximated at that with a stepped line, in such a way, that the temperature in each individual section is assumed to be constant.

The method proposed by A. Willmott [7, 8] happens to be general and can be used to solve most non-linear problems of calculating the heat exchange in regenerators with a digital computer. The author complicated his model by introducing a variable velocity of the gas flow and used the dependence of the thermo-physical parameters of the gas and the checkerwork material upon temperature.

P. Razelos [10] applied the heat exchange coefficient, which includes convective and radiant components. The developed computer program for calculating heat exchange in the checkerwork included in the form of tables the thermo-physical properties of the checkerwork material and gases, that depend on the temperature.

The authors K. Muske and F. Minet [11 - 13] divided the HBS checkerwork into several zones corresponding to different checkerwork materials, and the dependences of the heat capacities and viscosities for each of the gas components and the material thermal conductivity were used as approximating polynomials of table values. The disadvantages of the approach include the fact that the gas flow was considered uniform and the thermal conductivity in the radial direction was not taken into account. The obtained models identified on the basis of mill data and were used to optimize the operation of the HBSs block. The work [13] describes a model that uses for calculations the values of technological parameters measured with an interval of two minutes.

The author of [14] proposed a three-dimensional mathematical model that can simulate various complex phenomena such as turbulent mixing of fuel and air, combustion, convection floatation, thermal radiation, heat exchange between the gas and the heat-retaining brick.

To study the operation of new types of dome burners, the authors of [15] implemented a 3D mathematical model of HBS considering the basic regularities of heat exchange using the ANSYS Fluent program.

The authors of [16, 17] developed a finite-difference mathematical model of HBS based on an approximate thermal balance with regard to the gas flow, convective and radiant heat exchange. The model allows to take into account the design features and the technical condition of each of HBS.

The article by E. Kobysh [18, 19] presents the solution of a system of equations describing regenerative heat exchange considering the heating of the checkerwork throughout the thickness of a brick, allowing to solve problems of external and internal heat exchange.

For investigation of the heat exchange during the on-gas period from the balls, a mathematical simulation model was proposed [20] based on the finite difference method for iteratively solving the heat exchange problem in the checkerwork. In this case, only the on-gas period is considered.

Thus, at present, a sufficient number of models of the heat exchange process in HBSs has been created and implemented. In order to improve the accuracy of model-

ing and taking into account the basic regularities, the authors of this research decided to develop a simulation model based on an analytical solution of the heat exchange problem in a fixed bed.

4 Formalization and mathematical representation of the model

To analyze the operation of the actual HBS during the on-gas – on-blast cycle, it is necessary to solve the general problem under the following conditions:

- variable gas temperature at the inlet to the checkerwork;
- uneven initial temperature distribution of the material of the checkerwork along its height.

The general problem of the heat exchange in the checkerwork [1] in an analytical form possesses the following view:

$$\frac{\partial t_s}{\partial Z} = t_g - t_s, \quad -\frac{\partial t_g}{\partial Y} = t_g - t_s, \quad (4)$$

at border-line conditions

$$t_g(Y=0) = \varphi(Z), \quad t_s(Z=0) = f(Y), \quad (5)$$

where Z – dimensionless time; Y – dimensionless checkerwork height.

The solution to problem (4) is presented as a combination of two solutions with different border-line conditions:

1. Uneven temperature distribution of the checkerwork along its height at a constant temperature of the exchange gas at the checkerwork inlet:

$$t_g(Y=0) = 1, \quad t_s(Z=0) = f(Y). \quad (6)$$

2. Uniform temperature distribution of the checkerwork along its height at a variable in time temperature of the exchange gas at the checkerwork inlet:

$$t_g(Y=0) = \varphi(Z), \quad t_s(Z=0) = 0. \quad (7)$$

The solution of the problem under conditions (6) has the following form [3]:

$$t_g = - \int_0^{Y_H} f(Y) \frac{\partial \theta(Y_H - Y)}{\partial (Y_H - Y)} dY, \quad t_s = f(Y_H) \cdot e^{-Z} - \int_0^{Y_H} f(Y) \frac{\partial g(Y_H - Y)}{\partial (Y_H - Y)} dY, \quad (8)$$

where θ – relative dimensionless gas temperature; g – relative dimensionless solid temperature; Y_H – dimensionless checkerwork height at $H_0 = 30$ m;

Dividing the height of the checkerwork Y_H into n sections, assuming the checkerwork temperature on the section is equal to the average of its height, we replace the analytical formulas (8) with finite amounts, and get the following expressions:

$$t_g = \sum_{i=1}^n (f_i(Y_H) \cdot [\theta(Y_H - Y_i, Z) - \theta(Y_H - Y_{i-1}, Z)]), \quad (9)$$

$$t_s = f(Y_H) \cdot e^{-Z} + \sum_{i=1}^n (f_i(Y_H) \cdot [\vartheta(Y_H - Y_i, Z) - \vartheta(Y_H - Y_{i-1}, Z)]). \quad (10)$$

Taking the initial change in the material temperature of the checkerwork along its height as linear

$$f_i = a_i + b_i \cdot y, \quad (11)$$

and inserting (11) into equations (9) and (10), we obtain the following expressions for determining the temperatures of the gas and the material in the checkerwork at any time [21]:

$$t_g = \sum_{i=1}^n ([a_i + b_i \cdot Y_H] \cdot [\theta(Y_H - Y_i, Z) - \theta(Y_H - Y_{i-1}, Z)]), \quad (12)$$

$$t_s = f(Y_H) \cdot e^{-Z} + \sum_{i=1}^n ([a_i + b_i \cdot Y_H] \cdot [\vartheta(Y_H - Y_i, Z) - \vartheta(Y_H - Y_{i-1}, Z)]), \quad (13)$$

where a_i , b_i – linear equation coefficients; n – the number of iterations affecting the accuracy.

Similar to the above, the solution to the problem with the border-line conditions (7) has the form [3]

$$t_g = \varphi(Z_\Sigma) e^{-Y} - \int_0^{Z_\Sigma} \varphi(Z) \frac{\partial \theta(Z_\Sigma - Z)}{\partial (Z_\Sigma - Z)} dZ, \quad t_s = - \int_0^{Z_\Sigma} \varphi(Z) \frac{\partial \vartheta(Z_\Sigma - Z)}{\partial (Z_\Sigma - Z)} dZ, \quad (14)$$

determining the temperature of the exchange gas for the consecutive time periods

$$\varphi_j = \alpha_j + \beta_j Z, \quad (15)$$

and inserting expression (15) into (14) and making mathematical simplifications, we obtain the solution in numerical form

$$t_g = \varphi(Z_\Sigma) e^{-Y} + \sum_{j=1}^m (\varphi_j(Y) [\theta(Y, Z_\Sigma - Z_j) - \theta(Y, Z_\Sigma - Z_{j-1})]), \quad (16)$$

$$t_s = \sum_{j=1}^m (\varphi_j(Z_\Sigma) [\vartheta(Y, Z_\Sigma - Z_j) - \vartheta(Y, Z_\Sigma - Z_{j-1})]), \quad (17)$$

where Z_Σ – dimensionless time at which the maximum temperature of the exchange gas is reached; α_j , β_j – linear equation coefficients; j – time section number ($j \in [1, m]$); m – the number of time partitions.

By combining expressions (13), (14) and (18), (19), the authors of this research obtained a solution to the general problem of the heat exchange in the following form:

$$t_g = \sum_{i=1}^n (f_i(Y_H) [\theta(Y_H - Y_i, Z) - \theta(Y_H - Y_{i-1}, Z)]) + \varphi(Z_\Sigma) e^{-Y} + \sum_{j=1}^m (\varphi_j(Y) [\theta(Y, Z_\Sigma - Z_j) - \theta(Y, Z_\Sigma - Z_{j-1})]) \quad (18)$$

$$t_s = f(Y_H) e^{-Z} + \sum_{i=1}^n (f_i(Y_H) [\vartheta(Y_H - Y_i, Z) - \vartheta(Y_H - Y_{i-1}, Z)]) + \sum_{j=1}^m (\varphi_j(Z_\Sigma) [\vartheta(Y, Z_\Sigma - Z_j) - \vartheta(Y, Z_\Sigma - Z_{j-1})]) \quad (19)$$

In works [3], nomograms were used to determine the relative temperatures at the dimensionless layer height and time, which in its turn reduces the accuracy of calculations and makes it impossible to use a PC. At the same time, the author of [5] proposed the following analytical expressions for determining relative temperatures:

$$\vartheta = e^{-Y} \int_0^Z e^{-\delta} \sum_{k=0}^{\infty} \frac{(2\sqrt{Y\delta}/2)^{2k}}{(k!)^2} d\delta, \quad \theta = 1 - e^{-Z} \int_0^Y e^{-\delta} \sum_{k=0}^{\infty} \frac{(2\sqrt{Z\delta}/2)^{2k}}{(k!)^2} d\delta. \quad (20)$$

After a number of simplifications and integration, solutions (22) and (23) will have the following form [9]:

$$\vartheta = e^{-Y} \left(\sum_{i=0}^n \left(\frac{Y^i}{i!} \left(1 - e^{-Z} \sum_{k=0}^i \frac{Z^k}{k!} \right) \right) \right), \quad \theta = 1 - e^{-Z} \left(\sum_{i=0}^n \left(\frac{Z^i}{i!} \left(1 - e^{-Y} \sum_{k=0}^i \frac{Y^k}{k!} \right) \right) \right). \quad (21)$$

The solutions obtained by the authors of this research (18) and (19) and expressions (21) allow further calculations to be made with application of a PC, it improving the accuracy of the calculations.

5 Development of the algorithm for program implementation of the model

Fig. 3 represents a general flow-chart of the program. Let us consider the algorithm in further detail.

In block 1 the input and values assignment are carried out: of the initial temperature distribution of the checkerwork material along its height; of the exchange gas temperature at the checkerwork inlet in the on-gas period; the cold blast air temperature at the checkerwork inlet in the on-blast period; time step (by default - 100 s); number of partitions throughout the checkerwork height; checkerwork height, diameter, channel diameter and porosity; correction coefficients that consider the checker-

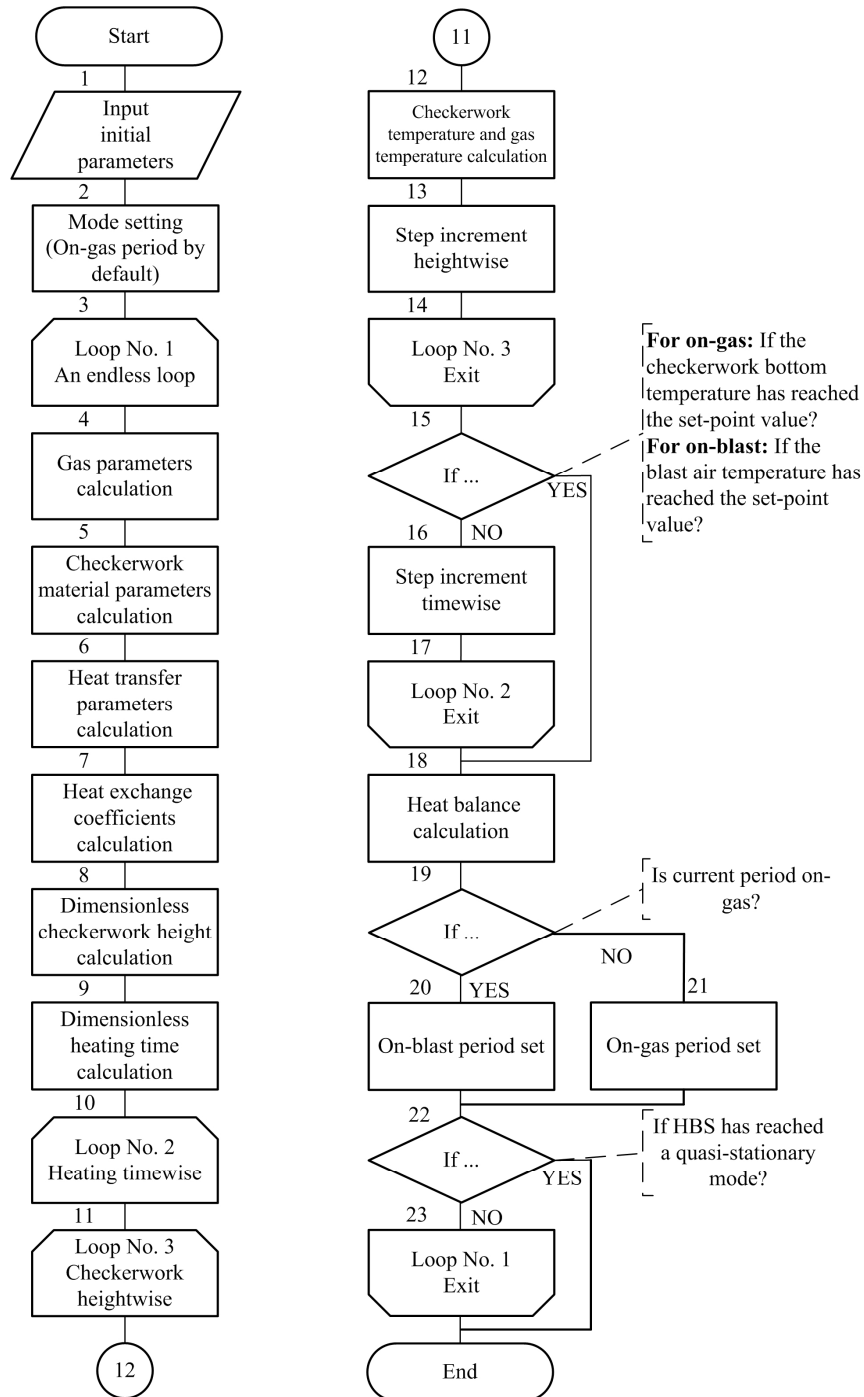


Fig. 3. General flow-chart of the algorithm of the program

work state; accuracy of calculations ($2.5 - 5\text{ }^{\circ}\text{C}$); flow rate, pressure and chemical composition of the exchange gas; flow rate, pressure and chemical composition of the blast air; set-point temperature value of the checkerwork bottom while heating of the checkerwork ($300 - 400\text{ }^{\circ}\text{C}$), at which the on-gas period ends; set-point temperature value of the blast air at the checkerwork outlet during the heating period of the blast air ($1100 - 1250\text{ }^{\circ}\text{C}$), at which the on-blast period ends.

In block 2 the mode of operation of HBS stove is set (on-gas or on-blast). The on-gas is automatically set when the program starts.

In block 3 an endless loop No. 1 starts, it ends when the HBS operation reaches a quasi-stationary mode – similar on-gas and on-blast periods are repeated three times.

In block 4 the following parameters of the exchange gas or blast air are calculated depending on the average temperature and overpressure: average temperature during on-gas or on-blast; density; mass heat capacity; volumetric heat capacity; kinematic viscosity; dynamic viscosity; thermal conductivity; velocity actual and reduced to normal conditions.

In block 5 the following parameters of material of the checkerwork are calculated depending on the average temperature: average temperature during the on-gas or on-blast; heat capacity; thermal conductivity.

In block 6 the parameters of heat exchange are calculated: the criteria of Reynolds, Peclet, Prandtl, Nusselt, Biot and thermal conductivity coefficient.

In block 7 heat exchange coefficients relative to volume and surface area are calculated. Correction for thermal massiveness is introduced [22, 23].

In block 8 the dimensionless height of the checkerwork is calculated.

In block 9 the dimensionless heating time of the checkerwork is calculated.

At the region of the program (blocks Nos. 10–17), a loop No. 2 is performed with iterations on the checkerwork heating timewise. The loop exit is performed on reaching the set-point accuracy of the set-point checkerwork bottom temperature value for the on-gas period or the set-point blast air temperature value at the checkerwork outlet for the on-blast period.

At the region of the program (blocks Nos. 11–14) a loop No. 3 is performed with iterations on the checkerwork heightwise at the current heating time. The loop exit is performed on reaching the end of the checkerwork.

In block 12 the calculation of the temperature of the checkerwork and of gas at the current time of heating according to (18) and (19) are performed. In block 13 the transition to the next layer of the checkerwork along its height is performed. In block 14 the exit from the loop 3 is performed.

Block 15 carries out a check: For on-gas period: if the bottom of the checkerwork temperature has reached the set-point value? For the on-blast period: if the blast air temperature has reached the set-point value? If it did, then an exit from the loop No. 3 and the transition to block 18 is performed, if NOT - to block 16.

In block 16 the time step is incremented.

In block 17 the exit from the loop No. 3 is performed.

In block 18 the calculation of the heat balance of the heating cycle is performed.

Block 19 carries out a check: current period – on-gas? If YES – transition to block 20, if NOT - transition to block 21.

In block 20 the HBS operation mode – on-blast period is set.

In block 21 the HBS operation mode – on-gas period is set.

Block 22 carries out a check: if HBS has reached a quasi-stationary mode? If YES – transition to the end of the program, if NOT - transition to block 3.

Based on the dependencies (21) and (22) generated by the authors, an application was developed using MS Visual Studio, the object-oriented programming environment, the application allows to analyze the operation of HBS in the on-gas – on-blast cycle. The program runtime varies from 0.2 to 120 s depending on the number of partitions throughout the checkerwork height and the configuration of PC. Part of the source data is entered using dialog boxes. The initial data on the initial temperature distribution throughout the checkerwork height, the intermediate and final results of the application operation are presented as Excel files with an option to view input and output data in the form of values and graphs.

6 The experiment and results

To determine the model operation adequacy, the parameters of HBSs of the existing block (hereinafter referred to as HBS1, HBS2, HBS3 and HBS4) were taken: checkerwork dimensions, checkerwork block parameters, parameters of BFG and air taken for combustion, as well as graphs of temperature changes of the checkerwork bottom and the dome of both periods. The study block operated in a sequential mode, and this mode was maintained by the sequential operation of HBS1 and HBS4, which are in a good condition, and by the parallel operation of HBS3 and the worn-out HBS2.

The on-blast period was simulated at the completely closed mixing valve – the hot blast air temperature was not stabilized before being supplied into the blast furnace. The exchange gas parameters (density, heat capacity, thermal conductivity, viscosity) in the model are presented as functions of pressure and temperature in the form of polynomials obtained from tabular data of reference books [24, 25]. The program implementing the model was completed when HBS was put into quasi-stationary mode of operation.

The authors of this research processed the technological parameters databases for the HBS block for the 4 months period. Periods that were significantly less or longer than the average duration of the periods, as well as periods for which the values of the technological parameters do not correspond to the working range, are excluded from the processing.

In order to further study the HBS operating periods, the simulation model was adjusted with application of tuning coefficients, which allow considering the current state of HBSs, primarily – of the regenerative checkerwork.

Fig. 4 shows the graphs of the variation of the dome temperatures (at the left) and the grid space (at the right) for each HBS of block, operating in the on-blast (a black line) and on-gas (a gray line) periods, which were obtained using simulation results (a dashed line) and production data (a continuous line).

In Fig. 4 lines, corresponding to the operating HBS, are built on the basis of data processing for more than 350 on-blast – on-gas cycles. The calculated maximum con-

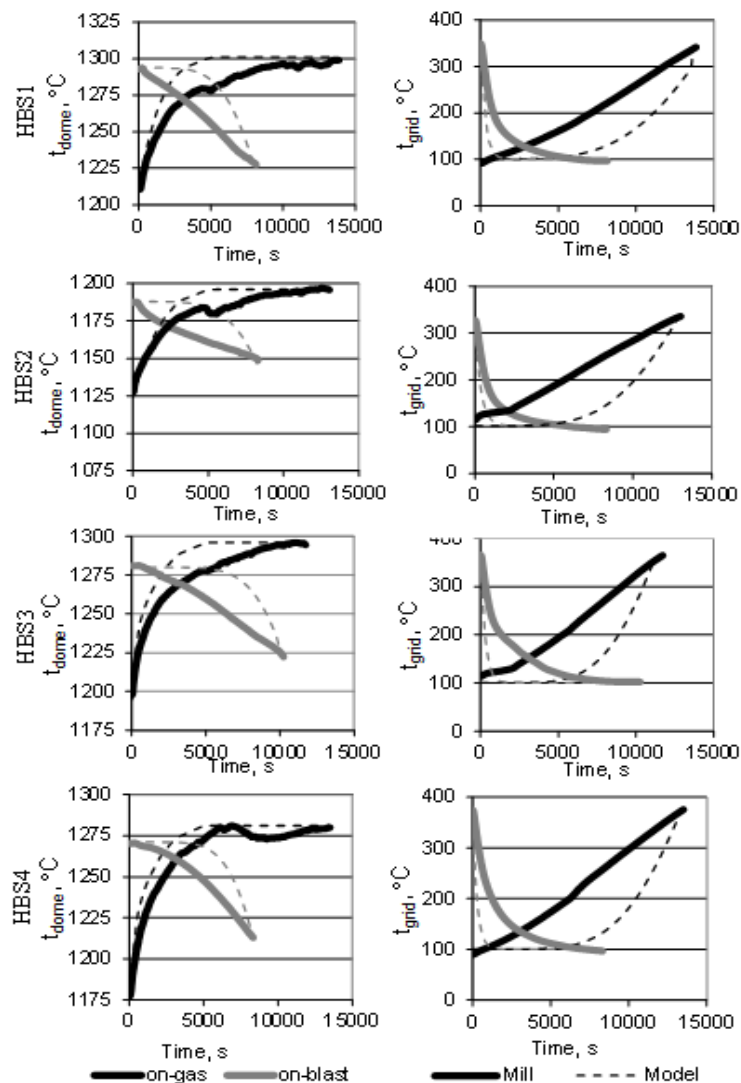


Fig. 4. Changes in the temperature of the dome t_{dome} and the grid device of HBS t_{grid} during the on-gas and on-blast periods

fidence interval for each value is ± 5 °C. According to the mill technological instruction, the maximum measurement error of the technical means complex for measuring the dome temperature is 12 °C, and the grid device temperature is 4 °C. The adequacy of the model is confirmed by the calculation of the Fisher criterion for each of the temperatures: the experimental values of the Fisher criterion are less than the tabulated ones.

The simulation results correspond to the processed values from the automation sys-

tem database. It should be noted that of the flue gases temperature is measured in the grid device space and, likewise it does not correspond to the gas temperatures in the checkerwork bottom layer, calculated by the model.

To assess the correct model functioning, the heat balance (Table 1) was simultaneously calculated for the on-blast and on-gas periods in accordance with the procedure [26]. The data (Table 1) indicate that certain balance sheet items represent the numbers of the same order and differ in the range from 1 to 8%.

Table 1. Heat balance of an operating HBS and simulation results

Parameter, TJ	Model results				Mill data				Deviation, %			
	HBS1	HBS2	HBS3	HBS4	HBS1	HBS2	HBS3	HBS4	HBS1	HBS2	HBS3	HBS4
On-gas period												
Initial checker- work heat Q_{ch}^{i}	0.494	0.217	0.307	0.392	0.513	0.237	0.332	0.418	4	8	8	6
Checkerwork heat after heating	1.109	0.453	0.757	1.028	1.128	0.488	0.806	1.105	2	7	6	7
Accumulated heat Q_{ch}^{a}	0.614	0.235	0.450	0.636	0.601	0.237	0.443	0.595	2	1	2	7
Flue gas heat Q_{g}^{f}	0.078	0.034	0.061	0.086	0.084	0.037	0.064	0.093	7	6	4	8
Exchange gas heat Q_{g}^{e}	0.698	0.270	0.517	0.728	0.718	0.282	0.517	0.721	3	4	1	1
On-gas period heat balance												
$(Q_{\text{ch}}^{\text{a}} + Q_{\text{g}}^{\text{f}})/Q_{\text{g}}^{\text{e}}$	0.99	1.00	0.99	0.99	0.95	0.97	0.98	0.95	-			
On-blast period												
Blast air heat Q_{g}^{b}	0.63	0.24	0.46	0.65	0.61	0.23	0.45	0.63	4	4	1	3
“On-gas – On-blast” cycle heat balance												
$Q_{\text{ch}}^{\text{a}}/Q_{\text{g}}^{\text{b}}$	1.02	1.02	1.01	1.02	1.01	0.96	1.02	1.06	-			

Furthermore the adequacy of the model was also confirmed: in work [27] during an on-site research of the regenerative heat exchanger operating at overpressure; by using the computer application described in work [19]. The results of the simulation obtained with the method of the authors of the research and with the method described in the article [19] are almost identical (the error is less than 2%).

7 Conclusion

Based on the review of the existing models describing the heat exchange in HBS, using the dependencies (18, 19) obtained by the authors, a computer model of HBS heat operation during different operating modes of HBS was proposed and implemented as a software application.

The adequacy of the simulation results was confirmed by the actual mill data and by the results of the model implemented on the basis of the different approach.

The model will be applied with the objective of increasing the blast air temperature: to optimize the implementation sheet of the HBS block mode; to study new types of refractories for the production of checkerworks; to assess the possibility of using the high-calorific fuel gas; to study new types of checkerworks while designing HBSs.

In automation systems, the model will be applied as an informational one, besides: it can be used to calculate the set-point values of the fuel-air parameters, dome temperature and to output these values in the control loops; to control the HBS block on the basis of estimation of the duration of the on-blast and on-gas periods.

References

1. Hausen, H.: Wärmeübertragung im Gegenstrom, Gleichstrom und Kreuzstrom. 2nd ed. Springer-Verlag, Berlin (1976)
2. Gres, L.P.: High-efficiency heating of the furnace blast [Высокоэффективный нагрев доменного дутья]. Porogi, Dnepropetrovsk (2008)
3. Kitayev, B. I., Timofeyev, V. N., et al.: Heat and mass transfer in a dense layer [Тепло- и массообмен в плотном слое]. Metallurgiya, Moscow (1970)
4. Tarasov V.P.: Gas dynamics of blast furnace process [Газодинамика доменного процесса]. Metallurgiya, Moscow (1990)
5. Schumann, T.E.: Heat transfer: a liquid flowing through a porous prism. Journal of the Franklin Institute. 208 (3): pp. 405-416 (1929)
6. Hausen, H., Binder, J.A.: Vereinfachte Berechnung der Wärmeübertragung durch Strahlung von einem gas an eine wand. International Journal of Heat and Mass Transfer. 5 (3-4): pp. 317-327 (1962)
7. Willmott, A.J.: Simulation of a thermal regenerator under conditions of variable mass flow. International Journal of Heat and Mass Transfer. 11 (7): pp. 1105-1116 (1968) doi:10.1016/0017-9310(68)90028-8
8. Willmott, A.J.: The regenerative heat exchanger computer representation. International Journal of Heat and Mass Transfer. 12 (9): pp. 997-1014 (1969) doi:10.1016/0017-9310(69)90111-2
9. Koifman, A.A., Simkin, A.I., Tomash, A.A.: Heavy pressure gases - brick checkerwork heat exchange intensification possibility [Возможность интенсификации процессов теплообмена между кирпичной насадкой и газами под давлением]. Izvestiya VUZ. Chernaya Metallurgiya, vol. 5: pp. 39-42 (2008)
10. Razelos, P., Benjamin, M.K.: Computer model of thermal regenerators with variable mass flow rates. International Journal of Heat and Mass Transfer 21 (6): pp. 735-744 (1978) doi:10.1016/0017-9310(78)90035-2
11. Muske, K.R., Howse, J.W., et al.: Hot blast stove process model and model-based controller (No. LA-UR-98-3183; CONF-980923-). Los Alamos National Lab., Computational Science Methods Group, NM (United States) (1998)
12. Muske, K.R., Howse, J.W., Hansen G.A., Cagliostro D.J.: Model-based control of a thermal regenerator. Part 1: Dynamic model. Computers & Chemical Engineering 24 (11): 2519-2531 (2000) doi:10.1016/S0098-1354(00)00614-1

13. Minet, F., Heyen, G., et al.: Dynamic data reconciliation of regenerative heat exchangers coupled to a blast furnace. *Computer Aided Chemical Engineering*. Elsevier, vol. 9: pp. 1053-1058 (2001) doi:10.1016/S1570-7946(01)80169-3
14. Kimura., Y., Takatani, K., Otsu, N.: Three-dimensional mathematical modeling and designing of hot stove. *ISIJ International* 50 (7): pp. 1040–1077 (2010) doi:10.2355/isijinternational.50.1040
15. Qi, F., Liu, Zh., Yao, Ch., et al.: Numerical Study and Structural Optimization of a Top Combustion Hot Blast Stove. *Advances in Mechanical Engineering*. 7(2): 10-19 (2015) doi:10.1155/2014/709675
16. Zetterholm, J., et al.: Model Development of a Blast Furnace Stove. *Energy Procedia*. Vol. 75: pp. 1758-1765 (2015)
17. Zetterholm, J., et al.: Dynamic modelling for the hot blast stove. *Applied Energy*, 185 (2): pp. 2142-2150 (2017) doi:10.1016/j.apenergy.2016.02.128
18. Kobyshev, E.I., Simkin, A.I., Kravchenko, V.P.: Heating of the packing in a blast-furnace air heater with an internal combustion chamber. *Steel in Translation*. 44 (1): pp. 38-42 (2014) doi:10.3103/S0967091214010094
19. Simkin, A.I., Kobyshev, E.I.: Control model of the heating hot blast stove regenerative chamber based on fuzzy knowledge with training set. *Metallurgical and Mining Industry*. 7 (6): pp. 96-101 (2015)
20. Zhang, H., Jiang, Z., Cheng, Z., Gui, W., et al.: Two-Stage control of endpoint temperature for pebble stove combustion. *IEEE Access*. Vol. 7: pp. 625-640 (2019) doi:10.1109/ACCESS.2018.2885581
21. Koifman, A.A., Simkin, A.I., Tomash, A.A. Modeling of heating of checkerwork of hot blast stove operating under pressure [Моделирование нагрева насадки доменного воздушнонагревателя, работающего под давлением] Reporter of Priazovskyi State Technical University: Collection of scientific papers – Part 2. Vol. 19: pp. 203-206 (2009)
22. Solomentsev, S.L.: Rational types of checkerworks and hot blast stoves [Рациональные типы насадок и доменных воздушнонагревателей]. Lipetsk State Technical University, Lipetsk (2001)
23. Wegman, E.F. (eds): Blast Furnace Production: Reference Edition [Доменное производство: Справочное издание в 2-х томах]. Vol. 1, Metallurgy, Moscow (1989)
24. Vargaftik, N.B., Vinogradov, Y.K., Yargin, V.S.: Handbook of physical properties of liquids and gases. Pure Substances and Mixtures. Begell House (1996)
25. Hilsenrath, J., Beckett, C. W., Benedict, W. S., et al: Tables of thermal properties of gases. Technical Report 564, National Bureau of Standards, (1955)
26. Koifman, A.A., Simkin, A.I.: Thermal balance of regenerative heat exchanger operating under pressure [Тепловой баланс регенеративного теплообменника, работающего под давлением]. Reporter of Priazovskyi State Technical University: Collection of scientific papers –Part 2. Vol. 18: pp. 141-144 (2008)
27. Koifman, A.A., Simkin, A.I.: Experimental determination of the effect of pressure on the gas-coolant heat regenerative checker. *Metallurgical and Mining Industry*, 4(2): pp. 107-112 (2012)

Expert system for assessing the labor professions complexity

Oleh Zaritskyi¹[0000-0002-6116-4426], Petro Pavlenko²[0000-0002-2581-230X]

¹National aviation university, avenue Kosmonavta Komarova, 1, Kyiv, 03680, Ukraine
zaritskyi@nau.edu.

²National aviation university, avenue Kosmonavta Komarova, 1, Kyiv, 03680, Ukraine
uapetrpav@ukr.net

Abstract. The problem of building a model of professional activity (working specialties) is an urgent scientific task due to the need for a clear formalization of tariff assessment levels, built on indicators of work complexity. The article deals with topical issues of formalizing the factors characterizing the complexity of the work, the task of developing an expert system designed to automate the calculation of estimates due to the weak formalization of the developed factors of the work model. Also motivational indicator of work formalized.

Keywords: work complexity, expert system, motivational indicator of work.

1 Introduction

Labor (professional activity) - conscious deliberate activities of people, aimed at changing the subject of labor with the tools of work in certain organizational forms work. The main content and components of labor are work performed by the employee. Work - certain tasks and duties that are performed, are been performing or should be performed by one person.

The job (work) complexity is an objective characteristic that reflects the content of the labor process and manifests itself in the requirements for worker qualification. Analytical evaluation of professional activity involves the allocation of the upper level characteristics of the model structural elements (entities) (Fig. 1) [1, 2], which, in turn, combine the characteristics of individual operations with certain features: educational-qualification level, methods, means of operations etc., i.e. expand and combine elements of models.

Professional activity is seen as the application of efforts and knowledge, skills and abilities to achieve the goals of the organization. The meaningful analytical evaluation of professional activity involves the presentation and processing of quantitative data using the theory of information and methods of mathematical apparatus and modeling, which involves the development of an ideal design, the so-called conceptual (content) model. The conceptual model is used to formalize the boundaries of the research object and needs further detail for the transition directly to the subject of research within the framework of the corresponding structural and functional-logical (information) models, detailed to a level sufficient and necessary for the analysis of professional activity - the level of entities and their attributes [3, 4].

2 Professional activity model

The proposed method of allocating professional activity structural elements and combining them with a clearly defined feature in groups allows to obtain a professional activity structural model at a more detailed level, in which the main emphasis is on the entities characteristics and the connection between them as elements of a complex system (Fig. 1).

Thus, professional activity consists of the number of successive or parallel operations and procedures, which in general constitute the technology of work, determined during the division of labor.

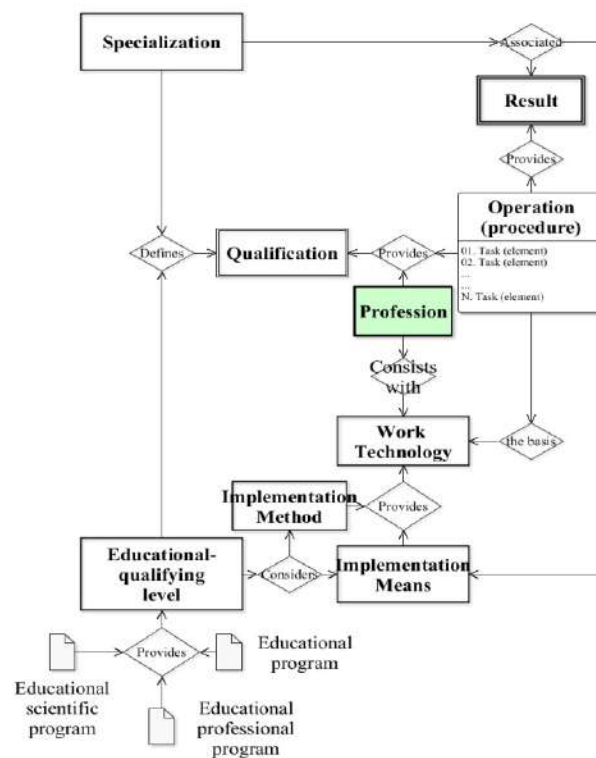


Fig. 1. Professional activity model.

The technology of work is a rational sequence of operations and procedures that are performed using specific means and methods of production in order to influence the object of management. The basis of the technology of work is the operation (Fig. 1).

Operation is any action, event (or system of events) that is united by a single plan and aimed achieving a corresponding goal. The complex of actions performed within the framework of one operation may consist of several elements (tasks), therefore the operation can be considered as a set of elements (tasks) of labor processes performed

by employees. Target completeness of action with respect to the performer is considered as the basis for the construction of the operation.

Thus, the operation (task) is a basic element of professional activity, which is subject of formalization in the form of a structural model for further functional-logical modeling within the framework of developing the theoretical foundations of analytical evaluation of professional activity.

The generalized structural model of professional activity in the annotation "entity – connection" [5] (Figure 2) describes the conceptual structure of a complex socio-technical system, which is used for further detail through parametric studies.

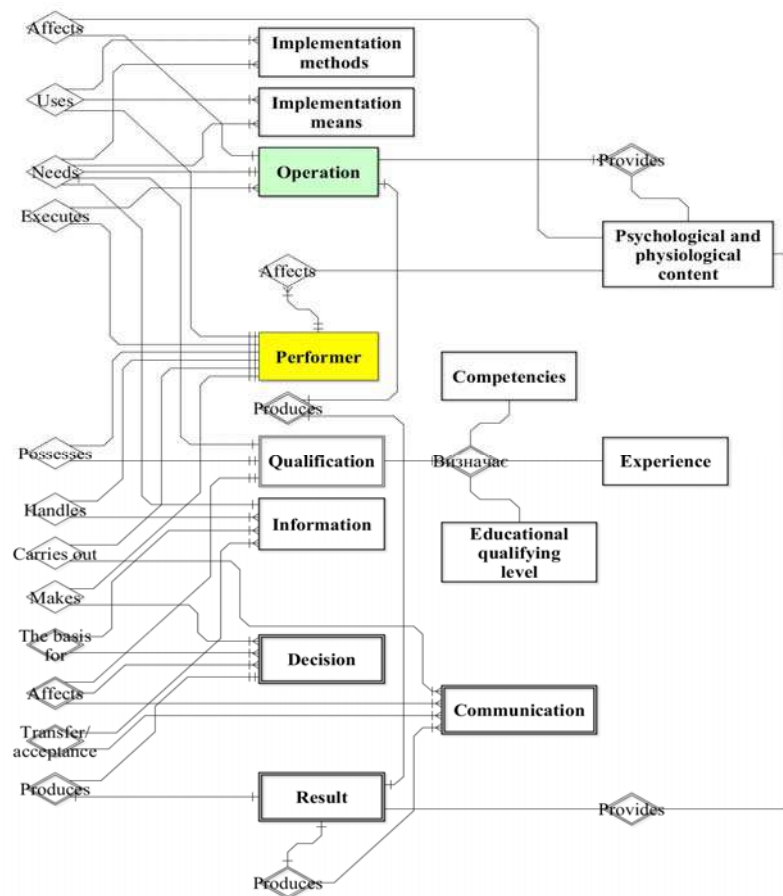


Fig. 2. Conceptual model of professional activity.

Existing methods for work complexity assessing consider professional activity in terms of certain efforts aimed at achieving the technological goal, and is determined by the combination of technical and organizational factors of the labor process: the complexity of equipment, its maintenance; the complexity of tools, etc.

The main factors determining the complexity of workers' work were grouped into four main groups [6]:

1. Technological factors.
2. Organizational factors.
3. Factors of responsibility.
4. Specific factors.

In addition, the complexity of labor is determined by the type of labor process, including the degree of its mechanization for the working professions:

Manual labor processes - executed manually with the help of mechanized tools of labor. Within the framework of this type of labor processes, it is necessary to distinguish separately hand-made creative work, which differs from most hand-made works by the increased content of creativity elements, artistic imagination, individual author's character of performance and other qualities inherent in, for example, work on the manufacture of folk arts products, jewelry-filigree production etc.

Machine labor processes - the main work is performed by machines, and elements of auxiliary work - manually or with the help of mechanisms. For example, machining of parts on mechanized feeders.

Machine-manual labor processes - performed by machines or mechanisms with the direct participation of the worker, when both the energy of the machine and the efforts of the performer are used simultaneously.

Automated labor processes - the main work is mechanized completely, and the auxiliary - partly (semiautomatic) or completely (automatic). In this case, the management of the work of the mechanisms is carried out automatically. In such cases, the function of the workers is reduced to the installation of machines, monitoring their work, eliminating defects, and on semi-automatic machines - to periodically present raw materials and the removal of finished products.

Hardware (instrumental) labor processes - carried out with special equipment (apparatus) by influencing the subject of work of thermal, electrical or chemical energy. In this case, the workers regulate the course of such processes. For example, iron melting, heat treatment of parts, most processes in the chemical and petroleum industries, etc.

The distinction between the types of labor processes affects the evaluation factors for each specialty. The essence of the assessing the complexity of labor analytical method is in a consistent, factor-to-factor assessment of the complexity of labor, the result of which is the integral index of complexity, - is the sum of evaluations by individual factors. Using this approach allows comparing the complexity of various, technologically unrelated works with a sufficient level of accuracy.

Each of these groups is divided into lower level factors that have a verbal description. To provide an objective assessment of the labor complexity, it is necessary to formalize the result in a digital form, taking into account the criteria of complexity for each factor, and their influence.

The methodology involves a combination of research methods and strategies, as well as certain hypotheses and ideas that need to be clarified and confirmed by relevant research:

1. It is assumed that in any profession it is possible to identify and evaluate a set of common factors that describe the main functional areas in its content, which will ensure the requirement for universality of the model of activity.

2. In principle, it is possible to allocate the minimum and sufficient number of its structural elements to ensure the adequacy of the model.

3. It is assumed that presentation of professional activity in the form of a graph (graphic model) will allow to explore it as a system of interconnected operations and to calculate their weight characteristics.

4. It is assumed that the use of fuzzy set theory methods to describe the weakly structured elements of the model (factors, characteristics), will allow to ensure its universality by extending the boundaries of modeling.

The proposed hypotheses meet the requirements of the ability, in particular explain all the phenomena and processes for which they have been used; can describe a broader class of phenomena and processes (new types of professional activities) and are fundamentally simple.

An analytical evaluation of professional activity is based on the analysis and evaluation of clearly formalized characteristics of operations (tasks) performed by the employee within their responsibilities. Characteristics of operations (tasks) are formalized exclusively with the help of structural modeling with further detailing of the main attributes within the functional-logical (informational) model, the main assumption during the development of which is given by the hypothesis 2.

Consider the structural model of the field of knowledge. Functional-logical (information) model (Fig. 3) - model of the research object, presented in the form of information describing the structural elements and variables of the object essential for this consideration, the characteristics of the elements, the connections between them, the inputs and the outputs of an object, which allows, through the input of information on the change of input values, to model the possible states of the object, that is to classify the type of professional activity and carry out its evaluation using analytical methods [7 - 9]. Thus, the conceptual structure of the subject area (Fig. 3 – 5) can be represented by a plurality of structural elements, each of which is described by the basic characteristics.

2.1 Technological factors

Depending on the labor process mechanization level, the influence of technological factors on its complexity varies significantly, as the degree of contact between the worker and the subject of work also changes. On the other hand, since the growth of mechanization leads to complicated equipment management and the complexity of the technological process and requires higher qualifications, the complexity of labor is greatly increased. Technological complexity factors are divided into components: the complexity of managing tools, the complexity of labor objects, the complexity of technological processes (Fig. 3).

2.2 Organizational factors

Organizational factors characterize the employee's autonomy in the process of performing the work and the breadth of performed works set (Fig. 4).

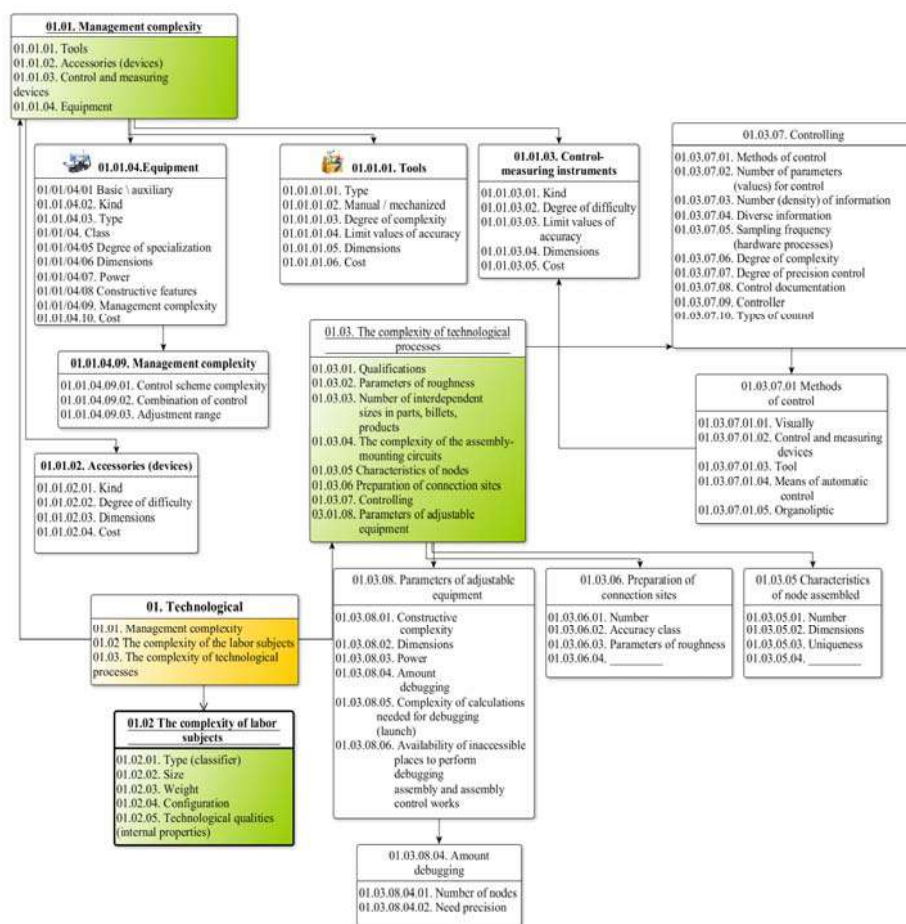


Fig. 3. Technological factors of the activity model.

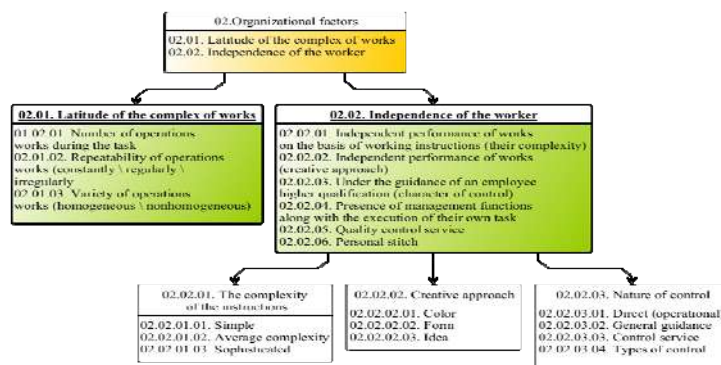


Fig. 4. Organizational factors of the activity model.

2.3 Factors of responsibility

The factors of responsibility allow us to consider work from the point of view of several types of responsibility (Fig. 5): for safety of others and one's own life, for the final result (work completion) and its influence on other units (employees) and the enterprise as a whole, as well as the place of work (its significance) in the technological process. The significance of the work characterizes the degree to which the employee is aware of its significant impact on the organizational system, the organization as a whole, and possibly on society. Completeness means that it requires and allows the completion of an entire and explicit cycle of operations, operations or work from beginning to end with a visible (understandable) result.

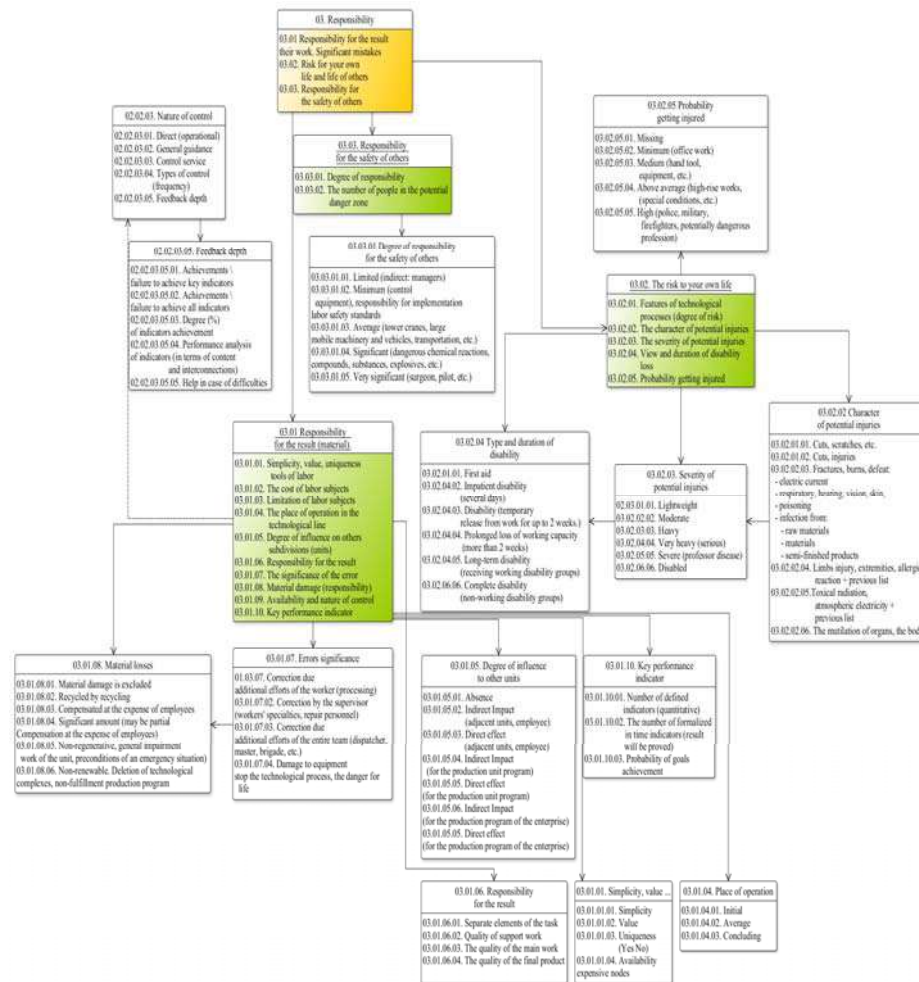


Fig. 5. Factors of responsibility in the model of activity.

3 Indicator of the work motivation potential

The F.Gerzberg's theory of the work characteristics addresses five main factors that determine the motivation and satisfaction of the employee in the work: variety (VAR), completeness (end result, COM), significance (degree of influence, SIG), independence (IND), feedback (control, FDB). All these factors and their attributes were considered in the models (Fig. 3-5).

In practice, using the procedure for diagnostic evaluation of the work characteristics, indicators related to the five mentioned aspects are combined and a single indicator – motivational potential indicator (MPI) is received.

$$MPI = \left(\frac{VAR + COM + SIG}{3} \right) \times IND \times FDB \quad (1)$$

The variety of the operation (work) is described by aspects of 02.01.02 and 02.01.03, which are evaluated by the corresponding ordinal scales (Table 1, 2).

Table 1. Repeatability of operations (works), k_1

<i>Attribute value</i>	<i>Scores</i>
Constantly	1
More regularly than constantly	2
Regular	3
Regularly with variable schedule	4
Irregularly	5

Table 2. Variety of operations (works), k_2

<i>Attribute value</i>	<i>Scores</i>
Monotonous homogeneous	1
Monotonous heterogeneous	2
Varied homogeneous	3
Varied heterogeneous	4

Completeness of work is determined by aspects of 03.01.04., 03.01.06. (Table 3, 4). The coefficient of diversity is used together with the index of homogeneity of the profession, which is determined by the coefficient of clustering k_{kl} of the work graphic model [10 - 11].

Table 3. The location of the operation in the technological line, k_3

<i>Attribute value</i>	<i>Scores</i>
Initial	1
Average	2
Concluding	3

Table 4. Responsibility for the result, k_4

<i>Attribute value</i>	<i>Scores</i>
Separate elements of the task	1
Quality of support work	2
The quality of the main work	3
The quality of the final product	4

The significance of the work is presented by aspects of 03.01.05., 03.01.07., 03.01.08. (Table 5-7).

Table 5. Degree of influence on other units (employees), k_5

<i>Attribute value</i>	<i>Scores</i>
Missing	0
Indirect impact on adjacent units (employees)	1
Direct influence on adjacent units (employees)	2
Indirect impact on the production program of the unit	3
Direct impact on the production program of the unit	4
Indirect impact on the production program of the enterprise	5
Direct impact on the production program of the enterprise	6

Table 6. Error significance, k_6

<i>Attribute value</i>	<i>Scores</i>
Correction due to additional employee efforts (processing)	1
Correction by the supervisor	2
Correction due to additional efforts of the whole team (dispatcher, master, brigade, etc.)	3
Damage to the equipment, stopping of the technological process, danger to life	4

Table 7. Material damage, k_7

<i>Attribute value</i>	<i>Scores</i>
Material damage is excluded	0
Recovery due to recycling	1
Compensation at the expense of the employee	2
Significant amount (compensation is possible partly at the expense of the employee)	3
Non-recoverable (violation of the general work of the unit, precondition of the emergency situation)	4
Non-recoverable (failure of process equipment, failure of production program)	5

Independence of work is determined by aspects of the factor 02.02. (Table 8,9).

Table 8. Independence at work, k_8

<i>Attribute value</i>	<i>Complexity</i>	<i>Scores</i>
Under the direction of the head of the highest category (leader)	-----	1
Independently based on working instructions (complexity of instructions)	simple	2
	medium	3
	complexity	4
	complex	5
Independently (creative approach)	color	6
	form	7
	plan (idea)	8
Presence of management functions	-----	8

Table 9. The nature of control, k_9

<i>Attribute value</i>	<i>Scores</i>
Direct (urgent)	1
General management (subdivision)	2
General control (functional direction)	3
Operations management (enterprise), statement of general objectives	4

The feedback is determined by the aspects included in the 02.02.03 factor. (Table 10, 11)

Table 10. Depth and volume of feedback, k_{10}

<i>Attribute value</i>	<i>Scores</i>
Achievement \ not achievement key indicators	1
Achievement \ not achievement of all indicators	2
Degree (%) achievement of all indicators (by periods)	3
Analysis of the indicators performance (in terms of content and interconnections)	4
Help in case of difficulties	5

Table 11. Types of control, k_{11}

<i>Attribute value</i>	<i>Scores</i>
Finish	1
Selective	2
Operative	3
Constant	4

Using the assessment scales (Table 1-11), the formula (1) can be submitted (2):

$$\overline{MPI} = \left[\frac{\left(\frac{k_1 + k_2 * k_{kl}}{2} \right) + \left(\frac{k_3 + k_4}{2} \right) + \left(\frac{k_5 + k_6 + k_7}{3} \right)}{3} \right] \times \left(\frac{k_8 + k_9}{2} \right) \times \left(\frac{k_{10} + k_{11}}{2} \right) \quad (2)$$

The value of the motivational potential of work is theoretically in the range of 1 - 100%, however, there is practically no work with motivational potential that accepts minimum and maximum values. The practical application of evaluating the motivational potential of the work consists in the development of recommendations for the improvement, reorganization of some aspects of work in order to improve the indicators of these factors (Table 1 - 10), since their degree of manifestation determines the probability of obtaining high results of work both in personal and organizational sense.

4 An expert system for work complexity assessing

The existing methodology involves evaluating the factors and their characteristics using a grading scale, which somewhat complicates the evaluation itself, as many factors can't be described in digital form. That is why the task of developing the concept of a knowledge-based information system was solved and a prototype expert system for the task evaluation of complexity of work was developed. This approach involves reviewing expert assessments of the task significance in terms of the fuzzy sets theory and the use of expert knowledge.

Formalization of expert's knowledge is related to the description of qualitative characteristics, which are usually not structured and can't be unambiguously interpreted. In addition, in problems solved with the help of intelligent systems, it is necessary to use inaccurate knowledge that can't be accurately interpreted as true or false, for example, the importance of the operation, its implementation time - the characteristics that determine the weight of each operation.

Using the conceptual apparatus of fuzzy logic, linguistic variables were introduced, which are part of the models of professional activity. Linguistic variable is a variable whose value is determined by the set of verbal characteristics of some property [12].

Each value of a linguistic variable is defined as a fuzzy set, which, in turn, is described by the base ordinal scale X and by the membership function $\mu(x)$.

The main tasks within the development of an expert system for the analytical evaluation of professional activities are the task of developing an appropriate knowledge base, mechanisms for its filling and integrating the work memory block with the module of algorithms for working with entities. The terminology of CLIPS [13 - 15] is used in the presentation of the material in connection with the free distribution of its code and the syntax typical of many common programming languages of expert systems.

The main elements of the expert system are represented by facts, production rules and rules within the framework of IT development of professional activity analytical evaluation. Within the scope of the study, a field of knowledge was defined - a model of knowledge about the subject field, that is, the transition from structured knowledge to its formalization was carried out. Production model N_p of knowledge (products) filed is presented by cortege (3):

$$N_p = (S_{p_d}; R_{a_p}; A^r \rightarrow A^s; R_{b_p}), \quad (3)$$

where S_{p_d} – description of the situations class; R_{a_p} – the condition under which the products are activated; $A^r \rightarrow A^s$ – the products core (3); R_{b_p} – condition (action) to be performed after the core is implemented (4):

$$P_t = \bigwedge_{r=1,r} A^r \rightarrow \bigwedge_{s=1,s} A^s, \quad (4)$$

where P_t – production designation; t – number of products in the production model; A^r – variable predicate (fact) from r variables, which is called parcel; A^s – variable terminal predicate from s variables, which is called the conclusion.

Input variables predicates that describe the current state of the model of professional activity and are part of the production model solely as parcels (5):

$$A^r \left([a_{ij}]_1, [a_{ij}]_2, \dots, [a_{ij}]_r \right). \quad (5)$$

The terminal predicates that determine the new states of the model parameters and are part of the production model solely as conclusions (6):

$$A^s \left([a_{ij}]_1, [a_{ij}]_2, \dots, [a_{ij}]_s \right). \quad (6)$$

Using the conceptual apparatus of fuzzy logic, fuzzy sets are introduced that describe the input and terminal predicates defined by the membership function. Linguistic variables with the help of sets of standardized functions are given slightly structured parameters of the professional activity model. The templates for presentation of knowledge in the expert system taking into account (3-6) and the syntax of the productive programming language CLIPS [16] are presented as designs (7):

$$P_t = \left(\begin{array}{c} \text{defrule} \left(\langle N_p \rangle [S_{p_d}] \right) \\ \left(\begin{array}{c} R_{a_p} \\ \left(\langle a_{ij} \rangle \langle value \rangle \right) \\ (...) \\ \left(\langle a_{ij} \rangle \langle value \rangle \right) \end{array} \right) \\ \Rightarrow \\ \left(\begin{array}{c} \text{assert} \left(\langle a_{ij} \rangle \langle value \rangle \right) \\ R_{b_p} \end{array} \right) \end{array} \right) \quad (7)$$

The rule header defines conditional elements (CE), or LHS-rule - Left Hand Side - the left side of the rule or the input predicates (5). If the rules templates do not conflict with the active facts and the condition is fulfilled R_{a_p} , the core of the product is activated and moved from the knowledge base to the working list of rules of the machine of logical output [16]. Terminal (6) part of the rule is given by a list of actions (RHS-rule - Right Hand Side), which must be executed when the rule is activated

with the assert operator.

The generalized syntactic structure of the knowledge field, realized within the expert system, is given by the expression:

$$F_k = \langle A, M, E \rangle,$$

where $A = \langle A, A_v, B \rangle$ – the structure of the input data to be interpreted in the system; $M = \langle S_k, S_f \rangle$ – data interpretation model; S_k – conceptual structure of the subject area; S_f – a functional structure that simulates the scheme of reasoning during the interpretation; $E = \langle E_o, E_a \rangle$ – output data structure.

The conceptual structure S_k is based on the revealed conceptual structure of the subject industry with the help of the conceptual analysis paradigm and the principles of hierarchy of concepts constructing. S_k represented by a structural model of professional activity developed within the framework of the research. S_f is presented in accordance with the information model and rules of the knowledge base, the mechanisms of its filling and integration of the working memory block with the module of algorithms for working with entities, that is, the integration of data and algorithms implemented on different bases: imperative and declarative programming languages.

Knowledge-based expert system [17] provides for the activation of relevant rules for the determination and calculation of numerical estimates of factors.

5 Conclusions

1. Patterns of weak structural characteristics of professional activity model structural elements (facts) and rules of the knowledge using the syntax of the declarative programming language CLIPS, which provided the opportunity for practical implementation of program modules, taking into account the characteristics of structural and information models of professional activity are developed.

2. The description of the model's structural elements with the help of the knowledge base developed rules, allowed to expand the boundaries of modeling, which was a significant disadvantage of existing systems and to practically implement the model data within the production information systems.

3. The developed templates of the basic rules allowed to determine the characteristics of the structural elements of the professional activity model and the weight of each operation in terms of linguistic variables, which required the transition to numerical estimates in order to ensure interaction with the calculation module for the evaluation of all activities implemented through object-oriented programming languages.

4. The issue of restrictions on the use of models requires additional research, but the preliminary results give grounds for claiming that they can be used for a wide range of occupations, due to the compliance of the developed model with the requirements of universality.

6 References

1. Brannick, M.: Job Analysis: Methods, Research and Applications for Human Resource Management in the New Millennium. Sage Publishers, Thousand Oaks, CA (2002).
2. Wilson, M.: A history of job analysis. Historical perspectives in industrial and organizational psychology. Lawrence Erlbaum Associates, NJ. Mahwah (2007).
3. Zaritskyi, O.: Functional modeling of the basic elements of professional activity within the framework of the "Essence - Communication" model. Problems of Informatization and Management. 2(50), 70–75 (2015).
4. Zaritsky, O.: Modeling of professional human activity. Integrated quality assurance of technological processes and systems: materials of abstracts V international. scientific practice conference (2015), Chernigov NTU, p. 202.
5. Zaritskyi, O.: Developing analysis of job complexity for needs of national trade market. The sixth world congress "Aviation in the XXI century" "Safety in Aviation and Space Technologies" (2014), National academy of sciences of Ukraine, Ministry of education and science of Ukraine, National aviation university, vol. 1, pp. 1.9.46 – 1.9.48.
6. Luchaninova, S.: Methodology for assessing the labor complexity of workers of different professional groups in order to distinguish levels of the tariff assessment of the scale of qualification digits. Scientific research work. Kramatorsk: Center of Productivity, (2007).
7. Zaritskyi, O.: Data representing and processing in expert information system of professional activity analysis. Modern Problems of Radio Engineering, Telecommunications and Computer Science, Proceedings of the 13th International Conference on TCSET (2016), Lviv Polytechnic National University. pp. 718 – 720.
8. Zaritskyi, O.: Development of professional activity mathematical model. Eastern European Journal of Advanced Technology / PP "Technological Center". 1\4(79), pp.10-18 (2016).
9. Zaritsky, O.: Mathematical modeling of professional activity in the framework of assessing the complexity of the work performed. XIII international. scientific practice conference "Mathematical and simulation of systems simulation (2017), Chernigov NTU, pp. 160–163.
10. Girvan, M., Newman, M.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences. 99(12), pp. 7821–7826 (2002).
11. Pons, P., Latapy, M.: Computing communities in large networks using random walks. In Computer and Information Sciences-ISCIS (2005), pp. 284–293.
12. Zaritskyi, O.: Decision-making information model as a part of aircraft control system. 4th International conference. Methods and systems of navigation and motion control (MSNMC-2016), (2016), Ministry of education and science of Ukraine, pp.228 – 231.
13. CLIPS Reference Manual. Volume II. Advance programming guide. Version 6.22. Washington: Johnson Space Center (2007).
14. CLIPS Reference Manual. Volume I. Basic programming guide. Version 6.22. Washington: Johnson Space Center (2004).
15. CLIPS Reference Manual. Volume III. Interfaces Guide. Version 6.22. Washington: Johnson Space Center (2004).
16. Zaritskyi, O.: Theoretical bases, methods and technologies of development of the professional activity analytical estimation intellectual systems. 2nd IEEE International Conference on Advanced Information and Communication Technologies (2017), Polytechnic National University, pp. 101 – 104.
17. Zaritskiy, O.: Data representing and processing in expert information system of professional activity analysis. Modern Problems of Radio Engineering, Telecommunications and Computer Science, Proceedings of the 13th International Conference on TCSET (2016). Lviv Polytechnic National University, pp. 718 – 720.

Application of information technologies for automation of metallurgical melts surface properties calculations

Tetiana Levytska¹[0000-0003-3359-1313]

¹Priazovsky State Technical University, Universytets'ka str.7, Mariupol 87555, Ukraine
tlevitiisys@gmail.com

Abstract. An experimental facility has been developed for researching the surface properties of metallurgical melts, which, unlike the known ones, is equipped with an optical system and a PC, that allow a digital image of a lying drop, obtained in the experiment, to be processed using a high-speed software package, which significantly reduces the time of the experiment, allows to store data and simplifies the work of the experimenter.

Keywords: surface properties, physico-chemical experiment, experimental facility, digital image, computer, segmentation, recognition, linear filtering, pixel, intensity levels

1 Introduction

The laborious and expensive high-temperature physicochemical measurements are characterized by a change in the studying object during the experiment due to the interaction with the structural materials of the measuring cell and the furnace atmosphere [1]. Therefore, it is very important to reduce the duration of the experiment, to measure the maximum possible number of properties during this time and to strive to automate the work of the experimenter [2-3]. With the advent of digital technology for recording measurement information and processors for its processing, such a task becomes real.

The high speed and accuracy of digital recorders allow multiple measurements of each characteristic in a period of time until they change, which reduces systematic error and makes it possible to effectively apply statistical methods to reduce random error [4-5]. Naturally, the transition to digital technology and partial automation of the experiment should be accompanied by appropriate changes in the calculation methods and approaches to their implementation.

Recently, more attention has been paid to methods based on the analysis of a digitized image of a two- and three-phase contact zone during an experimental research of both surface and interfacial tension, and contact angles. Obviously, the further development of the fixed-drop method is directly related to the development of new methods and specially developed algorithms that allow automatic computer calculation of surface properties. To date, there are no well-tested, modern, high-

precision, experimental methods for measuring surface and interfacial tension, therefore the development of this direction is relevant and promising.

2 Formal problem statement

It is known [6] that the contour of a drop is defined by the Laplace's equation of a capillary and it has to be written like the differential second-order equation:

$$\frac{y''}{[1+(y')^2]^{\frac{3}{2}}} + \frac{y'}{x[1+(y')^2]^{\frac{3}{2}}} = \frac{(y+y_0)(\rho_1-\rho_2) \cdot g}{\sigma}, \quad (1)$$

where

σ – surface tension, N/m²;

ρ_2 – density of the medium, where the drop is, kg/m³;

ρ_1 – density of the medium, which is form the drop, kg/m³;

g – acceleration of gravity, m/s²;

y – coordinate of the point on the surface along the vertical axis;

y_0 – coordinate of the point on the top along the vertical axis.

3 Literature review

There are different methods of the experimental definition the surface tension of liquids: method of capillary rise, ring or slab detachment method, method of a recumbent drop, method of a hanging drop, drop weight method, method of the maximum pressure in gas bubble. Method of a recumbent drop gives the most exact result and nowadays it is widely used in the high-temperature searches [7-8].

In this method the metal drop melts on the horizontal refractory substrate or forcibly formed under the sharp edge of the crucible. Drop are photographed at a temperature of shaping and then it is measured to know its maximum diameter ($2r$) and inches under it (h).

All known methods described in the literature [6-20] are based on the use of the system parameters connection with some characteristic dimensions of the experimental drop profile. According to the tables with theoretical droplet shapes that have been calculated in advance, the connection between the characteristic sizes and the droplet parameters are made, or using the formulas that approximate the table values, the system parameters are found.

However, the calculation of capillary characteristics using tables is inconvenient and time consuming. It is not possible to fully automate such a calculation.

The inconvenience associated with the tabular description of the function led S.I. Popel, Y.P. Nikitin and S.M. Ivanov to the graphical integration of the equation [17]. The accuracy of the graphs published by them is 2–3%. The method presented by them is applicable to drops of arbitrary size. The calculation is carried out by the

experimenter manually. The method of graphical integration gives results that are close enough to those obtained by Bashfort and Adams, who calculated them in a more accurate way [7]. Manual construction of graphs for calculating σ is a complex and lengthy process, it is not possible to improve the accuracy of the method and make any changes to those already published by S.I. Poppel and other graphs. The method of calculation proposed in the thesis allowed us to significantly speed up data processing and to fully automate the calculation of the surface tension of melts with a high degree of accuracy (less than 0.5%) compared to known methods.

Photographic methods of image registration, which are traditionally used in the practice of studying the capillary properties of melts according to the shape of a lying drop, don't allow obtaining final results directly in the course of the experiment. The tasks that confront the experimenter in this case — the processing of photographic materials, the manual measurement of photographs and the calculation of capillary characteristics using tables are very laborious and time consuming. Inevitably, by measuring the image, and especially by carrying out a tangent, the researcher introduces elements of subjectivity into this procedure, and the measurement error depends on the quality of the image. For example, tracing the maximum diameter per eye can give an initial error of 1.5%.

Analysis of the digitized image of the zone of two-and three-phase contact allows you to fully automate the calculations. Most of the known measurement methods are done manually. Although some automated systems using cameras [9] and cameras on a CCD matrix [8] are described. A number of calculation algorithms have been developed [9-11]. According to the authors of the article [12-13], such systems have not received wide distribution, and now further research is needed on methods, algorithms and systems for measuring capillary properties. The raster image of a large drop is formed using a digital camera[14]. The direct digital image is then processed on a PC. In this work, the Dorsey method is used. Most of the known measurement methods are done manually. Although some automated systems using cameras [7] and cameras on a CCD matrix [8] are described. A number of calculation algorithms have been developed [9-11]. According to the authors of the article [12-13], such systems have not received wide distribution, and now further research is needed on methods, algorithms and systems for measuring capillary properties. The raster image of a large drop is formed using a digital camera[14]. The direct digital image is then processed on a PC. In this work, the Dorsey method is used.

The analytic solution this differential equation doesn't have. An approximate solution can be obtained with a predetermined degree of accuracy. All known techniques described in the literature [15-19] are based on using of the coupling of the system parameters with certain characteristic dimensions of the experimental drop profile. According to the sheets with theoretical drop shapes, which are calculated in advance, a connection is made between the characteristic dimensions and parameters of the drop, or by the formulas approximating the tabulated values, find the parameters of the system. Calculating the capillary characteristics by using tables is inconvenient and time-consuming.

An analysis of the theoretical aspects of the recumbent drop method has shown that the previously developed formulas and tables are either difficult to apply or not

applicable at all for PC calculations, which requires transformations of the Laplace's equation to a form convenient for computer mathematical processing.

4 The numerical solution of the Laplace differential equation by empirical formulas

To find the particular solutions of the differential equation (1) that satisfies the initial conditions $y(0) = 0$; $y'(0) = 0$ for different values of the capillary constant, let's use the geometric meaning of the first and second derivatives.

As a result, was obtained the expression:

$$\frac{y''}{[1 + (y')^2]^{3/2}} = \frac{1}{R}, \quad (2)$$

which determines the plane curvature of the curve and the first derivative is equal to the tangent of the angle of inclination of the tangent to the axis 0x. Substituting these values into equation (1), was introduced the substitution:

$$A = \frac{\sigma}{(\rho_1 - \rho_2) \cdot g}. \quad (3)$$

So it was observed:

$$\frac{1}{R} + \frac{\frac{tg \varphi}{x \cdot \frac{1}{\cos \varphi}}}{A} = \frac{y + y_0}{A}, \quad (4)$$

$$\frac{\frac{tg \varphi}{x \cdot \frac{1}{\cos \varphi}}}{A} = \frac{\sin \varphi \cdot \cos \varphi}{x \cdot \cos \varphi}, \quad (5)$$

$$R = \frac{A}{(y - y_0) - A \cdot \frac{\sin \varphi}{x}}. \quad (6)$$

Equation (6) defines the radius of curvature at any point of the meridian section of the drop, but contains an undefined ratio $(\sin \varphi)/x$. When approaching the top of the drop of $x \rightarrow 0$, $y \rightarrow 0$, $\varphi \rightarrow 0$, the vertex both main droplet radius R_1 and R_2 drops of curvature equal to each other and equal to R_0 . From the Laplace's equation (1) for this case was observed

$$R_0 = \frac{2A}{y_0}.$$

Having $R_1 \rightarrow R_2 \rightarrow R_0 = 2 \cdot A / y_0$, this was substituted into equation (2) and found the limit of the right side of the resulting expression.

$$\frac{2 \cdot A}{y_0} = \lim_{x \rightarrow 0} \frac{A}{(y + y_0) - A \cdot \sin \varphi / x} = \frac{A}{y_0 - A \cdot \lim_{x \rightarrow 0} \frac{\sin \varphi}{x}}, \quad (7)$$

$$\lim_{x \rightarrow 0} \frac{\sin \varphi}{x} = \frac{y_0}{2 \cdot A}. \quad (8)$$

Were chosen on the meridional section line three sufficiently close to each other points M_{i-1} , M_i and M_{i+1} and drawn through them the normal lines (Fig. 1).

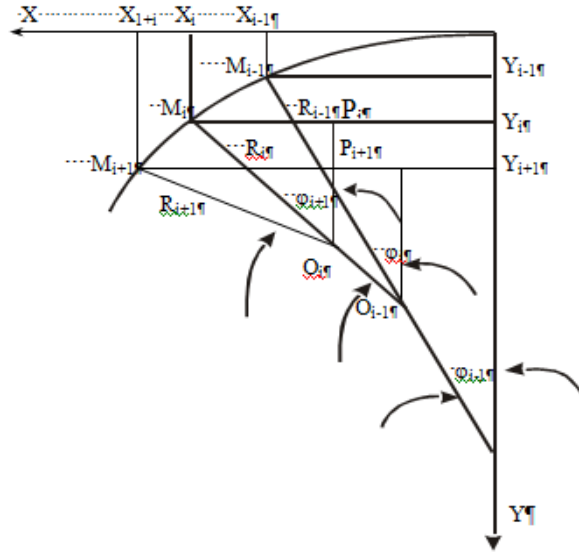


Figure 1. The scheme for constructing the drop elements for numerical integration.

The normal line passing through the point M_{i-1} intersects the normal line passing through the point M_i at the point O_{i-1} , and the normal lines passing through the points M_i and M_{i+1} intersect at the point O_i . It was denoted the angles formed by the normal lines $M_{i-1}O_{i-1}$, M_iO_i and $M_{i+1}O_i$, respectively, φ_{i-1} , φ_i , φ_{i+1} . To numerically integrate the differential equation (1) and calculate the droplet shape, it was assumed that for a sufficiently small change in the current angle φ in the range from φ_{i-1} to φ_i , the radius of curvature R_{i-1} (the segment $M_{i-1}O_{i-1}$ or the segment M_iO_{i-1}) of the initial angle φ_{i-1} of the original abscissa x_{i-1} and the initial ordinate y_{i-1} does not change. In the next interval of variation of the angle φ from φ_i to φ_{i+1} , the radius of curvature R_i (of the segment M_iO_i or $M_{i+1}O_i$) or the other is calculated for a new value of the angle φ_i , the new abscissa x_i and the new ordinate y_i .

From the rectangular triangle $O_iM_{i+1}P_{i+1}$ and $O_iM_iP_i$ was found:

$$\Delta y_i = |P_i - P_{i+1}| = R_i \cdot \cos \varphi_i - R_i \cdot \cos \varphi_{i+1} = R_i [\cos \varphi_i - \cos \varphi_{i+1}],$$

$$\Delta x_i = |M_{i+1}P_{i+1}| - |M_iP_i| = R_i \cdot \sin \varphi_{i+1} - R_i \cdot \sin \varphi_i = R_i [\sin \varphi_{i+1} - \sin \varphi_i],$$

$$y_{i+1} = y_i + \Delta y_i$$

$$x_{i+1} = x_i + \Delta x_i.$$

Thus, we have the following algorithm for calculating the meridional section line coordinates of the drop and volume (Fig. 2), which is implemented in the Delphi visual programming system.

We built dependences graphs of the coordinates $x = x(\varphi)$, $y = y(\varphi)$ of the meridional section of the drop on the current angle φ , as well as the volume of the part of the drop between its top and plane $y = y_0 = \text{const}$, shown schematically in Fig. 3

The graphs of the dependences of the coordinates $x = x(\varphi)$, $y = y(\varphi)$ of the Meridional section of the drop from the current angle φ , and also the volume of the part of the drop between its vertex and the plane $y = y_0 = \text{const}$ have functions at the corresponding values of the parameters $a_x, b_x, c_x, a_y, b_y, c_y, a_v, b_v, c_v$.

$$x = a_x \cdot \varphi^{b_x} \cdot \exp(c_x \cdot \varphi) \quad (9)$$

$$y = a_y \cdot \varphi^{b_y} \cdot \exp(c_y \cdot \varphi) \quad (10)$$

$$V = a_v \cdot \varphi^{b_v} \cdot \exp(c_v \cdot \varphi) \quad (11)$$

Therefore, for the empirical description of the numerical solution of the differential equation of the drop form (1), these dependences [4]. The coefficients of the obtained equations are found by the method of rectifying the obtained graphs and introducing new variables. Was found the ratio of the next to the previous value of any of the coordinates

$$\frac{x_{i+1}}{x_i} = \frac{a_x \varphi_{i+1}^{b_x} \exp(c_x \varphi_{i+1})}{a_x \varphi_i^{b_x} \exp(c_x \varphi_i)} = \left[\frac{\varphi_{i+1}}{\varphi_i} \right]^{b_x} e^{c_x [\varphi_{i+1} - \varphi_i]} \quad (12)$$

Take the natural logarithms from the left and right sides of this relation:

$$\ln(x_{i+1}/x_i) = b_x \ln(\varphi_{i+1}/\varphi_i) + c_x [\varphi_{i+1} - \varphi_i] = b_x \ln(\varphi_{i+1}/\varphi_i) + c_x \Delta \varphi$$

Denote $\ln(x_{i+1}/x_i)$ by $\tilde{x}$, $\ln(\varphi_{i+1}/\varphi_i)$ trough $\tilde{\varphi}$, then was obtained a linear relationship between the variables $\tilde{x}$ and $\tilde{\varphi}$.

$$\tilde{x} = b_x \cdot \tilde{\varphi} + c_x \cdot \Delta \varphi \quad (13)$$

While investigation unknown dependencies it is possible random errors associated with the measurement process. To reduce the effect of random measurement errors, was applied the least squares method, which allows to determine the parameters of the

chosen dependence, in which the deviation from the experimental data (in this case, the calculated data) is minimal.

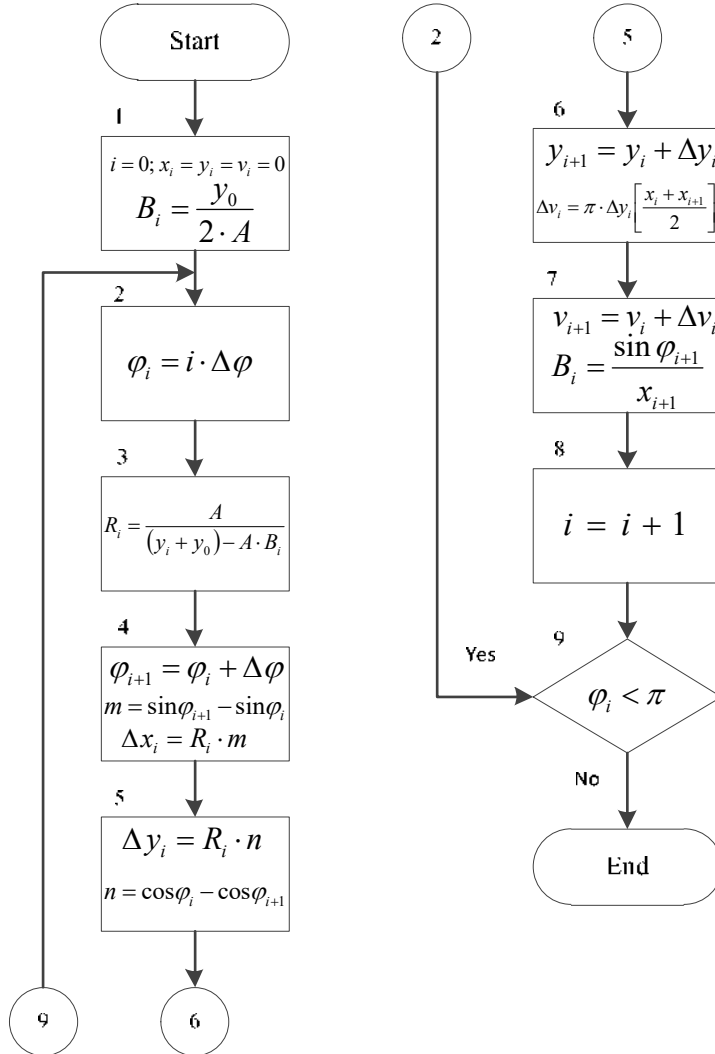


Figure2 .Algorithm for calculating the meridional section line coordinates of the drop and volume

There are calculation results $(x_1, \varphi_1), (x_2, \varphi_2), \dots, (x_n, \varphi_n)$ and the form of the function $\tilde{x} = b_x \cdot \tilde{\varphi} + c_x \cdot \Delta \varphi$. It is necessary to choose b_x and c_x so that the sum of the squares of the differences between the empirical and calculated values (deviations) is minimal:

$$\Phi(b_x, c_x) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [x_i - \tilde{x}]^2 = \min. \quad (14)$$

When the expression (13) is substituted into condition (14) it was obtained:

$$\Phi(b_x, c_x) = \sum_{i=1}^n [b_x \cdot \tilde{\varphi} + c_x \cdot \Delta\varphi - x_i]^2 = \min$$

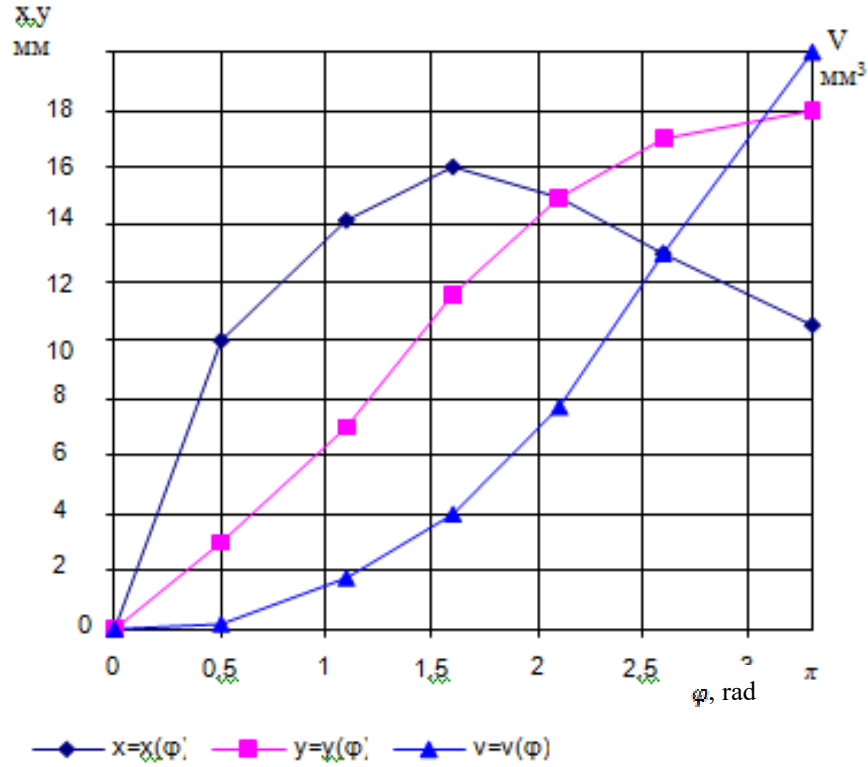


Figure 3. The dependence of the x, y coordinates and the volume of the drop v on the current angle φ

To find the values b_x and c_x that convert the left-hand side of the resulting expression to a minimum, it is necessary to equate the derivatives with respect to zero. A function can have an extremum (min) if all its partial derivatives are zero or nonexistent.

$$\frac{\partial \Phi(b_x, c_x)}{\partial b_x} = 0$$

$$\frac{\partial \Phi(b_x, c_x)}{\partial c_x} = 0$$

$$\begin{cases} 2\sum_{i=1}^n [b_x \cdot \tilde{\varphi} + c_x \cdot \Delta\varphi - x_i] \cdot \tilde{\varphi} = b_x \cdot \sum_{i=1}^n \tilde{\varphi}^2 + c_x \cdot \sum_{i=1}^n \tilde{\varphi} \cdot \Delta\varphi - \sum_{i=1}^n x_i \cdot \tilde{\varphi} = 0 \\ 2\sum_{i=1}^n [b_x \cdot \tilde{\varphi} + c_x \cdot \Delta\varphi - x_i] \cdot \Delta\varphi = b_x \cdot \sum_{i=1}^n \tilde{\varphi} \cdot \Delta\varphi + c_x \cdot \sum_{i=1}^n \Delta\varphi^2 - \sum_{i=1}^n x_i \cdot \Delta\varphi = 0 \end{cases}$$

$$\begin{cases} \underline{b_x} \cdot \sum_{i=1}^n \tilde{\varphi}^2 + \underline{c_x} \cdot \Delta\varphi \cdot \sum_{i=1}^n \tilde{\varphi} = \sum_{i=1}^n x_i \cdot \tilde{\varphi} \\ \underline{b_x} \cdot \Delta\varphi \cdot \sum_{i=1}^n \tilde{\varphi} + \underline{c_x} \cdot \Delta\varphi^2 \sum_{i=1}^n 1 = \Delta\varphi \cdot \sum_{i=1}^n x_i \end{cases}$$

This is the final form of the normal least-squares method [5]. It was solved the system and found empirical values according to Cramer's formulas:

$$c_x = \frac{\Delta c_x}{\Delta},$$

$$b_x = \frac{\Delta b_x}{\Delta},$$

$$\Delta = \begin{vmatrix} \sum \varphi^2 & \Delta\varphi \sum \varphi \\ \Delta\varphi \cdot \sum \varphi & \Delta\varphi^2 \cdot n \end{vmatrix}, \quad \Delta_{b_x} = \begin{vmatrix} \sum x_i \cdot \tilde{\varphi} & \Delta\varphi \sum \varphi \\ \Delta\varphi \cdot \sum x_i & \Delta\varphi^2 \cdot n \end{vmatrix},$$

$$\Delta_{c_x} = \begin{vmatrix} \sum \varphi^2 & \sum x_i \cdot \tilde{\varphi} \\ \Delta\varphi \cdot \sum \varphi & \Delta\varphi \cdot \sum x_i \end{vmatrix},$$

$$b_x = \frac{\Delta b_x}{\Delta} = \frac{\begin{vmatrix} \sum_{i=1}^n x_i \cdot \tilde{\varphi} & \Delta\varphi \sum_{i=1}^n \varphi \\ \Delta\varphi \cdot \sum_{i=1}^n x_i & \Delta\varphi^2 \cdot n \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n \varphi^2 & \Delta\varphi \sum_{i=1}^n \varphi \\ \Delta\varphi \cdot \sum_{i=1}^n \varphi & \Delta\varphi^2 \cdot n \end{vmatrix}} = \frac{\Delta\varphi^2}{\Delta\varphi^2} \frac{\begin{vmatrix} \sum_{i=1}^n x_i \cdot \tilde{\varphi} & \sum_{i=1}^n \varphi \\ \sum_{i=1}^n x_i & n \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n \varphi^2 & \sum_{i=1}^n \varphi \\ \sum_{i=1}^n \varphi & n \end{vmatrix}} = \frac{\begin{vmatrix} \sum_{i=1}^n x_i \cdot \tilde{\varphi} & \sum_{i=1}^n \varphi \\ \sum_{i=1}^n x_i & n \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n \varphi^2 & \sum_{i=1}^n \varphi \\ \sum_{i=1}^n \varphi & n \end{vmatrix}},$$

$$c_x = \frac{\Delta c_x}{\Delta} = \frac{\left| \begin{array}{cc} \sum \varphi^2 & \sum_{i=1}^n x_i \cdot \tilde{\varphi} \\ \Delta \varphi \cdot \sum_{i=1}^n \varphi & \Delta \varphi \cdot \sum_{i=1}^n x_i \end{array} \right|}{\left| \begin{array}{cc} \sum \varphi^2 & \Delta \varphi \sum_{i=1}^n \varphi \\ \Delta \varphi \cdot \sum_{i=1}^n \varphi & \Delta \varphi^2 \cdot n \end{array} \right|} = \frac{1}{\Delta \varphi} \frac{\left| \begin{array}{cc} \sum \varphi^2 & \sum_{i=1}^n x_i \cdot \tilde{\varphi} \\ \sum_{i=1}^n \varphi & \sum_{i=1}^n x_i \end{array} \right|}{\left| \begin{array}{cc} \sum \varphi^2 & \sum_{i=1}^n \varphi \\ \sum_{i=1}^n \varphi & n \end{array} \right|}.$$

After determination the coefficients c_x, b_x was found

$$a_x = \frac{1}{n} \sum_{i=1}^n \left[\frac{x_i}{\varphi_i^{b_x} \exp(c_x \varphi_i)} \right]$$

that is, determined all the coefficients of the empirical dependence (3).

Similarly, we find the coefficients $a_y, b_y, c_y, a_v, b_v, c_v$ of dependences (4) and (5), which allows us to determine the volume.

The analytical description of the numerical solution of the differential equation (1) by the empirical formulas (9-11) can be considered quite accurate [6-7]. Thus, on the basis of the empirical dependences obtained, prototypes of the drop contours are obtained.

To conduct experimental research of the density and surface properties of metallurgical melts, we have developed a facility made of low alloy steel. To obtain an image of a droplet, the installation is equipped with an optical system capable of moving in horizontal and vertical planes. The optical image of the object under research is formed by a Stemi SV 11 horizontal optical microscope and transmitted using an SONY SPT-M308CE monochrome video camera as an analog signal to a DATA TRANSLATION image grabber (frame grabber) (DT-3155). The digitized image with a spatial resolution of 768x576 pixels and 256 levels of gray is placed in the memory of an Intel computer for further processing.

The detection of droplet boundaries was carried out in three stages, according to the method described in works [20-23]. At the first stage, the operator

$$\begin{aligned} G_x &\equiv \frac{\partial f(x, y)}{\partial x} = (f(x+1, y+1) + 2f(x+1, y) + f(x+1, y-1)) - \\ &\quad - (f(x-1, y+1) + 2f(x-1, y) + f(x-1, y-1)), \\ G_y &\equiv \frac{\partial f(x, y)}{\partial y} = (f(x+1, y-1) + 2f(x, y+1) + f(x-1, y+1)) - \\ &\quad - (f(x+1, y-1) + 2f(x, y-1) + f(x-1, y-1)) \end{aligned}$$

is used in the entire field of the image under research to calculate the gradients in columns and in rows. Pixels with the largest intensity gradient value are selected for further processing. At the second stage, the magnitudes of the intensity gradients for the pixels selected at the first stage are compared in the horizontal, vertical and diagonal directions, and the maximum value is selected. In its direction, an approximation is made by a natural cubic spline under certain boundary conditions

The coordinate of the intersection point of the gray intensity level (half sum of the upper and lower plateaus) and the approximating curve is chosen. At the third stage, using the least squares method for each of five consecutive points along the profile of the lying drop, a second-order polynomial is selected. The pixels of the drop border, which are used for further mathematical processing, are calculated as the midpoint of such a polynomial curve.

5 Experiments and results

Comparison of geometric drops of pure metal melts and those obtained by calculation show their identity, which guarantees the accuracy of the measurements made (Fig. 4).

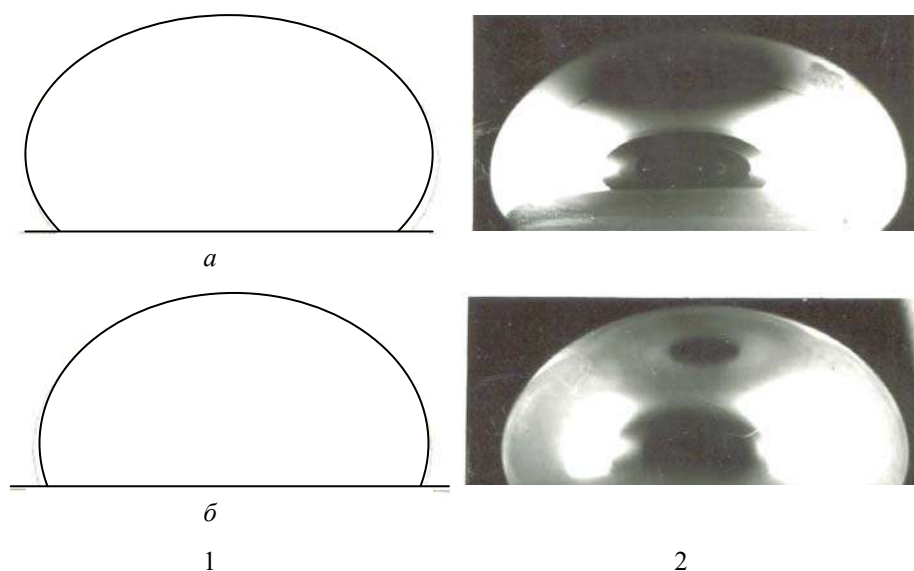


Figure4. Geometrical parameters of melt drops:

a – melt pure iron, at a temperature 1600°C ;

b – melt pure iron, at a temperature 1500°C .

The study of the wetting angle of the studied slag mixtures on steel 09G2S at 1600°C using the proposed method was carried out. The measurement results are shown in Table 1, where by positive experience is meant such that it is possible to clearly define the point of contact and build a tangent line. Accuracy increased by 2-2.5 times.

Table 1. Contact angle of SFM at the interface with steel 09G2S

Slagmixture	The proposed method		Known method		Amount of experiments			
	θ , grade	Measu rement error	θ , grade	Measu rement error	prototype		The proposed method	
					Total number	Positive experience	Total number	Positive experience
1	54°40′	1%	52°	2,2%	3	1	3	3
2	54°	0,5%	52°	2,6%	4	2	4	3
3	53°20′	0,8%	51°	2,4%	3	2	3	3
4	59°	1,1%	56°	3%	5	2	5	4
5	57°20′	0,9%	55°	2,1%	3	1	3	3
6	54°	0,75%	52°	1,9%	4	1	4	4
7	54°58′	1%	52°	2,7%	3	1	3	2
8	55°	0,5%	55°	2,9%	4	2	4	4
9	55°60′	1,01%	53°	3,1%	5	2	5	4
10	56°36′	0,95%	54°	2,4%	2	1	2	2

6 Conclusion

Techniques for solving the basic equation of a drop surface based on an improved mathematical apparatus with the implementation of Delphi visual programming in the system was improved, which significantly speeds up data processing and fully automates the calculation of the surface tension of melts with a high degree of accuracy (less than 0.5%) compared to known techniques.

A high-temperature experimental facility has been developed for the research of the surface properties of metallurgical melts, which, unlike the known facilities, is equipped with an optical system and a personal computer that allow the digital image of a lying drop, obtained in the experiment, to be processed using a high-speed software package, which significantly reduces the time for experiment and allows to store data and simplifies the work of the experimenter.

References

1. Streinrück, H., Rudische, C.: Numerical investigation of the entrainment of flux into the lubrication gap in continuous casting of steel: Fifth World Congress on Computational Mechanics July 7–12, 2002, Vienna, Austria, 256-271 (2002).
2. Thomas. B. G., Jenkins, M. S., Mahapatra, R. B.: Investigation of strand surface defects using mould instrumentation and modeling. Ironmaking & Steelmaking. Processes, Products and Applications, 18, 485-494(2013). doi.org/10.1179/030192304225019261

3. Maeda, M., Nakamura, N., Takeoka N.: Improvement on Mold Powder for Higher Speed Continuous Casting of Low Carbon Al-killed Steel. *Nissin Seiko Giho*, 57, 42-53 (1987).
4. Jonayat, A., Thomas, B.G.: Transient thermo-fluid model of meniscus behavior and Slag consumption in steel continuous casting. *The Minerals, Metals & Materials Society and ASM International* 2014, 45, 1842-1864 (2014). doi: 10.1007/s11663-014-0097-9
5. Irvin, W.R., Perkins, A., Brooks, M.G.: Effect of chemical, operational, and engineering factors on segregation in continuously cast slabs. *Iron and Steelmaking*. 11, 152-162 (1984).
6. Carlà, M., Cecchini, R.: An automated apparatus for interfacial tension measurements by the sessile drop technique. *Review of Scientific Instruments* 62, 1088-1097 (1991). doi.org/10.1063/1.1142011
7. Bashforth F., Adams J.C.: An attempt to test the theories of capillary action: by comparing the theoretical and measured forms of drops of fluid. *Cambridge Univ. Press Cambridge* (1883).
8. Smolders, C. A., Duyvis, E. M.: Contact angles; wetting and de-wetting of mercury: Part I. A critical examination of surface tension measurement by the sessile drop method. *Recueil des Travaux Chimiques des Pays-Bas*. 80(6), 588-680 (1961). doi.org/10.1002/recl.19610800607
9. Girault, H. H.J., Schiffrin, D. J., Smith, B.D.V.: The measurement of interfacial tension of pendant drops using a video image profile digitizer. *Journal of Colloid and Interface Sci.* 101(1), 257-266 (1984). doi.org/10.1016/0021-9797(84)90026-2
10. Carla, M., Cecchini, R., Bordini, S.: An automated apparatus for interfacial tension measurements by the sessile drop. *Review of Scientific Instruments*. 62(4), 1088 -1092 (1991). doi.org/10.1063/1.1142011
11. Goicochea, J., Garcia-Cordovilla C., Louis, E., Pamies, A.: Surface tension of binary and ternary aluminium alloys of the systems Al-Si-Mg and Al-Zn-Mg. *Journal of Materials Science*. 27(19), 5247-5252 (1992).
12. Liggieri, L., Passerone A.: An automatic technique for measuring the surface tension of liquid metals. *High Temperature Technology*. 7(2), 82-86 (1989). doi.org/10.1080/02619180.1989.11753417
13. Rotenberg, Y., Boruvka, L., Neumann, A. W.: Determination of surface tension and contact angle from the shapes of axisymmetric fluid interfaces. *Journal Colloid and Interface Sci.* 93(1), 169-183(1993). doi.org/10.1016/0021-9797(83)90396-X.
14. Faour, G., Grimaldi, M., Richou, J., Bois, A.: Real time pendant drop tension meter using image processing with interfacial area and interfacial tension control capabilities. *Journal Colloid and Interface Sci.* 181(2), 385-392(1996). doi.org/10.1006/jcis.1996.0395.
15. Hond-Kee, L., Froberg, G., Martin, P., Haira J.: Surface tension measurements of liquid iron-nickel-sulphur ternary system using the electromagnetic oscillating droplet technique. *ISIJ International*. 33(8), 833-838 (1993). doi: 10.2355/isijinternational.33.833
16. Andrien, C., Sykes, C., Brochard, F.: Average spreading parameter on heterogeneous surfaces. *Langmuir*. 10(7), 2077-2080 (1994). doi: 10.1021/la00019a010
17. Emelyanenko, A.M., Boinovich, L.B.: The role of discretization at the video image processing in sessile and pendant drop methods. *Colloids and surfaces A: Physicochemical and Engineering Aspects*. 189(1), 197-202(2001).doi: 10.1016/S0927-7757(01)00585-4
18. Berry, J., Neeson, M. J., Dagastine, R. R., Chan D. Y.C., Tabor R. F.: Measurement of surface and interfacial tension using pendant drop tensiometry. *Journal of Colloid and Interface Science*. 454(15), 226-237 (2015). doi: 10.1016/j.jcis.2015.05.012
19. Moser, Z., Gasior, W., Pstrum, J.: Surface tension measurements of the Bi-Sn and Sn-Bi-Ag liquid alloys. *Journal of Electronic Materials*. Mater. 30(9), 1109-1111(2001).

20. Hansen, F.K., Rodsrud, G.: Surface tension by pendant drop. *J. Colloid Interface Sci.* 141 (1), 1-9 (1991).
21. Thiessen, D.B, Chione, D.J., McCreary, C.B., Kiantz, W.B. Robust digital image analysis of pendant drop shapes. *Journal of Colloid and Interface Science.* 177(2), 658-665(1996). doi.org/10.1006/jcis.1996.0080
22. Duncan, D., Li, D., Gaydos, J., Neumann, A.W.: Correlation of line tension and solid-liquid interfacial tension from the measurement of drop size dependence of contact angles. 169(2), 212-217(1995). doi.org/10.1006/jcis.1995.1032
23. Rio, O.I., Neumann, A.W.: Contact angle measurements by axisymmetric drop shape analysis and an automated polynomial fit program. *Colloids and Surfaces A: Physicochemical and Engineering Aspects.* 143(2-3), 197-210 (1998). doi.org/10.1016/S0927-7757(98)00257-X

The Controlling of Transmission of Chaotic Signals in Communication Systems Based on Dynamic Models

Olexander Belej¹ [0000-0003-4150-7425], Iryna Artyshchuk² [0000-0001-7287-8451], Wojciech Sitek³ [0000-0002-6564-7177]

¹ Lviv Polytechnic National University, 5 Mytropolyt Andrei str., Building 4, Room 324, Lviv 79000, Ukraine

oleksandr.i.belei@lpnu.ua

² Lviv University of Trade and Economics, 10, Tuhon-Baranovskiy Str., Lviv 79008, Ukraine

irusia-13@ukr.net

³ Silesian University of Technology, Konarskiego 18A St., 44-100 Gliwice, Poland
wojciech.sitek@polsl.pl

Abstract. The article is devoted to the calculation of the characteristics of dynamic chaos based on the traffic of the corporate computer network. An algorithm for the transmission of chaotic information using dynamic chaos based on the model of chaotic masking and nonlinear mixing of the information signal is proposed. The problem of installing chaotic synchronization between two chaotic generators under conditions of phase signal filtration is considered. Based on the characteristics of the phase filter, the channel characteristics of the multichannel information signal filter are determined.

Keywords: dynamic chaos, algorithm, signal, traffic, computer network, transmission.

1 Introduction

At present, the development of telecommunication technologies determines the growth of research in the field of search for new solutions and innovative approaches to the mathematical description of the processes of transmission of information signals. One of the areas in the field of describing traffic in computer networks is the dynamic models that describe the real-time transmission of information in the form of differential or finite difference equations. A significant amount of work on modeling traffic in computer networks is based on the theory of mass service.

The transmission of messages with the help of a modulated chaotic signal has several advantages compared with the air conditioning modulation of the harmonic signal. Indeed, if there are only three controlled characteristics in the case of harmonic signals, then in the case of chaotic oscillations even a small change in the parameter gives a reliably detectable change in the nature of the oscillations. This means that the sources of chaos with changeable parameters have a wide range of information signal input circuits in the chaotic one. In addition, chaotic signals are broadband in princi-

ple. In communication systems, a wide frequency band of carrier signals is used both to increase the speed of information transmission and to increase the stability of the systems in the presence of disturbances. The noise similarity and the self-synchronization of chaos-based systems give them potential advantages over traditional spectrum expansion systems based on pseudo-random sequences.

Currently, it is known that chaotic signals generated by nonlinear deterministic dynamic systems, the so-called dynamic chaos, have a number of properties that facilitate the use of these signals for information transmission. A number of specific information transmission schemes have been proposed that use dynamic chaos, in particular, a scheme for the chaotic masking of an information signal; schemes with nonlinear mixing of the information signal into random messages. The possibilities of creating direct communication systems, in which chaotic oscillations act as a carrier of information generated directly in the frequency domain where information is transmitted, are discussed.

For modern networks characterized by a chaotic distribution of computing resources and a variety of end users, modeling for the creation of control systems for transmission of information signals is a particularly urgent task.

2 Formal problem statement

Chaotic synchronization of dynamic systems interconnected through a channel in which filtering of a chaotic signal transmitted from a leading dynamic system to a slave one is carried out is a basic model by which various methods of transmitting information based on chaotic synchronization are tested. Currently, there are a number of approaches aimed at solving the problem of establishing chaotic synchronization in the presence of signal filtering. In this case, as a rule, the filtering of a chaotic signal is carried out by a filter of the first or second order, which does not always adequately reflect the characteristics of a real information transfer channel. There are a number of works in which the influence of filtering on chaotic synchronization is considered in the context of a wireless information transmission channel, which is much closer to possible practical applications of chaotic synchronization.

Achieve synchronization, filters are applied that are inverse to the channel filter, which allows compensating for signal distortion at the input of the slave system, caused by the imposition of several replicas of the same signal on it.

The problem of developing methods and models of chaotic processes in communication systems based on the reflection of a nonlinear dynamic system for encryption and transmission of information has been studied. Conduct a study on this cryptographic algorithm for all required parameters.

The main purpose of the study is the use of chaotic algorithms for the transformation of information signals. The article deals with the characteristics of dynamic chaos for the transmission of messages and it is shown that the value of the highest Lyapunov index does not guarantee the randomness of the dynamics of information signals.

3 Literature review

Currently, it is known that chaotic signals generated by nonlinear deterministic dynamic systems, the so-called dynamic chaos, have a number of properties that facilitate the use of these signals for information transfer. The possibilities of creating direct communication systems, in which chaotic oscillations act as a storage medium generated directly in the frequency domain where information is transmitted, are discussed [1].

Fundamentals of the dynamic chaos oscillator operation are disclosed including optical ones, mathematical models are discussed, and principles of its application for information protection are described in [2]. The structure of a communication system is synthesized and its mathematical model that employs the invariant properties of the GCO is proposed. Computer simulation of the communication system is used to estimate the noise immunity of the system that works in the communication channel with distortions and noise in [3].

The experimental simulation shows that using the chaotic pilot signal for synchronization control can help achieve chaotic synchronization between communication transmitting and receiving systems and further enhance security and confidentiality of exchange of information and transfer of data [4]. In order to attain better synchronization, this TS fuzzy modeling is combined with the robust H_∞ observer theory based on linear matrix inequalities [5].

Once the pair is synchronized, the states can be used to secure the communication channel in one of four ways: chaotic modulation schemes, chaotic multicarrier schemes, chaotic multiple access schemes, and chaos-based encryption schemes [6]. The communication channel with a $2N$ -level logarithmic quantized is described, and the transmission delay of the communication channel is taken into account [7].

In [8] lay out a quantitative cryptanalysis approach for symmetric key encryption schemes that are based on the active/passive decomposition of chaotic dynamics. We end this chapter with a summary and suggestions for future research. The efficiency of data encryption at the transmitter and the recovery performance from an authorized receiver are also presented through diverse fiber transmission experiments. In these experiments, the security discrimination level between authorized and eavesdropping receivers is discussed [9].

The corresponding results demonstrate an ability to achieve initial synchronization. Furthermore, it is shown that in terms of code acquisition the PRBS outperforms the logistic and Bernoulli chaotic maps when used as pilot signals [10]. A smart IoT information transmission and security optimization model based on chaotic neural computing model are proposed. Simulation and analysis show that the proposed algorithm can ensure the availability and confidentiality of data at the same time [11].

Meanwhile, as improved teaching-learning-based optimization algorithm, teaching-learning-feedback-based optimization algorithm is proposed to optimize the parameters more excellently [12]. The subnets in the transmitter and receiver are one-to-one correspondence and form a pair of matching subnets, but the node size of each subnet can be inconsistent [13].

Authors in [14] show that our method can stabilize the chaotic oscillations in the user-centric cognitive radio networks. By comparisons with traditional parameter tuning methods, we confirm that our method is more efficient and faster to stabilize the cognitive radio systems. Meanwhile, the transmission error decays to zero exponentially. This implies that the synchronization error converges to zero under a limited communication channel [15].

The high quality of extraction of the hidden information signal is achieved due to the use of digital elements in the scheme, which ensures the identity of the parameters and high stability to noise [16].

Many articles are devoted to the transmission of messages with the help of a modulated chaotic signal. This modulation method has several advantages compared with the air conditioning modulation of the harmonic signal. Indeed, if there are only three controlled amplitudes in the case of harmonic signals, then in the case of chaotic oscillations even a small parameter change gives a reliably detectable change in the nature of oscillations [16]. Noise-like and self-synchronizing systems based on chaos give them potential advantages over traditional spectrum spreading systems based on pseudo-random sequences.

4 Analysis of the distribution of information transmission in chaotic traffic

The basis for the analysis was the data on the Internet channel load, obtained in the course of monitoring the work of the university's corporate network, measured during the year. Statistics are obtained when removing information from the router interfaces about the amount of information transmitted and the port loading, using the SNMP-protocol, using the Paessler Router Traffic Grapher package, which forms data tables and download schedules (Fig. 1).

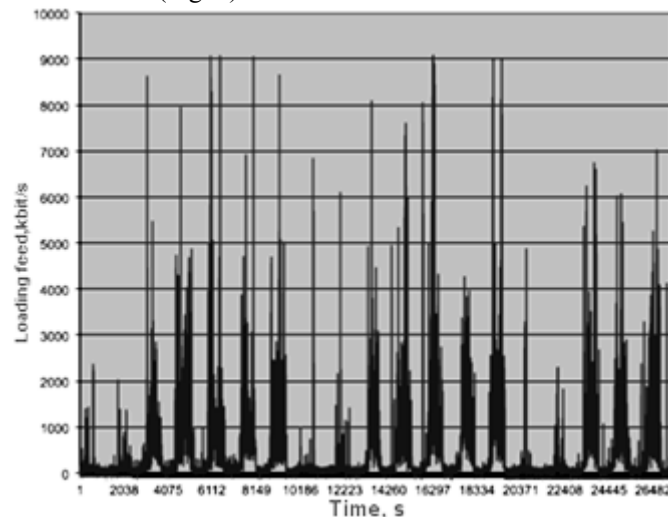


Fig. 1. Fragment of uploaded charts in the data link.

An empirical histogram of channel loading frequencies is shown in Fig. 2. The empirical histogram has a “ponderous tail”, indicating the presence of peak network load moments, in which there is a strong increase in delays and packet loss.

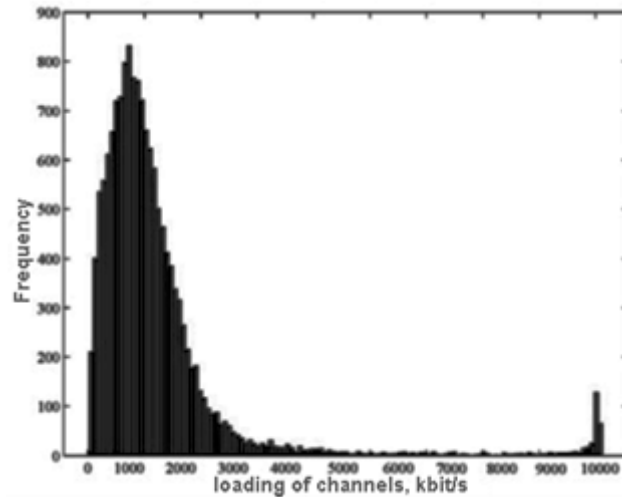


Fig. 2. The histogram of empirical in a network of output channel loading.

Information on the loading of communication channels was also obtained by monitoring the external communication channels of one of the provider companies and the site optimization company. The obtained histograms also have heavy tails (Fig. 3), indicating the presence of peak network load points, in which there is a strong increase in delays and loss of information.

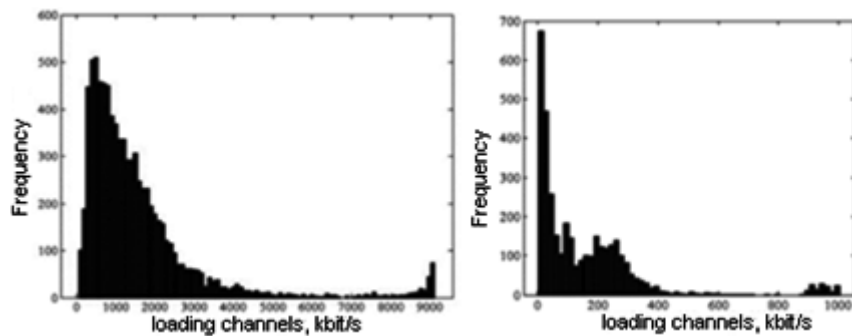


Fig. 3. The empirical distributions function of loading channels of the transmission information signal.

Due to the fact that the distribution function has a heavy tail and is not consistent with the Poisson distribution, queuing theory for the networks in question cannot provide an adequate mathematical description.

For TCP/IP, the distribution with heavy tails makes the main contribution to the self-similar nature of traffic and, therefore, the chaotic nature of the dynamics. A number of studies are devoted to the study of traffic randomness. The values of the Lyapunov senior indicator were estimated based on the traffic generated on the experimental bench, Internet traffic was for calculating various characteristics. The dynamic properties of chaos were used to solve telecommunication data exchange problems, but the study of chaotic properties remained outside the scope of publications. It is assumed that the time series is generated by a discrete:

$$x_{k+1} = f(x_k x_0), \quad (1)$$

or continuous system:

$$\frac{dx(t)}{dt} = F(x(t), x(0)). \quad (2)$$

Where $x=(x_1(t), \dots, x_n(t))$; n – the dimension of the phase space; t is time; k - discrete time (number); F, f - function vector. The phase trajectory of a continuous system is an n -dimensional curve, which is the solution of the system in the coordinates of the state space under given initial conditions x_0 . For discrete systems, states are connected by lines in accordance with the sequence of samples $k=1, 2, \dots$.

An important concept of dynamical systems is an attractor. For systems in the equilibrium position, the attractor is a point, for oscillatory systems, closed cycles. For chaotic systems, there is an attractor, which is called strange, in this case, the trajectories are tightened, but not to a point, a curve, a torus, but to some subset of the phase space. An attractor is an invariant characteristic of the system; it is preserved under the actions of transformations.

The unambiguous characteristics of the randomness of the signal are the spectrum of Lyapunov indicators. A positive maximum Lyapunov exponent is an indicator of chaotic dynamics, a zero maximum Lyapunov exponent indicates a limiting cycle or a quasiperiodic orbit, and a negative maximum Lyapunov exponent represents a fixed point. The system of dimension n has n Lyapunov exponents: $\lambda_1, \lambda_2, \dots, \lambda_n$, ordered in descending order. By definition, introduced by Lyapunov:

$$\lambda_i(x_0) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{|\delta_i(t)|}{|\delta_i(0)|} \quad (3)$$

where $|\delta_i(t)|$ - the fundamental solutions of the system, linearized in the neighborhood of x_0 .

Dynamical systems for which the n -dimensional phase volume decreases are called dissipative. If the phase volume is conserved, then such systems are called conservative. Conservative systems always have at least one conservation law. The presence of a conservation law often entails the existence of a Lyapunov zero exponent corresponding to it. For dissipative dynamic systems, the sum of Lyapunov indices is al-

ways negative. In dissipative systems, the Lyapunov exponents are invariant with respect to all initial conditions.

The positives of the highest Lyapunov exponent cannot be a necessary condition for the existence of chaos. Even in the Lorentz system, with the positive index being positive, as is known, under a number of conditions, a limit cycle takes place.

As an additional criterion, it is proposed to use the property of the absence of trivial conservation laws. Note that compression of the phase volume does not mean compression conversion.

To check for the presence of transformations of trajectory fragments, a genetic algorithm was developed and a program for MATLAB. At the same time, the following assumption has checked the system allows transformations, under conditions of weak symmetry breaking; there is some small value that slightly deviates from the symmetric mapping. Clearly, geometrically, this can be seen with almost similar loops on the attractor. Obviously, under such a test, under different initial conditions, for systems with regular dynamics, the presence of identical symmetry will be revealed, for a more complex, but not chaotic, translation, for systems seeking a stable equilibrium position, compression, for chaos - almost repetitive areas of phase trajectories.

The attractor reconstructed according to the traffic load is shown in Fig. 4.

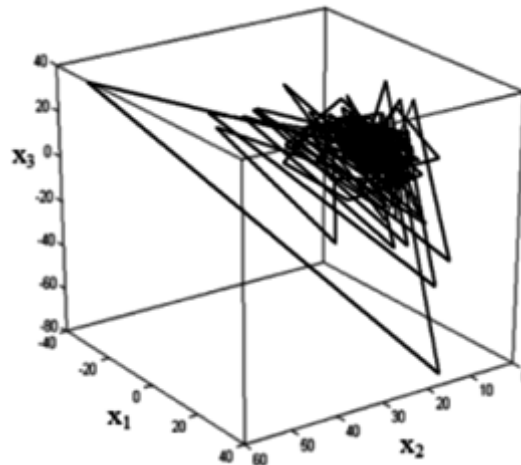


Fig. 4. Attractor, built on the basis of network traffic.

Confirmation of chaos can be the basis for building dynamic models: in the form of an ensemble of pendulums, affinity systems with control or in the form of series.

5 The construction of the modulation algorithm of the chaotic signal information

In most modern communication systems, harmonic oscillations are used as the information carrier. The information signal from the transmitter modulates these oscillations.

tions in amplitude, frequency, or phase; in the receiver, information is extracted by means of the reverse operation demodulation. The modulation of the carrier can be carried out either by modulating the already formed harmonic oscillations or by controlling the parameters of the generator in the process of generating oscillations.

Similarly, it is possible to produce a modulation of a chaotic signal with an information signal. However, the possibilities here are much wider. Indeed, if there are only three in the case of harmonic signals of controlled characteristics, then in the case of chaotic oscillations even a small parameter change gives a reliably detectable change in the nature of oscillations. This means that the sources of chaos with changeable parameters, there is a wide range of information signal input schemes in the chaotic one. In addition, chaotic signals are broadband in principle, the interest in which is associated with higher information capacity. In communication systems, a wide frequency band of carrier signals is used both to increase the speed of information transmission and to increase the stability of the systems in the presence of disturbances.

In Fig. 5 shows the simplest communication scheme using chaos. Re-sensor and receiver include all the same non-linear or linear systems as the source. In addition, the transmitter is switched on the adder and the receiver and subtracted. In the accumulator, the chaotic signal of the source of the information signal is added; the receiver is designed to extract the information signal. The signal in the channel is chaos-like and does not contain visible signs of the transmitted information, which allows transmitting confidential information. The signals at points A and A', B and B' are pairwise equal. Therefore, if there is an input information signal S at the input of the transmitter adder, the same signal will be allocated at the output of the receiver's subtracted.

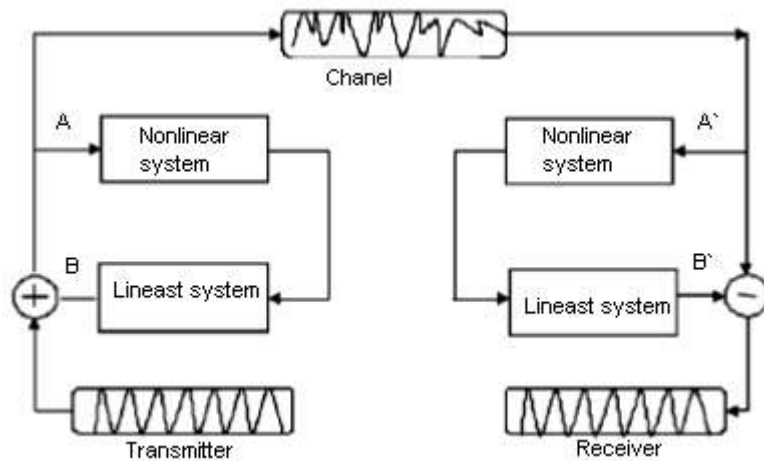


Fig. 5. The simplest communication scheme using the chaos of transformation information signal.

filter 3 and channel III filter; c) bifurcation diagram for the parameter M of the leading system I

The leading system consists of a nonlinear subsystem 4 and linear subsystems: low-pass filters of the first and second orders and filter 3, which is equivalent to the filter of channel III. The slave system consists of the same elements. If the subsystem and channel III filter are excluded from the model, then the proposed scheme is reduced to the traditional scheme for obtaining a chaotic synchronous response.

In the presence of a chaotic synchronous response, the signal z_2 at the output of the slave system is identical to the signal z_1 of the leading system. Indeed, the signal at the output of the slave system z_2 is a copy of the signal y_1 , passed through filter 3.

Let the transient response of the phase filter 3 be described by the function:

$$h(t) = \sum_{k=0}^N a_k \delta(t - \tau_k), \quad (4)$$

where $h(t)$ is the impulse response of the phase filter to the δ pulse, a_k is the amplitude of the k -th replica, τ_k is the delay in the transmission through the filter of the k th replica relative to the first replica of the input signal, for which, by definition, $\tau_0=0$. Physically, this is equivalent to forming at the output of the filter a sum of replicas of a signal having different delays. The implementation of the signal $s(t)$, passed the filter with the characteristic (1), is described by the expression:

$$\Phi(s(t)) = \int_{-\infty}^t h(t) s(t - \tau) d\tau, \quad (5)$$

in the form of a response function (5) with the signal $s(t)$ at the input of the filter.

As a result, the model shown in Fig. 6, described by the following system of equations:

$$\begin{aligned} T\dot{x}_1 + x_1 &= F(z_1(t)), \\ \ddot{y}_1 + \alpha_1 \dot{y}_1 + y_1 &= x_1, \end{aligned} \quad (6)$$

$$z_1(t) = \int_{-\infty}^t h(\tau) y_1(t - \tau) d\tau,$$

$$\begin{aligned} U(t) &= \int_{-\infty}^t h(\tau) y_1(t - \tau) d\tau, \\ T\ddot{x}_2 + x_2 &= F(U(t)), \\ \ddot{y}_2 + \alpha_1 \dot{y}_2 + y_2 &= x_2, \end{aligned} \quad (7)$$

$$z_2(t) = \int_{-\infty}^t h(\tau) y_2(t - \tau) d\tau. \quad (8)$$

The first two equations in (8) describe the filters of the first and second orders, respectively, the third equation (8) - the channel filter. Thus, the leading system in Fig. 6, a is an oscillatory system with a first-order low-pass filter, a second-order low-pass filter, and an N-th order phase filter that implements N different delays with respect to the signal at its input. The slave system is passive, so there is a reason to believe that in the system as a whole, it will be possible to obtain a stable chaotic synchronous response in the presence of phase filtering. To prove this possibility, a numerical simulation of the synchronization process was carried out. Numerical simulation was carried out in the framework of the discrete representation of signals and equations.

The discrete representation eliminates the need to transform the discrete response-function of the phase filter into a dynamic system with continuous time, which upon such transformation will have a high dimension with the number of phase variables equal to the number of discrete samples, the amplitudes of which are approximated by numerical modeling.

When switching to the discrete representation of the response function (6), the delays τ_k in the arrival of the replicas of the signal, taking a continuous spectrum of values in the real channel, in the discrete representation will be defined as

$\tau_k = n_i \Delta t$, where n_i is the number of the reference on which i - replica with respect to the first replica (for the first replica $n_0=0$), $\Delta t = \frac{1}{f_s}$ is the sampling time step, f_s is the sampling frequency. Thus, the channel response function (8) takes the form:

$$h(t_k) = \sum_{k=0}^N a_k \delta(t_k - n_k \Delta t). \quad (9)$$

Transformation (9) in a discrete form with a response function (11) takes the form:

$$\Phi(s(t_k)) = \alpha_0 s(t_k) + \alpha_1 s(t_k - n_1 \Delta t) + \alpha_2 s(t_k - n_2 \Delta t) + \dots + \alpha_N s(t_k - n_N \Delta t) \quad (10)$$

In fact, the conversion is a non-recursive N-th order digital filter with the impulse response (10).

Discrete analogs of first and second order filters included in equation (8) can also be represented as recurrence relations by replacing the derivative operator $p = \frac{d}{dt}$

with a delay operator by one sample $q^{-1} = s(t_k - \Delta t)$ using bilinear transform:

$$p = \gamma \frac{1 - q^{-1}}{1 + q^{-1}} \quad (11)$$

where $\gamma = \frac{2}{\Delta t}$.

After substitution (11) in (8), the first equation takes the form:

$$x_1(t_k) = b_0^{(1)} F(z_1(t_k)) + b_1^{(1)} F(z_1(t_k - \Delta t)) + a_1^{(1)} x_2(t_k - \Delta t) \quad (12)$$

The second equation (9) - the filter of the second order - takes the form:

$$x_1(t_k) = b_0^{(2)} x_1(t_k) + b_1^{(2)} x_1(t_k - \Delta t) + b_2^{(2)} x_1(t_k - 2\Delta t) - \dots - a_1^{(2)} y_1(t_k - \Delta t) - a_2^{(2)} y_1(t_k - 2\Delta t) \quad (13)$$

The coefficients in (14) and (15) are defined as:

$$b_0^{(1)} = b_1^{(1)} = \frac{1}{1 + \gamma T}, a^{(1)} = (1 - \gamma T) b_0^{(1)}, b_0^{(2)} = b_2^{(2)} = \frac{1}{\gamma^2 + \alpha_1 \gamma + 1}, \quad (14)$$

$$b_1^{(2)} = 2b_0^{(2)}, \alpha_1^{(2)} = 2(1 - \gamma^2) b_0^{(2)}, \alpha_2^{(2)} = (1 - \alpha_1 \gamma + \gamma^2) b_0^{(2)}.$$

The transformation to the discrete representation in equations (10) is carried out similarly.

The first three relations describe the leading system, the fourth the phase filter of the Nth order channel, the last three the driven system. The first equation describes a sequentially connected inertialess nonlinear and a first order low pass filter.

The inertialess nonlinear transformation is given by an expression of the same kind:

$$F(z) = M \left[|z + E_1| - |z - E_1| + \frac{1}{2} (|z - E_2| - |z + E_2|) \right] \quad (15)$$

where E_1, E_2 — constants, M — the gain. The specific choice of E_1, E_2 , and M uniquely determine the characteristic.

The result of the numerical solution of system is the implementation of the phase variables $x_1(t_k), y_1(t_k), z_1(t_k), U(t_k), x_2(t_k), y_2(t_k), z_2(t_k)$, calculated for the moments of time $t_k = \frac{k}{t}$, where k is the reference number. This implementation is equivalent

to the implementation of the phase trajectory obtained by numerically solving the equations (9).

During the simulation, the parameter values were $M=7.3, T=3, \alpha_1=0.1, E_1=0.5, E_2=1, \Delta t=0.05$. The channel phase filter was simulated using the impulse response (4), including $N = 10$ replicas (Fig. 6,b). The delays in the arrival of the replicas of the

signal are in the range from one to ~20 quasi-periods of chaotic oscillations. Model

(8) is dimensionless, the quasi-period T_0 is equal to the reciprocal of the resonant frequency of the second-order filter (8), that is $T_0=1/f_0$, $f_0=1-(\alpha/2)/2\pi$. Since $\alpha \ll 1$, the value of $f_0 \approx 1/(2\pi)$.

The bifurcation diagram of the system with respect to the parameter M is shown in Fig. 6, c. The leading system retains the property to generate chaotic oscillations, the ring generator remains a source of chaos with the introduction of an additional filter 3 (Fig. 6, a).

In fig. 7 shows the results of numerical simulation of chaotic synchronization in the system (19). In fig. 7 shows a fragment of the implementation of the signal before, and fig. 7, b - after the phase filter.

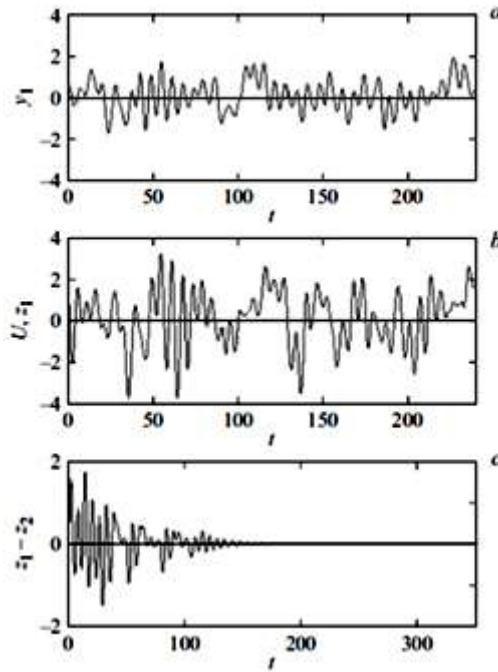


Fig. 7. The algorithm of transmitting information signals at different points of the chaotically synchronized dynamic system.

In Fig. 7 is presented the signal y_1 at the input of the channel filter, b is the signals U and z_2 at the output of the channel III filter and the output of the slave system II, respectively, c is the difference signal $z_1(t) - z_2(t)$ between the signal $z_1(t)$ of the master system and the output signal of the slave system $z_2(t)$.

Since the system is passive and the signal z_2 at its output undergoes the same transformations as the signal z_1 in the master system, the output signal z_2 of the slave system asymptotically tends to the signal z_1 of the master system (Figure 7, c); thus, a chaotic synchronous response in the system can be achieved.

The signal amplitude after the phase filter (Fig. 7, b) increases since the signal

passing through the phase filter is equivalent to forming at its output an incoherent sum $N=10$ replicas of a chaotic signal with delays determined by the response function.

Thus, it is shown in the paper that the elimination of the influence of the channel filter on the synchronization of the master and slave dynamic systems is achieved by introducing into their structure a phase filter equivalent to the channel filter. Due to this, the leading dynamic system forms a signal whose structure is not disturbed after passing through the channel filter.

The advantage of the considered method is the possibility of establishing chaotic synchronization for an arbitrary number of signal replicas entering the slave system, there are no restrictions on the order of the channel filter.

7 Conclusion

The advantage of the considered method is the possibility of establishing chaotic synchronization for an arbitrary number of signal replicas entering the slave system, there are no restrictions on the order of the channel filter.

The paper deals with chaotic phenomena in computer data networks. Based on the chaotic properties, mathematical models of the dynamic behavior of traffic can be constructed. Models can be used to provide guaranteed quality of service, analyze bottlenecks in the structure of a corporate network, and exchange data in cloud infrastructures.

At the same time, the indicators of chaos themselves, the structure of the attractor may have practical value. The change in the values of the Lyapunov senior indicator, the change in the topology of the attractor, is an indicator of the change in network activity. Computer attacks, failure of corporate data exchange systems or being the basis for a change in the administration policy - expansion of communication channels or replenishment of the list of prohibited network resources. The latter was observed with the growth of the popularity of social networks and video sharing resources.

References

1. Abualnaja, K.M., Mahmoud, E.E.: Analytical and Numerical Study of the Projective Synchronization of the Chaotic Complex Nonlinear Systems with Uncertain Parameters and Its Applications in Secure Communication. *Mathematical Problems in Engineering* (2014).
2. Izmailov, I., Poizner, B., Romanov, I., Smolskiy, S.: Deterministic Chaos Phenomenon from the Standpoint of Information Protection Tasks. In: *Cryptology Transmitted Message Protection. Signals and Communication Technology*. Springer, Cham (2016).
3. Leonov, K.N., Potapov, A.A., Ushakov, P.A.: Application of Invariant Properties of Chaotic Signals in the Synthesis of NoiseImmune Broadband Systems for Data Transmission. *Radiotekhnika i Elektronika*, Vol. 59, No. 12, pp. 1209–1229 (2014).
4. Lai, H., Huang, Y.: Chaotic Secure Communication Based on Synchronization Control of Chaotic Pilot Signal. *ISICA Communications in Computer and Information Science*, vol 575, Springer (2016).

5. Mozelli, L., Campos, C., Palhares, R.: Chaotic Synchronization and Information Transmission Experiments: A Fuzzy Relaxed H_∞ Control Approach. *Circuits Syst Signal Process*, Volume 26, Issue 4, pp 427–449 (2007).
6. Bendoukha, S., Abdelmalek, S., Ouannas, A.: Secure Communication Systems Based on the Synchronization of Chaotic Systems. *Studies in Systems, Decision, and Control*, vol. 200, Springer (2019).
7. Zheng, J., Cui, B.: State Estimation of Chaotic Lurie Systems via Communication Channel with Transmission Delay. *Circuits Syst., Signal Process*, 37, pp. 4568–4583 (2018).
8. Tenny, R., Tsimring, L.S., Abarbanel, H.D.I., Larson, L.E.: *Security of Chaos-Based Communication and Encryption*. Institute for Nonlinear Science. Springer, New York, NY (2006).
9. Argyris, A.: Hardware Implementation of Chaos-Secured Optical Communication Systems. *Studies in Computational Intelligence*, vol. 354. Springer, Berlin, Heidelberg (2011).
10. Jovic, B.: A Robust Sequence Synchronization Unit for Multi-user Chaos Based DS-CDMA Communication Systems. *Signals and Communication Technology*. Springer, Berlin, Heidelberg (2011).
11. Deng, L., Li, D., Cai, Z.: Smart IoT information transmission and security optimization model based on chaotic neural computing. *Neural Comput & Applic*, Springer-Verlag London Ltd., part of Springer Nature (2019).
12. Li, RG., Wu, HN.: Secure communication on fractional-order chaotic systems via adaptive sliding mode control with teaching–learning–feedback-based optimization. *Springer Nature B.V.*, Volume 95, Issue 2, pp. 1221–1243 (2018).
13. Zhou, L. & Tan, F.: A chaotic secure communication scheme based on synchronization of double-layered and multiple complex networks. *Nonlinear Dyn*, Springer Netherlands (2019).
14. Nonaka, Y., Hasegawa, M.: Controlling Chaotic Oscillations in User-Centric Cognitive Radio Networks. *Lecture Notes in Computer Science*, vol. 6883, Springer, Berlin, Heidelberg (2011).
15. Liu, W., Wang, Z., Zhang, W.: Controlled synchronization of discrete-time chaotic systems under communication constraints. *Nonlinear Dyn*, vol. 69, Springer Science Business Media B.V. (2012).
16. Ponomarenko, V.I., Prokhorov, M.D., Karavaev, A.S.: An experimental digital communication scheme based on chaotic time-delay system. *Nonlinear Dyn*, Springer Netherlands, Volume 74, Issue 4, pp. 1013–1020 (2013).

Software complex for automated diagnostics of internal parameters of technical systems

Kostiantyn Kasian¹[0000-0002-1257-156X], Mykola Kasian²[0000-0001-8999-2218],

¹Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine
konst_k@yahoo.com

²Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine
kasmiko48@gmail.com

Abstract. We offer an information technology for diagnosing technical systems by monitoring the values of their internal parameters and comparing with the maximum allowable values. Diagnosed parameters of the internal elements of the technical system are determined by the mathematical calculations based on the well-known mathematical model of the system and the measured values of its output characteristics. Optimization methods are used for calculating the values of the internal parameters of the system.

Keywords: diagnostics, technical system, output characteristics, internal parameter, permissible value, optimization

1 Introduction

The reliability of any technical device is determined by the quality of its development, is ensured in the manufacturing process and maintained during operation. The impossibility of creating absolutely reliable products makes relevant the research, development and applying of principles, ways, methods and means, which increase reliability by timely detection and elimination of equipment failures. The main ways to prevent the failure of technical systems include the effective monitoring and diagnosis of their technical condition. Timely detection and elimination of defects increases the probability of failure-free operation, and also reduces the cost of operating controlled objects. The time spent on device diagnostics can be significantly reduced by automation of the diagnosis and using rational diagnostic procedures. There are different variants of the strategy of finding defects for both automated and non-automated diagnostics [1-5].

Failures occurring in the system due to strong changes in the values of its internal parameters are easier to detect because they lead to significant changes in the output characteristics. At the same time, the drift of the parameters of the system elements, for example, in the course of its operation under the influence of ageing and external factors, leads to small variations in the output characteristics with a slow deterioration of the properties. In addition, it is not always the overrange of the parameter of the internal element of the system beyond the permissible limits that leads to an unac-

ceptable deviation of the controlled output characteristic. However, this may change for the worse the modes of operation of the system elements, which after a certain period of time will lead to an even greater change in the parameters and, as a result, to a defect. Such deviations of the internal parameters of the system are called latent defects, and they, as a rule, require the development of special diagnostic methods, since they are not detected by traditional methods. Also, during the operation of the system, there is a need not only to respond on the fly to its failures, but also to monitor the status of parameters and modes of operation of its elements and analyze trends in their work. This will allow not only to anticipate failures, but also to issue recommendations for their prevention by assessing the approximation of monitored parameters to their maximum allowable values, which can also vary depending on the time of operation, the effect of temperature and other external factors.

These tasks can be effectively solved by automated determination of the values of parameters and modes of operation of system elements based on stimulating input actions and experimentally obtained output characteristics.

2 Formal problem statement

The mathematical model should allow to determine the values of the parameters of the internal elements of the system being diagnosed from its known (measured) output characteristics.

In practice, there is only access to a limited number of circuit nodes, to which stimulated signals can be sent and output characteristics can be recorded. In addition, the number of internal circuit parameters is much greater than the number of obtainable nodes. Therefore, to determine the values of the parameters of internal elements from the measured output characteristics it is proposed to apply the optimization method. In the optimization process, the values of internal parameters are optimized so as to maximally bring the calculated values of the output characteristics of the system under diagnosis to the measured values.

As a criterion for the correspondence of the calculated values of the output characteristics to the measured values, we use the criterion of the minimum of the root-mean-square error (1):

$$f(q) = \sum_{j=1}^M \sum_{i=1}^N v_{ji} \left[1 - \frac{Y_{ji \text{ calc}}(q)}{Y_{ji \text{ meas}}(q)} \right]^2 \Rightarrow \min, \quad (1)$$

where $Y_{ji \text{ calc}}(q)$, $Y_{ji \text{ meas}}(q)$ are respectively the calculated and measured value of the j -th output characteristic at the i -th point, which depends on the vector of parameters of the elements of the scheme q ; v_{ji} are weighting indices of measurement accuracy of the j -th characteristic at the i -th point, which are calculated

$$v_{ji} = \frac{1 - \gamma_{j \text{ calc}}}{\sum_{j=1}^M \sum_{i=1}^N (1 - \gamma_{j \text{ meas}})}, \quad (2)$$

where γ_{jmeas} is the relative measurement error of the i -th sample on the j -th output characteristic.

3 Literature review

To minimize the mean-square error function, an optimization method is used.

The methods used to solve optimization issues are rather numerous [6–8]. Among them, there is no universal method that would be the best in all or most cases. Choosing a method conforming to the special features of a particular task increases the probability of its successful solution with minimal expenses. At the heart of the diagnostic model is a convex functional that is a composition of convex functions. For such functions, gradient methods of the second order, for example, the Newton method, are the best in terms of the rate of convergence and stability. However, this method requires the calculation of the Hessian matrix of the second partial derivatives and its inversion [6].

It is most expedient to use the Davidon-Fletcher-Powell method [6], which does not require the calculation of the inverse Hessian $G^{-1}(q_i)$ at every step, because the search direction at step i is the direction $-H_i g(q_i)$, where H_i is positively defined symmetric matrix, which is updated at every step. At the boundary, the matrix H becomes equal to the inverse Hessian. This method combines both the ideas of Newton's method and the property of conjugate gradients, and when applied to the minimization of convex quadratic functions l variables converge in no more than n iterations. It, like Newton's method, is based on the correlation

$$q_{i+1} = q_i - \lambda_i G^{-1}(q_i) g(q_i), \quad (3)$$

where q is the vector of the diagnosed parameters; i is the iteration number of optimization; λ is an optimization step parameter; g is the vector of the sensitivity functions of the target function to a change in the parameter values (gradient); G is the matrix of sensitivity functions of the second order, the so-called Hessian matrix.

The search for the minimum of the objective function (1) begins at the starting point q_0 (usually, the nominal values of the diagnosed parameters), and the initial matrix H_0 is taken as the identity one. The iterative procedure can be represented as follows:

1. At step i , there is a point q_i and the symmetric matrix H_i is positive defined.
2. As direction of the search, to take a direction $d_i = -H_i g_i$.
3. To find the function λ_i , which minimizes the function $f(q_i + \lambda_i d_i)$, to do a one-dimensional search along the straight line $q_i + \lambda_i d_i$.
4. To calculate the increase in the parameters $v_i = \lambda_i d_i$.
5. To calculate the new values of the parameters $q_{i+1} = q_i + v_i$.
6. To calculate the new values of the objective function $f(q_{i+1})$ and its gradient g_{i+1} for the values of the parameters q_{i+1} . If the values of $|g_{i+1}|$ or $|v_i|$ small enough, complete the optimization, otherwise continue.
7. To calculate the increase in the gradient $u_i = g_{i+1} - g_i$.
8. To calculate the matrix A_i

$$A_i = \frac{v_i \cdot v_i^T}{v_i^T \cdot u_i},$$

where v_i is the vector of increment of parameter values, v_i^T is the transposed vector of increment of parameter values, u_i is the vector of increment of the gradient of the objective function.

9. To calculate the matrix $B_i = -H_i u_i u_i^T H_i / (u_i^T H_i u_i)$,

where H_i is the matrix H at the i -th iteration, u_i is the gradient increment vector, u_i^T is the transposed gradient increment vector.

10. To update the matrix H in the following way: $H_{i+1} = H_i + A_i + B_i$,

where H_{i+1} is the matrix H at the $i+1$ -st iteration, H_i is the matrix H at the i -th iteration, A_i is the matrix A at the i -th iteration, B_i is the matrix B at the i -th iteration.

11. To increase i by one and to go back to step 2.

4 Requirements for the software complex

Functional capabilities and organization of software determine the basic requirements for the software system under development for automated diagnostics of the parameters of elements of technical systems.

The requirements to the functional capabilities of the complex can be formulated as follows.

The complex should allow:

- to analyze the characteristics of the technical system at the design stage;
- to obtain, according to the results of the analysis, the necessary test input effects and measurement points of the controlled output characteristics;
- by measured output characteristics to determine the actual values of the parameters of the elements of the system;
- automatically form the boundaries of the rejection tolerances for the parameters of the elements of the system, taking into account its lifetime, temperature and other external factors;
- by comparing the actual values of the parameters of the elements with the maximum permissible values, give information about the causes of system malfunction to the element level.

The practical implementation of the functions of constructing a mathematical model of the system being diagnosed and calculating its output characteristics can be performed using the methods and algorithms described in [9-11].

To implement the functions of calculating the values of the internal parameters of the system from the known values of its output characteristics, the additional use of the diagnostic model of the system [12] is necessary, on the basis of the objective function (1).

From the point of view of software organization, the complex should meet the following requirements:

- have a block-modular structure that allows you to effectively complement the complex with other functional blocks and software patterns that extend its functional capabilities;
- to be able to function as part of an integrated CAD-system used in the design of the system being diagnosed;
- be open to the emergence of new numerical analysis methods used in the operation of the complex in order to increase the efficiency of its application in the process of diagnosing technical systems;
- to incorporate a database of maximum permissible values of parameters of system elements.

5 The structure and algorithm of the software complex functioning

The structure of the developed software complex is shown in Figure 1 [13], where the dotted line outlines the software modules that are part of the CAD for designing the diagnosed system and are used by the developed software complex. The arrows in the diagram show the direction of data transfer between the modules.

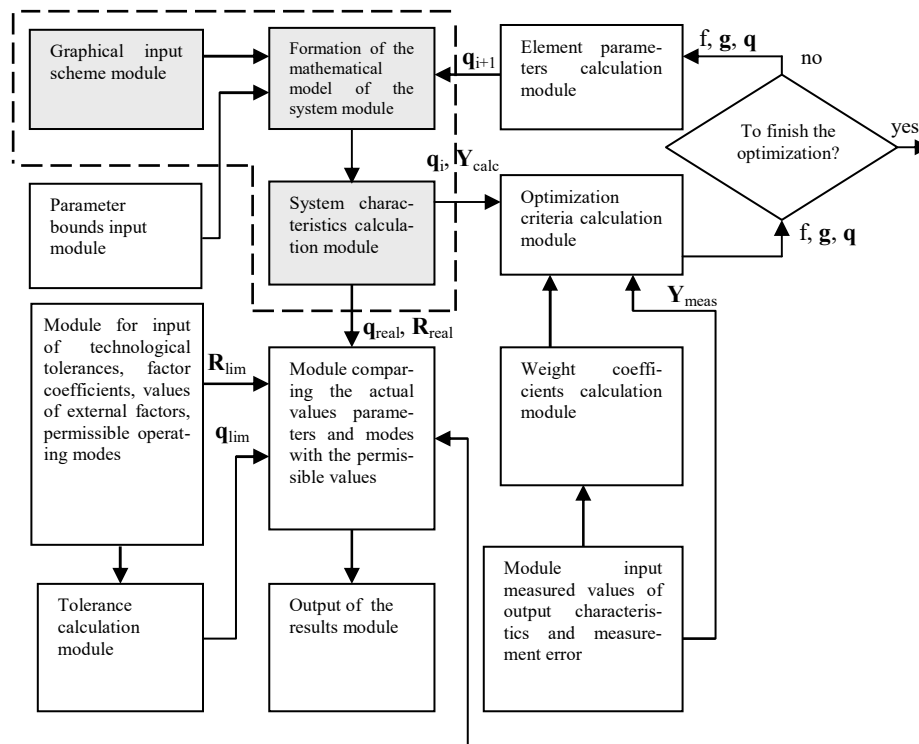


Fig. 1. The structure of the software complex for automated diagnosis of parameters of technical systems elements

Another important task of developing an automated diagnostics software complex is the development of an algorithm for its functioning. The algorithm of functioning of the complex is implemented by the managing program, which coordinates the interaction of program modules and ensures the implementation of certain procedures.

The algorithm of functioning of the complex, the block diagram of which is shown in Figure 2, is based on the method presented in Section 3, the requirements for the program complex and the structure of the complex presented above.

Block 1. The start of the algorithm of the software complex functioning.

Block 2. Introduction of the necessary source data: a basic diagram, technological tolerances on the parameters of elements, factor coefficients, maximum allowable modes of operation of elements, values of external factors, the parameter value of the calculation process cessation, the choice of diagnosable parameters and restrictions on them.

Block 3. Set of the input effects, measurement points and measurable values of the output characteristics.

Block 4. Formation of a mathematical model of the system being diagnosed with the data inserted.

Block 5. Calculation of output characteristics and functions of parametric sensitivity of output characteristics to changes in the values of the parameters of the elements.

Block 6. On the basis of the parametric sensitivity functions of the output characteristics, the formation of a test matrix with respect to the diagnosed parameters and the calculation of the rank of the matrix ρ .

Block 7. Determination of the degree of the possibility of a solution with respect to the diagnosable parameters $\mu = n_q - \rho$, where n_q is the number of diagnosable parameters. If the condition of diagnosability $\mu \leq 0$ is fulfilled, then the weighting factors are calculated (block 8). If the condition $\mu \leq 0$ is not fulfilled, then additional input actions and control points of measurement are needed, that is, a return to the fulfillment of block 3.

Block 8. The calculation of the weighting coefficients of the objective function on the basis of the measurement accuracy of the corresponding output characteristics.

Block 9. The calculation of the objective function to determine diagnosable parameters.

Block 10. Calculation of the gradient of the objective function.

Block 11. Checking the completion of the process of optimizing the values of diagnosable parameters.

Block 12. The inclusion of additional controlled characteristics.

Block 13. Calculation of parameters and modes of operation of the elements of the system being diagnosed.

Block 14. Calculation of the maximum permissible values of the parameters of the elements, taking into account the temperature and time of operation.

Block 15. Comparison of the calculated values of parameters and modes of operation of the elements of the system being diagnosed with the maximum allowable values.

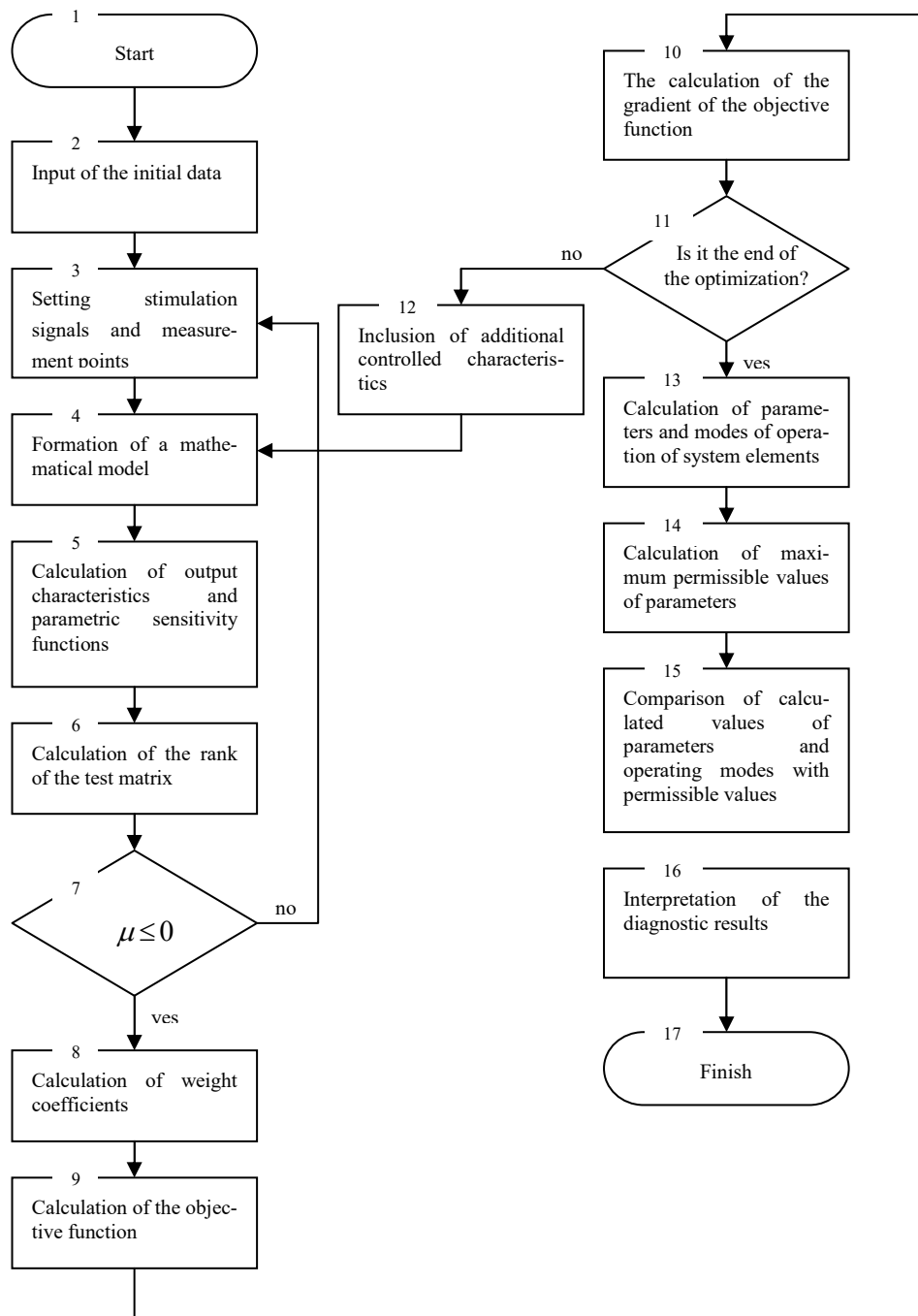


Fig. 2. Algorithm of the software complex for automated diagnosis of the parameters of the technical systems

Block 16. Interpretation of the results of calculations and diagnostics, that is, the derivation of the calculated values of the parameters and modes of operation of the elements and the results of their comparison with the maximum allowable values of the classification of the technical state of the system being diagnosed.

Block 17. The end of the algorithm of the software complex functioning.

If according to the results of diagnosing any one or several internal parameters of the system go beyond their maximum permissible limits, the presence of a defect is ascertained. If any internal parameter of the system has not reached its maximum permissible value, but is close to it, we can speak of a possible defect in the near future. In this case, it is necessary to predict the behavior of the parameter [14] in order to take action in advance without allowing the parameter to go beyond the tolerances.

6 Experiments and results

To test and confirm the practical application of the developed diagnostic method, we conduct the experimental studies of simple technical systems representing basic analog electrical circuits.

The passive RC low-pass filter was taken as the simplest test case [12]. Its basic electrical circuit diagram is shown in Figure 3.

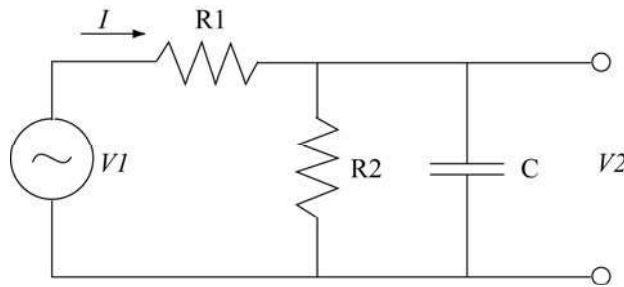


Fig. 3. The basic electrical circuit of the filter

The electrical parameters of the device are the resistance of resistors $R1$, $R2$ and the capacitance of capacitor C . Their nominal values:

$C1 = 100 \text{ pF}$, $R1 = 10 \text{ k}\Omega$, $R2 = 10 \text{ k}\Omega$.

To ensure the diagnosability of the filter, it is necessary to select input influences and output characteristics sufficient for unambiguous determination of the parameters of the filter elements. Such a linear frequency-dependent device is advisable to diagnose in the frequency domain. If the stimulating signal is the input harmonic voltage $V1$, and the output controlled characteristic is $V2$, then in accordance with section 5 the test diagnostic matrix will have a rank of 1. Thus, when monitoring the voltage $V2$, it is impossible to ensure the diagnosability of three circuit parameters. In this case, the scheme can be diagnosed only with a single fault. If the input action is voltage $V1$, and the output controlled characteristic is alternating current I , then the test matrix will also have a rank equal to 1. If the output controlled characteristics are

current I and voltage V_2 , then the rank of the resulting matrix is three, and the scheme will be diagnosed over all parameters. Therefore, the amplitude-frequency dependences of the output voltage V_2 and the input current I are used as output characteristics. The test input is a harmonic voltage with amplitude of 1V and a frequency of 0, 100 kHz, 200 kHz, 300 kHz, 400 kHz, 500 kHz.

First, using CAD MAES-P [12] according to the specified scheme, parameter values and test input effects, the selected output characteristics were calculated. Then, using the diagnostic program, these parameters were used to calculate the values of the parameters of the elements, which were compared with their previously specified values. Thus, verification of the accuracy of identification of parameter values was carried out on the calculated output characteristics, which were taken as measured and corresponded to the values of the parameters of the elements, both within tolerances and beyond their limits [13]. The results of the calculation of the parameters of the filter elements are shown in Table 1.

Table 1. Values of the diagnosed parameters (Modeling)

Parameter	Nominal Values		
	Set	Calculated	Error, %
R1, k Ω	10	9.990	0,01
R2, k Ω	10	9.999	0,01
C1, pF	100	99.998	0,002
	In the tolerance range		
	Set	Calculated	Error, %
R1, k Ω	9	8.999	0,011
R2, k Ω	11	10.999	0,009
C1, pF	90	89.999	0,001
	One parameters are out of tolerance		
	Set	Calculated	Error, %
R1, k Ω	15	14.999	0,007
R2, k Ω	10	9.999	0,01
C1, pF	100	99.992	0,008
	Two parameters are out of tolerance		
	Set	Calculated	Error, %
R1, k Ω	15	15.0009	0,006
R2, k Ω	15	15.0002	0,0013
C1, pF	100	99.993	0,07
	Three parameters are out of tolerance		
	Set	Calculated	Error, %
R1, k Ω	15	15.001	0,001
R2, k Ω	15	15.002	0,002
C1, pF	50	50.0005	0,001

From Table 1 it can be seen that in all cases, the parameters of the elements are uniquely identified with an error not exceeding 0.01%.

Then three filter layouts were assembled, the parameters of the elements were measured before installation. Elements with nominal values of parameters were installed in one layout, and elements with values outside tolerances were installed in other models. Test inputs were submitted and output characteristics measured.

Further, the measured values of the output characteristics of the filter using the diagnostic program were calculated values of the parameters of the elements. According to the results of their comparison with the maximum permissible values, defective elements were identified. The values of the parameters obtained as a result of the calculation, the measured values and the maximum permissible values are given in Table 2.

From Table 2 it can be seen that the error in calculating the parameter values does not exceed 0.3% for all three cases.

Table 2. Values of the diagnosed parameters (Experiment)

Parameter	Maximum permissible values		Mock-up 1 One parameters are out of tolerance		
	Lower	Upper	Calculated value	Measured value	Error, %
R1, k Ω	9	11	15.21	15.2114	0,0066
R2, k Ω	9	11	10.14	10.1095	0,3
C1, pF	80	120	103.6	103.385	0,2
			Mock-up 2 Two parameters are out of tolerance		
R1, k Ω	9	11	15.21	15.2112	0.0079
R2, k Ω	9	11	6.84	6.8464	0,094
C1, pF	80	120	103.6	103.75	0,144
			Mock-up 3 Three parameters are out of tolerance		
R1, k Ω	9	11	15.21	15.2102	0,013
R2, k Ω	9	11	6.84	6.84608	0,087
C1, pF	80	120	151.2	0,102	0,067

Thus, the accuracy of diagnosing the parameters of the elements is almost determined by the accuracy of the measuring devices.

For more complex schemes, the error in diagnosing, in addition to the accuracy of measuring devices, also depends on the accuracy of mathematical models of the systems being diagnosed.

7 Conclusion

Thus, in the course of automated diagnostics, the problem inverse to the system design problem is solved — the values of the internal parameters of its mathematical model are calculated from the known (measured) output characteristics of the system.

For a successful diagnosis, the main condition is the presence of an adequate (fairly accurate) mathematical model of the system being diagnosed because the higher the

accuracy of the system model, the higher the accuracy of diagnosis. Therefore, it will be most effective to diagnose a system using CAD-systems that are used during its design.

At the same time, it makes no difference to which particular area the diagnosed system belongs. It can be of any nature, provided it has a sufficiently accurate mathematical model and the possibility to apply test stimuli and measure the output characteristics necessary for a successful diagnosis.

So, as it is not possible to determine the values of all the internal parameters of its model by the usually available small set of output characteristics of the system, the key point in the diagnosis is to find and select those input test influences and control points for measuring the output characteristics that would uniquely determine the values of all internal parameters of the system being diagnosed. That is, they would ensure the single-extremes of the objective function (1), at least in a certain range of possible values of the internal parameters (usually 2-3 times exceeding the allowable spread of the parameter values according to the system specification).

To speed up this procedure and increase its efficiency, it is advisable to use methods based on neural networks and artificial intelligence [15, 16]. At the same time, the constantly increasing processing power of modern computers allows increasing the complexity of diagnosable systems (and their mathematical models) and approaching the solution of real practical problems.

References

1. Palade, V., Bocaniala, C. D., Jain, L.: Computational Intelligence in Fault Diagnosis. Springer (2006).
2. Blanke, M., Kinnaert, M., Lunze, M. J.: Diagnosis and Fault-Tolerant Control. Springer (2006)
3. Witczak, M.: Modelling and Estimation Strategies for Fault Diagnosis of Non-Linear Systems. From Analytical to Soft Computing Approaches. Springer (2007)
4. Ding, S. X.: Model-based Fault Diagnosis. Springer (2008)
5. Halderman, J. D.: Diagnosis and Troubleshooting of Automotive Electrical, Electronic, and Computer Systems (6th Edition). Pearson Education, Inc. (2012)
6. Bunday, B.D. Basic Optimization Methods. Hodder Arnold, London (1984)
7. Gill, P.E., Murray, W., Wright, M.H. Practical Optimization. Elsevier-Academic Press, London (1981)
8. Ravindran, A., Ragsdell, K. M., Reklaitis, G. V.: Engineering optimization: methods and applications. John Wiley & Sons, New Jersey (2006)
9. Wilson, P., Mantooth, H.A.: ModelBased Engineering for Complex Electronic Systems. Elsevier Science, Oxford (2013)
10. Fitzpatrick, D.: Analog Design and Simulation using OrCAD Capture and PSpice. Newnes (2012)
11. Attia, J. O.: SPICE and MATLAB for Electronics: An Integrated Approach, Second Edition. CRC Press (2010)

12. Kasian, M., Kasian, K.: Diagnostic mathematical model of radio-electronic devices. In: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET2018 - Proceedings, Slavske, 20-24 February 2018, pp. 123-127 (2018)
13. Kasyan, K., Kasyan, N.: Diagnosing of radio-electronic equipment with the help of traditional CAD systems. In: International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET2006 - Proceedings, Slavske, 28 February - 4 March 2006, pp. 585-587 (2006)
14. Kasian, K., Kasian, M.: Elementwise Diagnosing of Technical Systems with High Reliability Requirements. In: XIII International Conference Modern problems of radio engineering, telecommunications, and computer science TCSET2016 - Proceedings, Slavske, 23-26 February 2016, pp.558-561 (2016)
15. Kruse, R., Borgelt, C., Klawonn, F. et. al.: Computational intelligence: a methodological introduction. Springer-Verlag, London (2013)
16. Ruan, D. (ed.): Intelligent hybrid systems: fuzzy logic, neural networks, and genetic algorithms. Springer, Berlin (2012)

Diagnostic signal nonstationarity reduction to predict the helicopter transmission state on the basis of intelligent information technologies

Ihor Lymariiev¹[0000-0001-7871-0596], Sergey Subbotin¹[0000-0001-5814-8268],
Andrii Oliinyk¹[0000-0002-6740-6078], Ihor Drokin²

¹Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine limarevigor@gmail.com

²“Motor Sich” JSC, Motorostroiteley avenue, 15, Zaporizhzhia, 69068, Ukraine igor.dv.mail@gmail.com

Abstract. Analysis process and transformation of non-stationary diagnostic signals to reduce the degree of their non-stationarity and the subsequent synthesis of neural network predictive model based on them are considered. The software which allows to apply the developed methods for processing of signals of diagnostics of transmission of the helicopter and on their basis to build the forecasting model for the solution of practical problems of technical diagnostics is developed.

Keywords: nonstationary signal, diagnosis, Dickey - Fuller criterion, neural network, dimension reduction

1 Introduction

Preventing machinery failure is an important component of the maintenance activities of most engineering systems. For quality control of a technical product, at the end of its production and during design tests technical diagnostics is carried out.

Automation of decision-making in the process of technical diagnosis is an urgent task, as it helps to reduce the load on the human operator, as well as provides efficiency and speed of decision-making and reduces the dependence of complex technical systems on the negative impact of the human factor.

A promising basis for solving the problems of automation of diagnosis are intelligent information technologies that are able to build models for examples – experimental observations "input - output" and extract knowledge from the data in the learning process.

The paper considers the process of diagnosing the helicopter transmission. The helicopter is a complex and valuable technical product. Therefore, after designing the test sample is tested on the stand for a large number of hours. During these tests, changes in the vibration level of the sample are monitored. As a result of such tests, a large amount of data is accumulated in the form of diagnostic signals.

A sharp increase in vibration negatively affects the sample, which can lead to its failure. This entails material costs and the need to suspend testing. Therefore, having a system that can predict the vibration level for a certain time in advance will help to improve the testing process.

The basis of such a system can be used an intelligent model of data-driven [1, 2], for example, as one of the most powerful ones - the neural network model, by teaching it to predict the future level of vibration based on historical data collected. However, incoming signals are characterized by a high degree of non-stationary, which makes complex and long-time synthesis and training of such models.

The goal of the work is to develop methods for converting diagnostic signals to reduce their dimension, allocating the necessary component and reducing the degree of their non-stationary for further synthesis of neuromodels based on them.

The objective of this study is to predict these signals using neural network models.

2 Formal problem statement

Consider the problem of diagnosing the helicopter transmission [4]. Preventing machinery failure is an important component of the maintenance activities of most engineering systems. Helicopters are constantly exposed to periodic loads and vibrations that initiate and propagate the occurrence of damage in many components of the equipment. This is due to the design of the helicopter and the presence of complex mechanical systems, such as the inventive rotor, control rotor, main gearbox and other transmission elements. In most cases, the failure of these systems lead to catastrophic situations.

For monitoring the technical condition of the helicopter used systems like Health and Usage Monitoring Systems (HUMS) [5-6]. These systems make it possible to detect damage in the transmission components and predict their residual life.

Important in HUMS is the ability to assess the technical condition of critical elements of the transmission, using the data of vibration signals that were recorded during the flight or ground tests.

A neural network model will be used as a predictive model. Formally, the problem of neuromodels synthesis can be presented in this form.

Suppose given the original sample as a set of prece-dents (instances) $\langle x_b, y_i \rangle$ is a set of S precedents characterizing dependence $y_i(x_t)$, at the moment t , $t = 1, 2, \dots, T$, where $x_t = \{x_{tj}^s\}$, $y = \{y_t^s\}$, $s = 1, 2, \dots, S$, characterized by the set of N input features $\{x_{tj}^s\}$, $j = 1, 2, \dots, N$, where j is a number of feature, and output feature y . Each s -th precedent can be noted as $\langle x_b^s, y^s \rangle$, where $x_t^s = \{x_{tj}^s\}$.

Then the problem of model synthesis of dependence $y_i(x_t)$ will be considered in search of such structure $F()$ and adjusting such values of parameters w of a model which will satisfy the model quality criterion $f(F(), w, \langle x_b, y_i \rangle) \rightarrow opt$, where opt – is a symbol of optimum. Usually, the criterion of learning quality of neuromodels is defined as a function of model error (1):

$$E = \frac{1}{2} \sum_{s=1}^S (y^s - F(\omega, x^s))^2 \rightarrow \min. \quad (1)$$

3 Analysis of the level of nonstationarity of the signals of the diagnostic process

In most research methods, it is assumed that the signal, and in General, any time series is stationary, and the time dependence is considered possible to take into account a variety of non-statistical methods. For example, the series is decomposed into three components: trend, cyclic and random. A trend is a component of a series that changes over a long period of time due to the influence of fundamental factors. The cyclic component changes in time with a certain period according to the course of repetitive processes. The random component includes all the last components that can not be assigned to one of the first two groups. The use of such separation for technical diagnostic signals may not provide optimum results due to the following factors.

First, the component can be quite complicated. For example, its trend characteristics may be nonlinear and non-linear, as predicted in most trend allocation methods. Secondly, in the tasks of technical diagnostics there is always a "white noise", the smallest error of forecasting which is equal to its dispersion. Thus, the series is divided into three parts, although this separation is fuzzy, the cyclic component can become a trend at intervals of less than a period, and the trend and the cyclic component can go into the category of random component.

The solution of the problem of stationary signaling of technical diagnostics is possible, but in most cases it entails a significant change in the structure of the signal itself. And the evaluation of the technical product on the received signal can not provide the necessary level of its quality. Metrics that are based on the components of its internal structure, for example, based on its wavelet decomposition [11], can be applied to a non-stationary signal to assess the quality of a technical product on a diagnostic signal. However, before switching to the application of such methods of constructing forecasting models, it is necessary to make sure that all possible procedures have been carried out to minimize the unsteadiness of the series, which does not lead to a significant change in its initial structure.

To estimate the degree of unsteadiness of the signal, we will use the dilated Dickey-Fuller criterion [13], the number of shifts will be calculated based on the data minimization of the information criterion (AIC - metric). This criterion is used to test the hypothesis of stationarity of the time series. Also, this criterion can be used to estimate the degree of nonstationarity of the time series. The closer the criterion value to the boundary, the lower the degree of nonstationarity has the signal. The values of the Dickie-Fuller criteria for the obtained values are given in Table 1.

Table 1. The value of the Dickie-Fuller criterion for processed signals.

Name of the transmission element	Place of measurement of the input signal	Direction of measurement	Controlled parameter	Value of the extended Dickey-Fuller criterion	Critical value of the extended Dickey-Fuller criterion (for Mac Kinnon [14]) for $p = 0.95$
Rotor blade	Rotor shaft cover	Vertical	Vibration speed	-8.37	-2.86
First gear drive of the intermediate gearbox	Intermediate gearbox	Horizontal	Vibration overload	-9.76	-2.86
		Axial	Vibration overload	-10.44	-2.86
First gear drive of the tail gearbox	Tail gearbox	Vertical	Vibration overload	-13.55	-2.86
		Axial	Vibration overload	-9.29	-2.86

From the results it is possible to observe a sufficiently high degree of unsteadiness of the studied signals. To find the relationship between the current and the previous signal values, we use the auto-correlation function. Autocorrelation is a correlation of the function itself with a certain variable of an independent variable. The autocorrelation function is defined as:

$$R_f(\tau) = \int_{-\infty}^{+\infty} f(t)f^*(t-\tau)dt, \quad (2)$$

where the function $f(t)$ integrates into a product with a complex conjugate and a function displaced by a certain value τ .

The graph of the signal received on the intermediate gearbox (axial direction of measurement), which characterizes the state of the first gear drive of the intermediate gearbox is shown in Fig. 1. The graph also contains a 95% confidence interval to determine the statistical significance of the autocorrelation coefficient.

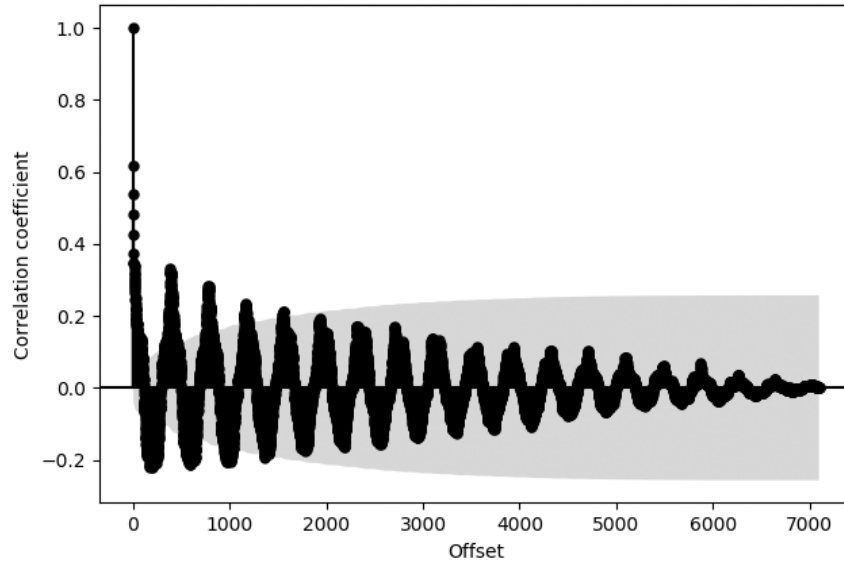


Fig. 1. Autocorrelogram of signal received on the intermediate gearbox (axial direction of measurement)

The carried autocorrelation analysis confirmed the proposed assumption of the presence of cyclicity in the signal. The presence of auto-correlation complicates the use of a number of methods of analysis of time series. Therefore, to reduce autocorrelation, elimination is used to shift from correlation of levels to correction of deviations from trends - residues (for example, the conversion of a time series into a number of values of differences between its adjacent members). This approach can not be applied in this case due to the above reasons for impossibility to change the initial structure of the diagnostic signal.

Another solution to this problem is to split the signal into several components, which will eliminate cyclicity and maintain the original signal strength.

4 Reduction of the nonstationarity of the signals based on information about the operating modes in the cycle of diagnosis

To reduce the non-stationary time series, methods can be applied that will lead to its breakdown into a certain number of smaller time series, in such a way that they are absent, explicitly, as a component of cyclicity [12]. When applying such a breakdown, the problem is to find the period of the cycle (which can change over time). For qualitative partitioning of a series it is necessary to use knowledge of the subject area of the investigated process. Apply expert knowledge in the field of aircraft engineering for breaking the signal of diagnosis. Fig. 2 shows a graph of signal of rotation of the turbocharger rotor left engine.

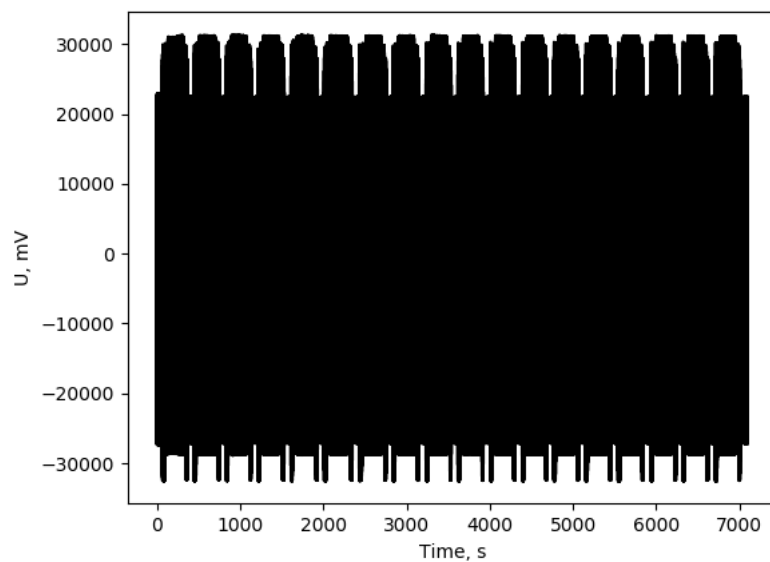


Fig. 2. Raw signal of rotation of the turbocharger rotor left engine

We process the processing of these signals to obtain the values of the rotational frequency by counting the number of times the signal changes in 1 second (the discretization frequency of the rough signal 7200 Hz). The graph of the signal received on the intermediate gearbox (axial direction of measurement), which characterizes the state of the first gear drive of the intermediate gearbox is shown in Fig. 3.

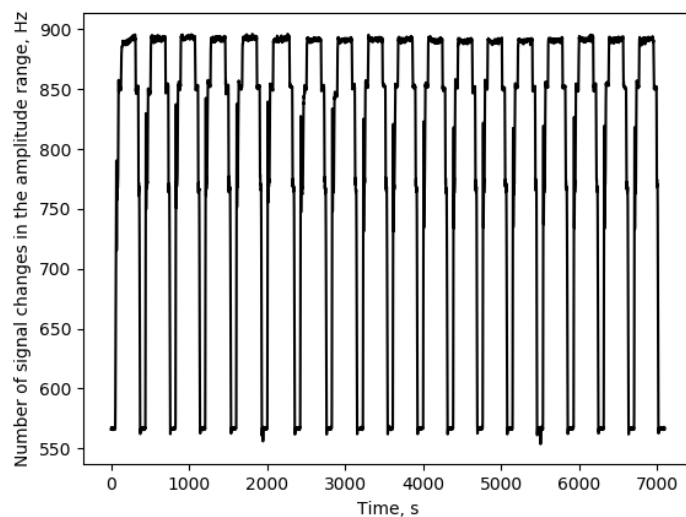


Fig. 3. Processed signal of rotation of the turbocharger rotor left engine

After processing the signal, you can observe the progress of 18 cycles of diagnosis. You can also observe the change in operating modes. In one mode of operation, the speed does not change significantly. The processed signals of rotation of the turbo-charger rotor right engine demonstrate similar results. This is because the left and right engines duplicate each other to increase reliability.

It is proposed to allocate from the general signal of diagnosing the intervals during which the diagnosis took place in one mode of operation. For example, we isolate from the general signal for diagnosing the intervals during which the diagnosis took place in the second mode of operation. It is proposed to perform autocorrelation analysis for the received signals.

The graph of the signal received on the intermediate gearbox (axial direction of measurement), which characterizes the state of the first gear drive of the intermediate gearbox in the second mode of operation of turbochargers is shown in Fig. 4. The graph also contains a 95% confidence interval to determine the statistical significance of the autocorrelation coefficient.

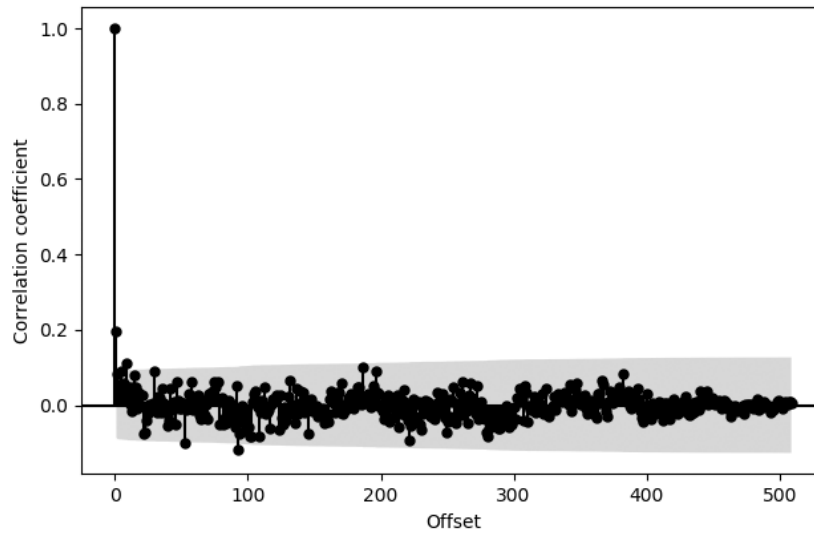


Fig. 4. Autocorrelogram of signal received on the intermediate gearbox (axial direction of measurement) in the second mode of operation of turbochargers

The performed autocorrelation analysis demonstrates the absence of expressed cyclicity in the signal. This indicates a qualitative breakdown of diagnostic signals, which should help to reduce the unsteadiness of newly formed signals.

To quantify the degree of non-stationarity of the signals, we calculate the values of the Dickey - Fuller criterion for them. Table 2 contains the values of the extended Dickey - Fuller criterion for part of the signals obtained from the signals of diagnosing the helicopter transmission by the proposed methods gearbox in the second mode of operation of turbochargers.

Table 2. The values of the extended Dickey - Fuller criterion for the signals obtained from the signals of diagnosing the helicopter transmission in the second mode of operation of turbochargers.

Name of the transmission element	Place of measurement of the input signal	Direction of measurement	Controlled parameter	Value of the extended Dickey-Fuller criterion	Critical value of the extended Dickey-Fuller criterion (for Mac Kinnon [14]) for $p = 0.95$
Rotor blade	Rotor shaft cover	Vertical	Vibration speed	-5.42	-2.86
First gear drive of the intermediate gearbox	Intermediate gearbox	Horizontal	Vibration overload	-5.65	-2.86
		Axial	Vibration overload	-4.78	-2.86
First gear drive of the tail gearbox	Tail gearbox	Vertical	Vibration overload	-5.44	-2.86
		Axial	Vibration overload	-5.56	-2.86

There is a significant decrease in the values of the Dickey - Fuller criteria, which indicates a decrease in the degree of signal non-stationarity after the proposed partition. Also, there is a much smaller scope of the test values than in table 1, this makes it possible to assume that the degree of non-stationarity of the signals corresponds to the non-stationarity of the diagnosis process in this mode. All this makes it possible to better use the signals for the synthesis of predictive models.

5 Experiments and results

A four – layer recurrent neuron network with an input layer, two hidden layers with GRU-cell (Fig. 5) and an output layer with one linear neuron was chosen as a neural network model. GRU [15] is a simplified model of a well – known LSTM cell with significantly fewer parameters. Through this, learning GRU is easier than LSTM, so it is gaining popularity in many real-world tasks.

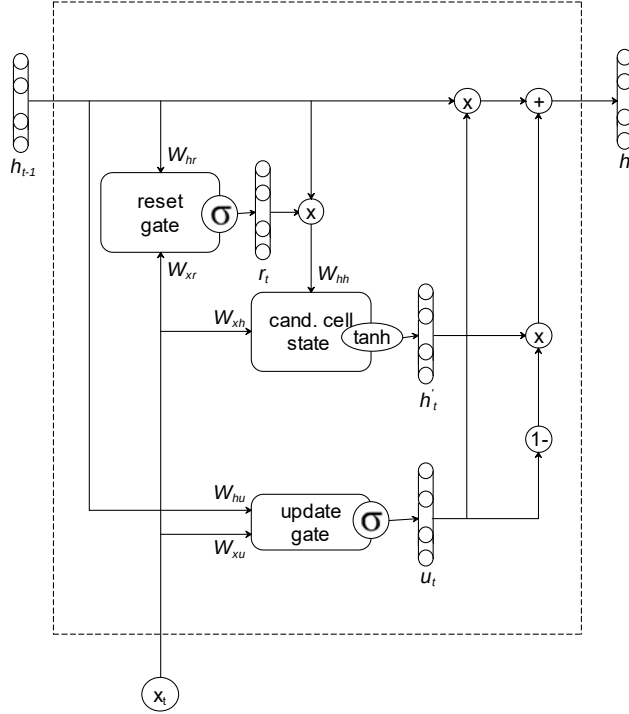


Fig. 5. The structure of the GRU-cell

At the input of the GRU-cell receives a vector x , which contains the current values of the signals. The output of the cell is calculated by the following formulas (3)-(6):

$$u_t = \sigma(W_{xu}x_t + W_{hu}h_{t-1} + b_u), \quad (6)$$

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r), \quad (7)$$

$$h'_t = \tanh(W_{xh'}x_t + W_{hh'}(r_t \otimes h_{t-1})), \quad (8)$$

$$h_t = (1 - u_t) \otimes h'_t + u_t \otimes h_{t-1}. \quad (9)$$

The synthesized neural network model was tested and optimized on a sampling of signals during the diagnosis of the helicopter transmission at the second mode of the turbochargers of the engines. The total initial sample contained data collected in total for 1010 seconds for 22 diagnostic signals. She was divided into a training sample - 1 - 800 seconds, and the test - 801 - 1010 seconds.

As the input data of the models, the data for all signals for a given time window width were used. Data on one of the 22 signals with a given forecasting horizon, that is, a forward shift in the time scale of 120 seconds, were used alternately as initial data. The neural network error was calculated by formula 1. Neural network training was conducted over 250 epochs.

To improve the quality of the forecast, optimization of neural network hyperparameters was carried out. The results are given in tables 3 - 5.

Table 3. Results of the optimization of the batch size

Batch size	Error on test sample	Standard deviation on the test sample
5	0,0195	0,0121
10	0,0219	0,0127
20	0,0208	0,0156
40	0,0178	0,0105
80	0,0180	0,0123
100	0,0244	0,0291

Table 4. Results of optimal optimization algorithm selection

Name of the algorithm	Error on test sample	Standard deviation on the test sample
SGD	0,0207	0,0183
RMSprop	0,0179	0,0103
Adagrad	0,0184	0,0116
Adadelat	0,0182	0,0114
Adam	0,0211	0,0184
Adamax	0,0185	0,0122
Nadam	0,0193	0,0128

Table 5. Results of the optimization of the learning rate

Learning rate	Error on test sample	Standard deviation on the test sample
0,001	0,0187	0,0125
0,01	0,0179	0,0089
0,1	0,0216	0,0096
0,2	0,1030	0,2297
0,3	0,1176	0,1969

Fig. 6 shows a graph of the true and predicted in the neural network model values of the signal received on the intermediate gearbox (axial direction of measurement) and processed with parameter vibration overload that characterizes the state of the first gear drive of the intermediate gearbox during the second mode of turbochargers for the interval 801 - 1010 seconds (test sample).

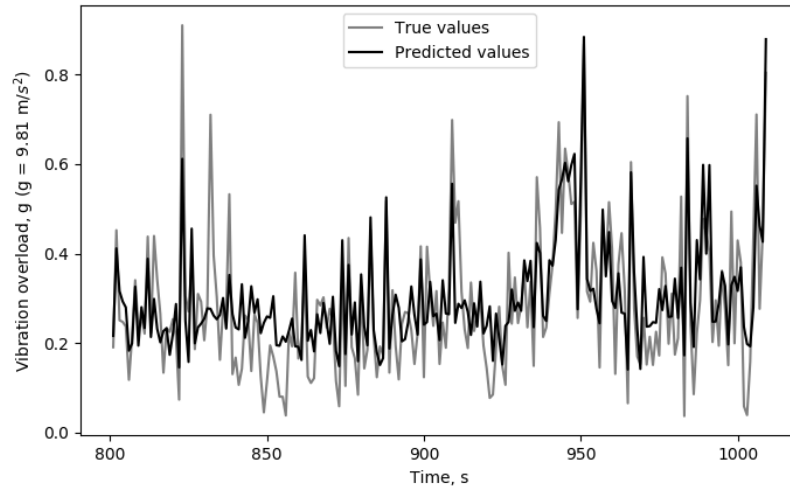


Fig. 6. True and predicted values of the signal

The prediction error is 0.0167 for the second mode of operation of turbochargers, which is 13.45 % in percentage terms. The forecast error for a signal without division into operating modes is 0.1096, which is 28.56% in percentage terms.

Therefore, if it is necessary to predict the state of the helicopter transmission for all modes, it is recommended to divide the signal by the proposed method, to predict each part of the signal using a predictive model and combine them on the basis of information about the duration of the operating modes.

6 Conclusion

In the conducted research of methods of analysis and transformation of non-stationary signals of diagnosing the method of reducing the degree of non-stationarity of the received signals based on expert information about the modes of operation during the diagnosis cycle was proposed.

The effectiveness of the method of reducing the degree of unsteadiness of signals based on expert information about the modes of operation during the diagnosis cycle is evidenced by a significant optimization of the value of the Dickey-fuller criterion after its application. This method, unlike most, does not require changing the internal structure of the signal, and works only by splitting the signal into several other signals, which can then be combined.

On the basis of processed signals, a neural network was synthesized to predict the state of the helicopter transmission and its hyperparameters were configured.

The results are planned to be used to improve the diagnostic quality of the helicopter transmission.

The method of transformation of nonstationary signals and the construction of neural network models developed and applied in the investigated method can be used to solve problems in which it is necessary to predict the future state of the object, by its diagnostic signals.

References

1. Shin Y. C., Xu C.: Intelligent systems: modeling, optimization, and control. CRC Press, Boca Raton (2009)
2. Bishop C. M.: Pattern recognition and machine learning. Springer, New York (2006)
3. Giurgiutiu, V., Cuc, A., Goodman, P.: Review of vibration-based helicopters health and usage monitoring methods. In: 55th Meeting of the Society for Machinery Failure Prevention Technology - Proceedings, Virginia Beach, VA, 2–5 April 2001, pp. 1–10 (2001)
4. Lymariev I. O., Subbotin S.A., Oliinyk A. A., Drokin I. V.: Methods of large-scale signals transformation for diagnosis in neural network models. Radio Electronics, Computer Science, Control, vol. 4, pp. 63–71 (2018)
5. Kazlas P. T., Monsen P. T., LeBlanc M. J.: Neural network-based helicopter gear-box health monitoring system. In: Neural Networks for Signal Processing, Proceedings of the IEEE-SP Workshop, 6–9 September 1993 : proceedings, pp. 431–440 (1993)
6. Dempsey P., Lewicki D., Le D.: Investigation of current methods to identify helicopter gear health. In: 7 IEEE Aerospace Conference, 3–10 March 2007 : proceedings. – Montana, pp. 81–87 (2007)
7. Subbotin S. A., Oliinyk A. A.: The dimensionality reduction methods based on computational intelligence in problems of object classification and diagnosis. Advances in Intelligent Systems and Computing, vol. 543, pp. 11–19 (2017)
8. Jensen R., Shen Q.: Computational intelligence and feature selection: rough and fuzzy approaches. John Wiley & Sons, Hoboken (2008)
9. Lee J. A.: Nonlinear dimensionality reduction. Springer, New York (2007)
10. Guyon I., Elisseeff A.: An introduction to variable and feature selection. Journal of machine learning research, vol. 3, pp. 1157–1182 (2003)
11. Tverdohleb J., Limarev I., Logominov V., Dubrovin V.: Wavelet analysis of complex nonstationary oscillatory signals. In: 4-th IEEE International Scientific-Practical Conference “Problems of Infocommunications. Science and Technology.” - Proceedings, Ukraine, Kharkiv, NURE, 2017. pp. 17-20 (2017)
12. Ukil A.: Intelligent systems and signal processing in power engineering. Springer, Berlin (2007)
13. Dickey, D. A., Fuller, W. A.: Distribution of the Estimators for Autoregressive Time-Series with a Unit Root. Journal of the American Statistical Association, vol. 4, pp. 427–431 (1979)

14. MacKinnon, J. G.: Critical Values for Cointegration Tests. In: R. F. Engle and C. W. J. Granger, eds., Long-Run Economic Relationships: Readings in Cointegration, Oxford University Press, pp. 267–276 (1991)
15. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning, MIT Press (2016)
<http://www.deeplearningbook.org>

The Throughput of Technical Channels as an Indicator of Protection Discrete Sources from Information Leakage

Igor Korobiichuk¹[0000-0002-5865-7668], Serhii Ivanchenko²[0000-0003-1850-9596], Oleksandr Roma²[0000-0001-9074-6137], Anatolij Golishevsky³[0000-0001-9981-7771], Ruslan Hryshchuk⁴[0000-0001-9985-8477]

¹Warsaw University of Technology, Institute of Automatic Control and Robotics, Warsaw, 02-525, Poland

i.korobiichuk@mchtr.pw.edu.pl

²National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Institute of special communication and information protection, Kyiv, 03056, Ukraine
soivanch@ukr.net, proffesor69@ukr.net

³National scientific research institute of special communications and information protection of Ukraine, Kyiv, 03142, Ukraine
tolyan207@ukr.net

⁴Sergey Korolyov Zhytomyr Military Institute, Cybersecurity Department of the Research Center, Zhytomyr, 10004, Ukraine
dr.hryshchuk@i.ua

Abstract. Considered the theoretical conditions of discrete information sources security, analyzed the throughput as an indicator of its protection from information leaks. Introduced the notion of throughput limit values of information leaks channel, proposed its use as a security norms. The resulting values establish a connection of specified index with the desired probability of error in the channel.

Keywords: information, information security, security risk, safety of information, technical channels of information leakage, technical information protection.

1. Introduction

One of the tasks of information security management [1, 2], as determined Deming-Shewhart cycle, is a security risk analysis, including risk of a breach of information security from leaking by technical channels. It is known, that work of technical means and systems of processing and transmission of information is almost always accompanied by a list of adverse effects [3-12]. These are unwanted radiation of electromagnetic fields to the environment, emission and leakage of electric current to taking off conductors. These carriers may arise from information sources and uncontrolled spread the dangerous signals outside the controlled area.

It is known, that the criterion of protection is the condition of impossibility extract semantic content from intercepted message [1, 2], that at the point of interception is

determined certain of energy terms – norm the signal/noise ratio. In turn, a quantitative norm of this measure must provide authentically needed probability of error in technical leakage channel and, in accordance, previously specified security risk [1, 12-14]. For substantiation of dependence signal/noise ratio with probability of error in work [13, 14] was based on the Kotelnikov's optimum receiver [15-18]. In its relation was found the assessment of lower limit of the probability of errors in the channel, determined by allowable signal/noise ratio and does not depend from choose of interception receiver. As is evident, determining the desired probability of error in the technical leakage channels requires certain studies and solutions.

So, the main task of this article is justification the technical channel throughput as indicator of protection discrete sources from information leakage, and also justification the dependency of this throughput with the necessary probability of error in the technical leakage channel, that will provide a given state of information security risk by given parameters.

2. Materials and methods of research

To resolve this specified task, consider the channel of leakage, as discrete channel, which connects a source of information and enemy receiver with some non-fixed, but different from zero probability of error (Fig. 1).

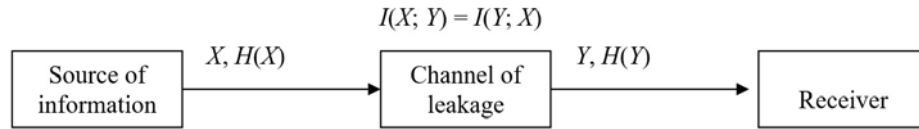


Fig. 1. Discrete channel of information leakage

Suppose at the input channel of information sources gets some information sequence data X_k^n , for ease, binary symbols, where k – number of combination ($k = 1, 2, 3, \dots, 2^n$) with probability $p(X_k^n)$. The specified probability $p(X_k^n)$ – is a priori probability, which characterizes the chances of guessing by enemy before it is transmitted through the channel.

In the channel under the influence of noise sequence X_k^n transformed into some consistency Y_l^n the same length and the same alphabet symbols, where l – number of combination ($l = 1, 2, 3, \dots, 2^n$). Noise is random nature, that's why this transformation can be described by conditional probability $p(Y_l^n / X_k^n)$, which is a measure that determines possibility of transition sequences X_k^n to Y_l^n . In this case a posteriori probability, which determines the chances of guessing the information signs by enemy, after taking its continuous implementation, can be expressed by the Bayes formula:

$$p(X_k^n / Y_l^n) = \frac{p(X_k^n) p(Y_l^n / X_k^n)}{p(Y_l^n)}, \quad (1)$$

where

$$p(Y_l^n) = \sum_{k=1}^{2^n} p(X_k^n) p(Y_l^n / X_k^n) \quad (2)$$

– the probability of sequence Y_l^n at the output of channel.

For the given channel its throughput is defined as the maximum amount of information for all possible its implementation for all probability allocations, which can be transmitted through the channel, [19 - 22]:

$$C = \max_X I(X; Y), \quad (3)$$

where

$$I(X; Y) = H(X) - H(X/Y) = H(Y) - H(Y/X) = I(Y; X) \quad (4)$$

– mutual channel information quantity,

$$H(X) = M \left[\frac{1}{n} \log_a \frac{1}{p(X_k^n)} \right] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{2^n} p(X_k^n) \log \frac{1}{p(X_k^n)} \quad (5)$$

– the average amount of information that produces source for one binary sign - absolute source entropy X ,

$$H(X/Y) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{2^n} \sum_{l=1}^{2^n} p(X_k^n, Y_l^n) \log \frac{1}{p(X_k^n / Y_l^n)} \quad (6)$$

– conditional source entropy for a binary sign,

$$H(Y) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^{2^n} p(Y_l^n) \log \frac{1}{p(Y_l^n)}, \quad (7)$$

– unconditional entropy of channel output Y for one binary sign,

$$H(Y/X) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{2^n} \sum_{l=1}^{2^n} p(Y_l^n, X_k^n) \log \frac{1}{p(Y_l^n / X_k^n)} \quad (8)$$

– conditional entropy of channel output for one binary sign or conversion $X_k^n \rightarrow Y_l^n$.

If $k = l$, then the combination of sequences $X_k^n = Y_l^n$ and, in accordance, the transmission in the channel occurred unmistakably. It is not difficult to make sure, that $H(Y/X) = 0$ bit, and $C = 1$ bit. In other cases there is an error and if it is probability higher, than higher the difference between X_k^n and Y_l^n , than greater distorted of symbols and information lost. In this case $H(Y/X) \rightarrow 1$ bit, and $C \rightarrow 0$ bit. It should be noted, that equality of latter achieved by conditions of equal probability and statistic errors independence, but statistical dependence of mistakes, even i by conditions of Its equal probability will result to increased throughput [15-18].

Shannon's information theory also known, that information can freely pass through the channel only on condition, if $C \geq H(X)$. If $C < H(X)$ information through the channel may partially pass and partially lost. On condition, if $C = 0$ bit, than channel condition is provided as "channel breakage" or "no channel".

Condition of "channel breakage" is theoretically idealized condition, which is desirable for all technical information leakage. But in practice it is difficult to reach, and even at all impossible. Typically, ensure of that conditions for the safe processing and transmission of information requires the creation specific objects and building constructions, it is associated with the creation of large-scale controlled areas, using highly efficient surge protector, or autonomous power sources, etc. All it requires large financial and energy costs, which sometimes may not be appropriate because of their dominance over most information precious. Because safety of information should include a balance between deliberately spent and acceptable lost.

In order to optimize the costs of securing technical information leakage requires a list of scientific-technical justifications and solutions, associated with providing small but nonzero values of through. This indicator must take into account the variety of modes of technical processing and transmission items, of which information leakage, features of group transmission channels that combine information from different sources with different properties, the possibility of automated regulation of noise immunity channels, etc.

Considering this throughput of technical leakage channels C_{tlc} can be attributed to indicators of the evaluation of information security. For example, if asked acceptable information security risk, then it possible to put accordance of the limit allowable maximum throughput of technical leakage channel $C_{tlc \text{ l.a.}}$. The presence of of this rule allows to find the probability of error in the leakage channel and power conditions in the point of possible interception – the signal/noise ratio, which will ensure the security of information with given permissible risk.

3. The results of research

Let's analyze the nature of throughput for the general case and justify communication of maximum throughput and needed for ensure probability of errors. In this case will suppose, that data, which simultaneously can be processed by technical means and flow by technical channels, can have very different origins, that is generated from Q different information sources with a different syntax and different semantics (Fig. 2).

Obviously, that due to the fragmentation of sources for each one, dangerous signal spreading by the same environment, which provides the same signal/noise ratio at the intercept receiver input, technical leakage channel can create different throughput. It is also obvious, if leakage of information from each source consider separately, then shown on Fig. 2 technical channel may consist of several technical channels, generated from different sources with its throughput. But for all of these channels takes place only one receiver, for which need to justify condition of impossibility of information interception.

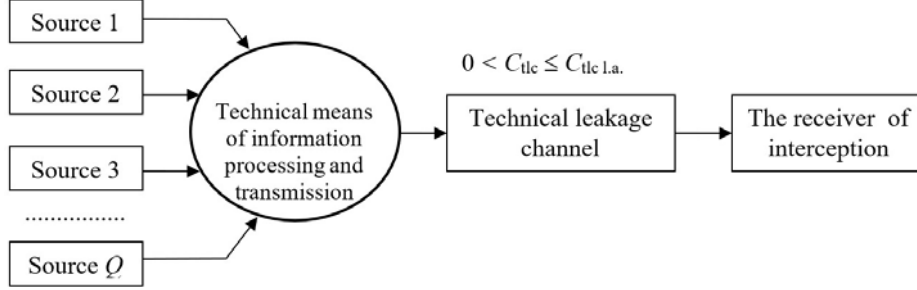


Fig. 2. Typical technical channels of information leakage for modern technical means of information processing and transmission with the given limit allowable throughput

Suppose that in leakage channels have a place an errors, the presence of which is guaranteed by noise of environment spreading of dangerous signals. All errors in the channel are statistically independent and are characterized by probability, which is not below some certain of intercept receiver features.

Based on the used in article the definition of throughput, which is expressed by ratios (3), (4), (5), (6), (7) and (8), throughput of technical leakage channel, that calculated on average per one binary sign, may be expressed by the formula [15, 16]:

$$C_{\text{tlc}} = 1 - H(Y/X) \text{ [bit]}. \quad (9)$$

One of the approaches for ensure the guaranteed safety, is calculation the indicators of security for the worst case in terms of security. In this case protection will be ensured and in the the worst case and in all other cases, with the creation of a certain reserve.

Determining of worst case is requires the inspection of conditions security for all of leakage sources. Therefore, will expression the technical leakage channels throughput as averaging of each source throughput and select among from last limit allowable maximum, which will apply as an indicator of security:

$$C_{\text{tlc}} = \frac{1}{Q} \sum_{r=1}^Q C_r. \quad (10)$$

On the other hand, considering (9) and (8) specified throughput will look like:

$$C_{\text{tlc}} = 1 - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{2^n} \sum_{l=1}^{2^n} p(Y_l^n, X_k^n) \log \frac{1}{p(Y_l^n / X_k^n)}. \quad (11)$$

Is necessary to find identity of ratios (10) and (11) and select in throughput of technical leakage channels a maximum component C_r and to accept it as the norm for C_{tlc} .

To do this will analyze of probability in ratio (11).

In view of the independence of errors, mentioned probabilities can be expressed in the form of next ratios.

Conditional probability, which characterizes the possibility of sequence at the output of channel Y^n , if at the input came X^n :

$$p(Y^n / X^n) = p(y_1 / x_1) p(y_2 / x_2) \times \dots \times p(y_i / x_i) \times \dots \times p(y_n / x_n). \quad (12)$$

where $p(y_i / x_i)$ – the probability of in i - grade of sequence Y^n the sign - y at the output of channel, If at the input came sign x_i of sequence X^n in the same grade.

It is not difficult to notice, that the probability $p(y_i / x_i)$ of additive channel can be expressed as

$$p(y_i / x_i) = p(y_i - x_i) = p(e_i). \quad (13)$$

In view of the (13) the ratio (12) for conditional probability can be expressed in terms of the multiplication of errors probability $p(e_i)$, which taking values within the limits $0 < p(e_i) < 1/2$ [7].

$$p(Y^n / X^n) = p(e_1) p(e_2) \times \dots \times p(e_i) \times \dots \times p(e_n). \quad (14)$$

Compatible probability, which describes the possibility of simultaneous existence on the channel input a sequence X^n and on the channel input Y^n , equals:

$$p(Y^n, X^n) = p(X^n) p(Y^n / X^n) = [p(x_1) p(x_2) \times \dots \times p(x_i) \times \dots \times p(x_n)] \times [p(y_1 / x_1) p(y_2 / x_2) \times \dots \times p(y_i / x_i) \times \dots \times p(y_n / x_n)] \quad (15)$$

and considering (13) will view as:

$$p(Y^n, X^n) = [p(x_1) p(x_2) \times \dots \times p(x_i) \times \dots \times p(x_n)] \times [p(e_1) p(e_2) \times \dots \times p(e_i) \times \dots \times p(e_n)] \quad (16)$$

where $p(x_i)$ – probability of x_i - sign on the channel output.

Inserting (14) and (16) to (11) and simplifying its, we get next:

$$C = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left[1 + \sum_e p(e_i) \log_2 p(e_i) \right] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n C_i, \quad (17)$$

where C_i – channel throughput for i - information sign.

Equating $Q = n$, we get the required identity, where

$$C_{\text{tlc l.a}} = \max_{i,n,Q} C_i = 1 + \sum_e p(e_i) \log_2 p(e_i) \quad (18)$$

It should be noted, that the ratio (14), which expresses technical leakage channels throughput, is multidimensional convex downwards function of the probability of errors $p(e_i)$. It is not difficult to verify, taking by all arguments the second derivative and identifying a sign. The specified can be shown graphically for arguments $p(e_i)$, having fixed everyone else in the range from 0 to 1/2 (Fig. 3):

$$p(e_j) = \text{const}, 0 < p(e_j) < 1/2 \text{ for all } j = 1, 2, \dots, n, j \neq i. \quad (19)$$

As seen from the graphical representation of throughput, ensuring its limit allowable value $C_{\text{l.a.}}$ for different values i can be determined different values of errors probability.

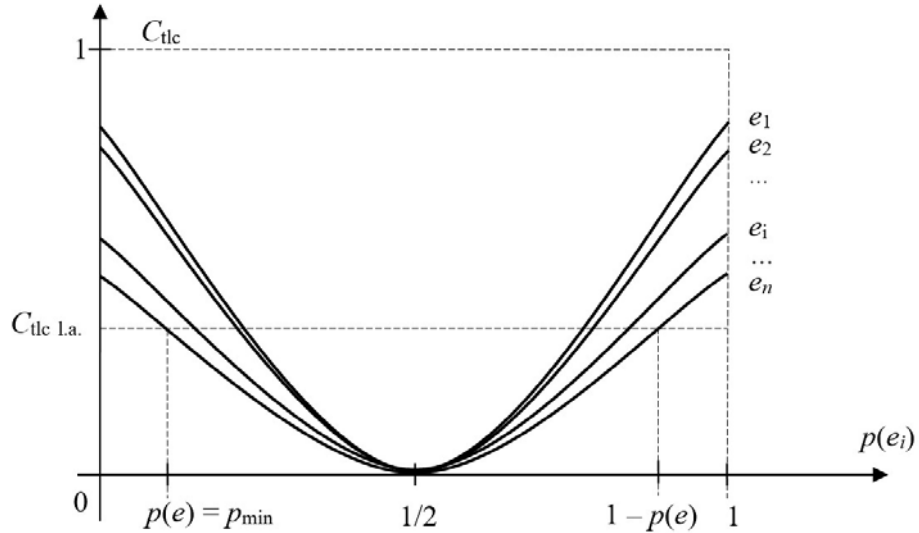


Fig 3. Graphs of dependence of throughput leakage sources from errors probability in the technical channel

Assertion. If asked a discrete binary channel for which presence of errors is guaranteed, all errors are statistically independent and their probabilities for all i ($i = 1, 2, \dots, n$) are within the limits $p_{\min} \leq p(e_i) \leq 1 - p_{\min}$, then its throughput by average does not exceed limit allowable value as expressed by the ratio:

$$C_{\text{l.a}} = 1 + p_{\min} \log_2 p_{\min} + (1 - p_{\min}) \log_2 (1 - p_{\min}) \geq C \quad (20)$$

The proof. If you know the errors probability $p(e_i)$, then throughput channel for i -information signs may be expressed by ratio:

$$C_i = 1 + \sum_e p(e_i) \log_2 p(e_i) = 1 + p(e_i) \log_2 p(e_i) + (1 - p(e_i)) \log_2 (1 - p(e_i)) = 1 - h(p(e_i)) \quad (21)$$

where $h(p) = p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{(1 - p)}$ – entropy function of the argument p .

From the convexity properties of entropic function $h(p)$ and its symmetry from $p = 1/2$ it follows that for argument $p \in (p_{\min}, 1 - p_{\min})$ its value is limited from below:

$$h(p) \geq h_{\min} = p_{\min} \log_2 \frac{1}{p_{\min}} + (1 - p_{\min}) \log_2 \frac{1}{(1 - p_{\min})}, \quad (22)$$

and throughput from top

$$C_i \leq C_{\text{gr.d}} = 1 + p_{\min} \log_2 p_{\min} + (1 - p_{\min}) \log_2 (1 - p_{\min}). \quad (23)$$

Inserting inequality (19) to the value (14) we obtain upper limit, or limit allowable channel throughput for a given $p_{\min}$:

$$C = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n C_i \leq C_{\text{gr.d}} = 1 + p_{\min} \log_2 p_{\min} + (1 - p_{\min}) \log_2 (1 - p_{\min}) \quad (24)$$

that had to prove.

Therefore, based on Shannon information theory provided considered theoretical security of discrete sources analysis throughput as an indicator of security from leaking by technical channels. The concept of limit allowable throughput leakage channel and suggested its to use as a security norms.

Conclusions

Obtained analytical ratios, establishing communication of norms given index with the necessary of errors probability in the channel. In establishing consistency between the set of security risks and necessary limit allowable throughput of the technical leakage channel, using of these ratios allows determining the desired probability of error in the channel and in the next -desired signal/noise ratio in the point of possible information interception.

Obtained results in the article is the development of the theory of information security from leaking by technical channels, and also could form the basis for the implementation in standards of information security series ISO/IEC 27000.

References

1. Information technology. Security techniques. Information security management systems. Requirements [ISO/IEC 27001:2013].
2. The Law of Ukraine "About protection of information in telecommunication systems". <http://www.zakon.rada.gov.ua/go>
3. Buzov G.A., Kalinin S.V., Kondratev A.V. Protection of information from leaks through technical channels, Goryachaya liniya: Moskva, Telecom (2005)
4. Parshutkin, A.V., Levin, D.V., Zaytsev, S.A., Egin, A.V.: Application of structural interference for data protection from information Leakage in the stray electromagnetic radiations channel. SPIIRAS Proceedings, vol 3 (58), pp. 160-181. (2018) doi: 10.15622/sp.58.7
5. Kuhn G.: Compromising emanations: eavesdropping risks of computer displays. This technical report is based on a dissertation submitted June 2002 by the author for the degree of Doctor of Philosophy to the University of Cambridge, Wolfson College (2002) <http://www.cl.cam.ac.uk/techreports>
6. Kuhn, M.G.: Electromagnetic eavesdropping risks of flat-panel displays. Lecture Notes in Computer Science, 3424, pp. 88-107 (2005)
7. Kuhn, M.G.: Optical time-domain eavesdropping risks of CRT displays. Proceedings - IEEE Symposium on Security and Privacy, 2002-January, art. no. 1004358, pp. 3-18 (2002) doi: 10.1109/SECPRI.2002.1004358
8. Korobiichuk, I., Dobrzahansky, O., Kachniarz, M.: Remote control of nonlinear motion for mechatronic machine by means of CoDeSys compatible industrial controller. Tehnički vjesnik/Technical Gazette, Vol. 24/No. 6, pp. 1661-1667 (2017) doi: 10.17559/TV-20151110164217
9. Lenkov, S.V., Peregudov, D.A., Horoshko, V.A.: Methods and means of information protection. Tom I. Unauthorized receipt of information, Ariy: Kyiv (2008)
10. Qiu, J., Li, H., Zhao, C.: Cancelable palmprint templates based on random measurement and noise data for security and privacy-preserving authentication. Computers and Security, 82, pp. 1-14 (2019) doi: 10.1016/j.cose.2018.12.003
11. Korobiichuk I.V., Hryshchuk R.V., Horoshko V.O., Hokhlacheva Yu.E. Self-diagnostics of complex systems with a software-configurable structure. Informatics and Mathematical Methods in Simulation, vol 8, No. 1, pp. 36-47 (2018)
12. Korobiichuk, I., Hryshchuk, R., Mamarev, V., Okhrimchuk, V., Kachniarz, M.: Cyberattack Classifier Verification. International Conference on Diagnostics of Processes and Systems DPS 2017: Advanced Solutions in Diagnostics and Fault Tolerant Control, pp. 402-41 (2018) doi: 10.1007/978-3-319-64474-5_34
13. Ivanchenko, S.O.: Justification safety risk information about its security from leaking by technical channels. Scientific and technical digest "Legal, regulatory and metrological support of information security in Ukraine", Kiev, NTUU "KPI" SRC "Tezis", vol 1 (31), pp. 9 – 13 (2016)
14. Duc, A., Faust, S., Standaert, F.-X.: Making masking security proofs concrete: Or how to evaluate the security of any leaking device. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9056, pp. 401-429 (2015) doi: 10.1007/978-3-662-46800-5_16
15. Fink L. M.: The theory of transfer of discrete messages [2-d edition], Sov. Radio: Moskva, (1970)
16. Kulkarni, A.N., Bukate, R.R., Nanaware, S.D.: Study of Various Attacks and Routing Protocols in MANETS. 2018 International Conference on Information, Communication, En-

- gineering and Technology, ICICET 2018, art. no. 8533696 (2018) doi: 10.1109/ICICET.2018.8533696
17. Burachenko, D.L., Zavarin, G.D., Klyuyev, N.I. et. al.: General theory of communication, VAS: Leningrad (1970)
 18. Lee, M., Neifeld, M.A., Ashok, A.: Capacity of electromagnetic communication modes in a noise-limited optical system. *Applied Optics*, 55 (6), pp. 1333-1342 (2016) doi: 10.1364/AO.55.001333
 19. Ivanovsky, R.I.: Theory of probability and mathematical statistics, BHV: Petersburg (2008)
 20. Niyato, D., Hossain, E.: A queuing-theoretic and optimization-based model for radio resource management in IEEE 802.16 broadband wireless networks. *IEEE Transactions on Computers*, 55 (11), pp. 1473-1488 (2006) doi: 10.1109/TC.2006.172
 21. Gallager R. G.: Information theory and reliable communication, Sovetskoe radio: Moskva (1974)
 22. Qian, Y., Zhou, X., Li, J., Shu, F., Jayakody, D.N.K.: A Novel Precoding and Impulsive Noise Mitigation Scheme for MIMO Power Line Communication Systems. *IEEE Systems Journal* (2018) doi: 10.1109/JSYST.2018.2880962

A Structure of Semantic Service in a Distributed Knowledge Based System

Nataliia Kulykovska¹[0000-0003-4691-5102], Artur Timenko²[0000-0002-7871-4543]

¹Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine
natalya.gontar@gmail.com

²Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine
timenko.artur@gmail.com

Abstract. The main difference in distributed systems based on knowledge is the use of the service approach and ontologies in knowledge engineering. Semantic service acts as a software agent that regulates the interaction of all components of the system. Semantic service provides: semantic principle, which defines a formal description of information and allows you to define the following characteristics of services: scalability, semantic interoperability, formal models of services and ontologies; the principle of decision making; the principle of distribution, which allows you to aggregate the capabilities of several computing objects through cooperation. The communication and collaboration mechanism of the distributed systems based on knowledge covers the message transfer between services, information exchange between semantic service and distributed knowledge bases, and the semantic service behaviors in the entire process from request submission, request handle, to result return. The knowledge base is formed from three types of ontologies. The rules for recognizing performance problems are converted into ontology and the system knowledge base is placed. The working memory of a semantic service contains facts that correspond to simple events, composite events, and identified problems and recommendations. The algorithm for using semantic service is discussed in the article. To assess the performance of the SS attached, create an information system of two components: event trace generator and event trace analyzer.

Keywords: distributed system, semantic service, components of a distributed system, knowledge engineering, data, ontology, services, semantic web, structure of semantic service

1 Introduction

Economic globalization and global networks have become the trend of world development. Especially with the development of information technology and the innovation of management theory, modern enterprises need to go beyond the boundaries of traditional companies to achieve for a rapid and effective integration of resources on a global scale, thus can produce high-quality products that meet the needs of the market or provide services the customer need in a very short period of time [1].

In this case, a new distributed computer system in terms of development of modern technologies put the emphasis on the properties of interoperability and scalability. This direction is connected with the rapid growth of blockchain technology and the Internet of things (IoT). As analysts predict that, by 2025, the share of block chain apps will account for 10% of world gross domestic product [2]. By 2020 will create more than 30 billion IoT devices. IoT will affect every industry, from retail to health care [3].

At the moment in the literature there are a large number of definitions of the concept of "distributed system". The most comprehensive definition proposed by AS Tanenbaum [4]: "Distributed system (DS) is a set of independent computers, which is perceived by its users as the only consistent system." Another definition is proposed in [5]: DS are software and hardware systems, in which execution of operations (actions, calculations) necessary to ensure the target functionality of the system is distributed (physically or logically) between different performers. In the computing field, under our computer, in our study, we will understand the software and hardware system created for a specific practical application, the functionality of which is distributed on various nodes.

On the one hand, DS are tools that allow solving a large number of complex tasks, most of which are not solved by other methods. DS can eliminate the main drawback of centralized systems - the limitations of increasing computing power. At the same time, the analysis revealed a number of problems that need to be solved. The limited use of DS is complicated by the use of equipment from different manufacturers with different types of architectures. Due to the wide variety of aspects of building computing systems, as well as the variety of existing operating systems, it becomes necessary to create methods for adaptive planning of distribution of flows in a distributed system, which will significantly speed up the processing of incoming requests for services and increase the overall system performance.

Modern representations of data are changing the ways and forms of communication, production and consumption of information [6]. The dominance of horizontal relations, structure-role of information decentralization of all types of data available at any time on every device. The user should not care about the specific technology used to provide computing capacity or data storage, so you can say that a user has some information about the remote resource. Distributed systems based on knowledge (DKBS) in order to study the data, their processing and use evolving technologies of the semantic web.

The formal model of the DKBS consists of a set of ontologies; lots of services; a set of events that describe the processes of the system; semantic service (SS); a set of composite services and knowledge base. The completeness and effectiveness of the system is determined by a multitude of ontologies. The main difference in DKBS is the use of the service approach and ontologies in knowledge engineering. SS acts as a software agent that regulates the interaction of all components of the system.

2 Literature review

Software agents originally were discussed in the 70's, and in the mid 90's briefly gained some momentum but then stalled. The "software agent" term has found its way into a number of technologies and has been widely used, for example, in artificial intelligence, databases, operating systems and computer networks literature. Although there is no single definition of an agent [7, 8, 9] all definitions agree that an agent is essentially a special software component that has autonomy that provides an interoperable interface to an arbitrary system and/or behaves like a human agent, working for some clients in pursuit of its own agenda. Most discussions on agents focus on their autonomy, intelligence, mobility and interaction [10, 11, 12, 13, 14]. Agent-based systems [15, 16] claim to be next generation software capable of adapting dynamically to changing business environment and of solving a wide range of knowledge processing application. Although sophisticated software agents can be difficult to build from scratch due to the skills and knowledge needed, the widely available agent construction toolkits may provide a quick and easy start to building software agents without much agent expertise. Significant research and development into multi-agent systems (MAS) has been conducted in recent years [17, 18, 19], and there are many architectures available today [20, 21]. Nevertheless, several issues still need to be faced to make the multi-agent technology widely accepted: secure and efficient execution supports; standardization; appropriate programming languages and coordination models.

Decentralization and openness are inherent properties of multiagent systems (MAS). The technologies they provide are thus the right abstraction for developing Web-oriented applications. Moreover, different works have been proposed to use Semantic Web technologies (SWT) for representing various dimensions of MAS (e.g., interaction protocols, norms, organizations).

Consequently, recent research in MAS have seen an intensive use of Knowledge representation together with increasing use of SWT. We envision that SWT will ultimately play a central role in all parts of MAS. Thus far, work combining MAS and SWT have only been concerned about addressing one dimension of MAS at a time. Also, they were mostly tackling the agent [22, 23] and interaction dimensions [24, 25] to ease communications, especially on domain knowledge. Other works have used those technologies to model part of the organization structure [26], norms and commitment [27], reputation [28] and more. The situation shows that it is time to go beyond these ad hoc solutions and integrate the pieces into a complete Semantic-Web-based infrastructure. We observe too that none of the mentioned contributions were really taking advantage of the Web aspect of these technologies, except some Web service integration. We also want to provide models and specifications for the SS so that services can uniformly query and reason about DKBS, web services, data, and ontologies.

3 A Distributed Knowledge Based System

The proposed architecture is based on services and is designed to support all the processes of the system life cycle. That is, the proposed architecture is meant to include all the concepts necessary to perform all activities related to the life cycle. We shall adopt a generic, knowledge-based architecture. The enterprise members are represented by autonomous agents, geographically scattered, which are able to cooperate to achieve a common business goal.

The various persons constituting the DKBS, in order to process distributed knowledge bases, assume the following roles:

- Knowledge Manager: is in the top of the hierarchy, and acts as a project manager, but in higher levels. It's like a knowledge strategist, cooperating, defining, and distributing the knowledge to coordinate all the other roles;
- Knowledge Provider: is the owner of human knowledge. It is typically an expert in the application domain, but could be another person in the organization who does not have the expert status;
- Knowledge Analyst: uses a range of methods and tools that make the analysis of a standard knowledge-intensive task relatively straightforward;
- Knowledge System Developer: is responsible for the design and implementation. The developer must have a basic background of analysis methods. In knowledge system development, the main knowledge problems have been solved by the knowledge analyst. Therefore, this role must have some skills of software designers;
- Knowledge User: makes use, directly or indirectly, of a knowledge system. Its interaction with the knowledge base system is important for the project development and validation;
- Project Manager: manages the project, specially the knowledge engineer and the knowledge system developer.

In turn, the structure of the system based on knowledge, is shown in Fig. 1. The main components are [29, 30,31]:

- base of knowledge. Base of knowledge intended for storage of knowledge about the subject. Its concrete form depends strongly on the chosen model of knowledge representation. The presence of this component is the main difference between the systems based on knowledge from other types of programs;
- output machine. Output machine generates a response to the user request using the knowledge base. Its principle of operation is also dependent on the chosen model of knowledge representation;
- editor knowledge base – program to change the contents of the knowledge base;
- user interface – the mechanism by which the communication user and the system.

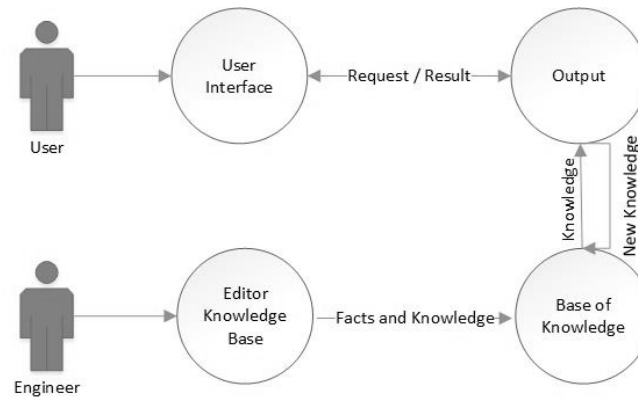


Fig. 1. Structure of systems based on knowledge.

In addition to the agents representing the distributed knowledge base, a SS is introduced in the MAS community. This agent reacts by seizing deal opportunities present in the DKBS, and proceed thus to establish the corresponding virtual entities.

The principle of the Distributed Knowledge Base proposed in this paper is to depends on a number of ontologies that provide semantic principle: ontologies of events DKBS (Task Ontology); ontologies of services (Application Ontology); ontology specialized areas of data systems based on (Domain Ontology). (fig. 2).

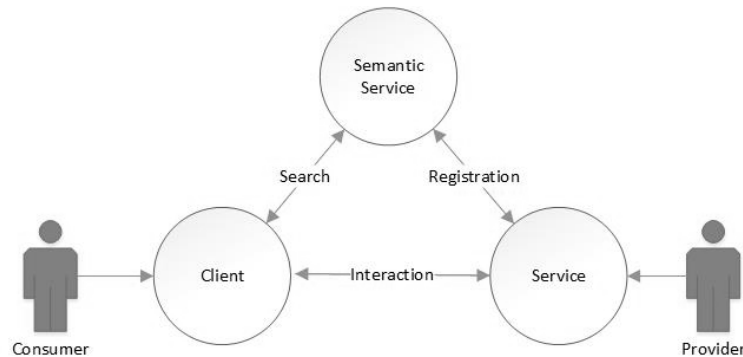


Fig. 2. Structure of DKBS.

DKBS can be represented as a model of interaction of the main artifacts: clients, services and SS. Each component in the system is associated with a specific ontology. SS provides:

- semantic principle, which defines a formal description of information and allows you to define the following characteristics of services: scalability, semantic interoperability, formal models of services and ontologies;
- the principle of decision making;

- the principle of distribution, which allows you to aggregate the capabilities of several computing objects through cooperation.

Considering all the types of ontologies [32, 33, 34], we chose it for three, because the Application Ontology is a model of a service describes its functional and non-functional characteristics. Domain Ontology is a certain base set of knowledge for each subject area. Task Ontology keeps information about all the functions and actions of the system.

Each ontology has its own expressive capabilities depending on their functional purpose. Ontology-oriented for subject area (Domain Ontology), describe the dictionary of terms concrete and formal set four end subsets: concepts, relations, axioms and interpretation functions. Task Ontology consists of a Glossary of terms, specialize tasks and actions in subject area. Every task has a different status and stages of its implementation. The main feature for all concepts of the ontology problem is the time. Application Ontology is the most specific ontology, in addition to all the basic concepts, contains specialized terms and instances subject area.

SS includes domain ontology and a set of modules to operate the services and their ontological description a lot of events we have formed as a set of claims that can be applied to DKBS and a set of axioms of their appearance.

4 Structure of Semantic Service in a DKBS

In each interaction of the SS, the ontology functionality is necessarily preserved. Each function is characterized by the input (1) and output arguments (2). Correspondingly, each action SS (3) is characterized by a function of the work with the ontology, the values of its arguments when you call and the completion of the system.

$$f_{ss} \rightarrow FA_k^{IN} = (fa_1^{IN}, \dots, fa_k^{IN}); k = 1, \dots, K. \quad (1)$$

$$f_{ss} \rightarrow FA_n^{OUT} = (fa_1^{OUT}, \dots, fa_K^{OUT}); k = 1, \dots, K. \quad (2)$$

$$a_{ij} = \{f_k, FAval_i^{IN}, FAval_j^{OUT}\}; i = 1, \dots, K; j = 1, \dots, K; f_k \in F. \quad (3)$$

In the DKBS, the communication and collaboration mechanism between services is a problem that must be well considered in the design of the SS structure. The communication and collaboration mechanism covers the message transfer between services, information exchange between SS and distributed knowledge bases, and the SS behaviors in the entire process from request submission, request handle, to result return. These procedures can be summarized as follows:

- SS deploys appropriate services on its local Knowledge Base, performs the registration on the knowledge bases;
- The retrieval request accesses the system periodically. SS gets retrieval requests and searches its Knowledge Base.
- SS returns the retrieval result to the system after the search finishes.

- The SS implies a classification and a sort order to the retrieval result according to its corresponding retrieval request.
- User views the retrieval results on the interface and selects the needed Knowledge. With the information about the Knowledge Base embedded in the retrieval result, the user then contacts the relevant enterprise to obtain the Knowledge.

The principle of operation of the SS is illustrated in Fig. 3. The knowledge base is formed from three types of ontologies. The rules for recognizing performance problems are converted into ontology and the system knowledge base is placed. The working memory of a SS contains facts that correspond to simple events, composite events, and identified problems and recommendations. The algorithm for using SS to perform an analysis is as follows. The initial data is the sequence of events of the system, let's call it the trace. Trace contains simple events. *E* The algorithm processes the events of the route in the order of the time of their occurrence and performs the following cycle:

1. Read the next event from the trace *E*.
2. Convert it to the concept of ontology and add it to the working memory of the SS.
3. Start the SS. If the new event is a search query, then pick up the relevant answers in the knowledge base. For an indefinite concept of a simple event, the output engine consistently executes the rules for constructing composite events, for constructed fact-composite events, the rules for identifying performance problems. The result of triggering a SS can be:
 - (a) adding a new fact-simple event to some of the constructed fact-composite events;
 - (b) identification of fact-composite events of working memory for compliance or non-compliance with the performance problem;
 - (c) delete or save a new event in the working memory;
 - (d) search for a saved event.
4. Retrieve the search results from the working memory or offer recommendations to solve the problem.

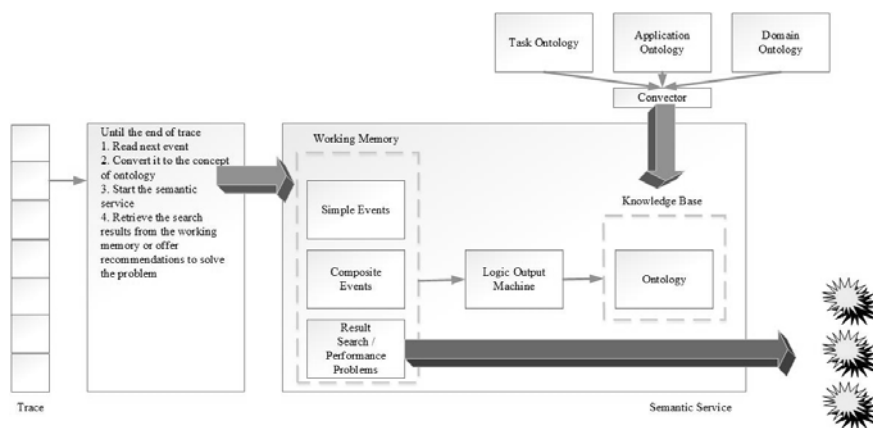


Fig. 3. Structure of Semantic Service.

To assess the performance of the SS attached, create an information system of two components:

- event trace generator DKBS;
- event trace analyzer for SS.

Events are recorded in the track when calling action functions. A simple event is represented as:

$$e = \langle f, et, EPval, t, d, cs \rangle, \quad (4)$$

where $f \in F$ - action function;

et - the type of event that sets its parameters $et \Rightarrow EP = (ep_1, \dots, ep_K)$;

$EPval = (v_1, \dots, v_K)$ - event parameter values;

t - time of the event;

d - event duration;

cs - service code.

The implementation scheme of the tracer generator is presented in fig. 4. The tracer generator consists of the following components:

- *OntologyGenerator* program converts messages into ontology concepts;
- *library SM* is designed to track messaging between services;
- *library Timer* provides functions for obtaining current system time with high accuracy;
- *library TracerWriter* saves simple events created by the tracer to a trace file.

Presumably all generator modules are implemented in the Java programming language.

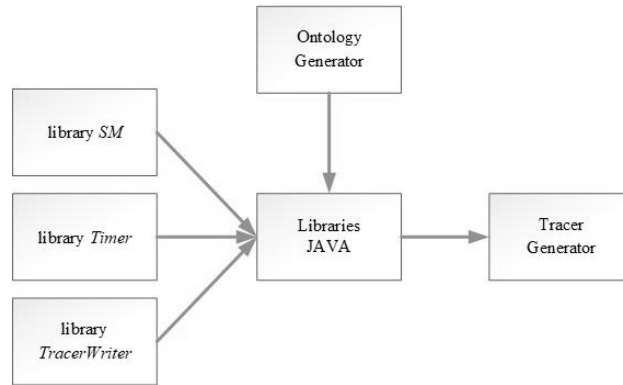


Fig. 4. The tracer generator.

The implementation scheme of the SS trace analyzer is shown in fig. 5. The trace analyzer consists of the following components:

- *library UI* implements the CC user interface.
- *library TraceReader* provides reading simple events from trace files.
- *library Ontology* implements event conversion to ontology concepts.
- *library Query* generates queries to the SS.

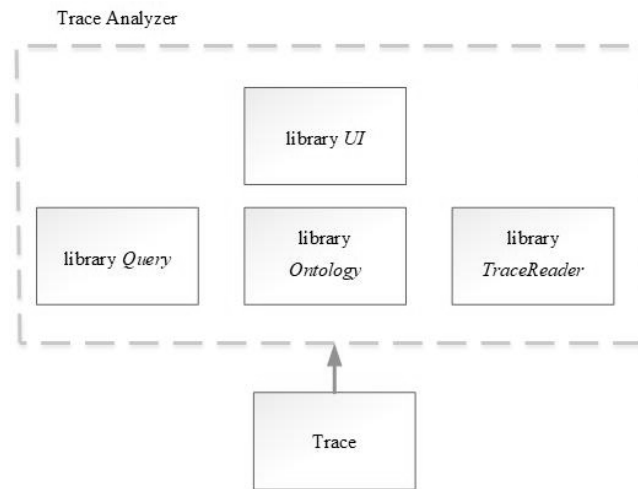


Fig. 5. The trace analyzer.

5 Conclusion

Modern representations of data are changing the ways and forms of communication, production and consumption of information. The dominance of horizontal relations, structure-role of information decentralization of all types of data available at any time on every device. The user should not care about the specific technology used to provide computing capacity or data storage, so you can say that a user has some information about the remote resource. DKBS in order to study the data, their processing and use evolving technologies of the semantic web. The formal model of the DKBS consists of a set of ontologies; lots of services; a set of events that describe the processes of the system; SS, a set of composite services and knowledge base. The completeness and effectiveness of the system is determined by a multitude of ontologies. The main difference in DKBS is the use of the service approach and ontologies in knowledge engineering. SS acts as a software agent that regulates the interaction of all components of the system. SS provides: semantic principle, which defines a formal description of information and allows you to define the following characteristics of services: scalability, semantic interoperability, formal models of services and ontologies; the principle of decision making; the principle of distribution, which allows you to aggre-

gate the capabilities of several computing objects through cooperation. The communication and collaboration mechanism of the DKBS covers the message transfer between services, information exchange between SS and distributed knowledge bases, and the SS behaviors in the entire process from request submission, request handle, to result return. The knowledge base is formed from three types of ontologies. The rules for recognizing performance problems are converted into ontology and the system knowledge base is placed. The working memory of a SS contains facts that correspond to simple events, composite events, and identified problems and recommendations. The algorithm for using SS is discussed in the article. To assess the performance of the SS attached, create an information system of two components: event trace generator and event trace analyzer.

References

1. Ruzhi Xu, Peiguang Lin, Cheng Liu: Research on Distributed Knowledge Base System Architecture for Knowledge Sharing of Virtual Organization. Atlantis Press, pp. 349-357 (2010)
2. WEB 3.0 budet pitat'sya stekom tekhnologii blokcheyn [WEB 3.0 will be powered by blockchain technology stack]. Available at: <https://101blockchains.com/ru/web-3-0/> (accessed 01.03.2019)
3. Deloitte projects by 2025 that 10% of global GDP to be built on blockchain applications. Available at: <https://www.freightwaves.com/news/2017/9/2/deloitte> (accessed 02.09.2017)
4. Tanenbaum A., Van Steen M.: Distributed systems. Pearson Prentice Hall (2007)
5. Burns B.: Designing Distributed Systems: Patterns and Paradigms for Scalable, Reliable Services. O'Reilly Media, Inc (2018)
6. Allemang D, Hendler J.: Semantic web for the working ontologist modeling in RDF, RDFS and OWL. Elsevier Inc. (2008)
7. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall, Englewood Cliffs (1995)
8. Genesereth, M.R., Ketchpel, S.P.: Software Agents. Communications of the ACM 37(7): 48–53 (1994)
9. Wooldridge, M.J., Jennings, N.R.: Intelligent Agents: Theory and Practice. Knowledge Engineering Review 10(2): 115–152 (1995)
10. Stoian, G., Popirlan, C.I.: A proposal for an enhanced mobile agent architecture (EMA). In: Annals of the University of Craiova, Mathematics and Computer Science Series, vol. 37(1), pp. 71–79 (2010)
11. Tandareanu, N., Popirlan, C.I.: A Mobile Agents Approach for Knowledge Bases Processing. In: Samarati, P., Yung, M., Martinelli, F., Ardagna, C.A. (eds.) ISC 2009. LNCS, vol. 5735, pp. 27–32. Springer, Heidelberg (2009)
12. Popirlan, C.I.: Knowledge Processing in Contact Centers using a Multi-Agent Architecture. WSEAS Transactions On Computers 9(11), 1318–1327 (2010)
13. Popirlan, C.I., Stefanescu, L.: Mobile Agents System for Intelligent Data Analysis. In: Proceedings of WSEAS Applied Computing Conference (ACC 2009), Athens, Greece, pp. 663–668 (2009)
14. Popirlan, C.I., Dupac, M.: An Optimal Path Algorithm for Autonomous Searching Robots. In: Annals of University of Craiova, Mathematics and Computer Science Series, vol. 36(1), pp. 37–48 (2009)

15. Sycara, K.P.: Multiagent Systems. *AI Magazine*, 79–92 (1998)
16. Wooldridge, M.: Introduction to Multiagent Systems. John Wiley and Sons, UK (2002)
17. Kabbaj, A.: Development of Intelligent Systems and Multi-Agents Systems with Amine Platform. In: Schärfe, H., Hitzler, P., Øhrstrøm, P. (eds.) ICCS 2006. LNCS (LNAI), vol. 4068, pp. 286–299. Springer, Heidelberg (2006)
18. Balachandran, B., Enkhsaikhan, M.: Developing Multi-agent E-Commerce Applications with JADE. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part III. LNCS (LNAI), vol. 4694, pp. 941–949. Springer, Heidelberg (2007)
19. Rodriguez, S., Fernandez, V., Julin, V., Corchado, J.M., Ossowski, S., Botti, V.: A THOMAS Based Multi-Agent System for Recommendations and Guidance in Malls. *Journal of Physical Agents* 3(2), 21–26 (2009)
20. Hagras, H., Callaghan, V., Colley, M.: Intelligent Embedded Agents. *Journal of Information Sciences*, 289–292 (2005)
21. Bouchachia, A., Prosssegger, M.: A Bi-Clustering Agent-based Approach for Map Segmentation. In: Proceedings of the IEEE Symposium on Intelligent Agents, IA 2009, Part of the IEEE Symposium Series on Computational Intelligence, pp. 99–105 (2009)
22. T. Klapiscak and R. Bordini. JASDL: A Practical Programming Approach Combining Agent and Semantic Web Technologies. In Proc. of DALT, pp 91–110, (2009)
23. M. Hadzic, E. Chang, and P. Wongthongtham: Ontology-Based Multi-Agent Systems. Springer (2009)
24. OWL-DL as a FIPA-ACL content language. In Proc. of FOCA 2006, (2006).
25. Y. Zou, T. Finin, Y. Peng, A. Joshi, and R. S. Cost: Agent Communication in DAML World. In Proc. of WRAC 2002, pp. 347–354 (2002)
26. D. Okouya, L. Penserini, S. Saudrais, A. Staikopoulos, V. Dignum, and S. Clarke: Designing MAS Organisation through an Integrated MDA/Ontology Approach. In Proc. of TWOMDE: pp. 55–60 (2008)
27. N. Fornara and M. Colombetti: Representation and monitoring of commitments and norms using OWL. *AI Comm.*, 23(4):341–356 (2010)
28. L. Nardin, A. Brandão, and J. Sichman: Experiments on semantic interoperability of agent reputation models using the SOARI architecture. *Eng. Appl. of AI*, 24(8), pp. 1461–1471 (2011)
29. Dzharratano Dzh.: Ekspertnyye sistemy: printsipy razrabotki i programmirovaniye: per. s angl. K. A. Ptitsyna. – 4-ye izd. [Expert systems: principles of development and programming]. Moskva, Vil'yams (2007)
30. Smolin D. V.: Vvedeniye v iskusstvennyy intellekt: konspekt lektsiy [Introduction to Artificial Intelligence]. Moskva (2004)
31. Behnam Azvine, Nader Azarmi, Detlef D. Nauck: Intelligent Systems and Soft Computing: Prospects, Tools and Applications. Springer (2006)
32. Gruber Thomas R.: A translation approach to portable ontology specifications. Appeared in *Knowledge Acquisition*, vol. 5(2), pp. 199–220 (1993)
33. Yu Liyang: Introduction to Semantic Web and Semantic Web services, CRC Press (2007)
34. John Davies, Rudi Studer, Paul Warren: Semantic Web technologies: trends and research in ontology-based systems. John Wiley & Sons (2006)

Collective Based on EC-GDSA Digital Signature Protocol to Protect the Doctors' Medical Conclusion of the Consilium

Hanna Nelasa^{1[0000-0002-3708-0089]}

¹Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine

nelasa@zntu.edu.ua

Abstract. Modification digital signature protocol EC-GDSA on elliptic curves to collective form for telemedicine is proposed. This form may be used for security authentication of general conclusion of the doctors' concilium. Computational algorithm is given.

Keywords: telemedicine, cryptography, collective digital signature protocol, elliptic curve, EC-GDSA

1 Introduction

The fundamental task of each state is to ensure the right of its citizen to health and safety. There is a high concentration of technical and human resources in large cities and metropolitan areas in Ukraine. In such a situation, qualitative, timely and qualified medical care is not available to the majority of the population living in rural areas, what leads to a deterioration of the health of the population and average life expectancy decrease. The main task of modern medicine is to raise the quality of life of the population, increase the level of availability of medical services at the expense of the latest medical technologies. Increasing demands of Ukrainian people regarding the quality of medical goods and services compel the society look for new ways and forms for its satisfaction, or at least to improve existing ones. One such way is telemedicine.

Telemedicine is a complex of actions, technologies and measures used during the medical care using the means of remote communication for the exchange of information in electronic form. Medical assistance with application of telemedicine involves the opportunity of providing the patient with medical services for counseling, diagnosis, treatment using means of remote communication in the form of information exchange in electronic form, including by sending electronic messages, holding video-conferences [1]. Telemedicine services have already been developed and popularized abroad.

Telemedicine services are very comfortable and the most effective in remote areas. On-line consultation with physicians is much cheaper than a trip to a doctor and allows one to get a consulate faster. Implementation of telemedicine technologies will

also reduce the number of physical visits to the medical institution by patients, as well as early discharge from the hospital and further undergo certain procedures and monitoring of a health status at home.

Conditionally, telemedicine can be divided into three general directions related to medical education, the interaction “doctor-to-doctor” and the interaction “doctor-to-patient”. An important aspect here is the issue of transmission, processing, analysis, storage, protection of medical information and safe identification of the users of telemedicine services. So, telemedicine is a direction of medicine simultaneously combines telecommunications, information technologies and medical education.

In order to provide telemedicine services, it is necessary to comply with the requirements of the Law of Ukraine [2], the Order of the Ministry of Health of Ukraine [3] and the requirements of the current legislation on information security.

Confidential medical information is transmitted through computers and open networks, and requires the protection and authentication.

The urgency of the issues of protection of telemedicine information and ensuring the proper level of cryptographic security is undoubted. This is due to the fact that the latest medical technologies and measures, including telemedicine counseling, are being implemented and used in the practice of healthcare institutions. The provision of telemedicine services should ensure the preservation of personal, medical and other secrets provided by the legislation of Ukraine, as well as the confidentiality of personal data. Particularly, software used for telemedicine counseling should provide the appropriate level of protection of information and its authenticity by using an electronic digital signature.

Mandatory services to be provided to clients and owners are services such as integrity, integrity, authenticity, availability, confidentiality and reliability. The quality of providing the specified services and the level of guarantees to a large extent is determined by the methods, mechanisms and protocols of cryptographic protection of information. In order to provide the specified services with the necessary level of guarantees, it is necessary to use asymmetric and symmetric cryptographic transformations as well as cryptographic authentication protocols and the establishment of keys based on them.

Immediate implementation of medical information systems (telemedicine, electronic medical cards, "mobile medicine", cloud technologies, etc.) to the practice of the physicians will lead to increase in the competitiveness of domestic health facilities and their financial and economic security.

Directions of activity of large medical clinics and centers of private and state ownership concerning telemedicine are connected with “virtual clinics”, telemedicine consultations, “another thought” service and a consultation of the physicians.

2 Formal problem statement

The following main types of telecommunications are used for the transmission of medical data: stationary and mobile telephony; Internet (IP - protocol); GSM and CDMA networks (SMS). Individual telemedicine is often used for mobile communi-

cation and wireless Internet (GPRS, CDMA, 3G, 4G). The protection of transmitted medical information in home telemedicine systems is carried out with the help of various software and hardware, primarily cryptographic, by encryption.

The development and convenient implementation of digital signature cryptographic protocols based on elliptic curves can be a tool allows medical personnel transmit and receive relevant messages and medical information with adequate cryptographic stability and the required speed when telemedicine tools are used.

Concilium is a meeting of several doctors of one or different specialties. Concilium may be necessary to establish the state of health of the examinee, diagnosis, determination of medical forecast, tactics of further examination and treatment, expediency referral to a specialized department or other, profile, medical institution.

The key difference between the doctors' congress is the collective responsibility for the content of the medical conclusion. To ensure the authenticity of such a document in the telemedicine system, authors are encouraged to use the collective digital signature protocols. In contrast to multiple signatures [4], where signing and checking of an electronic document is important for the sequence of signatures / signatures of each participant and, in addition, the size of the signature increases proportionally to the number of participants.

For the time being in Ukraine, with the aim to ensure the functioning of telemedicine networks, it is planned to use the specialized software, which corresponds to: the DICOM standard. One of the advantages of the DICOM standard is the integration of visual data with medical data in a single file, which is why the formation of a DICOM file/package is crucial. A separate DICOM file includes, in addition to the title and image (which may contain three-dimensional information) a set of graphic and text data (relative to a patient).

The high speed of IT development forces government authorities to learn how to use them to improve the efficiency of work and be prepared for rapid change. Germany is one of the leading countries in the field of Internet development in applying cryptographic primitives and protocols and is guided by the US and EU standardization system. In this regard, it is important for Ukraine to study the German experience and determine the possibilities for its application in terms of cryptographic protection of information in general and in particular in the medical sphere.

The goal of the work is modifying the EC-GDSA electronic digital signature protocol into a collective form for the purpose of cooperation with clinics in Germany and other countries to create telemedicine services for the physician consultation and "second opinion" services, as well as its software implementation in the form of client-server application.

3 Literature review

Theoretical and theoretical foundations concerning the functioning of telemedicine systems, as well as information security in computer networks and systems, are considered in the works of domestic scientists, Godlevsky L.S. [5], Kovalenko O.S.[6], Lyakh Yu.E. [7], Martsenyuk V.P. [8, 9], Mayorov O.Yu. [10], Minzer O.P. [11], Khimzon I.I. [12]. Questions about the approaches to protecting data protection in

medical systems using cryptosystems, and increasing the reliability of cryptographic keys are also considered by such foreign experts as Yudin O.K. [13], Venkatasubramanian K.K. [14], Cherukuri S. [15] and many others.

There is a problem of collective responsibility for the contents of a document (in this context, a medical conclusion) in telemedicine systems. From an IT point of view, it can be solved with a multiple signature [4]. But in this case the order of signers and verifiers is important. And the size of the signature increases proportionally to the number of signers.

To eliminate these shortcomings, a new approach to the formation of a collective public key based on users' public keys was proposed [16]. Based on this approach, collective digital signature protocols are developed, including those based on modern electronic digital signature standards in different countries [17-21]. These protocols use a common (collective) public key, which is formed on the basis of the individual public keys of the group of users. The practice of electronic digital signature systems enables the use of standard open source directories and / or standard public key certificates (Internet accessibility), which facilitates the practical application of a new approach to generating a collective electronic digital signature.

At this time, cryptography on elliptical curves [22-24], the principles of which were independently proposed by N. Koblicz and V. Miller, have been widely used to offer open-source cryptographic protection systems. In Ukraine, the research of electronic digital signature on elliptical curves is also engaged in a large number of specialists, including I.D. Gorbenko, M.M. Savchuk, G.Z. Khalimov V.I. Dolgov, V.K. Zadiraka, O.O. Kuznetsov, A.V. Bessalov and others.

4 Threat analysis

Telemedicine systems, as well as other information and telecommunication systems have problems with the security of information. This is tied to the fact that telemedicine and telecommunication services are perceived separately from information security services, but should be considered complexly.

To a greater extent teleconsultation is carried out between doctors, and much more rare between a doctor and a patient. In the first case of telemedicine consultation, the opportunity of mutual exchange of electronic messages with files of various formats is very important (results of diagnostic examination, results of analyzes conducted, etc.). In this case, the most practical in terms of efficiency and economic feasibility is the sharing of files that are confidential, so the necessary protection measures should be applied to this information. In telemedicine systems, which include local networks (central subsystem with workstations, servers and data warehouses), to remotely interact with its users, remote connections of the global network are used.

Separate clinical and informational security when providing telemedicine services. Clinical security in telemedicine is related to telemedicine systems, to what extent they affect and to the extent that they are safe for the life and health of people involved in the provision of telemedicine services (patient, doctor, engineer, etc.). Information security is about medical and personal secrets, it is the security of network

infrastructure, as well as information and telemedicine systems from interference (intentional or accidental, internal or external), theft of medical or personal information, blocking processes that impede the work of users.

Loss and falsification of medical data may occur as a result of the negative impact of software, computer viruses, improper performance and equipment failure, negligent or intentional user actions. The threat and disadvantage of access to telemedicine systems are DoS attacks directed at intermediate and end nodes, as well as failures in equipment, damage to communication channels, substitution of an authorized object, listening of communication channels; loss of data (in full or in part), systematic falsification that is typical of most remote access systems.

Information security is provided through the use of the following telecommunication and hardware-software solutions: corporate medical networks, closed channels (Virtual Private Network); cryptographic protection and electronic digital signature; spam and antivirus protection; application of authorized access to servers, separate databases, workstations, etc.; transmission of information in anonymous form; the use of international standards (DICOM, SCP-ECG, etc.) in the transmission of information.

Security policy should be formed depending on the particular situation and tasks, in order to protect the personal data of a patient, it is proposed to use a variety of cryptographic algorithms, to use electronic digital signature for the transmission of medical information.

Regarding the ways of organizing telemedicine services within the hospital itself, it is possible to carry out diagnostic manipulations (e.g., ECG) by middle medical personnel, and the reading of the diagrams and the diagnosis are performed by qualified doctors in the functional diagnostic department. In case of the need for telemedicine counseling involving simultaneously two or more doctors-consultants, a telemedicine consultation (concilium) is carried out.

5 The modified method of digital signature based on EC-GDSA

In the event that a doctor encounters a complicated case in the process of treating patients in his medical practice, he usually needs additional advice (advice) from another specialist (the "another opinion" service) or several colleagues (a physician's consultation) at his or her medical institution, or else, within the city, country or abroad. In those cases where distance is a critical factor, remote counseling is provided through telemedicine.

Germany is one of the most developed states of the European Union with a high level of development of the market of telemedicine services, which is potentially the most ready to introduce new technologies.

Let's consider the national standard of electronic digital signature of Germany EC-GDSA [table 1], and we modify it to a collective form for the purpose of organizing the interaction of several doctors in consulting patients.

Let's illustrate the diagram of the deployment of the software complex (Figure 1), which ensures the authenticity, confidentiality and integrity of the medical conclusion when carrying out the consultation of physicians.

Table 1. - The EC-GDSA Digital Signature Standard

Formation of a digital signature	Verification of digital signature
Input: secret key d , public key $Q = d^{-1} \times P$, general system parameters. Output: digital signature $\langle r, s \rangle$ for message M .	Input: public key Q , general system parameters, digital signature $\langle r', s' \rangle$ for message M' . Output: The signature is valid or not.
1. $h = H(M) \bmod n$; 2. Choose random $k \in \{1, \dots, n-1\}$; 3. $k \times P = (x, y)$; 4. $r = \pi(x, y) \bmod n$; 5. $s = (kr - h)d \bmod n$.	1. $h = H(M) \bmod n$; 2. $w = (r')^{-1} \bmod n$; 3. $u_1 = h'w \bmod n$; 4. $u_2 = s'w \bmod n$; 5. $(x, y) = u_1 \times P + u_2 \times Q$; 6. $v = \pi(x, y)$; 7. $r' = v$

The system consists of:

- *Server* - leading doctor;
- *Client* - doctors - consultant from 1 to l .

Diagnostic equipment is directly connected to the host's doctor's server.

Consider the modification of the current EC-GDSA Digital Signature Standard for the implementation of a collective signature, which may be useful for ensuring the safe authentication of a general medical conclusion when the consultation is carrying out.

The protocol consists of three parts: the generation of a public key, the formation of a collective signature, the verification of the collective signature.

The designation:

P - the base point of the elliptic curve (the general parameter of the cryptosystem);

l - number of signatories (consultants);

n - the order of the cyclic subgroup of the elliptic curve points;

d_i - the secret key of the i -th signer (doctor-consultant);

h - the hash of the document (medical conclusion);

Q - general public key (point of the elliptic curve);

Q_i - share of the public key of the i -th doctor-consultant (EC point).

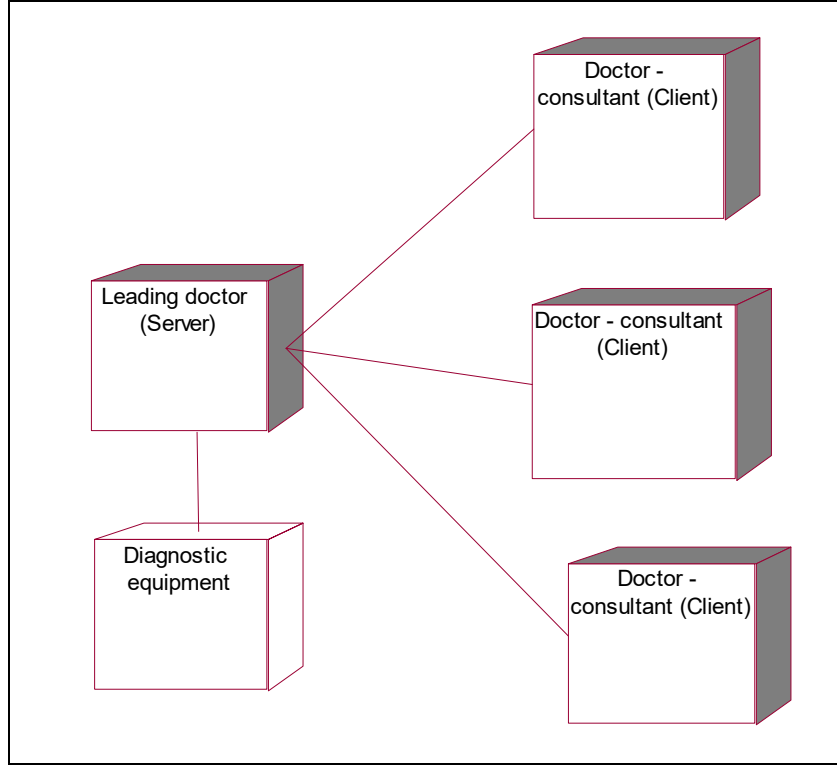


Fig. 1. - Diagram of the deployment of a software for concilium

Part 1. Generation of a public open collective key.

1. The leading doctor (Server) initiates the calculation of the general public key (point Q). After that, each doctor-consultant (Client) calculates the share of the public key, which is then transferred to the leading doctor.

$$Q_i = d_i^{-1}P.$$

2. The leading doctor (Server) calculates the total secret key Q , as the sum of the shares of the public keys of the group of l signers (consultants of the concilium).

$$Q = \sum_{i=1}^l Q_i = \sum_{i=1}^l d_i^{-1}P.$$

Part 2. Formation of a collective signature.

1. A leading physician (Server) calculates a hash of a document (medical conclusion) and transmits it to all doctors-consultants (Client).

2. Each i -th subscriber (doctor - consultant, Client) calculates a point as follows:

a) selects a random parameter (random variable) k_i , $1 < k_i < n$;

b) calculates the point $R_i = k_i P$

c) sends it to the *Server*.

3. The leading doctor (*Server*) calculates the total point R as the sum of R_i :

$$R = \sum_{i=1}^l R_i = (X_R, Y_R)$$

4. The leading doctor (*Server*) calculates the first part of the collective signature

$$r = \pi(R) = X_R \bmod n$$

and sends it to all signers.

5. Each subscriber (doctor-consultant, *Client*) calculates its part S_i :

$$s_i = (k_i r - h) d_i \bmod n$$

using its private key d_i , and sends it to *Server*

6. The leading doctor (*Server*) calculates the second part of the collective signature

$$s = \sum_{i=1}^l s_i .$$

7. Collective signature is a pair of numbers – (r,s).

Part 3. Checking the collective signature.

1. The verifier calculates the hash of the document (medical conclusion).

2. The verifier calculates the values:

$$w = (r')^{-1} \bmod n ,$$

$$u_1 = h' w \bmod n ,$$

$$u_2 = s' w \bmod n .$$

3. and, using an open collective key Q , forms a point

$$(x, y) = u_1 P + u_2 Q$$

and calculates the value

$$v = \pi(x, y) = x \bmod n .$$

5. If $r' = v$, then the signature is authentic, in the other part it is forgery.

Justification of the correctness of the submitted protocol.

Since the value of r is determined by the formula $r = \pi(R) = \pi(\sum_{i=1}^l k_i P)$ when forming a signature, and when checking the signature verifier $u_1 P + u_2 Q$ is the point R , then

$$\begin{aligned} u_1 P + u_2 Q &= h' w P + s' w Q = h' w P + s' w \sum_{i=1}^l d_i^{-1} P = w(h' + s' \sum_{i=1}^l d_i^{-1}) P = \\ r^{-1}(h' + \sum_{i=1}^l [(k_i r - h) d_i] \sum_{i=1}^l d_i^{-1}) P &= r^{-1}(h + \sum_{i=1}^l k_i r - h) P = \sum_{i=1}^l k_i P = R \end{aligned}$$

as a result we have:

$$v = \pi(u_1 P + u_2 Q) = \pi(R),$$

which corresponds to the value r when the signature is formed.

Note that the Server function can act as a separate individual of the protocol, and one of the consultants. In this case, it combines Server and Client functions, which does not interfere with the correct operation of the protocol. If the number of participants equals two, we have the implementation of the "another opinion" service, which from the point of view of technical implementation is a special case of a concilium.

6 Conclusion

Thus, modern economic, scientific and social changes, rapid development and introduction of innovative technologies, and a significant increase in data exchange information processes require timely response to changes in the medical sphere, and oblige comprehensive protection and confidentiality of medical data in modern telemedicine systems.

The authors carried out an analysis of threats and modern methods of protecting confidential medical information. The analysis of possible variants of confidentiality and identification of the authorship of electronic communications for the provision of telemedicine services, as well as analysis of critical, in terms of information security, elements and communication channels infrastructure of telemedicine systems.

It is established that in the context of the reform of the health care of Ukraine, a significant reduction of the bed fund, important issues are the improvement of medical diagnostic processes, the implementation of information and telecommunication technologies, the organization of interaction between medical, preventive and specialized health care institutions through telemedicine, remote provision of qualified medical services, maximum utilization of the intellectual potential of the best clinics.

In the case of the need for telemedicine counseling with the participation of several consultants, a telemedicine consultation is conducted. It has been established that some issues, including those concerning the practice of the use of a collective digital signature in domestic health care institutions, for example, for the consultation of physicians, remain unresolved. Germany is one of the most developed states of the

European Union with a high level of telemedicine market development, potentially the most ready to introduce new technologies. Therefore, the possibility of modifying the current EC-GDSA electronic digital signature standard for Germany for the implementation of a collective signature is considered, which may be useful for ensuring the safe authentication of a general medical conclusion when physician consultation is carrying out.

The development of software for the proper level of cryptographic protection of information in telemedicine, the implementation of the EC-GDSA collective digital signature protocol, including to solve the problem of conducting a consultation of doctors, can also be used for cooperation with medical institutions in Germany and other European countries and the world to create a telemedicine service "second thought". For this purpose, the analysis and research of modern digital signature standards based on cryptographic transformations in the groups of points of elliptic curves has been carried. The proposed collective digital signature protocol is based on the EC-GDSA standard using a method similar to another one used in the collective signing protocol DSTU4145.

A software package with a client-server architecture is developed. It allows to ensure the preservation of the authenticity, integrity of the non-contradiction of electronic medical documents by protecting them with collective digital signature methods, based on the standards DSTU4145, EC-GDSA. To develop the project, the C++ programming language and the Miracle cryptographic library were selected, which includes the functions of long arithmetic and the implementation of the arithmetic of elliptic curves defined over the simple and extended finite fields of Galois.

References

1. Vladzimirskiy A.V., Telemedicine: Curatio Sine Tempora et Distantia. Digital Printing, Moscow, 663p. (2016)
2. Information Law System "League of Law": Law of Ukraine 18.09.2017 No. 7117 "On Increasing the Availability and Quality of Medical Services in Rural Areas". http://search.ligazakon.ua/l_doc2.nsf/link1/JH5HP00V.html (2017)
3. Ukraine. Ministry of Health. On Approval of Regulatory Documents on the Application of Telemedicine in the Healthcare: Order dated 19.10.2015, No. 681., A Collection of Regulatory and Policy Documents on Public Health 1, pp. 12-26 (2016)
4. Min-Shiang Hawng, Cheng-Chi Le. Research issues and challenges for multiple digital signature, Int. J. of Network Security, vol.1 (1), pp.1-7 (2005).
5. Godlevsky L.S., Dets V.V., Stepanenko K.I. [and others]: Perspectives for the introduction of telemedicine technologies for the elderly in the Odessa region. Biophysical standards and information technologies in medicine. Proceedings of sciences conf. – Odessa, Astroprint, pp. 48-52 (2004)
6. Kovalenko O.S., Buryak V.I.: Standardization of medical information systems with taking into account pan-European integration. Clinical informatics and telemedicine (1), pp. 35-41 (2004)
7. Lyakh Yu.E., Vladzimirskiy A.V.: Introduction to telemedicine. Essays on biological and medical informatics, Donetsk: Lebed, 101 p. (1999)
8. Martsenyuk V. P. Development and application of information system of recording (self-recording) patients for consultation of university clinics specialists / V. P. Martsenyuk, P.

- R. Selsky, A. V. Semenets // Ukrainian journal of telemedicine and medical telematics. – 2013. – Vol. 11, No. 2. – P. 173–178.
9. Martsenyuk V. P. Performance of telemedicine technology usage for increasing quality of treatment and diagnostical work of primary care / V. P. Martsenyuk, P. R. Selsky // Actual problems of pharmaceutical and medical science and practice. – 2013. – Vol. 12, No. 3. – P. 53–54.
 10. Mayorov O. Yu., Bilov L.B., Nezhenytsky C.A.: Information systems of public health services (hospital information systems) - a tribute to the fashion or necessity (feasibility study of the introduction of the program complex "C-Hospital®"). Clinical informatics and telemedicine, vol.1, (1), pp. 1-12 (2004)
 11. Mintser O.P., Bolgov M. Yu.: Information mapping of medical diagnostic process at the level of data logic. Ukrainian Journal of Telemedicine and Medical Telematics, vol. 5 (2), pp. 128-138 (2007)
 12. Khimzon I.I.: Development and research of the efficiency of new information technologies for the management, processing and accounting of medical documentations in the conditions of the hospital department: abstract for the sciences degree of dr. Tech. Sciences: special 05.13.02 "Mathematical modeling in scientific researches". National Academy of Sciences of Ukraine, Institute of Cybernetics V.M. Glushkov, Kyiv, 30 p. (1995).
 13. Judin O.K., Korchenko O.G., Konakhovych G.F.: Protection of information in data transmission network, Kyiv: "NVP" INTERSERVICE ", 716 p (2009)
 14. Venkatasubramanian K. K., Banerjee A., Gupta S. K. S. PSKA: Usable and secure key agreement scheme for body area networks // IEEE Transactions on Information Technology in Biomedicine. 2010. Vol. 14, No 1. Pp. 60-68.
 15. Cherukuri S., Venkatasubramanian K., Gupta S. K. S. BioSec: A Biometric Based Approach for Securing Communication in Wireless Networks of Biosensors Implanted in the Human Body // Proceedings of Workshop on Wireless Security and Privacy. 2003. Pp. 432-439.
 16. Moldovian D.H., Moldovyan N.A. Two-key cryptosystems with a new mechanism of digital signature generation // Information Security Management. 2006. Vol. 10. No. 3. pp. 307–312.
 17. Hortinskaya, L.V., Moldovian, N.A., Kozina, G.L.: Implementation of protocols of collective signatures on the basis of GOST 34.310-95 and DSTU 4145-2002 standards. Legal, normative and metrological provision of the information security system in Ukraine. Kyiv, NTUU "KPI", No. 1., pp. 82-86 (2008)
 18. Nelasa H.V., Kozina G.L., Moldovian N.A.: Collective digital signature protocols on elliptic and hyperelliptic curves. Radio Electronics, Computer Science, Control. 1(19) Zaporizhzhia, ZNTU, pp. 127-133 (2008)
 19. Kozina, G.L., Moldovian, M.A., Nelasa, H.V.: Crypto Protocols: Digital Signature Schemes. Zaporizhzhia, ZNTU, 158 p. (2014)
 20. Shovgenyuk R.V., Software of cryptographic protection in telemedicine. Master's thesis, supervisor associate professor Nelasa H.V., ZNTU, 224 p. (2017)
 21. Nelasa H., Shovhenyuk R., Korolkova O.: Collective digital signature EC-GDSA based protocol for telemedicine. II All-Ukrainian Scientific and Practical Conference "Prospective Directions of Modern Electronics, Information and Computer Systems" (MEICS-2017) Dnipro, 22 November 24, pp. 123-124 (2017)
 22. Koblitz N. A Course in Number Theory and Cryptography. Berlin, Heidelberg, New York: Springer, 1994.
 23. Malhotra K., Gardner S., Patz R. Implementation of elliptic-curve cryptography on mobile healthcare devices // Networking, Sensing and Control. 2007. Pp. 239-244.
 24. Liu A., Ning P. Tinyecc: A configurable library for elliptic curve cryptography in wireless sensor networks // Information Processing in Sensor Networks. 2008. Pp. 245-256.

Investigation of the Longitudinal Track Profile Influence on the Forces Acting in the Train Inter-car Connections Using the MSC.ADAMS Software

Alexandr Shimanovsky¹[0000-0001-8550-1725], Pavel Sakharau²[0000-0003-0356-7750],

¹Belarusian State University of Transport, 34 Kirova St., Gomel, 246653, Belarus
tm.belsut@gmail.com

²Belarusian State University of Transport, 34 Kirova St., Gomel, 246653, Belarus
sahpa@mail.ru

Abstract. The simulation of a freight train movement in the MSC.ADAMS software package is considered. There are presented the investigation results for longitudinal forces in the inter-car connections of homogeneous and heterogeneous trains while moving through parts of the track longitudinal profile. It was established the effect of the train length and mass as well as the loaded and empty cars location along the train on the values of the coupling devices inner longitudinal forces.

Keywords: freight train, modeling of the solid body, longitudinal forces, track profile grade change.

1 Introduction

The electric traction introduction is one of the most important measures to reduce the costs of train traction. The efficiency of modern electric locomotives operation depends on the full usage of their power. Such locomotives show real energy savings compared to electric locomotives of previous generation only at high-mass trains transporting. The maintenance of long-compounded trains is associated with difficulties in the train controlling on way profile sections with grade changes. The insufficient knowledge of the heterogeneous train dynamics at locomotive electric braking has led to the need of its limited usage on the Belarusian railway. This fact caused the additional wear of the brake pads and wheel sets of cars due to the need of wagon brakes operation. Also it is missed the possibility of electrical energy returning to the contact network at regenerative braking. Therefore, it became necessary to analyze the longitudinal forces arising in the train moving along the grade changes of the longitudinal track profile in the locomotive electrodynamic braking mode.

2 Literature review

A detailed description of train motion mathematical models as a mechanical system is described in [1]. The main calculation schemes of the train cars are considered: cars with rigidly attached cargoes and cars with goods moving relatively to wagon body, including reservoirs with liquid. The presented train movement calculation schemes and equations are given for a solid rod or chain of solid bodies. Particular attention is paid to the analysis of the inhomogeneous train movement through the longitudinal profile grade changes with and without gaps in the harness devices. On the base of the research, it was concluded that the longitudinal forces in the train at impacts cause intense vertical oscillations of the car body, but their influence is insignificant on the longitudinal forces value. Therefore, the vertical oscillations of the cars can be neglected at analyzing the longitudinal forces in the train.

The study of a heterogeneous train motion along a complex track profile was considered in the work of Grebenyuk [2]. It is pointed that the heterogeneity of the train cars according to their types leads to an increase in longitudinal forces at braking by about 20 %, and their loading inhomogeneity – by 30-35 %. Placing one group of lighter cars into the train head increases the forces in the inter-car joints, and placing them in the train tail reduces the values compared to the forces arising in a uniform train of equal mass and length. It is noted that longitudinal forces arising in a train, moving along a complex profile, depend on a large number of factors. The most important of them are the magnitude and nature of changes in external forces (thrust, braking), the characteristics of inter-car connections and clearances in them, the position of automatic couplers at the beginning of the control actions realization.

The nature of the train motion, its velocity and acceleration substantially depend on the track profile. In the work of Vershinsky [3], there are presented the analytical calculations of longitudinal forces in harness devices of various train sections for the case of moving along a fixed track profile and along a broken profile. The calculation results show that if the profile does not change within the train length, the efforts in the inter-carriage connections are determined by the tractive forces, the distribution of masses and braking forces along its length. At the same time, when the train moves along the grade change of the track longitudinal profile, the magnitude of the greatest gradient has a great influence. The increase in the number of grade changes within the train length leads to a decrease in longitudinal forces. It was concluded that, the case of the electric braking by a locomotive moving in a recess, from descent to ascent is particularly difficult from the cars stability against squeezing point of view.

The dissertation work of Masleeva [4] considers the investigation of transient processes caused by the long train movement control on the longitudinal track profile grade changes. It has been established that the values of longitudinal forces having impact and quasistatic nature do not practically depend on the angle characterizing the asymmetry of the grade changes relative to the vertical, and they are determined only by the difference of the conjugating slopes and the radius of the connecting curve.

The study of longitudinal forces in non-stationary modes of train motion based on computer simulation was considered in [5, 6]. It is noted that the heterogeneity of the train composition according to the wagons and cargo types leads to significant

changes in the transition process nature and in the distribution of maximum efforts along the train. There is an increase in the values of maximal forces in the inter-car connections by 10–12 % compared with the intensity of forces in a homogeneous train.

Simulation of train movement across longitudinal profile grade changes was performed with the help of the Universal Mechanism software package and it was considered in [7-10]. There was investigated the possibility of using the increased longitudinal profile grade changes at the railway construction and repair to reduce financial costs. Thus, in [10], on the basis of studies of wagon vertical accelerations and forces in the inter-car connections, the necessity of using biclothoid vertical curves is proved for pairing elements of a longitudinal profile. It is shown that in this case the values of the train longitudinal forces decrease to 14 % and the vertical forces – to 45 %, compared with movement along vertical circular curves.

In work [11] the process of pneumatic braking is considered. The effect of the delay in the braking forces application along the train and the inter-car joints parameters on the longitudinal forces arising in coupling devices was investigated theoretically and experimentally. There was made the conclusion that the three waves of longitudinal elastic oscillations occur along the train. These forces cause the increase in the inter-car joints forces, they reach maximum values at the peak of the third wave in the first half of the train.

A comparison between various methods of braking calculations used in the EU and Russia is analyzed by Bureika [12]. It revealed that the differences between the values of the braked mass of Russian freight cars estimated by TSI (Europe Union) method and calculated by MPS (Russian) method for four-axle freight cars with cast-iron brake shoes are (0–2.2) %.

There can be found a great amount of articles dedicated to researches on electrodynamic braking. Thus, Pugi [13] presents a modular tool for prediction of train braking performance, with particular attention to accurate prediction of stopping distances. A comparison between braking forces in the cases of air and electrical braking is conducted in it. In article [14], there are investigated the main factors that lead to the empty wagons derailment in the long-compound trains. It is concluded that the main reasons for the cars stability loss are large longitudinal forces, deviation of automatic couplings, as well as the condition of the rolling surface and wheel flanges. There is a need to reduce the friction coefficient between the rail and the wheel flange to prevent the wheel from raking it on the rail and further derailment.

The problem of reducing of longitudinal forces between cars is considered in [15]. There are presented the main factors affecting the longitudinal forces in the automatic couplers: characteristics of amortizing devices, longitudinal vibrations of wagons, uneven height position of automatic couplers due to different loading of each car, misalignment of automatic couplers in the transverse direction. It is noted that the introduction of electropneumatic brake is most effective method to reduce the forces arising in freight trains.

To increase computational accuracy and to reduce time of computations it was proposed in [16] to use the global and local coordinate systems. This is especially important at train movements simulation on long track sections. The global coordinate

system is used to describe the locomotive motion, the local one takes into account the deviations in the inter-car connections. It is stated that the use of such an approach allows to reduce the computation time by 8.6 % and to maintain the accuracy of the calculations.

There are also developed the approaches related to the kinematic parameters determination for the several train sections in a real-time mode and their transfer of incoming data to the locomotive [17]. Analysis of the incoming information allows to make a decision about changing the control action (thrust or braking) in order to reduce the longitudinal forces in the train.

The results of the reviewed investigations confirm the scientific relevance of studies on the train dynamics at transient driving conditions. The purpose of the presented work is to analyze the longitudinal forces in the inter-car joints arising at movement along a track with variable inclination.

3 Mathematical model of train

For the mathematical description of the train movement it was taken into consideration the calculation scheme of the train considered as a chain of solid bodies with elastic-viscous inter-car connections. The pattern of the train movement over a complex profile section is shown in Figure 1. There, cars are considered to be absolutely rigid bodies with masses concentrated in their centers of mass. Vertical oscillations of wagons on springs, their angular displacements, and also gaps in the inter-car connections are not taken into account.

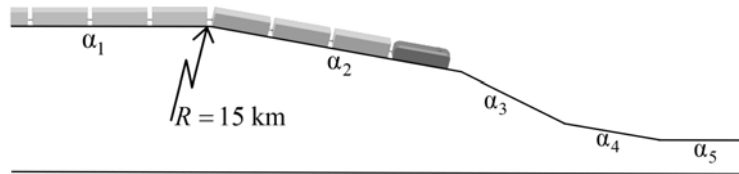


Fig. 1. Scheme of the train movement along a complex profile section

The train motion in the presented scheme is described by differential equations system [18, 19]:

$$\left. \begin{aligned} m_l \ddot{x}_l - T_1 - m_l g \sin \alpha_l + S_l + R &= 0; \\ m_k \ddot{x}_k + T_k - T_{k+1} + S_k - m_k g \sin \alpha_k &= 0; \quad k = 1, 2, \dots, n-1; \\ m_n \ddot{x}_n + T_n + S_n - m_n g \sin \alpha_n &= 0. \end{aligned} \right\}, \quad (1)$$

where n – number of the train cars; m_l, m_k – mass of the locomotive and the k -th car, respectively ($k = 1, 2, \dots, n$); $\ddot{x}_l, \ddot{x}_k$ – longitudinal acceleration of the locomotive and the k -th car; T_k – forces acting from the inter-car connections; g – acceleration of

gravity; α_l , α_k – slope of the track under the moving locomotive and the k -th car; S_l , S_k – resistance forces to locomotive and the k -th car movement; R – external force acting the locomotive (thrust or electrodynamic braking).

Forces in the inter-car coupling with elastic-viscous connections are determined in accordance with the expression

$$T_k = c(x_k - x_{k-1}) + K(\dot{x}_k - \dot{x}_{k-1}), \quad (2)$$

where c – stiffness coefficient of elastic elements in the inter-car connections; in the calculations the value is $c = 1,1 \cdot 10^7$ N/m; x_k , $\dot{x}_k$ – the movement and speed of the k -th car respectively; K – damping factor; the accepted value $K = 10^6$ N·s/m.

The resistant to movement forces depend on the type of the track and, in the case of a continuous railway are determined in accordance with the Rules [20] by the formulas:

– for the locomotive driving in traction mode

$$S_l = m_l g (1.9 + 0.008 \cdot 3.6 \dot{x}_l + 0.00025 \cdot (3.6 \dot{x}_l)^2) / 1000; \quad (3)$$

– for the locomotive driving in the idle mode

$$S_l = m_l g (1.9 + 0.008 \cdot 3.6 \dot{x}_l + 0.00025 \cdot (3.6 \dot{x}_l)^2) / 1000; \quad (4)$$

– for the empty cars ($q \leq 6$ ton/axis)

$$S_k = m_k g (1.0 + 0.042 \cdot 3.6 \dot{x}_k + 0.00016 \cdot (3.6 \dot{x}_k)^2) / 1000; \quad (5)$$

– for the loaded ones ($q > 6$ ton/axis)

$$S_k = m_k g (0.7 + (3 + 0.09 \cdot 3.6 \dot{x}_k + 0.002 \cdot (3.6 \dot{x}_k)^2) / q) / 1000; \quad (6)$$

where q – weight per one pair of wheels, ton.

4 A computer model of the train

To implement the presented mathematical description in the MSC.ADAMS software package, $n + 1$ parallelepipeds were created for the locomotive and the cars (in Figure 2 one of the model cars is shown). Their dimensions correspond to the dimensions of the existing rolling stock. They are located at a distance of the car length between the axles of automatic couplings (13.92 m) from each other. Through the curve "GCURVE_1" passes the cars centers of gravity (for the second car - "VAGON_2.cm"). It sets the parameters of the longitudinal profile of the track for the train displacement. The car is connected to the curve in two points "PTCV_5" and "PTCV_6" by the "Point-Curve Constraint" connection (the connections "PTCV_4" and "PTCV_7" belong to 1 and 3 cars, respectively). A resistant to the movement force "SFORSE_2" is applied to the car center of gravity. The force is determined by the formula (5) or (6) depending on the load of the car. The cars are interconnected by

springs "SPRING_2" and "SPRING_3", the forces in them are determined by the formula (2). The point of connection of the spring-car corresponds to the location of the coupler wedge drawbar.

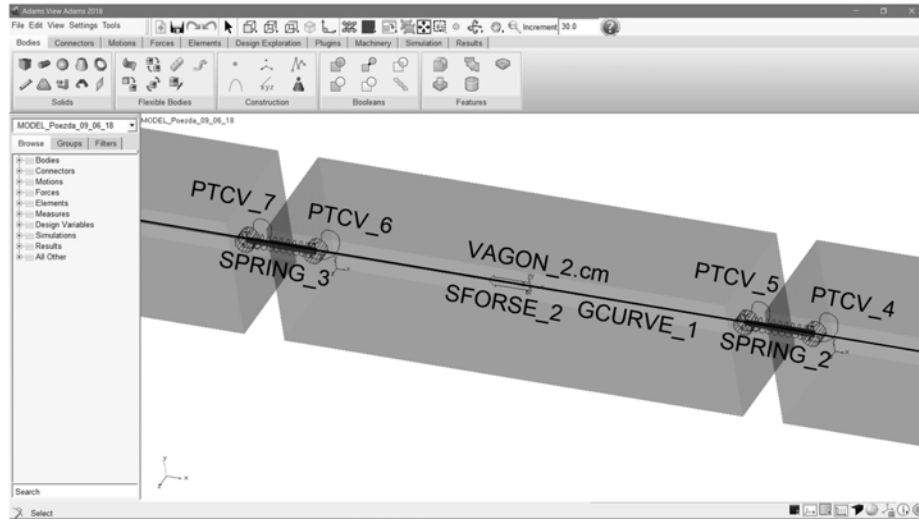


Fig. 2. The train model in MSC.ADAMS/View

The final model allows to change the initial speed and the mass of each wagon, the parameters of elastic connections, the railway profile, as well as to change the number of cars in the train by their activating or deactivating.

For the convenient analysis of the motion simulation computational results, they are presented in the form of graphs. An example of the traveled distance, the velocity and acceleration of the locomotive movement graph is presented in Figure 3.

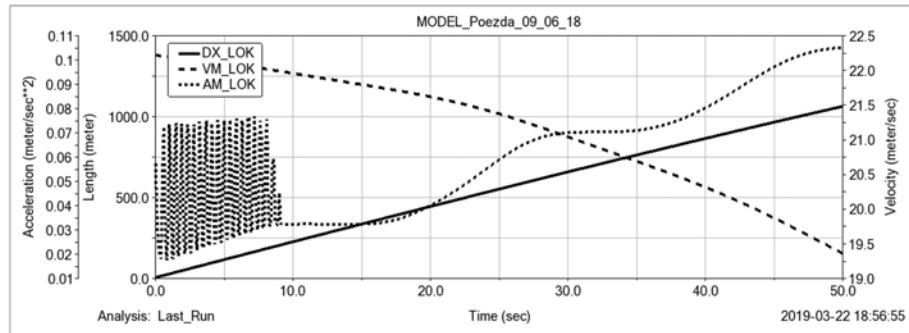


Fig. 3. Kinematical parameters of the locomotive movement in MSC.ADAMS

The earlier studies carried out by the presented model [21] made it possible to analyze the degree of the track profile grade changes influence on the magnitude of the maximal forces in the homogeneous train inter-car connections. Analysis of the train

movement from plane to ascent and from site to descent in the idle mode showed that an increase in the conjugation radius of elements in the vertical plane leads to a decrease in maximal forces in automatic couplers. The change of forces in the homogeneous train inter-car joints is almost the same at moving one slope up and down and it is determined only by the difference in slopes of adjacent elements and the radius of their conjugation.

Since it was established that the dependence of the maximal forces in the inter-car connections on the difference between the adjacent elements slopes is close to linear, the current investigation considers the movement only along the grade change of the constant steepness and conjugation radius. The presented below results were obtained for the case of the train moving from the horizontal way to a 13 ‰ grade steep rise

5 Computational results

Trains of the same mass can have different length and number of cars. With the help of the model, the influence of these factors on the longitudinal forces in the train was studied. Figure 4 shows the change in forces in the most heavily loaded inter-car connections of trains of the same mass (2500 tons) and different numbers of cars over time for case of motion through a profile break in the idle mode. The initial velocity was 80 km/h. In each train, the mass is distributed uniformly over the car. If there were more cars their mass was taken less.

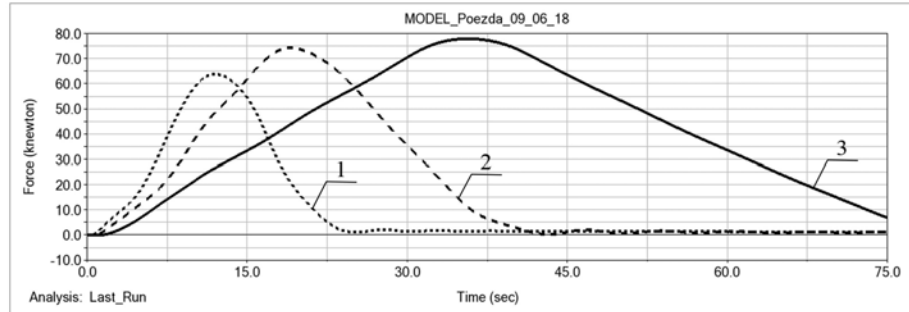


Fig. 4. Change in longitudinal forces in the most loaded automatic couplings of a homogeneous trains with the equal mass: 1 – 25 cars, 13 inter-car joint; 2 – 50 cars, 24 inter-car joint; 3 – 100 cars, 48 inter-car joint

With a decrease in the number of cars from 100 to 25, the maximal longitudinal forces decreased from 78 to 64 kN, and there could be observed a significant reduction in time of compressive forces action. The time of the compressed position of the inter-car connections is proportional to the train length and for the considered cases it reduced by 4 times. Consequently, a decrease in the homogeneous trains length due to a decrease in the cars number leads to a slight decrease in the maximum longitudinal forces and to a reduction in their action time, proportional to the number of cars. At the same time, the mass of the train remains unchanged.

For a detailed study of the maximum forces distribution along the train length, similar calculations were carried out for a train of 5000 tons mass. The results are presented in Figures 5 and 6.

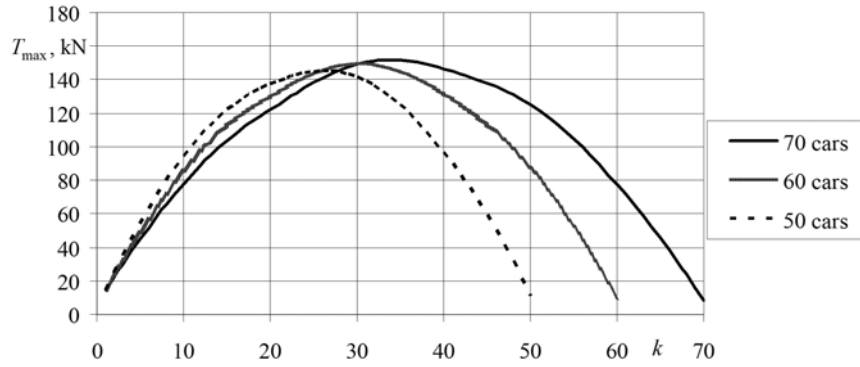


Fig. 5. The distribution of maximal forces $T_{\max}$ along the length of a train of equal mass, formed from 50, 60 and 70 cars (k – car number)

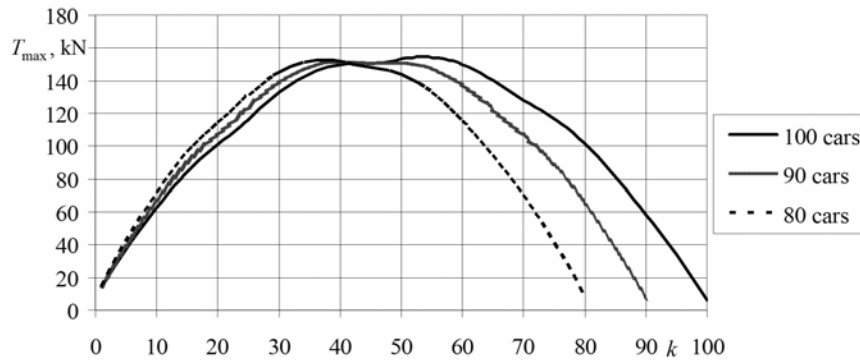


Fig. 6. The distribution of maximal forces $T_{\max}$ along the length of a train of equal mass, formed from 80, 90 and 100 cars (k – car number)

According to the presented figures, the maximal forces in the inter-car couplings appear in the central part of the train and change relative to it almost symmetrically. This is typical for homogeneous trains. There is a non-significant increase in forces from 145 to 155 kN when the number of cars changes from 50 to 100 and the train mass decreases from 100 to 50 tons correspondingly. The maximal forces increase for case of reducing car mass can be explained by an increase in specific resistant to movement forces (forces per unit mass) of less loaded cars compared to more loaded ones due to a decrease in q in formula (6).

There was investigated the effect of a homogeneous train mass on the longitudinal forces in automatic couplers. The train was formed from 100 cars and its initial velocity was equal to 80 km/h. There were considered four variants of common train mass: 2500, 5000, 7500, 10000 tons. There was simulated the movement of each train

through a broken profile. Curves of forces changes over time in the most loaded inter-car connections of these trains are shown in Figure 7, and the maximal forces distribution along the train is shown in Figure 8.

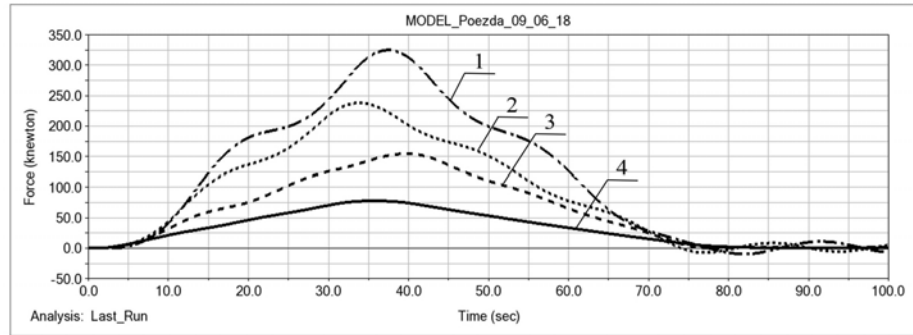


Fig. 7. Change of longitudinal forces in the most loaded inter-car connections of a homogeneous train formed from 100 cars: 1 – car weight 100 t, 52 inter-car joint; 2 – car weight 75 t, 47 inter-car joint; 3 – car weight 50 t, 54 inter-car joint; 4 – car weight 25 t, 48 inter-car joint

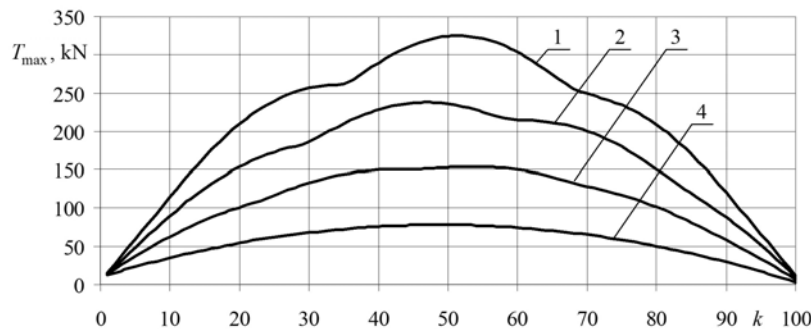


Fig. 8. The maximal forces $T_{\max}$ distribution along the length of a homogeneous train formed from 100 cars for gross car weights: 1 – 100 t; 2 – 75 t; 3 – 50 t; 4 – 25 t

The maximal forces in a homogeneous train with an unchanged number of cars are proportional to the train mass. The maximal longitudinal force appears in the central part of the train and, according to the Figure 7, it can be located in the inter-car connection of both the first half (connections 47 and 48; cars masses are 75 and 25 tons, respectively), and in the second half (connections 52 and 54; cars masses are 100 and 50 tons, respectively) of the train.

The curves of the maximal forces distribution along the length of the train of different mass (Figure 8) are similar to the dependences obtained for the train of constant mass passing the profile grade changes of different steepness. Analysis of the simulation results showed that the dependence of the maximal forces in the train on its mass is practically linear. This relationship is shown in Figure 9.

It should be noted that the above mention dependence is valid for the case of movement through a profile grade change with 13 % slopes difference and 15 km

radius of adjacent elements conjugation. For other slopes and conjugation radii, the maximal force values change proportionally. A slope angle increase leads to a shift of the presented graph above up, and a decrease – to a downward shift. An increase in the conjugation radius of adjacent elements in the vertical plane leads to a decrease in the angle of inclination of the line, and radius decrease - to an increase in this angle.

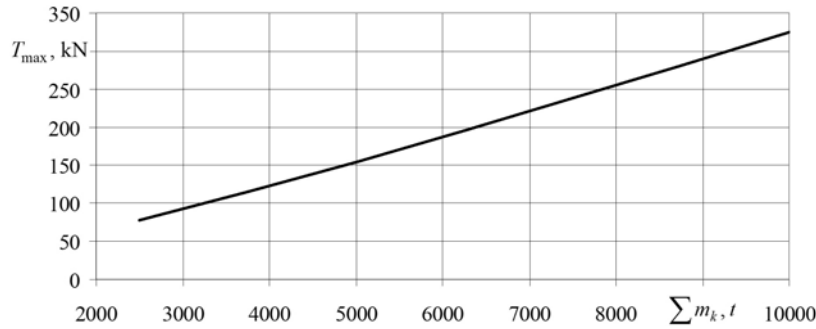


Fig. 9. The dependence of the maximal forces T_{max} in the train inter-car joints on its mass for the case of movement along the profile grade change

Figure 10 presents the simulation results demonstrating the possibility of using the dependence shown in Figure 9 for an approximate estimate of the maximal forces arising in the train of various mass and number of cars.

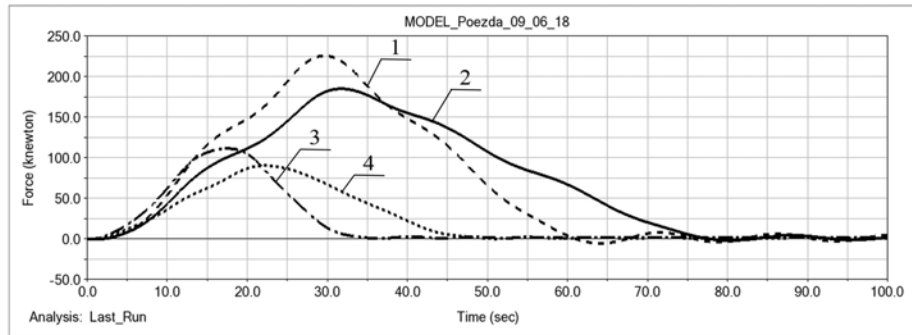


Fig. 10. Change of longitudinal forces in the most loaded inter-car connections of a homogeneous train of different masses and number of cars: 1 – 80 cars with mass 90 t; 2 – 100 cars with mass 60 t; 3 – 40 cars with mass 100 t; 4 – 60 cars with mass 50 t

In the presented graphs the maximal forces are 225, 185, 111, 90 kN. The forces appeared in the trains of 7200, 6000, 4000, 3000 t correspondingly. These results are in good agreement with the graph shown in Figure 9.

With the help of the model, there were performed the computations and their results allow us to determine the time moment and the location of the profile point where there I observed the maximal force in the inter-car connection at the train movement through the profile grade change in the idle mode.

The movement of a homogeneous train formed from 50 loaded cars weighing 100 tons is considered. The maximal forces in all automatic inter-car couplings and their position on the track profile section are determined. The results of the simulations are presented in Figure 11. Their analysis showed that the maximal forces in all inter-car connections appear directly at moving along a transition curve connecting adjacent elements of different steepness. The greatest forces values, corresponding to the automatic couplings of the train central part, arise in the center of this transition curve and they smoothly decrease at a distance. In this case, the interconnection between the coupler position along the train and the location of the profile point at which the maximum force occurs is observed: when the coupling is removed from the train tail, this point moves away from the beginning of the transition curve to its end.

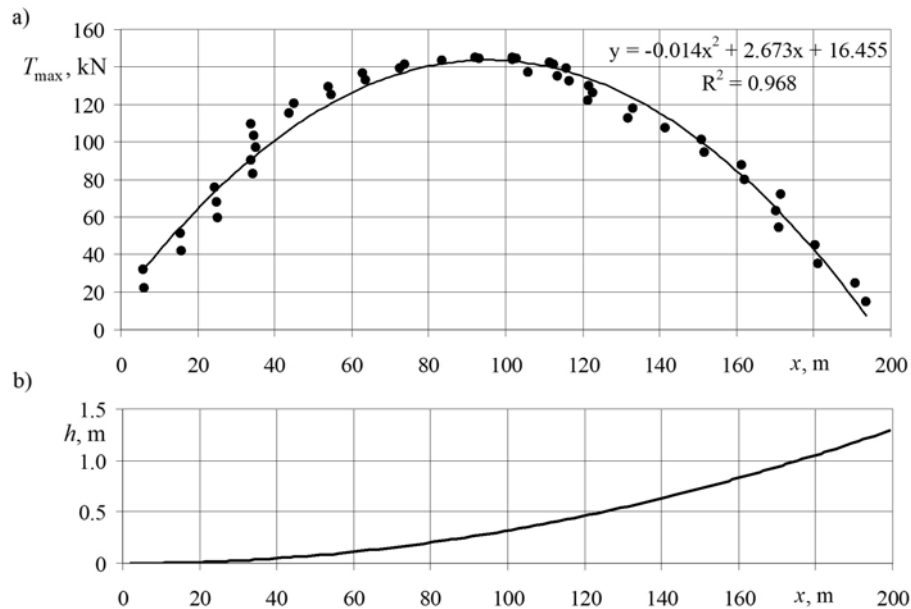


Fig. 11. The location of the inter-car joints at the moment when the maximal longitudinal forces appear in them in section (a) and the profile section level change (b)

There was performed the movement analysis of the heterogeneous by mass trains. At simulation there was taken the same number of cars – 90 and the train mass – 6300 t. The empty cars were arranged in groups of 30 in the head, center and tail of the train, and they were also distributed in groups of 10 cars located in different parts of the train. The mass of empty car is 24 t, and of the loaded one – 93 t. A homogeneous train formed of cars of 70 t each is considered for comparison. Figure 12 shows the dependence of the longitudinal forces change on time in the most loaded inter-car connections of homogeneous and heterogeneous trains.

Graphs of the maximal forces distribution for the inter-car connections with different empty cars locations are shown in Figures 13–15.

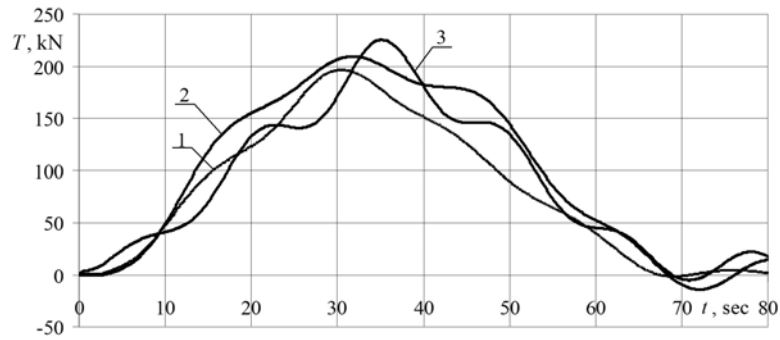


Fig. 12. The dependence between the forces in the most loaded inter-car connections of homogeneous and heterogeneous trains and the time of movement along the profile grade change: 1 – without empty cars; 2 – empty cars location from 31 to 60; 3 – empty cars are groups from 1 to 10, 21 to 30, 41 to 50

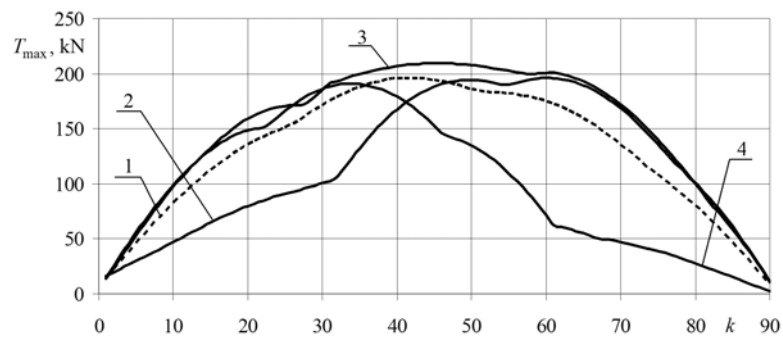


Fig. 13. Comparative characteristics of the maximal longitudinal forces distribution for couplings of homogeneous and inhomogeneous trains of the same mass: 1 – without empty cars; 2 – empty cars location from 1 to 30; 3 – empty cars location from 31 to 60; 4 – empty cars location from 61 to 90

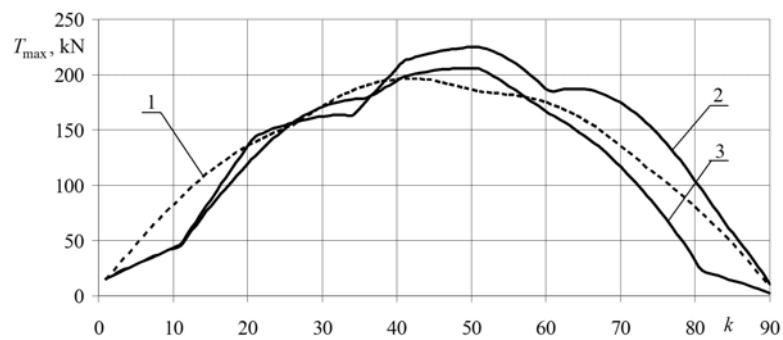


Fig. 14. Comparative characteristics of the maximal longitudinal forces distribution for couplings of homogeneous and inhomogeneous trains of the same mass: 1 – without empty cars; 2 – empty cars are groups from 1 to 10, 21 to 30, 41 to 50; 3 – empty cars are groups from 1 to 10, 41 to 50, 81 to 90

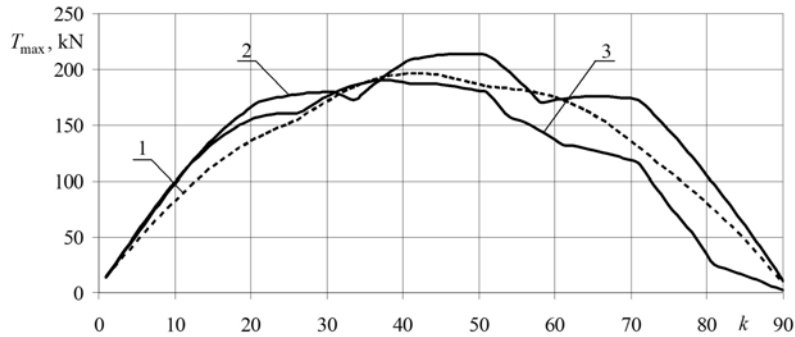


Fig. 15. Comparative characteristics of the maximal longitudinal forces distribution for couplings of homogeneous and inhomogeneous trains of the same mass: 1 – without empty cars; 2 – empty cars are groups from 21 to 30, 41 to 50, 61 to 70; 3 – empty cars are groups from 41 to 50, 61 to 70, 81 to 90

The graphs presented in Figures 12–15 show that when driving an otherwise heterogeneous train through a grade change of a longitudinal track profile in the idle mode, the arrangement of empty cars in the middle part of the train leads to an increase in the maximal longitudinal forces. The alternating arrangement of empty and loaded wagons in the head and central part of the train leads to an even greater increase in strength. So the location of 30 empty cars in groups of 10, located after the locomotive, 20th and 40th cars is the most unfavorable variant. In this case, the maximal force increases by 15 % compared with a homogeneous train of the same mass.

There was analyzed the movement of an inhomogeneous train with 6000 t mass. The train was composed of 100 cars, the number of empty cars varied from 10 to 50. At the same time, empty cars were grouped in the head, in the center or in the tail of the train. The simulation results are presented in Figure 16.

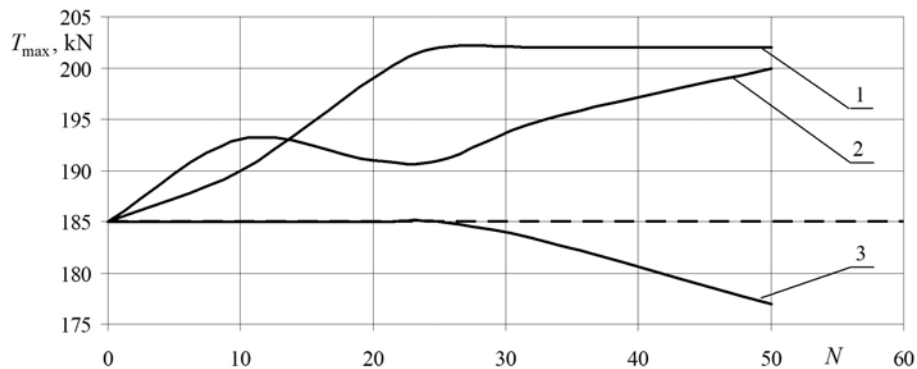


Fig. 16. Maximal forces in a heterogeneous train of the same mass with different numbers and location of empty cars: 1 – in the middle; 2 – in the head; 3 – in the tail

6 Conclusion

Simulation modeling made it possible to evaluate the effect of train track profile grade change on the forces between cars at train movement. It was determined that when the train moves in the idle mode through the profile grade change with 13 ‰ slope difference, the maximal forces change from 90 kN (train mass is 3000 t) to 320 kN (train mass is 10000 t).

The dependence of the maximal forces in a uniform train on its mass was obtained. It was established that the maximal forces in all inter-car connections are at the moment when the cars pass along a transition curve connecting the elements of different inclines, and the greatest of them is located in the central part of the train at the curve middle passing.

There was taken into consideration the effect of train mass heterogeneity on the maximal forces in the automatic inter-car couplers at train movement along the track profile grade change. The simulation results showed that the greatest increase in longitudinal forces is observed when the empty cars are located in the central part or they alternate with loaded in the head and central part of the train. At the same time, the maximal forces increase by 15 % or more compared to a uniform train of the same mass.

The obtained simulation results can be used for the development of recommendations for the forming and maintenance of trains.

References

1. Blokhin, E. P., Manashkin, L. A.: The dynamics of the train (non-stationary longitudinal oscillations) (in Russian). Transport, Moscow (1982)
2. Grebenyuk, P. T.: Longitudinal train dynamics (in Russian). Intekst, Moscow (2003)
3. Vershinskiy, S. V., Danilov, V. N., Khusidov, V. D.: Car dynamics (in Russian). Transport, Moscow (1991)
4. Masleeva, L. G.: Study of longitudinal forces in the train caused by grade changes of the path longitudinal profile and control actions: PhD Dissertation Thesis (in Russian). Dnepropetrovsk (1979)
5. Naumenko, N., Khizha, I., Sobolevska, Yu., Bogomaz, G.: Dynamic loading of a freight train at its movement on a longitudinal path profile grade change in modes of run-out and braking (in Russian). Technical mechanics **4**, 86–90 (2010)
6. Naumenko, N., Khizha, I., Nikitchenko, A.: Characterization of influence of power characteristics of the perspective draft gears on dynamics of the cargo train at the non-stationary modes of the movement (in Russian). Technical mechanics **2**, 27–31 (2009)
7. Islamov, A. R.: Simulation modelling of train movement via adjacent elements of grade (in Russian). Herald of the Ural State University of Railway Transport **4**, 77–82 (2011)
8. Akkerman, G. L., Islamov, A. R.: Studying Power Parameters of Train Movement Via Breaks in Adjacent Elements of Grade by Simulation (in Russian). Transport of the Ural **3**, 63–65 (2012)
9. Akkerman, G. L., Islamov, A. R.: The influence of the longitudinal profile pairing elements on the force in the coupler connection heavy freight trains (in Russian). Railway Track and Facilities **2**, 25–27 (2013)

10. Islamov, A. R.: Research of a coupling elements for a railway track longitudinal profile by simulation: PhD Dissertation Thesis (in Russian). Saint-Petersburg (2014)
11. Luca, P., Duccio, F. Andrea, R.: Modelling the longitudinal dynamics of long freight trains during the braking phase. In: 12th IFToMM World Congress, Besancon (France), June 18-21 (2007)
12. Bureika, G., Mikaliunas, S.: Research on the compatibility of the calculation methods of rolling-stock brakes. *Transport* **23** (4), 351–355 (2008)
13. Pugi, L., Malvezi, M., Papini, S., Vettori, G.: Design and preliminary validation of a tool for the simulation of train braking performance. *Journal of Modern Transportation*, **21** (4), 247–257 (2013)
14. Xin, G., Kaiyun, W., Lirong, G., Min, Y., Kaikai, L., Wanming, Z.: Investigation on Derailment of Empty Wagons of Long Freight Shock and Vibration Train during Dynamic Braking. *Shock and Vibration* **2018**, 1–18 (2018). doi: 10.1155/2018/2862143
15. Piechowiak, T.: Longitudinal Dynamics of the Rail Vehicles. *J. Mechanical and Transport Engineering* **69** (4), 47–61 (2017). doi: 10.21008/j.2449-920X.2017.69.4.04
16. Qing, W.: A new coordinate system for Longitudinal Train Dynamics simulations. *Advances in Vehicle Engineering* **3** (4), 161–166 (2017).
17. Davydov, Y.: Longitudinal Dynamics in Connected Trains. *Procedia Engineering* **165**, 1490–1495 (2016). doi: 10.1016/j.proeng.2016.11.884
18. Spiriyagin, M., Cole, C., Sun, Y. Q., McClanachan, M., Spiriyagin, V., McSweeney, T.: Design and Simulation of Rail Vehicles. CRTC Press, Boca Raton (2014)
19. Varazhun, I., Zavarotny, A.: Modeling of the interaction between cargo tiers and flatcar during the collision of cars. *Technolog* **5**, 195–198 (2013)
20. Rules of traction calculations for train work (in Russian). Transport, Moscow (1985)
21. Sakharau, P.: Investigation of Longitudinal Forces in the Freight Trains at Its Movement Along the Track Longitudinal Irregularities (in Russian). *Mechanics: Researches and Innovations* **11**, 209–219 (2018)

On Constrained Optimization of Polynomials on Permutation Set

Sergiy Yakovlev¹[0000-0003-1707-843X] and Oksana Pichugina²[0000-0002-7099-8967]

¹ National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
svsyak7@gmail.com

² National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
oksanapichugina1@gmail.com

Abstract. Constrained polynomial optimization problem on permutation set is explored. For the problem, an equivalent formulation with a convex objective function and functional constraints is formed based on forming convex extensions of all functions involved in the model. New approaches to the construction of convex extensions of polynomials from the permutation set are proposed. Original optimization methods using a polyhedral-spherical structure of the set are offered. Decomposition schemes underlying some global optimization techniques on permutations are described. Results of numerical experiments are presented.

Keywords: Combinatorial Optimization, Constraints, Permutation Set, Polynomial, Convex Extension.

1 Introduction

Various real-world problems can be modeled as optimization problems over the permutation set [1-5]. Among them are scheduling, assignment, routing problems and many others. They form a class of permutation-based problems (PBP) [6, 7]. Some of them allow embedding into Euclidean space and considering them as Euclidean combinatorial sets. Permutation set and its Boolean invariant - permutation matrices set - are well studied [8-10]. For instance, it is known H-representations of their convex hulls, main combinatorial characteristics, solutions of linear optimization and projection problems, etc. However, real problems require fulfilling many conditions on permutations. In mathematical models, they are represented by additional constraints, typically spoiled the original combinatorial structure of permutation set and requiring developing specific technics for solving the corresponding optimization and feasibility problems. From a theoretical point of view, permutation, as well as some of its subsets, are of interest. Among such subsets are even, odd, alternating, distance, cyclic, circular, complete permutations, etc. [11-14]. It is a separate problem to construct additional constraints to single out any of the specific subsets. When it is solved, we come to a constrained PBP again. The paper is dedicated to studying permutation set embedded into the Euclidean space from different points of view in order to offer new approaches to constrained permutation-based optimization.

2 Optimization Problem of Polynomials on Permutations

Consider the optimization problem on permutation set in the following formulation. Let Π be permutation space on which functional $\xi : \Pi \rightarrow R^I$ is defined. We are looking for $\pi^* \in \Pi$ such that

$$\pi^* = \arg \min_{\pi \in \Pi'} \xi(\pi), \quad (1)$$

where $\Pi' \subseteq \Pi$ is a feasible set.

We perform a mapping $\psi : \Pi \rightarrow E$ between elements $\pi \in \Pi$ and vectors $x = (x_1, x_2, \dots, x_n) \in E \subset R^n$, i.e., suppose

$$x = \psi(\pi), \quad \pi = \psi^{-1}(x) \quad \forall \pi \in \Pi, x \in E.$$

Combinatorial sets for which such mapping can be established are called Euclidean combinatorial sets. Permutation set is Euclidean combinatorial one, which is especially important for further research in this paper.

Denote $J_n = \{1, \dots, n\}$. Let $G = \{g_i, i \in J_n\}$ be a set of real numbers such that $0 < g_1 < g_2 < \dots < g_n$. To each $\pi = (\pi_1, \dots, \pi_n) \in \Pi$ we correspond vector $x = (x_1, x_2, \dots, x_n) \in R^n$, assuming

$$x_i = g_{\pi_i}, \quad i \in J_n.$$

The set of such vectors $x = (x_1, x_2, \dots, x_n)$ induces a finite point configuration $E_n(G) \subset R^n$. Let $\tilde{G} = \{g_1, g_2, \dots, g_n\}$ be a multiset of real numbers such that

$$0 < g_1 \leq g_2 \leq \dots \leq g_n.$$

Then $E_n(\tilde{G}) \subset R^n$ be permutation set with repetitions. For both sets $E_n(G)$ and $E_n(\tilde{G})$, we will use the same notation E .

Suppose function $f : E \rightarrow R^I$ is such that

$$f(x) = \xi(\psi^{-1}(x))$$

for all $x \in E$.

Consider a discrete optimization problem: to find

$$x^* = \arg \min_{x \in E' \subseteq E} f(x), \quad (2)$$

where $E' = \psi(\Pi')$.

The problems (1) and (2) are equivalent in the sense that if a solution $x^* \in E'$ of the problem (2) is found, then

$$\pi^* = \psi^{-l}(x^*)$$

is a solution of the problem (1). We reformulate (2) in the form

$$f(x) \rightarrow \min, \quad x \in X, \quad (3)$$

where set X is given by the constraints:

$$x \in E, \quad (4)$$

$$\varphi_i(x) \leq 0, \quad i \in J_k, \quad (5)$$

$$\varphi_i(x) = 0, \quad i \in J_m \setminus J_k. \quad (6)$$

and functions $f(x)$, $\varphi_i(x)$, $i \in J_m$, are defined on E .

Thus,

$$X = \{x \in E : \varphi_i(x) \leq 0, \quad i \in J_k, \quad \varphi_i(x) = 0, \quad i \in J_m \setminus J_k\}.$$

3 Properties of Polynomials on Permutation Set

Properties of the problem (3) - (6) are determined by properties of permutation set E and behavior of functions $f(x)$, $\varphi_i(x)$, $i \in J_m$ on R^n , as well as on $E \subset R^n$. Let functions $f(x)$, $\varphi_i(x)$, $i \in J_m$ be polynomials.

Theorem 1. For any polynomial $f(x)$, there exists a constant C and a polynomial $\tilde{f}(x)$, that can be represented as a positive linear combination of monomials, such that $f(x) = \tilde{f}(x) + C$ for any $x \in E$.

Proof of the theorem follows from the fact that any symmetric function takes a constant value on E . We choose a positive linear combination of monomials as the symmetric function. Then any monomial with a negative coefficient can be represented on E by a positive linear combination of monomials and a constant.

An important property of a polynomial function $f : E \rightarrow R^l$ is the possibility of constructing its convex extension to a convex set $X \supset \text{conv } E$.

Definition 1. A function $\tilde{f} : X \rightarrow R^l$ defined on a convex set X is called a convex extension of function $f : E \rightarrow R^l$ onto X if

$$\tilde{f}(x) = f(x)$$

for any $x \in E \subset X$.

To indicate that equality of functions holds on E , we will use a notation

$$f(x) \underset{E}{=} \tilde{f}(x).$$

The theory of convex extensions of functions defined on combinatorial sets mapped onto R^n , is considered in [15] and is based on the following theorem.

Theorem 2. Let $E = \text{vert conv } E$. Then for any function $f : E \rightarrow R^l$ there exists a convex function $\tilde{f} : \text{conv } E \rightarrow R^l$, such that $f(x) \underset{E}{=} \tilde{f}(x)$.

Definition 2. A finite set E satisfying a condition $E = \text{vert conv } E$ is called vertex-located.

A convex hull of any finite set $E \subset R^n$ is a polytope $P = \text{conv } E$ whose vertices form a vertex located set. Permutation set E is vertex located and coincides with a set of vertices of the permutohedron P described by the following linear system

$$\sum_{i=1}^n x_i = \sum_{i=1}^n g_i \quad (7)$$

$$\sum_{i \in \omega} x_i \geq \sum_{i=1}^{|\omega|} g_i, \quad (8)$$

where the summation is over various indexes from subsets $\omega \subset J_n$ of the cardinality $|\omega|$.

Note that for any τ , equality

$$\sum_{i=1}^n (x_i - \tau)^2 = \sum_{i=1}^n (g_i - \tau)^2 \quad (9)$$

holds for points of E .

Thus, set E belongs to a family (9) of hyperspheres centered at the point $\tau = (\tau, \tau, \dots, \tau) \in R^n$ with a radius

$$r = \left(\sum_{i=1}^n (g_i - \tau)^2 \right)^{1/2} \quad (10)$$

and a parameter $\tau \in R^1$. The minimum value of the radii is attained at

$$\tau = \frac{1}{n} \sum_{i=1}^n g_i. \quad (11)$$

Theorem 3 [16]. A point $x \in E$ if and only if it satisfies the system (7)-(9).

Definition 3. Finite set $E \subset R^n$ that belongs to both a hypersphere and a vertex set of a polytope is called polyhedral-spherical.

Properties of polyhedral-spherical sets were studied in [17–19] and adapted to permutation set.

4 Convex Extensions of Polynomials on Permutation Set

Let us present new approaches to constructing convex extensions of polynomials defined on permutation set E . The extensions will be formed by induction.

By assumption

$$E \subset R_{>0}^n = \text{int } R_+^n.$$

Clearly, any monomial with a positive coefficient is positive on $R_{>0}^n$. For convex extensions of monomials, we also require positivity on $R_{>0}^n$.

Remark 1. A square of any positive convex function is also a convex function. We will use this fact for a recursive construction of convex extensions monomials from already constructed ones of lower degrees.

A monomial of a degree k is represented in the form:

$$F_{k_1 \dots k_n}^{(k)}(x) = \prod_{i=1}^n x_i^{k_i}, \quad k_1 + k_2 + \dots + k_n = k \geq 1. \quad (12)$$

If $k = 1$, then

$$F_{k_1 \dots k_n}^{(1)}(x) = x_i,$$

where $k_i = 1, k_j = 0, j \in J_n, j \neq i$.

A convex extension of this function will coincide with the original one and positive on $R_{>0}^n$, i.e.,

$$\tilde{F}_{k_1 \dots k_n}^{(1)}(x) = F_{k_1 \dots k_n}^{(1)}(x).$$

For a quadratic monomial $F_{k_1 \dots k_n}^{(2)}(x)$ we offer the following convex extension onto R^n

$$\hat{F}_{k_1 \dots k_n}^{(2)}(x) = \begin{cases} x_i^2, & \text{if } k_i = 2, k_j = 0 \quad \forall j \in \mathbf{J}_n, j \neq i; \\ \frac{1}{2} \left((x_i + x_j)^2 + \sum_{\substack{k=1 \\ k \neq i, j}}^n x_k^2 - \sum_{k=1}^n g_k^2 \right), & \text{if } k_i = k_j = 1, \quad \forall i, j \in \mathbf{J}_n, j \neq i. \end{cases}$$

Positive convex extension of $F_{k_1 \dots k_n}^{(2)}(x)$ onto $R_{>0}^n$ has the form

$$\tilde{F}_{k_1 \dots k_n}^{(2)}(x) = \max\{0, \hat{F}_{k_1 \dots k_n}^{(2)}(x)\}.$$

We construct a positive convex extension onto $R_{>0}^n$ for monomial (12) of a degree k under the assumption that for any monomials of degree less k such extensions have been constructed. It is true

$$\begin{aligned} F_{k_1 \dots k_n}^{(k)}(x) &= F_{j_1 \dots j_n}^{(j)}(x) F_{l_1 \dots l_n}^{(k-j)}(x) = \\ &= \frac{1}{2} \left(\left(F_{j_1 \dots j_n}^{(j)}(x) + F_{l_1 \dots l_n}^{(k-j)}(x) \right)^2 - \left(F_{j_1 \dots j_n}^{(j)}(x) \right)^2 - \left(F_{l_1 \dots l_n}^{(k-j)}(x) \right)^2 \right), \end{aligned}$$

where

$$j = \lfloor (k+1)/2 \rfloor, \quad l_i = k_i - j_i, \quad i \in \mathbf{J}_n.$$

Based on the induction hypothesis for monomials of degree less k and taking into account Remark 1 and Theorem 1, we can construct a convex extension $\hat{F}_{k_1 \dots k_n}^{(k)}(x)$.

Then a positive convex extension on $R_{>0}^n$ is

$$\tilde{F}_{k_1 \dots k_n}^{(k)}(x) = \max\{0, \hat{F}_{k_1 \dots k_n}^{(k)}(x)\}.$$

The considered approach to constructing the convex extension is generally complicated and yields no smooth functions. Depending on a class of functions in the problem (3)-(6), simpler algorithms can be offered. In the article [17], such an approach to the construction of smooth convex extensions of monomials defined on permutation set was proposed.

Theorem 4. For any monomial $F_{k_1 \dots k_n}(x)$ of the form (10) defined on a set E , there exists a smooth convex extension onto $R_{>0}^n$ given by a formula

$$\bar{F}_{k_1 \dots k_n}^k(x) = \left(\prod_{i=1}^n g_i \right)^\mu \prod_{i=1}^n x_i^{k_i - \mu},$$

where

$$\mu = \max_{i \in J_n} k_i.$$

Corollary. For any polynomial defined on a set $E \subset R_{>0}^n$, there exists a smooth convex extension onto $R_{>0}^n$.

5 Optimization of Polynomials on Permutations with Constraints

Consider the optimization problem (3)-(6) and implement the following equivalent transformations to it. Represent equations (6) in the form

$$\varphi_i(x) \leq 0, \quad -\varphi_i(x) \leq 0, \quad i \in J_m \setminus J_k.$$

Next, for polynomials $f(x)$, $\varphi_i(x)$, $i \in J_m$, we construct the following convex extensions onto a convex set $P \supset E$:

$$\bar{f}(x) = f(x),$$

$$\tilde{\varphi}_i(x) = \varphi_i(x), \quad i \in J_m,$$

$$\tilde{\varphi}_i(x) = -\varphi_{i-m+k}(x), \quad i \in J_q \setminus J_m, \quad q = 2m - k.$$

Let us form an optimization problem

$$\bar{f}(x) \rightarrow \min, \quad x \in \tilde{X}, \quad (13)$$

where the set $\tilde{X}$ satisfies (3) and the system of inequalities

$$\tilde{\varphi}_i(x) \leq 0, \quad i \in J_q, \quad (14)$$

Problems (3)-(6) and (13),(14) are equivalent in the sense that

$$\arg \min_{x \in X} f(x) = \arg \min_{x \in \tilde{X}} \bar{f}(x).$$

Let us consider $\tilde{X}$ in the form

$$\tilde{X} = E \cap W,$$

where

$$W = \{x \in \Pi : \tilde{g}_i(x) \leq 0, i \in J_q\}.$$

As a result, we have constructed a polyhedral relaxation of the original problem as a convex optimization problem

$$\bar{f}(x) \rightarrow \min, x \in W, \quad (15)$$

One of the interesting properties of permutation set is its decompositions into permutation sets of lower dimension lying on parallel planes. For instance, if $x = (x_1, x_2, \dots, x_n) \in E_n(\tilde{G})$, then for any $x_i \in \tilde{G}$ holds

$$E_n(\tilde{G}) = E_{n-1}(\tilde{G}^i) \times x_i,$$

where

$$\tilde{G}^i = \tilde{G} \setminus \{x_i\}, i \in J_n.$$

Evidently, when solving the problem (13),(14), such a decomposition induces convex optimization problems again. In fact, the offered branching scheme, used recursively, allows organizing branch and bound schemes to solve the problem (15). In this case, lower bounds correspond to solutions of the induced convex programs, where the number of levels of a decision trees does not exceed n . Also, note that a specifics of the permutohedron and permutation set allows strengthening the lower bounds [20].

Let us describe some approaches to optimization on permutations based on its polyhedral-spherical structure. One of them uses an idea of optimization of convex objective function on the hypersphere described by equation (9). In the general case, the exact methods to solve this problem are unknown. An exception is a linear or convex quadratic function $f(x)$. This feature underlies an approach to solve the quadratic optimization problem on E offered in [20]. Clearly, an intersection of a hypersphere with a hyperplane yields a lower-dimension sphere. This allows performing a decomposition of the original problem into simpler ones in accordance to the decomposition of E into hyperplanes parallel to facets of permutohedron, followed by solving induced problems of quadratic function optimization on a sphere of lower dimension.

It is of interest to apply the penalty method to the problem (15). We offer an ap-

proximate method for solving this problem based on using of a polyhedral-spherical representation of the set E , i.e., its presentation as the intersection of polytope (7), (8) with hypersphere (9). Consider the following penalty function:

$$F(\alpha_k, x) = \bar{f}(x) + \alpha_k \sum_{i=1}^q \left(\max\{0, \tilde{\varphi}_i(x)\} \right)^2.$$

It is clear that, given that penalty coefficients $\alpha_k > 0$, the following sequence of optimization problems

$$F(\alpha_k, x) \xrightarrow[k \rightarrow \infty]{} \min, \quad x \in \Pi,$$

consists of convex programs, whose solutions, under rather general conditions and $\alpha_k \rightarrow \infty$, converge to a solution of the problem (15).

Let us introduce one more optimization approach based on the use of the following penalty function:

$$\Phi(\alpha_k, \beta_k, x) = F(\alpha_k, x) + \beta_k \left(\sum_{i=1}^n (x_i - \tau)^2 - r^2 \right)^2,$$

where $\alpha_k > 0$, $\beta_k > 0$ are penalty coefficients, r and τ are given by expressions (10), (11), respectively.

Then for sufficiently large α_k and β_k , a solution of the following problem

$$\Phi(\alpha_k, \beta_k, x) \rightarrow \min, \quad x \in \Pi \tag{16}$$

can be considered as an approximate solution to the problem (3)-(6).

When solving optimization problems on permutohedron P , auxiliary linear and quadratic optimization problems arise, which are solvable explicitly using the properties of permutation set [16,21,22]. For instance, the minimization of a linear function on E is reduced to the ordering of its coefficients. Similarly, the problems of finding the nearest vertex and facet of the permutohedron are solvable. This allows offering modifications of the conditional gradient method and the gradient projection method for solving nonlinear optimization problems on the permutohedron.

As a model constrained optimization problem on permutation set, the problem of balancing rotating discretely distributed masses was considered [23]. There are n blades with masses $\mu_i, i \in J_n$, that are placed on a balanced disk with a given angular step. It is required that masses of successively placed blades should differ by at least the value Δ_1 and at most Δ_2 . The goal is to minimize the total unbalance of the whole system. Mathematically, the problem can be represented as the optimization problem on the permutation set E of the following objective function

$$f(x) = \left[\left(\sum_{i=1}^n x_i \cos 2\pi i/n \right)^2 + \left(\sum_{i=1}^n x_i \sin 2\pi i/n \right)^2 \right]^{1/2}$$

subject to constraints

$$\Delta_I \leq |x_{i+I} - x_i| \leq \Delta_2, \quad i \in \mathbf{J}_{n-I},$$

$$\Delta_I \leq |x_n - x_I| \leq \Delta_2.$$

The set E is induced by $\tilde{G} = \{\mu_1, \mu_2, \dots, \mu_n\}$. A test problem of balancing was formed and solved with the following parameters:

$$n = 100, \quad \Delta_I = 0.2, \quad \Delta_2 = 1.0,$$

$$\mu_i = \sqrt{i}, \quad i \in \mathbf{J}_n.$$

PC with characteristics i3/8G/SSD 256G was used. A solution was obtained by the penalty method in the formulation (16) with parameters $\alpha_k = 10$, $\beta_k = 5$. Namely, the total unbalance of 0.31 was obtained for the running time 17 sec. In addition, the test problem of balancing rotating 96 blades masses presented in [23] was solved resulted in the total unbalance 0.07 obtained over 9 sec.

6 Conclusions

The paper presents new approaches to solving permutation-based problems allowing embedding into the Euclidean space. The main result consists in the reduction of an arbitrary problem of conditional optimization of a polynomial function defined on a set of permutations to optimization of a convex objective function with convex functional restrictions on a vertex located set. Thus, for any permutation-based optimization problem on permutations with arbitrary constraints, an equivalent problem is constructed, where convex extensions of its objective function and functional constraints participate.

Now, geometric properties of permutation set and permutohedron, as well as features of continuous and convex optimization methods, become applicable to the solution of new classes of combinatorial optimization problems. The results of this paper can be extended to optimization problems on other vertex-located and polyhedral-spherical sets, in particular, on Boolean and binary ones, known by a variety of practical and theoretical applications [2,5,8-11]. In addition, since the sets of even, odd, alternating, distance, cyclic, circular, complete permutations are subsets of the of permutation set, the above results also apply to them.

References

1. Korte, B., Vygen, J.: Combinatorial Optimization: Theory and Algorithms. 6th ed. 2018 edition. Springer, New York (2018).
2. Pardalos, et al.: Handbook of combinatorial optimization. Springer, New York (2013).
3. Papadimitriou, C.H., Steiglitz, K.: Combinatorial optimization: algorithms and complexity. Mineola : Dover Publications (2013).
4. Schrijver, A.: Combinatorial optimization: polyhedra and efficiency. Berlin Heidelberg: Springer-Verlag (2003).
5. Burkard, R.E.: Quadratic assignment problems. Handbook of combinatorial optimization, vol. 5(1), pp. 2741–2814 (2013).
6. Onwubolu, G.C., Davendra, D. (Eds.): Differential evolution: A Handbook for global permutation-based combinatorial optimization. Springer-Verlag, Berlin Heidelberg (2009).
7. Mehdi, M.: Parallel hybrid optimization methods for permutation based problems. PhD thesis. Université des Sciences et Technologie de Lille (2011).
8. Brualdi, R.A.: Combinatorial matrix classes. Cambridge: University Press (2006).
9. Kochenberg, G., et. al.: The unconstrained binary quadratic programming problem: a survey. Journal of Combinatorial Optimization, vol 1, pp. 58-81 (2014).
10. Bohn, A., et. al.: Enumeration of 2-Level Polytopes. Algorithms - ESA 2015, Berlin, Heidelberg: Springer, pp. 191–202 (2015).
11. Bona, M.: Combinatorics of Permutations. Boca Raton, FL: Chapman Hall-CRC (2012).
12. Korsh, J.F., LaFollette, P.S.: Loopless array generation of multiset permutations. Comput. Journal, vol. 47(5), pp. 612-621 (2004).
13. Skala, M.: Counting distance permutations. Journal of Discrete Algorithms, vol. 7, no. 1, pp. 49–61 (2009).
14. Weisstein, E.W.: CRC concise encyclopedia of mathematics. Second Edition. Chapman and Hall/CRC, Boca Raton (2002).
15. Yakovlev, S.: Convex extensions in combinatorial optimization and their applications. Springer Optimization and its Applications, vol. 130, pp. 567-584 (2017).
16. Pichugina, O.S., Yakovlev, S.V.: Functional and analytic representations of the general permutations. Eastern-European Journal of Enterprise Technologies, vol. 1(4), pp. 27-38 (2016).
17. Yakovlev, S.V., Pichugina, O.S.: Properties of combinatorial optimization problems over polyhedral-spherical sets. Cybernetics and Systems Analysis, vol. 54(1), pp. 99-109 (2018).
18. Pichugina, O., Yakovlev, S.: Optimization on Polyhedral-Spherical Sets: Theory and Applications. In 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pp. 1167-1175 (2017).
19. Yakovlev, S.V., et al.: Polyhedral spherical configuration in discrete optimization. J. of Autom. and Information Sci., vol. 51(1), pp.38-50 (2019).
20. Yakovlev, S.V.: Bounds on the minimum of convex functions on Euclidean combinatorial sets. Cybernetics and Systems Analysis, vol. 23(3), pp. 385-391 (1989).
21. Stoyan, Y.G., et al.: Quadratic optimization on combinatorial sets in R^n . Cybernetics and Systems Analysis, vol. 27(4), pp. 561–567 (1991).
22. Yakovlev, S.V., Valuiskaya, O.A.: Optimization of linear functions at the vertices of a permutation polyhedron with additional linear constraints. Ukrainian Mathematical Journal, vol. 53(9), pp. 1535-1545 (2001).
23. Stoyan, Y.G., et al.: Method of balancing rotating discretely distributed masses. Energomashinostroenie, vol. 2, pp. 4-5 (1982).

Methods for comparing numbers in non-positional notation of residual classes

Victor Krasnobayev ^[0000-0001-5192-9918], Alexandr Kuznetsov ^[0000-0003-2331-6326],
Mihael Zub ^[0000-0002-7085-797X], Kateryna Kuznetsova ^[0000-0002-5605-9293]

V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, 61022, Ukraine
v.a.krasnobaev@gmail.com, kuznetsov@karazin.ua,
mishazub007@gmail.com, kate.kuznetsova.2000@gmail.com

Abstract. Methods for optimizing computations in computer systems and components that are based on the use of arithmetic transformations in the system of residual classes are considered. Methods of comparing numbers in the system of residual classes based on the representation and processing of data without directly converting the compared numbers from a modular code to a positional code and back are investigated. Some methods are based on the principle of obtaining and comparing a unitary single-row code. Based on the proposed methods, algorithms for their implementation have been developed, in accordance with which a class of patentable devices has been developed for performing arithmetic and algebraic comparison of numbers in the system of residual classes.

Keywords. Non-positional number system, class of residues, arithmetic non-positional coding, computer systems and components.

1 Introduction

It is known that the use of non-positional number system in residual classes (RCS) significantly increases the reliability and performance of the computer system (CS) [1-8]. However, the need to determine the positional characteristics of numbers in the RCS reduces the overall effectiveness of the use of modular codes. Existing methods for processing positional data, in particular, methods for comparing numbers into RCS, have significant drawbacks, the main of which is the need to convert numbers from RCS to positional number system and vice versa, which reduces user productivity and reliability of CS [9-12].

In the article there are four methods for comparing numbers into RCS, based on the presentation and processing of data without directly converting the compared numbers from a modular code (an RCS code) into a positional code and backward.

2 Methods for arithmetic comparison of numbers in RCS

Let the RCS be given by ordered $(m_i < m_{i+1})$ mutually n pairs by simple natural numbers (bases) $m_1, m_2, \dots, m_n$, d let the compared operands be represented as:

$$A = (a_1, a_2, \dots, a_n), B = (b_1, b_2, \dots, b_n).$$

In the case, it is assumed that the source operands lie in the appropriate intervals:

$$\left[\frac{j_1 M}{m_n}, \frac{(j_1 + 1)M}{m_n} \right) \text{ and } \left[\frac{j_2 M}{m_n}, \frac{(j_2 + 1)M}{m_n} \right),$$

where $M = \prod_{i=1}^n m_i$, and the interval's number $j_k + 1$ is determined by the well-known

expression $j_k = \gamma_n \bar{m}_n (\text{mod } m_n)$, where the value of m_n is determined from the comparison solution $\bar{m}_n M / m_n \equiv l (\text{mod } m_n)$. When $j_1 \neq j_2$ the operation of arithmetic comparison can be implemented by comparing the number of intervals, namely: if $j_1 < j_2$, then $A < B$, if $j_1 > j_2$, then $A > B$. When $j_1 = j_2$ is determined by the number $j_2 + 1$ of the interval $\left[\frac{j_3 M}{m_n}, \frac{(j_3 + 1)M}{m_n} \right)$, in which the number $A - B$ is located. If

$0 \leq j_3 < (m_n + 1) / 2$, then $A < B$, and if $\frac{m_n + 1}{2} \leq j_3 < m$, then $A > B$.

The well-known [1] method of arithmetic comparison of numbers with RCS involves the conversion of numbers and type $A^{(H)} = (0, 0, \dots, \gamma_n)$, which requires $n - 1$ clock cycles of the null operation. In addition, it is necessary to make a positional comparison of the $(j_1 + 1)$ and $(j_2 + 1)$ intervals of the source operands A and B . All of this complicates the comparison algorithm and increases the comparison time of numbers, which leads to the need to develop a comparison method for RCS, which do not require determination of positional characteristics. Consider each of these methods.

2.1 The method of arithmetic comparison of numbers in the RCS (the method of arithmetic parallel subtraction)

We will consider comparable numbers in arbitrary intervals: $[jm_i, (j + 1)m_i)$, where

$$\overline{j = 1, N - 1} \left(N = \prod_{\substack{k=1 \\ k \neq i}}^n m_k \right).$$

In the case, the source operands A, B are reduced to numbers that are multiples to m_i , by modular subtraction of the following form:

$$A_{m_i} = A - a_i = (a_1^{(i)}, a_2^{(i)}, \dots, a_{i-1}^{(i)}, 0, a_{i+1}^{(i)}, \dots, a_n^{(i)}),$$

$$B_{m_i} = B - b_i = (b_1^{(i)}, b_2^{(i)}, \dots, b_{i-1}^{(i)}, 0, b_{i+1}^{(i)}, \dots, b_n^{(i)}),$$

where

$$a_i = (a'_1, a'_2, \dots, a'_i, a'_n), \quad b_i = (b'_1, b'_2, \dots, b'_i, b'_n).$$

Further, by means of a set of constants $0, m_i, 2m_i, \dots, (N-1)m_i$, represented by $(n-1)$ -y base of the RCS $m_1, m_2, \dots, m_{i-1}, m_{i+1}, \dots, m_n$, the construction of the so-called single-row code, respectively, in the form

$$K_N^{(n_A)} = \{z_N z_{N-1} \dots z_2 z_1\}, \quad z_{n_A} = 0 \quad (z_1 = 1; 1 = \overline{1, N}, 1 \neq n_A),$$

$$K_N^{(n_B)} = \{z'_N z'_{N-1} \dots z'_2 z'_1\}, \quad z'_{n_B} = 0 \quad (z'_1 = 1; 1 = \overline{1, N}, 1 \neq n_B).$$

The algorithm for constructing a single-row code in RCS can be represented as follows (1):

$$\left\{ \begin{array}{l} A_{m_i} - 0 = z_1, \\ A_{m_i} - m_i = z_2, \\ A_{m_i} - 2m_i = z_3, \\ \dots\dots\dots \\ A_{m_i} - \left(\prod_{\substack{\rho=1 \\ \rho \neq 1}}^{n-1} m_\rho - 1 \right) m_i = z_N, \end{array} \right. \quad \left\{ \begin{array}{l} B_{m_i} - 0 = z'_1, \\ B_{m_i} - m_i = z'_2, \\ B_{m_i} - 2m_i = z'_3, \\ \dots\dots\dots \\ B_{m_i} - \left(\prod_{\substack{\rho=1 \\ \rho \neq 1}}^{n-1} m_\rho - 1 \right) m_i = z'_N. \end{array} \right. \quad (1)$$

In the case, we have:

$$z_{n_A} = 0, \text{ at } A_{m_i} - n_A \cdot m_i = 0; \quad z_{n_A} = 1, \text{ at } A_{m_i} \neq n_A \cdot m_i;$$

$$z'_{n_B} = 0, \text{ at } B_{m_i} - n_B \cdot m_i = 0; \quad z'_{n_B} = 1, \text{ at } B_{m_i} \neq n_B \cdot m_i.$$

Geometrically, this number comparison method can be explained as follows. The interval $\left[0, \prod_{i=1}^n m_i\right)$ is divided into segments. The source operands A и B , by subtract-

ing the constants of the form $\prod_{\substack{\rho=1 \\ \rho \neq i}}^{n-1} m_\rho$, c and shift to the left edge of their hit interval.

This is the equivalent to bringing the compared numbers to numbers, which are compile to comparison m_i of the RCS. On this basis, the accuracy W c and comparison of the operands depends on the size of the base m_i , that is $W = W(m_i)$. However, with the maximum accuracy of the comparison $W_{max} = W(m_{min})$ (for an ordered of the RCS $W_{max} = W(m_1)$) the number of equipment for technical devices realizing the comparison operation in the RCS increases dramatically. Indeed, the number of equipment N_0 of the comparison device in the RCS significantly depends on the number of address N_i , performing the operation of parallel subtraction:

In this way, necessity of ensure a high degree of accuracy of comparison requires a significant amount of equipment, which reduces the efficiency of using existing methods of comparing numbers in the RCS. This situation determines the relevance and importance of finding more effective methods of comparing numbers in the RCS, ensuring a high accuracy of W comparison with a minimum numbers N_0 of equipment comparing devices.

In general, the task is interpreted as follows. It is necessary to find $N_0 = \min$ at W_{max} , i.e. $N_0(W_{max}) = \min$. As shown above $W_{max} = W(m_1)$. In this case, a change in the base m_i ($i = \overline{1, n}$) affects only the number N_i of the equipment of the group of adders. In this case, the problem is correctly formulated as a definition (2)

$$N_1(W_{max}) = \min. \quad (2)$$

Obviously, with high accuracy equal to the unit length of the interval, $W_{max} = W(m_i = 2)$. However, in this case $N_1(W_{max}) = \max$, и this result does not satisfy the condition (2). On the other hand – $N_1(W_{min}) = \min$. Thus, it is necessary to develop such a method of comparison (2). This problem is solved by the method described below.

2.2 The method of arithmetic comparison of numbers in the RCS (the method of parallel subtraction with the comparison of residues)

We introduce an additional operation of comparing the residues a_n and b_n the magnitude of the bases m_n of the RCS. In this case, the result of the comparison of the residuals simultaneously with the result of the comparison of the single-row code $K_N^{(nA)}$ and $K_N^{(nB)}$, is determined by the solution of the problem (2), i.e. W_{max} and with the minimum amount of equipment N_{min} is the result of solving the operation of

comparing two numbers in RCS. The algorithm for determining the result of an arithmetic comparison operation can be represented as follows:

$$\left\{ \begin{array}{l} \text{if } n_A > n_B, \text{ then } A > B; \\ \text{if } n_A < n_B, \text{ then } A < B; \\ \text{if } n_A = n_B, \text{ and at the same time} \\ \left\{ \begin{array}{l} a_n = b_n, \text{ at } A = B, \\ a_n > b_n, \text{ at } A > B, \\ a_n < b_n, \text{ at } A < B. \end{array} \right. \end{array} \right. \quad (3)$$

The set of relations (3) represents the general algorithm for the implementation of the operation of arithmetic comparison of numbers in the RCS. It is advisable to consider examples of specific performance of the operation of arithmetic comparison of numbers in the RCS. Let the RCS be given by bases, $m_1 = 2$, $m_2 = 3$ and $m_3 = 5$. Code words are given in Table 1. In Table 2 given constants $a_n(b_n)$ presented in a given RCS, and in Table 3, constants of a single-row code are $\Delta \cdot m_n$ $\left(\Delta = 0, \prod_{i=1}^{n-1} m_i \right) m_i$ on the basis of the RCS ($i = \overline{1, n-1}$).

Table 1. Table of code words in RCS

A	A in RCS			A	A in RCS		
	$m_1 = 2$	$m_2 = 3$	$m_3 = 5$		$m_1 = 2$	$m_2 = 3$	$m_3 = 5$
0	0	00	000	15	1	00	000
1	1	01	001	16	0	01	001
2	0	10	010	17	1	10	010
3	1	00	011	18	0	00	011
4	0	01	100	19	1	01	100
5	1	10	001	20	0	10	000
6	0	00	001	21	1	00	001
7	1	01	010	22	0	01	010
8	0	10	011	23	1	10	011
9	1	00	100	24	0	00	100
10	0	01	000	25	1	01	000
11	1	10	001	26	0	10	001
12	0	00	010	27	1	00	010
13	1	01	011	28	0	01	011
14	0	10	100	29	1	10	100

Table 2. Table of constants

γ_3	Constants
000	(0, 00, 000)
001	(1, 01, 001)
010	(0, 10, 010)
011	(1, 00, 011)
100	(0, 01, 100)

Example 1. Let the compared numbers be represented as $A_{23} = (1, 10, 011)$ and $B_{21} = (1, 00, 001)$. In this case, the values of the constants (Table 2) determine the values $A_{m_n} = A_{23} - a_n = (0, 10, 000)$, $B_{m_n} = B_{21} - b_n = (01, 00, 000)$, which corresponds to the shift of the operands A and B to the left edge of the interval $[20, 25)$. Then we use the constants of the single-row code (Table 3), we determine the single-row code for the input numbers in the form:

$$K_N^{(n_A)} = K_6^{(4)} = \{110111\}; K_N^{(n_B)} = K_6^{(n_B)} = K_6^{(4)} = \{110111\},$$

where

$$N_1 = \prod_{i=1}^{n-1} m_i = 6.$$

At the same time the comparison result $a_n = 011 > b_n = 001$ is determined in parallel with the $(n = \lfloor \log_2(m_n - 1) \rfloor + 1)$ -th bit comparison circuit in time. So, if $n_A = n_B = 4$, in accordance with the above algorithm, we determine that $A_{23} > B_{21}$.

Table 3. Table of constants

$(0 \div N - 1) \cdot m_n$	RCS		Number of zero position
	$m_1 = 2$	$m_2 = 3$	
$0 \cdot m_3 = 0$	0	00	1
$1 \cdot m_3 = 5$	1	10	2
$2 \cdot m_3 = 10$	0	01	3
$3 \cdot m_3 = 15$	1	00	4
$4 \cdot m_3 = 20$	0	10	5
$5 \cdot m_3 = 25$	1	01	6

Example 2. Let the compared numbers be represented as $A_{23} = (1, 10, 011)$, $B_3 = (1, 00, 011)$. In this case, the following differences are determined from the values of the constants (Table 2):

$$A_{m_n} = A_{23} - a_n = (0, 10, 00) \text{ and } B_{m_n} = B_3 - b_n = (1, 00, 000),$$

which corresponds to the shift of the operand A_{23} to the left edge of the interval $[20, 25)$, and the operand B_3 to the left edge of the interval $[0, 5)$. Next, using the constants of the single-row code (Table 3), we determine the single-row code for the considered input operands A_{23} , and B_3 :

$$K_N^{(n_A)} = K_6^{(5)} = \{101111\}, K_N^{(n_B)} = K_6^{(1)} = \{111110\}.$$

So, if $n_A = 5 > n_B = 1$, in accordance with the above algorithm, we determine that $A_{23} > B_3$.

2.3 The arithmetical method of comparison of numbers in the RCS (the method of comparison with a constant)

Consider the method of implementing the operation of arithmetic comparison of numbers in the RCS. The essence of this method is that not the operands A and B , are directly compared, but the quantities $\chi = (A - B) \bmod M = (\gamma_1, \gamma_2, \dots, \gamma_n)$ and m_1 .

In this case, the value is determined:

$$\chi_{m_1} = \chi - \gamma_1 = (0, \gamma'_2, \gamma'_3, \dots, \gamma'_n),$$

where constants $\gamma_1 = (\gamma_1, \gamma'_2, \gamma'_3, \dots, \gamma'_n)$ and $\gamma_1 = (a_1 - b_1) \bmod m_1$ are represented in a given RCS.

Then the general algorithm for comparison of the operands is presented in the form:

$$\begin{cases} A > B, \text{ at } \chi_{m_1} \leq m_1; \\ A < B, \text{ at } \chi_{m_1} > m_1; \\ A = B, \text{ at } \chi_{m_1} = 0. \end{cases} \quad (4)$$

By dialing constants

$$0, m_1, 2m_1, \dots, (N-1) \cdot m_1 \left(N = \prod_{i=1}^n m_i \right),$$

which were represented in the RCS with bases $m_2, m_3, \dots, m_n$, the construction of a single-row code in the form of:

$$K_n^{(n_\chi)} = \{z_N z_{N-1} \dots z_2 z_1\},$$

where

$$\begin{cases} \chi_{m_1} - 0 = z_1, \\ \chi_{m_1} - m_1 = z_2, \\ \chi_{m_1} - 2m_1 = z_3, \\ \dots \dots \dots \\ \chi_{m_1} - (N-1)m_1 = z_n. \end{cases}$$

In addition

$$z_{n_\chi} = 0, \text{ for } \chi_{m_1} - n_\chi m_1 = 0,$$

$$z_{n_\chi} = 1, \text{ for } \chi_{m_1} - n_\chi m_1 \neq 0.$$

The first module m_1 RCS is also represented by a single-row code of length N binary digits, in which the second place on the right will be zero ($m_1 - 1 \cdot m_1 = 0$), and on the rest - one, i.e. the single-row unitary code will be represented as: $K_n^{(2)} = \{11\dots 101\}$.

Further, by known methods, in accordance with algorithm (4), operands A and B , represented by a single-row code, are compared.

Example 3. For the above RCS we consider an example of the implementation of this method. Let the numbers be equal $A = (0, 01, 010)$ and $B = (1, 10, 011)$. We define the value:

$$\begin{aligned} \chi &= (A - B) \bmod M = [(0 - 1) \bmod 2, (01 - 10) \bmod 3, (010 - 011) \bmod 5] = \\ &= (1, 10, 100). \end{aligned}$$

By value $\gamma_1 = 1$ we define a constant in the form $\gamma_1 = (1, 01, 001)$ (Table 4). Then we perform the operation:

$$\chi_{m_1} = \chi - \gamma = (00, 01, 011).$$

The operand χ_{m_1} , which is a multiple of the module $m_1 = 2$ value, goes to the first inputs of the corresponding adders, the second inputs of which receive the corresponding constants (Table 5).

Table 4. Constants

γ_1	Constants
00	(0, 00, 000)
01	(1, 01, 001)

Table 5. Table of constants number of zero position

$(0 \div N-1) \cdot m_1$ $\Delta \cdot 2$	RCS		Number of zero position
	$m_2 = 3$	$m_3 = 5$	
0 $m_1 = 0$	00	000	1
1 $m_1 = 2$	10	010	2
2 $m_1 = 4$	01	100	3
3 $m_1 = 6$	00	001	4
4 $m_1 = 8$	10	011	5
5 $m_1 = 10$	01	000	6
6 $m_1 = 12$	00	010	7
7 $m_1 = 14$	10	100	8
8 $m_1 = 16$	01	001	9
9 $m_1 = 18$	00	011	10
10 $m_1 = 20$	10	000	11
11 $m_1 = 22$	01	010	12
12 $m_1 = 24$	00	100	13
13 $m_1 = 26$	10	001	14
14 $m_1 = 28$	01	011	15

Since $\chi_{m_1} - 14_{m_i} = 0$, then the single-row code will take the form:

$$K_n^{(n_\chi)} = K_{15}^{(14)} = \{011111111111111\}.$$

In accordance with algorithm (4), we determine that $A < B$.

The advantage of the considered method is to ensure maximum accuracy of comparison with an acceptable amount of equipment for its implementation.

2.4 The method of algebraic comparison of numbers in the RCS

It is easy to go from the implementation of the operation of arithmetic comparison of numbers to the algebraic comparison of numbers. In this case, the compared numbers A and B have one additional significant digit, i.e. the number is accompanied by a indication $\Omega_A(\Omega_B)$ of the sign $signA(signB)$, where:

$$\Omega_A(\Omega_B) = \begin{cases} 0, & \text{if } A(B) \geq 0, \\ 1, & \text{if } A(B) < 0. \end{cases}$$

In this way, the compared numbers are presented in the form:

$$A' = (\Omega_A; A) = \{\Omega_A; (a_1, a_2, \dots, a_n)\}, \text{ and } B' = (\Omega_B; B) = \{\Omega_B; (b_1, b_2, \dots, b_n)\},$$

and the method of comparing numbers A' and B' is determined by the set of operations (5).

We will conduct a comparative analysis of the implementation time of the comparison operation of two numbers A and B or the proposed method and the most well-known. The essence of the known method is to convert the numbers A and B from the system of residual classes to the positional number system A_{PNS} , B_{PNS} and the further comparison of the operands of A_{PNS} и B_{PNS} . Moreover, the transfer of numbers from RCS to PNS is made in accordance with the expression:

$$A_{PNS} = \left| \sum_{i=1}^n a_i B_i \right|_M,$$

where $a_i = [A_{PNS} / m_i] m_i$, where B_i – is the orthogonal basis over the m_i base of the RCS.

$$\left\{ \begin{array}{l} \text{if } n_A > n_B \left\{ \begin{array}{l} \Omega_A = 0, \Omega_B = 0, \text{ then } A' > B', \\ \Omega_A = 0, \Omega_B = 1, \text{ then } A' > B, \\ \Omega_A = 1, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 1, \Omega_B = 1, \text{ then } A' > B'; \end{array} \right. \\ \\ \text{if } n_A < n_B \left\{ \begin{array}{l} \Omega_A = 0, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 0, \Omega_B = 1, \text{ then } A' > B', \\ \Omega_A = 1, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 1, \Omega_B = 1, \text{ then } A' > B'; \end{array} \right. \\ \\ \text{if } n_A = n_B \left\{ \begin{array}{l} a_n = b_n \left\{ \begin{array}{l} \Omega_A = 0, \Omega_B = 0, \text{ then } A' = B', \\ \Omega_A = 0, \Omega_B = 1, \text{ then } A' > B', \\ \Omega_A = 1, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 1, \Omega_B = 1, \text{ then } A' = B'; \end{array} \right. \\ \\ a_n > b_n \left\{ \begin{array}{l} \Omega_A = 0, \Omega_B = 0, \text{ then } A' > B', \\ \Omega_A = 0, \Omega_B = 1, \text{ then } A' > B', \\ \Omega_A = 1, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 1, \Omega_B = 1, \text{ then } A' < B'; \end{array} \right. \\ \\ a_n < b_n \left\{ \begin{array}{l} \Omega_A = 0, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 0, \Omega_B = 1, \text{ then } A' > B', \\ \Omega_A = 1, \Omega_B = 0, \text{ then } A' < B', \\ \Omega_A = 1, \Omega_B = 1, \text{ then } A' > B'. \end{array} \right. \end{array} \right. \quad (5)$$

For the known and proposed methods of comparing numbers, the implementation time is determined by the corresponding mathematical relations

$$T_{RCS}^{(1)} = t^{(11)} + t^{(12)} + t^{(13)} + t^{(14)}, \quad (6)$$

$$T_{RCS}^{(2)} = t^{(21)} + t^{(22)} + t^{(23)}, \quad (7)$$

where

- $t^{(11)}$ – is the implementation time of the multiplication operation for the maximum m_n by the magnitude of the RCS module (the implementation time of the multiplication operation of two-digit binary numbers $K = \lceil \log_2(m_n - 1) \rceil + 1$);
- $t^{(12)}$ – is the time $(n-1)$ of the i th addition of two numbers of the type $a_i B_i + a_{i+1} B_{i+1} (i = \overline{1, n-1})$;
- $t^{(13)}$ – is the time for determining the deduction of the number of $A_{PNS} (B_{PNS})$;
- $t^{(14)}$ – is the comparison time of the positional operands of the A_{PNS} and B_{PNS} ;
- $t^{(21)}$ – is the time of implementation of the operation of subtraction in the RCS $A_{COK} - a_i$;
- $t^{(22)}$ – is the time of implementation of the operation of subtraction in the RCS $A_{m_i} - k \cdot m_i$;
- $t^{(23)}$ – is the time of comparison of two positional N -bit single-row unitary codes of two corresponding numbers.

It is known that the time of addition of t_c and t_y multiplication of two operands in the PNS is determined by the following relations:

$$t_c = \tau(2\rho - 1) \text{ and } t_y = 2\tau\rho^2, 0$$

where ρ is the bit width of the processed operands; τ – is the time of "shift" of one binary digit. In this case, the time of "triggering" of the logical element AND (OR) is determined by the expression: $t_{AND} \approx t_{OR} \approx \tau / 2$, and $t^{(23)} \approx 6t_n$ and expression

$$t^{(13)} \approx 2\tau \left\{ \left[\log_2 \left(\prod_{i=1}^n m_i - 1 \right) \right] + 1 \right\}^2.$$

Taking into account the above, relations (6) and (7), respectively, are presented in the form:

$$T_{RCS}^{(1)} = 2\tau k^2 + (n-1)\tau(2k-1) + 2\tau \left\{ \left[\log_2 \left(\prod_{i=1}^n m_i - 1 \right) \right] + 1 \right\}^2 + 3\tau; \quad (8)$$

$$T_{RCS}^{(2)} = 2\tau + 2\tau + 3\tau. \quad (9)$$

According to with expressions (8) and (9), values are calculated $T_{RCS}^{(1)}$, $T_{RCS}^{(2)}$ (table 6) for various l – byte bit CS grids ($l = \overline{1,4}$). From Table 6 it can be seen that with increasing length of the CS discharge grid, which is typical of the current trend in the development of systems and tools for processing digital information, the effectiveness of applying the proposed methods for comparing numbers, as compared to existing ones, increases.

Table 6. Comparative analysis of the implementation time of the comparison operation

T	l			
	1	2	3	4
$T_{RCS}^{(1)}$	324	870	1916	3334
$T_{RCS}^{(2)}$	6	6	6	6

On the basis of the proposed methods, algorithms for their implementation have been developed, in accordance with which a class of patentable devices has been developed for performing arithmetic and algebraic comparisons of numbers in an RCS [13-15]. Prospective direction of a further research is the argumentation of practical recommendations concerning a realization of the introduced method and the ways of its use in different mechanisms of an information security of telecommunications networks and systems [16-31].

3 Conclusion

In this paper, a method is proposed for comparing numbers in a non-positional number system of remainder classes, which is based on the principle of obtaining and comparing a unitary single-row code.

The developed method can be used to improve advanced computer systems and their components. In particular, its practical use allows to increase the performance of computer calculations and the reliability of information systems.

References

1. Shu, S., Wang, Y., Wang, Y.: A research of architecture-based reliability with fault propagation for software-intensive systems. In: 2016 Annual Reliability and Maintainability Symposium (RAMS), Tucson, AZ, 2016, pp. 1–6. (2016) doi:10.1109/RAMS.2016.7447984
2. Gokhale, S.S., Lyu, M.R., Trivedi, K.S.: Reliability simulation of component-based software systems. In: Proceedings Ninth International Symposium on Software Reliability Engineering (Cat. No.98TB100257), Paderborn, Germany, 1998, pp. 192–201. (1998) doi:10.1109/ISSRE.1998.730882

3. Tiwari, A., Tomko, K.: Enhanced Reliability of Finite State Machines in FPGA Through Efficient Fault Detection and Correction. *IEEE Transaction on Reliability*, vol. 54(3), pp. 459–467. (2005) doi:10.1109/TR.2005.853438
4. Singh, C., Sprintson, A.: Reliability assurance of cyber-physical power systems. In: *IEEE PES General Meeting*, Providence, RI, 2010, pp. 1–6. (2010) doi:10.1109/PES.2010.5590189
5. Krasnobayev, V.A., Koshman, S.A., Mavrina, V.A.: A Method for Increasing the Reliability of Verification of Data Represented in a Residue Number System. *Cybernetics and Systems Analysis*, 2014, vol. 50(6), pp. 969–976. (2014) doi:10.1007/s10559-014-9688-3
6. Reddy, M., Nalini, N.: FT2R2Cloud: Fault tolerance using time-out and retransmission of requests for cloud applications. In: *2014 International Conference on Advances in Electronics Computers and Communications*, Bangalore, 2014, pp. 1–4. (2014) doi:10.1109/ICAEECC.2014.7002396
7. Braun, C., Wunderlich, H.: Algorithm-based fault tolerance for many-core architectures. In: *2010 15th IEEE European Test Symposium*, Praha, 2010, pp. 253–253. (2010) doi:10.1109/ETSYM.2010.5512738
8. Radu, M.: Reliability and fault tolerance analysis of FPGA platforms. In: *IEEE Long Island Systems, Applications and Technology (LISAT) Conference 2014*, Farmingdale, NY, 2014, pp. 1–4. (2014) doi:10.1109/LISAT.2014.6845211
9. Akushsky, I., Yuditsky, D.: *Machine arithmetic in residual classes*. Moscow, Sov. Radio (in Russian) (1968)
10. Torgashov, V.: *System of residual classes and reliability of digital computers*. Moscow, Sov. Radio (in Russian) (1973)
11. Krasnobayev, V., Kuznetsov, A., Koshman, S., Moroz, S.: Improved Method of Determining the Alternative Set of Numbers in Residue Number System. In: *Recent Developments in Data Science and Intelligent Analysis of Information. ICDSIAI 2018. Advances in Intelligent Systems and Computing*, Springer, Cham, 2018, vol. 836, pp. 319–328. (2019) doi:10.1007/978-3-319-97885-7_31
12. Krasnobayev, V.A., Yanko, A.S., Koshman, S.A.: A Method for arithmetic comparison of data represented in a residue number system. *Cybernetics and Systems Analysis*, 2016, vol. 52(1), pp. 145–150. (2016) doi:10.1007/s10559-016-9809-2
13. Patent for utility model No. 73379, Ukraine, G 06F 7/04. Device for the popular danich, which is presented in the classroom extra. No. u 2012 01885. Appl. 02/20/2012. Publ. 09/25/2012, Bull. No. 18
14. Patent for utility model No. 73384, Ukraine, G 06F 7/04. Device for the same number of numbers in the classroom too. No. u 2012 01916. Appl. 02/20/2012. Publ. 09/25/2012, Bull. No. 18
15. Patent for utility model No. 73407, Ukraine, G 06F 7/04. Device for algebraic porous numbers with class numbers No. U 2012 02150. Appl. February 24, 2012. Publ. 09/25/2012, Bull. No. 18
16. Gorbenko, I.D., Zamula, A.A., Semenko, A.E., Morozov, V.L.: Method for complex improvement of characteristics of orthogonal ensembles based on multiplicative combining of signals of different classes. *Telecommunications and Radio Engineering*, 2017, vol. 76(18), pp. 1581–1594. (2017) doi:10.1615/TelecomRadEng.v76.i18.10
17. Si-ru, Z.: Application of Improved Genetic Algorithm in Optimization Computation. In: *2009 Fifth International Conference on Natural Computation*, Tianjin, 2009, pp. 144–148. (2009) doi:10.1109/ICNC.2009.393
18. Gorbenko, I., Zamula, A., Morozov, V.: Methods for implementing communications in info-communication systems based on signal structures with specified properties. In: *2017*

- 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2017, pp. 101–104. (2017) doi:10.1109/INFOCOMMST.2017.8246359
19. Kuznetsov, A., Kiyani, A., Uvarova, A., Serhiienko, R., Smirnov, V.: New Code Based Fuzzy Extractor for Biometric Cryptography. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 119–124. (2018) doi:10.1109/INFOCOMMST.2018.8632040
 20. Gorbenko, I., Kuznetsov, A., Tymchenko, V., Gorbenko, Y., Kachko, O.: Experimental Studies of the Modern Symmetric Stream Ciphers. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 125–128. (2018) doi:10.1109/INFOCOMMST.2018.8632058
 21. Eberhart, R.C., Shi, Y.: Tracking and optimizing dynamic systems with particle swarms. In: Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546), Seoul, South Korea, 2001, pp. 94–100. (2001) doi:10.1109/CEC.2001.9343760
 22. Ivanov, O., Ruzhentsev, V., Oliynykov, R.: Comparison of Modern Network Attacks on TLS Protocol. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 565–570. (2018) doi:10.1109/INFOCOMMST.2018.8632026
 23. Kandemir, M., Choudhary, A., Ramanujam, J., Kandaswamy, M.A.: A unified framework for optimizing locality, parallelism, and communication in out-of-core computations. IEEE Transactions on Parallel and Distributed Systems, 2000, vol. 11(7), pp. 648–668. (2000) doi:10.1109/71.877759
 24. Pousset, Y., Vauzelle, R., Combeau, P.: Optimising the computation time of radio coverage predictions for microcellular mobile systems. IEEE Proceedings - Microwaves, Antennas and Propagation, 2003, vol. 150(5), pp. 360–364. (2003) doi:10.1049/ip-map:20030543
 25. Xing, C., Wang, H.: Optimized similarity computation for nearest neighbor choosing. In: Proceedings of 2012 2nd International Conference on Computer Science and Network Technology, Changchun, 2012, pp. 2011–2014. (2012) doi:10.1109/ICCSNT.2012.6526313
 26. Munetomo, M., Bando, S.: A scalable infrastructure of interactive evolutionary computation to evolve services online with data. In: 2013 IEEE International Conference on Big Data, Silicon Valley, CA, 2013, pp. 28–28. (2013) doi:10.1109/BigData.2013.6691793
 27. Kavun, S.: Conceptual fundamentals of a theory of mathematical interpretation. Int. J. Computing Science and Mathematics, 2015, vol. 6(2), pp. 107–121. (2015) doi:10.1504/IJCSM.2015.069459
 28. Kavun, S.: Indicative-geometric method for estimation of any business entity. Int. J. Data Analysis Techniques and Strategies, 2016, vol. 8(2), pp. 87–107. (2016) doi:10.1504/IJDATS.2016.077486
 29. Kovtun, V., Kavun, S., Zyma, O.: Co-Z Divisor Addition Formulae in Homogeneous Representation in Jacobian of Genus 2 Hyperelliptic Curves over Binary Fields. International Journal Biomedical Soft Computing and Human Sciences, 2012, vol. 17(2), pp. 37–43. (2012) doi:10.24466/ijbschs.17.2_37
 30. Zamula, A., Kavun, S.: Complex systems modeling with intelligent control elements. Int. J. Model. Simul. Sci. Comput., 2017, vol. 08(01). (2017) doi:10.1142/S179396231750009X

31. Kavun, S., Zamula, A., Mikheev, I.: Calculation of expense for local computer networks. In: Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 2017 4th International, Kharkiv, Ukraine, 2017, pp. 146–151. (2017) doi:10.1109/INFOCOMMST.2017.8246369

Investigation of the Kolmogorov–Wiener filter for treatment of fractal processes on the basis of the Chebyshev polynomials of the second kind

Vyacheslav Gorev¹[0000-0002-9528-9497], Alexander Gusev²[0000-0002-0548-728X]

and Valerii Korniienko³[0000-0002-0800-3359]

Dnipro University of Technology, 19 Dmytra Yavornytskoho Ave, 49005 Dnipro, Ukraine

¹lordjainor@gmail.com

²gusev1950@ukr.net

³vikor7@ukr.net

Abstract. We consider the Kolmogorov–Wiener filter for continuous fractal processes with a power-law structure function. The corresponding filter is used for data forecast; the noiseless case is considered. The aim of the paper is to obtain the weight function for the corresponding filter based on the integral Wiener–Hopf equation. The problem under consideration is important, for example, for traffic forecast in telecommunication systems and for the forecast of the chemical composition of cast iron. An exact analytical solution for the corresponding equation meets difficulties, so an approximate solution is sought in the form of a truncated Chebyshev polynomial expansion. The Chebyshev polynomials of the second kind are used. The behavior of the polynomial solutions for different numbers of polynomials is investigated. The results are compared with the corresponding results of our previous paper where another polynomial set is used. It is found that the corresponding behavior is almost identical for different polynomial sets.

Keywords: Kolmogorov–Wiener filter weight function, continuous stationary fractal processes, power-law structure function, Chebyshev polynomials of the second kind.

1 Introduction

We consider the Kolmogorov–Wiener filter for data forecast for continuous fractal processes. Nowadays fractal processes take place in a huge variety of different systems, see, for example, [1–5] and various references in [5].

The problem of the Kolmogorov–Wiener weight function search for continuous fractal processes with a power-law structure function is stated in [4]. In that paper it is mentioned that such a model could be suitable for teletraffic description in IEEE 802.11b networks and for the routers between the internal networks and the Internet. In fact, the Wiener–Hopf integral equation is a Fredholm integral equation of the first kind. In [4] a simplified Volterra integral equation is used instead of the Fredholm one

and the idea of the solution for the Volterra integral equation is described. Finally, an exact analytical solution for the corresponding equation was obtained in [6].

Maybe, in some simplified cases the Volterra integral equation can indeed be applied to the investigation of data forecast in real systems. But in the general case it is not applicable, and the Fredholm integral equation should be solved instead of the Volterra one. In contrast to the Volterra integral equation, an exact analytical solution for a Fredholm integral equation meets difficulties. Thus, an approximate solution for the corresponding equation is sought. The method of a truncated orthogonal polynomial expansion is rather popular in the literature in order to obtain an approximate solution for the Fredholm integral equation of the first kind, see, for example, the corresponding investigation in the framework of statistical physics [7–10].

In paper [11] such a method was applied to the problem under consideration. A set of polynomials which are orthogonal without weight is used in [11]. It is shown that although the method can give reliable results in a rather wide range of parameters, it has some drawbacks in the case of a power-law structure function. The most significant drawback is the fact that the accuracy of the method does not necessarily increase with the number of polynomials. For some numbers of polynomials the method gives reliable results, but for other numbers it may fail. Most likely the reason is that the corresponding correlation function, which is the kernel of the Wiener–Hopf integral equation, is not a positively defined function, so the convergence of the method is not guaranteed, see a similar discussion in the framework of statistical physics in [12].

But, anyway, the question arises: may the results be better if we use another polynomial set? Is the behavior of the polynomial solutions identical for different sets of polynomials? This interesting question should be investigated because it is rather hard to propose another analytical method for the solution for the corresponding Wiener–Hopf equation. In this paper we use a set of the Chebyshev polynomials of the second kind. So, the aim of this work is to obtain the Kolmogorov–Wiener filter weight function on the basis of a truncated expansion in the Chebyshev polynomials of the second kind and to compare the results with the results of paper [11].

2 Description of the truncated polynomial expansion method

We consider stationary continuous fractal processes with a power-law structure function. The correlation function of such processes has the form [4]

$$R(t) = \sigma^2 - \frac{\alpha}{2} |t|^{2H} \quad (1)$$

where σ is the process variance, H is the Hurst exponent and α is a constant.

Let the filter input signal be defined for $t \in [0, T]$. As is known [13], in such a case the Kolmogorov–Wiener filter weight function $h(t)$ is the solution of the following Wiener–Hopf integral equation

$$\int_0^T d\tau h(\tau) R(t-\tau) = R(t+k) \quad (2)$$

where $k \leq T$ is the time interval for which the forecast is made. Such an equation can hardly be solved exactly, so an approximate solution should be found.

In paper [11] a truncated polynomial expansion method is used, and the following polynomials are taken:

$$S_n(\tau) = \frac{S'_n(\tau)}{\sqrt{\int_0^T dt (S'_n(t))^2}} \quad (3)$$

where

$$S'_n(\tau) = \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} \\ 1 & \tau & \tau^2 & \cdots & \tau^n \end{vmatrix}, \quad \mu_n = \int_0^T x^n dx = \frac{T^{n+1}}{n+1}. \quad (4)$$

Such polynomials are orthogonal without weight on $t \in [0, T]$:

$$\int_0^T dt S_n(t) S_m(t) = \delta_{mn} \quad (5)$$

where δ_{mn} is the Kronecker delta.

In this paper we take another polynomial set. We use the Chebyshev polynomials of the second kind. Their explicit expressions are [14]

$$U_n(x) = \sum_{k=0}^{[n/2]} C_{n+1}^{2k+1} x^{n-2k} (x^2 - 1)^k \quad (6)$$

where $[n/2]$ is the integer part of $n/2$. They are orthogonal on $x \in [-1, 1]$ with the orthogonality condition [14]:

$$\int_{-1}^1 U_n(x) U_m(x) \sqrt{1-x^2} dx = \frac{\pi}{2} \delta_{mn} = \begin{cases} \pi/2, & m = n \\ 0, & m \neq n \end{cases} \quad (7)$$

But we need a polynomial set that is orthogonal on $t \in [0, T]$. On the basis of (6) after making the following change of the variables:

$$z = x + 1, \quad y = zT/2 \quad (8)$$

one can derive the following expression:

$$\int_0^T U_n \left(\frac{2y}{T} - 1 \right) U_m \left(\frac{2y}{T} - 1 \right) \sqrt{1 - \left(\frac{2y}{T} - 1 \right)^2} dy = \frac{T\pi}{4} \delta_{mn}. \quad (9)$$

So the polynomials $U_n \left(\frac{2y}{T} - 1 \right)$ are orthogonal on $y \in [0, T]$ with the weight $\sqrt{1 - \left(\frac{2y}{T} - 1 \right)^2}$. So, an approximate solution of the integral equation (2) is sought in the form

$$h(\tau) = \sum_{s \geq 0} g_s U_s \left(\frac{2\tau}{T} - 1 \right). \quad (10)$$

After substitution of (10) into (2), one can obtain

$$\sum_{s \geq 0} g_s \int_0^T d\tau U_n \left(\frac{2\tau}{T} - 1 \right) R(t - \tau) = R(t + k) \quad (11)$$

which after multiplying by $U_n \left(\frac{2t}{T} - 1 \right)$ and integrating over t leads to

$$\sum_{s \geq 0} g_s \int_0^T \int_0^T dt d\tau U_n \left(\frac{2t}{T} - 1 \right) U_s \left(\frac{2\tau}{T} - 1 \right) R(t - \tau) = \int_0^T dt U_n \left(\frac{2t}{T} - 1 \right) R(t + k). \quad (12)$$

Denoting

$$G_{ns} = \int_0^T \int_0^T dt d\tau U_n \left(\frac{2t}{T} - 1 \right) U_s \left(\frac{2\tau}{T} - 1 \right) R(t - \tau), \quad b_n = \int_0^T dt U_n \left(\frac{2t}{T} - 1 \right) R(t + k) \quad (13)$$

one can rewrite (12) as

$$\sum_{s \geq 0} g_s G_{ns} = b_n, \quad n \geq 0. \quad (14)$$

As can be seen, (14) is an infinite set of linear equations in the unknown coefficients g_s . This set can hardly be treated, so it should be artificially truncated to a finite number of equations:

$$\sum_{s=0}^{l-1} g_s G_{ns} = b_n, \quad n = \overline{0, l-1}. \quad (15)$$

The Kolmogorov–Wiener filter weight function

$$h(\tau) = \sum_{s=0}^{l-1} g_s U_s \left(\frac{2\tau}{T} - 1 \right) \quad (16)$$

where the coefficients g_s are the solutions of (15) is the weight function in the l -polynomial approximation.

Here and in what follows the quantities G_{ns} are called the integral brackets. On the basis of (13) after making the following change of the variables:

$$x = \frac{2\tau}{T} - 1, \quad y = \frac{2t}{T} - 1 \quad (17)$$

one can obtain the following expression for the integral brackets:

$$G_{ns} = \frac{T^2}{4} \int_{-1}^1 \int_{-1}^1 dx dy U_n(x) U_s(y) R\left(\frac{T}{2}y - \frac{T}{2}x\right). \quad (18)$$

It should be stressed that such a choice of polynomials is rather convenient. As can be seen from (6), the polynomials $U_n(x)$ obey the property

$$U_n(x) = \begin{cases} U_n(-x), & n:2 \\ -U_n(-x), & n \not: 2 \end{cases} \quad (19)$$

By changing x to $-x$ and y to $-y$ in (18), on the basis of (19) it can be seen that $G_{ns} = 0$ if n and s are of different parity. This property takes place because the correlation function (1) is an even function. Also, the evenness of the correlation function leads to the fact that $G_{ns} = G_{sn}$. These two properties allow one to calculate G_{ns} by a straightforward calculation only for $n \geq s$ and n, s being of the same parity. Such a fact significantly reduces the computing time.

In the following section the numerical behavior of the l -polynomial approximation solutions is investigated.

3 Behavior of polynomial solutions

The behavior of the polynomial solutions is investigated for the parameters

$$T = 100, \quad k = 3, \quad \sigma = 1.2, \quad H = 0.8, \quad \alpha = 3 \cdot 10^{-3}. \quad (20)$$

First of all, this set does not contradict the inequality $|R(t)| \leq R(0)$. Secondly, this set is investigated in [11]. For the set (20) the numerical values for the coefficients in (16) in the l -polynomial approximations are given in Table 1.

Table 1. Numerical values for the coefficients of polynomials in (16) for parameters (20)

l	Coefficients $g_0, g_1, \dots, g_{l-1}$ rounded off to three significant digits
1	$4.86 \cdot 10^{-3}$.
2	$4.86 \cdot 10^{-3}, -1.46 \cdot 10^{-2}$.
3	$1.50 \cdot 10^{-2}, -1.46 \cdot 10^{-2}, -1.03 \cdot 10^{-1}$.
4	$1.50 \cdot 10^{-2}, -8.84 \cdot 10^{-3}, -1.03 \cdot 10^{-1}, -1.62 \cdot 10^{-2}$.
5	$-1.96 \cdot 10^{-3}, -8.84 \cdot 10^{-3}, 2.53 \cdot 10^{-2}, -1.62 \cdot 10^{-2}, 4.28 \cdot 10^{-2}$.
6	$-1.96 \cdot 10^{-3}, -5.84 \cdot 10^{-3}, 2.53 \cdot 10^{-2}, -1.37 \cdot 10^{-2}, 4.28 \cdot 10^{-2}, -1.64 \cdot 10^{-2}$.
7	$-1.41 \cdot 10^{-3}, -5.84 \cdot 10^{-3}, 1.22 \cdot 10^{-2}, -1.37 \cdot 10^{-2}, 2.71 \cdot 10^{-2}, -1.64 \cdot 10^{-2}, 3.05 \cdot 10^{-2}$.
8	$-1.41 \cdot 10^{-3}, -3.97 \cdot 10^{-3}, 1.22 \cdot 10^{-2}, -1.06 \cdot 10^{-2}, 2.71 \cdot 10^{-2}, -1.66 \cdot 10^{-2}, 3.05 \cdot 10^{-2}, -1.64 \cdot 10^{-2}$.
9	$-1.61 \cdot 10^{-3}, -3.97 \cdot 10^{-3}, 7.26 \cdot 10^{-3}, -1.06 \cdot 10^{-2}, 1.92 \cdot 10^{-2}, -1.66 \cdot 10^{-2}, 2.84 \cdot 10^{-2}, -1.64 \cdot 10^{-2}, 2.69 \cdot 10^{-2}$.
10	$-1.61 \cdot 10^{-3}, -2.68 \cdot 10^{-3}, 7.26 \cdot 10^{-3}, -8.16 \cdot 10^{-3}, 1.92 \cdot 10^{-2}, -1.44 \cdot 10^{-2}, 2.84 \cdot 10^{-2}, -1.85 \cdot 10^{-2}, 2.69 \cdot 10^{-2}, -1.63 \cdot 10^{-2}$.
11	$-1.90 \cdot 10^{-3}, -2.68 \cdot 10^{-3}, 4.40 \cdot 10^{-3}, -8.16 \cdot 10^{-3}, 1.42 \cdot 10^{-2}, -1.44 \cdot 10^{-2}, 2.40 \cdot 10^{-2}, -1.85 \cdot 10^{-2}, 2.95 \cdot 10^{-2}, -1.63 \cdot 10^{-2}, 2.51 \cdot 10^{-2}$.
12	$-1.90 \cdot 10^{-3}, -1.70 \cdot 10^{-3}, 4.40 \cdot 10^{-3}, -6.19 \cdot 10^{-3}, 1.42 \cdot 10^{-2}, -1.21 \cdot 10^{-2}, 2.40 \cdot 10^{-2}, -1.75 \cdot 10^{-2}, 2.95 \cdot 10^{-2}, -1.99 \cdot 10^{-2}, 2.51 \cdot 10^{-2}, -1.63 \cdot 10^{-2}$.
13	$-2.19 \cdot 10^{-3}, -1.70 \cdot 10^{-3}, 2.41 \cdot 10^{-3}, -6.19 \cdot 10^{-3}, 1.06 \cdot 10^{-2}, -1.21 \cdot 10^{-2}, 1.98 \cdot 10^{-2}, -1.75 \cdot 10^{-2}, 2.74 \cdot 10^{-2}, -1.99 \cdot 10^{-2}, 3.03 \cdot 10^{-2}, -1.63 \cdot 10^{-2}, 2.41 \cdot 10^{-2}$.
14	$-2.19 \cdot 10^{-3}, -9.32 \cdot 10^{-4}, 2.41 \cdot 10^{-3}, -4.59 \cdot 10^{-3}, 1.06 \cdot 10^{-2}, -9.94 \cdot 10^{-3}, 1.98 \cdot 10^{-2}, -1.56 \cdot 10^{-2}, 2.74 \cdot 10^{-2}, -1.99 \cdot 10^{-2}, 3.03 \cdot 10^{-2}, -2.10 \cdot 10^{-2}, 2.41 \cdot 10^{-2}, -1.62 \cdot 10^{-2}$.
15	$-2.46 \cdot 10^{-3}, -9.32 \cdot 10^{-4}, 9.03 \cdot 10^{-4}, -4.59 \cdot 10^{-3}, 7.78 \cdot 10^{-3}, -9.94 \cdot 10^{-3}, 1.62 \cdot 10^{-2}, -1.56 \cdot 10^{-2}, 2.44 \cdot 10^{-2}, -1.99 \cdot 10^{-2}, 3.01 \cdot 10^{-2}, -2.10 \cdot 10^{-2}, 3.10 \cdot 10^{-2}, -1.62 \cdot 10^{-2}, 2.34 \cdot 10^{-2}$.
16	$-2.46 \cdot 10^{-3}, -3.04 \cdot 10^{-4}, 9.03 \cdot 10^{-4}, -3.27 \cdot 10^{-3}, 7.78 \cdot 10^{-3}, -8.05 \cdot 10^{-3}, 1.62 \cdot 10^{-2}, -1.36 \cdot 10^{-2}, 2.44 \cdot 10^{-2}, -1.86 \cdot 10^{-2}, 3.01 \cdot 10^{-2}, -2.19 \cdot 10^{-2}, 3.10 \cdot 10^{-2}, -2.18 \cdot 10^{-2}, 2.34 \cdot 10^{-2}, -1.61 \cdot 10^{-2}$.
17	$-2.70 \cdot 10^{-3}, -3.04 \cdot 10^{-4}, -3.05 \cdot 10^{-4}, -3.27 \cdot 10^{-3}, 5.54 \cdot 10^{-3}, -8.05 \cdot 10^{-3}, 1.32 \cdot 10^{-2}, -1.36 \cdot 10^{-2}, 2.13 \cdot 10^{-2}, -1.86 \cdot 10^{-2}, 2.82 \cdot 10^{-2}, -2.19 \cdot 10^{-2}, 3.23 \cdot 10^{-2}, -2.18 \cdot 10^{-2}, 3.15 \cdot 10^{-2}, -1.61 \cdot 10^{-2}, 2.30 \cdot 10^{-2}$.
18	$-2.70 \cdot 10^{-3}, 2.22 \cdot 10^{-4}, -3.05 \cdot 10^{-4}, -2.15 \cdot 10^{-3}, 5.54 \cdot 10^{-3}, -6.41 \cdot 10^{-3}, 1.32 \cdot 10^{-2}, -1.16 \cdot 10^{-2}, 2.13 \cdot 10^{-2}, -1.69 \cdot 10^{-2}, 2.82 \cdot 10^{-2}, -2.12 \cdot 10^{-2}, 3.23 \cdot 10^{-2}, -2.35 \cdot 10^{-2}, 3.15 \cdot 10^{-2}, -2.24 \cdot 10^{-2}, 2.30 \cdot 10^{-2}, -1.61 \cdot 10^{-2}$.

The investigation is made up to the 18-polynomial approximation; the Wolfram Mathematica 11.0 package is used. The obtained weight function in each approximation is substituted into the integral equation (2) and the left-hand and the right-hand sides of the equation are numerically compared.

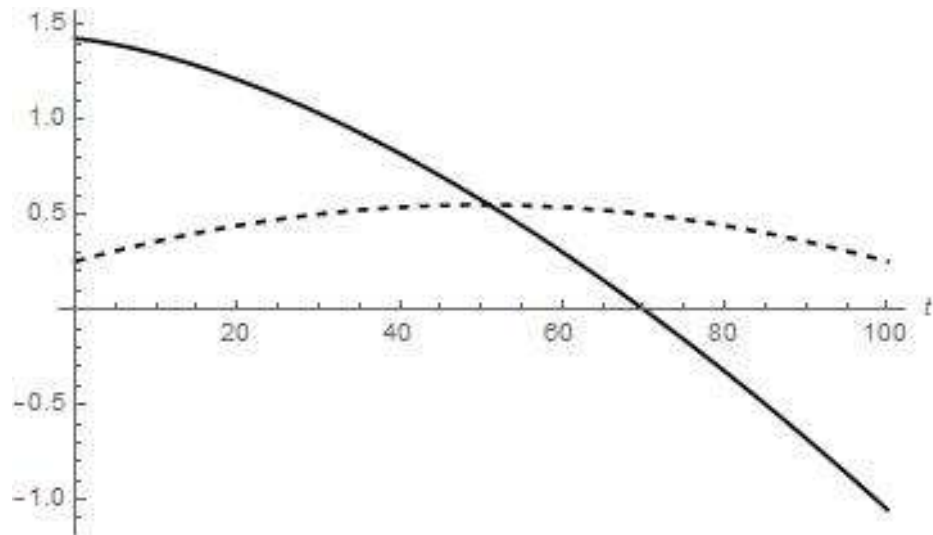


Fig. 1. Comparison of the left-hand and right-hand sides of eq. (2) for parameters (20) for the one-polynomial approximation.

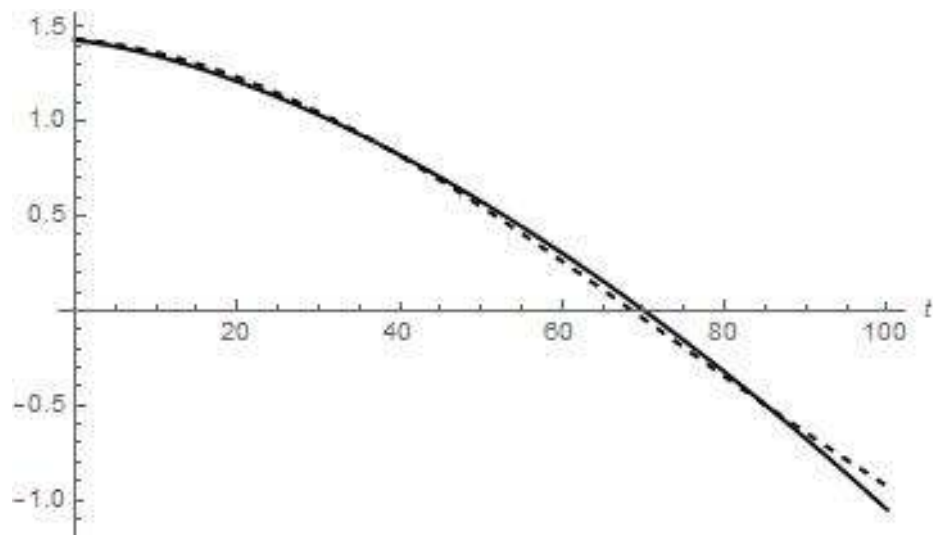


Fig. 2. Comparison of the left-hand and right-hand sides of eq. (2) for parameters (20) for the two-polynomial approximation.

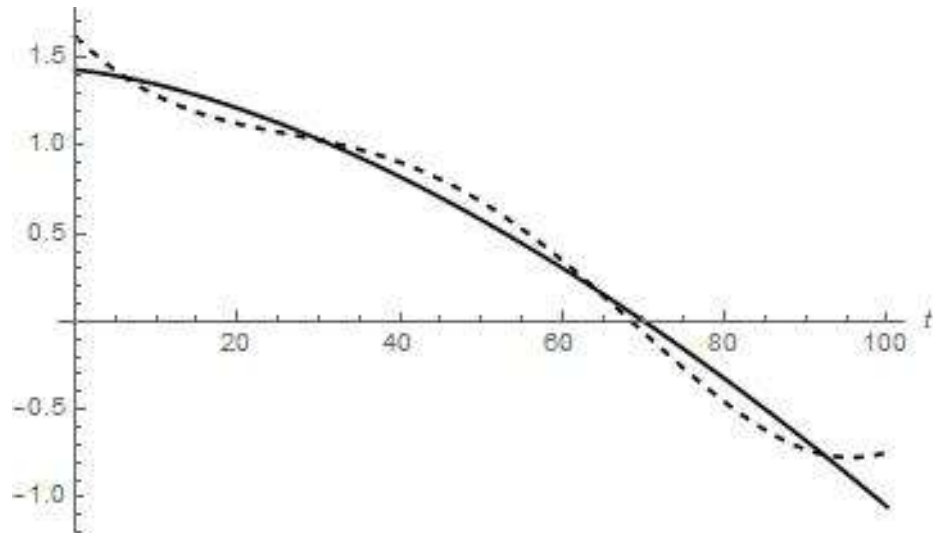


Fig. 3. Comparison of the left-hand and right-hand sides of eq. (2) for parameters (20) for the three-polynomial approximation.

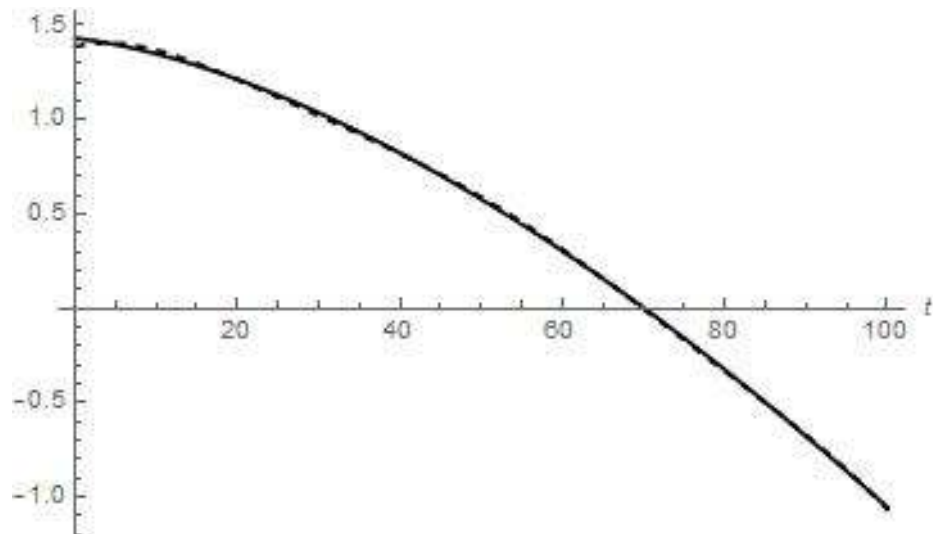


Fig. 4. Comparison of the left-hand and right-hand sides of eq. (2) for parameters (20) for the five-polynomial approximation.

As can be seen from Fig.1 – Fig.3, the one-polynomial approximation is not accurate, but the two-polynomial approximation is rather accurate. The corresponding graph for the four-polynomial approximation is not given because it is almost identical to the graph for the three-polynomial one. The four- and three-polynomial approximations

are worse than the two-polynomial one, but better than the one-polynomial one. The five-polynomial approximation is accurate (see Fig. 4). The accuracy slowly increases with the number of polynomials from the five- to the eight-polynomial approximations.

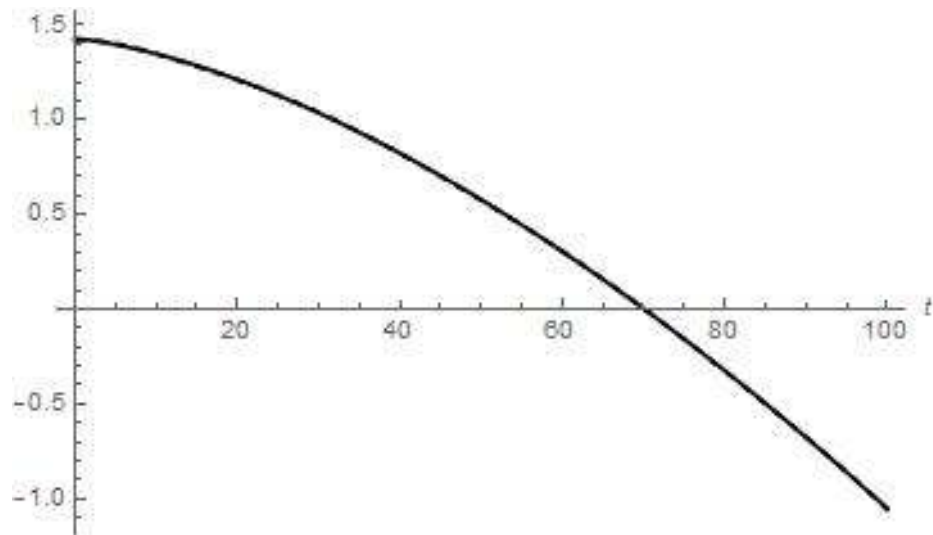


Fig. 5. Comparison of the left-hand and right-hand sides of eq. (2) for parameters (20) for the eight-polynomial approximation.

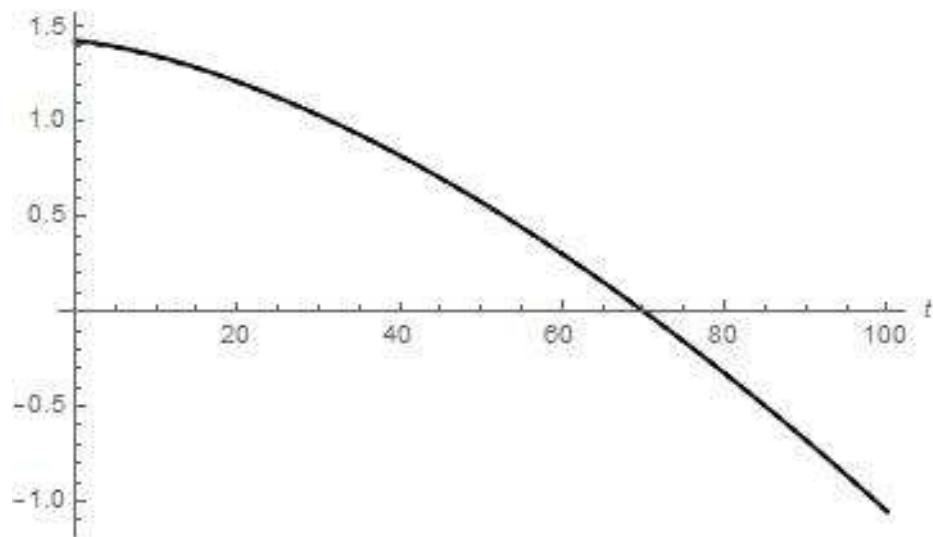


Fig. 6. Comparison of the left-hand and right-hand sides of eq. (2) for parameters (20) for the eighteen-polynomial approximation.

As can be seen from Fig. 5, the eight-polynomial approximation gives an almost ideal coincidence between the left-hand and the right-hand sides of the integral equation (2). But the approximations for the numbers of polynomials from nine to fifteen fail. For these approximations the curves for the left-hand and right-hand sides of eq. (2) are very far from each other. But the sixteen-, seventeen- and eighteen-polynomial approximations again give almost ideal results. The graphs for them are in fact identical and the graph for the eighteen-polynomial approximation is given in Fig. 6.

Such behavior of polynomial solutions is rather strange, but it can be explained as follows. The kernel of the integral equation (2) is not a positively defined function, so the convergence of the polynomial procedure is not guaranteed. In other words, the accuracy of the method does not necessarily increase with the number of polynomials.

It should be stressed that the behavior of the polynomial solutions described in [11] for the polynomial set (3) is, in fact, the same. The behavior of polynomial solutions is also investigated for the sets of parameters $T=10$, $k=3$, $\sigma=1.2$, $H=0.8$, $\alpha=10^{-1}$ and $T=1000$, $k=3$, $\sigma=1.2$, $H=0.8$, $\alpha=8\cdot 10^{-5}$. For the corresponding sets of parameters the behavior of the polynomial solutions for the polynomial sets (3) and (10) is almost identical. In [11] it is stressed that although the accuracy of the polynomial solutions may not increase with the number of polynomials and some approximations may fail, in a rather wide range of parameters (from $T=10$ to $T=1000$) some of the approximations give reliable results.

4 Conclusions

The Kolmogorov–Wiener filter weight function is investigated for continuous fractal processes with a power-law structure function. The method of a truncated orthogonal polynomial expansion is used in order to obtain an approximate solution of the corresponding Wiener–Hopf integral equation. In this paper the Chebyshev polynomials of the second kind which are orthogonal with weight on $t \in [0, T]$ are used. The numerical calculations are made on the basis of the Wolfram Mathematica 11.0 package.

It is found that the behavior of the polynomial approximations for the Chebyshev polynomials (10) and the behavior of the corresponding approximations for the polynomials (3), which is investigated in [11], are in fact the same. So, it may be concluded that the behavior of the polynomial solutions for the problem under consideration almost does not depend on the chosen polynomial set.

The proposed method of the approximate solution of the integral Wiener–Hopf equation for processes with a power-law structure functions has some drawbacks. The accuracy of the polynomial approximations may not increase with the number of polynomials, and some approximations may fail. This may happen because in such a case the kernel of the corresponding integral equation is not a positively defined function.

Nevertheless, in a rather wide range of parameters some polynomial approximations may give reliable results. Each approximation should be checked numerically before its further application to the investigation of data forecast in different systems.

References

1. Gusev, O., Kornienko, V., Gerasina, O., Aleksieiev O.: Fractal analysis for forecasting chemical composition of cast iron. In book "Energy Efficiency Improvement of Geotechnical Systems", Taylor & Francis Group, London, 225–231 (2016).
2. Korniienko, V., Matsyuk, S., Udovik. I., Alekseyev, A.: Method and algorithms of nonlinear dynamic processes identification. Scientific Bulletin of National Mining University, No. 1, 98–103 (2016).
3. Kornienko, V., Gerasina, A., Gusev, A.: Methods and principles of control over the complex objects of mining and metallurgical production. In book "Energy Efficiency Improvement of Geotechnical Systems", Taylor & Francis Group, London, 183–192 (2013).
4. Bagmanov, V. Kh., Komissarov, A. M., Sultanov, A. Kh.: Teletraffic forecast on the basis of fractal filters. Bulletin of Ufa State Aviation Technical University, Vol. 9, No. 6 (24), 217–222 (2007), in Russian.
5. Pipiras V., Taqqu, M.: Long-Range Dependence and Self-Similarity. Cambridge University Press (2017).
6. Gorev, V. N., Gusev, A. Yu., Korniienko, V. I.: On the analytical solution of a Volterra integral equation for investigation of fractal processes. Radio Electronics, Computer Science, Control, No. 4, 42–50 (2018)
7. Gorev, V. N., Sokolovsky, A. I.: Plasma kinetic coefficients with account for relaxation processes. International Journal of Modern Physics B, Vol. 29, No. 32, 1550233 (2015).
8. Gorev, V. N., Sokolovsky, A. I.: Hydrodynamic, kinetic modes of plasma and relaxation damping of plasma oscillations. Ukrainian Journal of Physics, Vol. 60, No. 3, 232–246 (2015).
9. Loyalka, S. K., Tipton, E. L., Tompson, R.V.: Chapman–Enskog solutions to arbitrary order in Sonine polynomials I: Simple, rigid-sphere gas. Physica A, Vol. 379, 417–435 (2007).
10. Loyalka, S. K., Tipton, E. L., Tompson, R.V.: Chapman–Enskog solutions to arbitrary order in Sonine polynomials II: Viscosity in a binary, rigid-sphere, gas mixture. European Journal of Mechanics B/Fluids. Vol. 28, 335–352 (2009).
11. Gorev, V. N., Gusev, A. Yu., Korniienko, V. I.: Polynomial solutions for the Kolmogorov–Wiener filter weight function for fractal processes. Radio Electronics, Computer Science, Control, No. 2 (2019), in press.
12. Ziman, J. M.: Electrons and Phonons. The Theory of Transport Phenomena in Solids. Oxford University Press (2001).
13. Miller, S., Childers, D.: Probability and Random Processes With Applications to Signal Processing and Communications. Second edition. Amsterdam, Elsevier/Academic Press (2012).
14. Gradshteyn, I. S., Ryzhik, I. M.: Table of Integrals, Series, and Products. Eighth edition. Edited by Daniel Zwillinger and Victor Moll. Amsterdam, Elsevier/Academic Press, (2015)

Modernized Mathematical Model of Text Document Classification

Tetiana Golub¹[0000-0001-6024-008X]

¹Zaporizhzhia National Technical University, Zhukovsky str.,64, Zaporizhzhia, 69063, Ukraine

golub.tv6@gmail.com

Abstract. The modernized mathematical model of the main stages of the text document classification is proposed. It takes into account the characteristics of certain categories. A mathematical description of the document data set creating stages, a document classification into categories is proposed. The principles of reducing the feature space dimension are described and the proposed method what used for determining the term weights is argued. The application of the method proposed in the article leads to reduce the analysis time of each document in order to make a decision about its category. This leads to decrease the resulting time for the analysis of the entire document set.

Keywords: text document classification, term vector, mathematical model, term weight, SLF parameter

1 Introduction

The information amount witch presented in text form increases continuously. Text information is accumulated in all areas of human activity. It is represented from data stored on personal computers to data in the form of Big Data. It covers such areas as business, research institutions, government and financial institutions that use technology intensively. Text information contains statistical data, control commands, reference information and principle laws of different processes. A feature of such information is the lack of its structuredness. It makes more complicated the process of its analysis [1].

Text analytics converts text into numbers. It allows organizing data and helps to identify patterns. Structured data are easier to analyze. Therefore, decisions made on their basis are more quality [2, 3].

If it is necessary to find the information in a data large amount, firstly it must be classified [4]. This process is the consideration subject in the proposed study.

Text classification refers to one of the computational linguistic tasks. It includes the definition of the text thematic affiliation, the text author, the statement emotional coloring and etc.

The task of organizing documents is solved to simplify the search for the necessary information. It is one of the most urgent tasks. Text classification is needed to solve

this problem. [5]. It is difficult to solve the classification problem because the data flow is constantly increasing. Therefore, its decision is relevant.

Many approaches to solving this problem are described in the literature. An overview and comparison of currently relevant methods are presented in accordance with the various stages of this process in [1, 6–8]. According to these sources one of the most important points of test classification is key feature selection. The works [9–15] were devoted to solving this problem. Various approaches, including statistical, frequency, latent-semantic and others are disclosed there. However, the described methods consider terms within the entire document collection. It is not possible to assess the importance of a separate term for each category separately.

The classification of text documents is the process of analyzing its content and automatically defining a document into one or several categories [16, 17]. Categories are sets of documents with a common theme. The set of categories is set by the expert or is determined automatically on the basis of the training sample. Automatic classifier is used in the information-analytical system at the stage of processing documents. An automatic classifier is a program that determines the subject of documents and assigns them to categories [6].

The inverse problem is also relevant. It consists of document selection from a document set according to the category defined by the user.

Presented in the literature mathematical models do not consider the term importance for certain categories. The author offers an improved mathematical model that takes into consideration this parameter.

The proposed in the article model considers this parameter which allows reducing the time for assessing the belonging of a document to certain categories by reducing the size of the term vector of certain categories for the text document classification.

2 Task formalization

The classifying document process in a formal form can be described as follows. The text document classification will be understood as the task of automatically defining a document into one or several categories based on its content. The category will be understood as a variety of documents with a general theme. Many categories are set by an expert or determined automatically using a training set. Automatic classifier is used in the information-analytical system at the document processing stage [6].

Mathematical models of the text document classification process given in [18, 19] are common. The author proposes the improvement of the existing variants of the term weight determining process as a part of the classification process with considering requirements of the task in this article.

It is proposed next designations to formally describe the process of text documents classifying:

- $T = \{t_1, \dots, t_{|A|}\}$ – document term set;
- $B = \{b_1, \dots, b_{|B|}\}$ – term set;
- $D = \{d_1, \dots, d_{|D|}\}$ – documents set;
- $C = \{c_1, \dots, c_{|C|}\}$ – category set;

— $E = \{e_1, \dots, e_{|E|}\}$ – category term set.

In the general case, the searching task of documents which corresponding to a particular category is following.

A set of documents D , from which it is necessary to choose those documents d_j , which most likely belong to the category c_i determined in advance from the set of categories C exists. The solution of this problem is considered in this article.

$$\Phi(d_j, c_i) = \begin{cases} 0, & \text{if } d_j \notin c_i \\ 1, & \text{if } d_j \in c_i \end{cases} \quad (1)$$

3 Document term set creating

The text document classification is performed using the analysis of the text document terms. A term is an intuitively defined expression of a formal language. It is the formal name of the object [2]. In this study, the term will be understood as the word obtained after stemming. Stemming is the reduction of a word to a certain normal form using the clipping of its endings and suffixes. The formation of terms is one of the tasks of the preprocessing stage.

The text is presented in the form of a document term set model for solving the classification problem. Each term has its own weight.

Text preprocessing is performed when determining whether a document belongs to any category, considering the importance of each term.

The preprocessing process has the following characteristics suggested by the author:

— $T \in B$ - all terms of the document are included in the set of possible terms;

— $E \in B$ - all terms of the category are included in the set of possible terms;

The set of elements of the sets T and E forms the set B . The sets T and E constitute the set B .

— $T_M = \langle n_1(t_1), n_2(t_2), \dots, n_{|T|}(t_{|T|}) \rangle$ – a multiset of the set T . It allows collecting the occurrence of the set elements several times.

The formation of a document multiset of one group, category, allows its power for each term determines. This parameter estimates the quantitative index of the term occurrence.

- $E_M = \langle n_1(e_1), n_2(e_2), \dots, n_{|E|}(e_{|E|}) \rangle$ – a multiset of the set E. It allows collecting the occurrence of the set elements several times.

Based on the category multiset, it is possible to determine the term indicators by the power of their occurrence. Analyzing this parameter it is possible to determine in how many categories of the collection the considered term occurs at least once. It allows distinguishing terms that are characteristic for all categories and are not characteristic to a particular category. These terms do not contain information for classification. Therefore it is possible to exclude them from the analyzed set.

Subsequent text processing is performed based on these characteristics.

All words that appear in the documents can be ordered in some way, for example alphabetically. Then, for each document it is possible to write out the entire set of weights matching the dictionary words. If some term is out of the document, then the weight will be zero. That is the vector will be:

$$d_i = (w_1, w_2, \dots, w_n), \quad (2)$$

where d_i - i -th document vector representation, w_i - weight of the i -th document term, n - the total number of different terms in all the documents of the collection – the power of set B [21]

4 Term weight identification

The term weight values of the set E for each category of the incoming set B are determined to assess the occurrence to the document category.

There are many methods to determine the term weights in the literature are presented. Some of them are:

- Boolean weight. $w = \text{sign}(\text{tf})$, i.e. 1 - if word occurs in the document, 0 - otherwise;
- $w = \text{tf}$ – number of word duplications in the document [3];
- $w = \text{tf}/\text{df}$ - the coefficient «tf•idf», i.e. the multiplication of the words occurrence frequency (tf), to the reciprocal value of the words occurrence frequency in all documents of the collection (inverse df). There are many options to define the weight value of the i -th term (w_{ij}) in the document d_j . One of the simplest options is the following: $w_{ij} = \text{tf} \cdot \log_{10}(1/\text{df})$. When the formulas «tf•idf» are used the problem of common words is solved – when the words with no meaning are of high weight;
- SLF parameter [3];
- Latent semantic analysis [7, 8].

The mentioned methods for determining the term weight values characterize that terms within a single document or within the entire collection as a whole. The importance and significance of a term within a single category is ignored in both cases.

The SLF parameter [3], used to determine the weight values of each term of the set E, compensates for this disadvantage. The parameter SLF is a coefficient that characterizes the assessment of terms with regard to their inclusion in the category. This

method considers the importance of each term for a particular category, unlike many other approaches to determining weight values.

The following parameters were defined to find the SLF parameter:

1. df_{tc} - the number of documents of category c , in which the term t occurs at least once;
2. N_d - the number of documents in the category c ;
3. NDF_{tc} - normalized frequency of occurrence of the term t in the category c . It is found as the ratio of the document number of the category c , in which the term t occurs at least once, to the number of documents in the category c . This estimate is local to the category.

$$NDF_{tc} = df_{tc} / N_c \quad (3)$$

4. SLF_t - logarithmic sum of the term t frequencies:

$$SLF_t = \log(|C| / \sum(NDF_{tc})) \quad (4)$$

The SLF_t indicator eliminates the imbalance between categories with small and with a large number of documents.

The SLF parameter for each term within the collection is determined according to the formula (5)

$$TFSLF_t = TF_t(E_{ti}) \cdot SLF_t, \quad (5)$$

where $TF_t(E_{ti})$ – the frequency of the term belonging to the set B . It is defined as the ratio of the certain term number occurrences to the total number of the document terms. Thus, the importance of the term t_i within a separate document d_j is estimated [9].

Vector B_T , with the weight coefficient values of the set T terms within the entire collection as a whole, will be obtained. In this case, the significance of terms of a particular category is not fully considered. It reduces the quality indicators of the classification implementation of texts belonging to similar in meaning and used words topics.

The SLF parameter considers the term importance for categories within the collection, but does not take into account the importance of terms for each category separately. The following modification of the term weight definition based on the given parameter is proposed by the author for solving this problem.

The author proposes a sequence of actions for defining non-informative terms for each category individually based on the SLF parameter and statistical data. And further removal of these terms from the term vector of a separate category.

The sequence of actions to determine the weight values of the terms of the set E of each category for each category term e_i :

- the coefficient tf/df for each category term e_i is determined;
- the value of the weight of each term by categories is determined;
- uncharacteristic terms for each category are identified and removed.

The coefficient TF for each category term e_i within the collection as a whole is defined as the ratio of the total number of each term within a separate category to the total number of each term within the collection as a whole (6).

$$TF(t_i, c_j) = \frac{fr_{ij}}{\sum_i fr_{ij}}, \quad (6)$$

where $0 \leq i \leq |E|$, $0 \leq j \leq |C|$.

The importance of the term τ_i within a single document d_j is evaluated. [14]

The weight of each term by category, taking into account its occurrence in collection categories (the set E , containing the $CTFSLF(t_i, c_j)$ values of each category term) is defined as the product of the TF coefficient for each term of individual categories and the SLF parameter:

$$CTFSLF(t_i, c_j) = TF(t_i, c_j) * SLF_k, \quad (7)$$

where $0 \leq i \leq |E|$, $0 \leq j \leq |C|$, $0 \leq k \leq |B|$.

The $CTFSLF$ method for determining the term weights makes it possible to take into consideration the term importance within a particular category.

5 The feature space dimension reducing

The computational complexity of various classification methods directly depends on the feature space dimension. Therefore, the stage of the used term number reducing, or the stage of reducing the dictionary size of the category $|B|$ to $|B'| < |B|$, often is performed for classification problem solving.

The purpose of this stage is to reduce the data set dimension. This goal is achieved by removing uninformative for classifying terms. It allows decreasing the data size, to reduce the computing power requirements of the algorithm [4].

In this case, each documents terms vector undergoes the following preliminary processing:

- elimination of stop words (often used and not carrying a semantic load such as unions) [5];
- performing a morphological analysis of words [5];
- using clustering methods [6].

The following method of terms vector size reducing on the basis of the modernization described previously is proposed by author. It consists of the stage of determining non-characteristic terms for separate categories and the stage of their remove.

The value of K_j is calculated to determine the threshold value. The value of K_j is calculated as the inverse value of the number of documents which belongs to the analyzed categories. It is used to remove non-informative terms.

The term weight describes the property of its belonging to certain category. Terms that are found in all categories are low weight. Terms whose weights are below threshold are excluded.

$$K_j = \frac{1}{|D_j|} \quad (8)$$

where $0 \leq j \leq |C|$.

Further, the weight value is compared with the threshold value for each collection term.

If the value of the term weight is less than the threshold, this value is equaled to zero:

$$\Psi(e_i, c_j) = \begin{cases} 0, & \text{if } e_i < k_i \\ CTFSLF(t_i, c_j), & \text{if } e_i \geq k_i \end{cases} \quad (9)$$

where $0 \leq i \leq |E|$, $0 \leq j \leq |C|$.

The given analysis allows us to identify and exclude from the analysis such terms with low informativeness, as often encountered in the categories in the document corpus, and which are not informative for classification.

Thus, the removal of the terms distinguished from the feature space as a result of the analysis will reduce the length of the analyzed set and simplify the classification task.

The resulting term vector is used to search for documents belonging to a particular category, using the classification process.

6 Document classification into categories

In general, the task of classifying documents into categories is to find the maximum sum value of the term weighted coefficients that coincide with the terms characterizing a separate category.

The following parameter is introduced by author to evaluate this indicator.

W – a set that indicates the degree which shows this document falls into a separate category. A set is defined as the intersection of the document set T and the corresponding categories set E . All terms that are included in both sets are included in the set W .

$$W = T \cap E \quad (10)$$

The estimated value of the belonging degree of document to a separate category can be defined as the sum of the products of the set W elements by the corresponding weight values Ψ for terms belonging to the set T .

Then the degree of document compliance to a separate category can be determined as follows.

$$NW_d = \sum_i W(t_i) \cdot TFSLF(t_i, e_j) \quad (11)$$

where NW_d – the normalized value, the degree of coincidence of the term set belonging to category T to the term set of category E.

When a document and category match, this parameter will have a maximum value relative to other categories, and when comparing a document with a foreign category, the match will be observed mainly only for common words that can be attributed to several categories and whose significance decreases with increasing number of these categories.

7 Classification stage time reducing

The application of the method proposed by the author will reduce the spent time at the classification stage.

According to the property of additivity, the resulting value of time spent on the classification of the n documents is equal to the sum of time spent on the classification of each document separately. That is, the resulting value of time is determined by adding the individual time spent on the classification of each document. It is proposed next designations:

- A –total number of documents for classification;
- $S = \{s_1, \dots, s_{|A|}\}$ –the set containing the time spent on the classification of each document analyzed sample;
- $S1$ - the set containing the time spent on the classification of each document analyzed sample using based method;
- $S2$ - the set containing the time spent on the classification of each document analyzed sample using proposed method.

The total time to perform classifications of all documents using the methods $S1$ and $S2$ is determined:

$$S_\Sigma = \sum_i s_i \quad (12)$$

According to the properties of commutativity and associativity for the addition operation, the elements of the sets $S1$ and $S2$ can be grouped into two groups. The first group consists of the sum of expenditure time equal in total value for both sets. The second group consists of the summands whose total values differ. If the different total values from the second group of the sample $S2_i$ are less than the different total values of the sample $S1_i$, then it can be argued that the sum of the sample $S2$ is less than the sum of the sample $S1$ that is presented in (13).

$$\text{if } s1_i > s2_i \text{ then } S1_\Sigma > S2_\Sigma \quad (13)$$

Thus, analyzing the obtained results, it can be argued that the shorter the time spent on implementing the classification process of each document separately, the shorter the time value of implementing the classification as a whole. Since reducing the time spent on the classification of a certain document leads to a decrease in the time spent on the classification as a whole. So, this task is relevant.

8 Proposed method testing

The task of document classifying by individual categories of class 004 " Computer science and technology. Computing. Data processing" of the UDC classifier was selected for testing the proposed method. Certain categories are:

- 004.0 " Special auxiliary subdivision for computing",
- 004.2 " Computer architecture",
- 004.4 "Software",
- 004.9 " Application-oriented computer-based techniques".

30 documents of each category were used as a training sample. Categories of documents were determined by their authors. Testing was conducted on unused for training documents for each category. The training and testing results are shown in tables 1-2.

Table 1. Term vector size after learning stage

Category	Words in docum	SLF		CTFSLF		Ex-cluded words	De-creasing part
		Terms in vector	Term part	Terms in vector	Term part		
004.0	148419	22118	14,90%	18450	12,43%	3668	16,58%
004.2	111213	12510	11,25%	8978	8,07%	3532	28,23%
004.4	108077	18752	17,35%	14652	13,56%	4100	21,86%
004.9	104207	17411	16,71%	13473	12,93%	3938	22,62%
Average result	—	—	15,05%	—	11,75%	3809	21,53%

Table 2. Spent time for testing stage

Category of document	Time for SLF, s	Time for CTFSLF, s	Decreasing time, s	Decreasing part of time
004.0	0,03125	0,02500	0,006251	20,00%
004.2	0,018751	0,01250	0,006249	33,33%
004.4	0,021877	0,021875	0,000002	0,01%
004.9	0,028126	0,015627	0,012499	44,44%
Summary /	0,100004	0,075003	0,025001	24,44%

average result				

As can be seen from table 1, the terms average proportion of the words in documents total number according to the original SLF method is 15.05%. The proposed CTFSLF method shows a result of 11, 75%. The average number of terms excluded from each category is 21.53%. As a result, the average time for determining the category of a document was reduced by 24.44% (table 2). This shows the promise of the proposed method.

9 Conclusions

Thus, this article a modernized mathematical model of the text document classification main stages taking into account the characteristics of certain categories proposed. A mathematical description of the document data set creating stages for a document classification into categories is proposed. The principles of reducing the feature space dimension are described and the proposed method using for determining the weights of terms is argued.

The purpose of the proposed approach is to identify and exclude non-informative terms for a particular category, i.e. leave inherent informative terms that characterize the category. The using of this approach leads to reduce the amount of computations performed for searching in the general collection of documents belonging to a particular category. As a result, the analysis time to classification of certain document is reduced. This leads to reduce the resulting time for analyzing the entire set of documents.

References

1. Thangaraj M., Sivakami M.: Text classification techniques: A literature review. *Interdisciplinary Journal of Information, Knowledge, and Management*, 2018, vol. 13, pp. 117-135 (2018)
2. Brindha S., Sukumaran S., Prabha, K.: A survey on classification techniques for text mining. 3rd International Conference on Advanced Computing and Communication Systems. IEEE. Coimbatore, Indi., (2016) doi: 10.1109/ICACCS.2016.7586371
3. Daud A., Li J., Zhou L.: Muhammad F. Knowledge discovery through directed probabilistic topic models: a survey. *Frontiers of computer science in China*, 2010, vol. 4, no. 2, pp. 280–301 (2010)
4. Korde V., Mahender N.: Text classification and classifiers: a survey. *International Journal of Artificial Intelligence & Applications (IJAIA)*, 2012, Vol. 3, no. 2, pp. 85–99 (2012)
5. Pankov S. V., Shebanin S. P., Ribakov A. A.: Thematic classification of text. *ROOKEE, ROMIP 2010, Kazan', Russia*, 2010, pp. 142-147 (2010)
6. Golub T. The Analysis of text documents classifiers constructing methods, *Modern problems of radio engineering, telecommunications, and computer science*, 2016, pp.742-745 (2016)
7. Yang Y., Zhang J., Kisiel B.: A scalability analysis of classifiers in text categorization. *ACM SIGIR'03*, (2003)

8. Sebastiani F.: Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 2002, vol. 34, pp. 1-47 (2002)
9. Karpovich S.N.: Multi-valued text documents classification using probabilistic thematic modeling ml-PLSI. *SPIIRAS Proceedings*. 2016. vol. 4(47), pp.92 – 104. (2016) doi: 10.15622/sp.47.5.
10. Kuralegov I.: Automatic classification of documents based on latent semantic analysis. 1st International Conference Digital Libraries: Advanced Methods and Technologies, Digital Collections, St-Petersburg, Russia, 1999, pp. 89-96. (1999)
11. Andreev A. M.: Automatic classification of text documents using the neural network algorithms and semantic analysis. *Advanced Methods and Technologies, Digital Collections*, St-Petersburg, Russia, 2003, pp. 76-86. (2003)
12. Krasnov A., Ilatovskiy A.S., Khomonenko A.D., Arsen'yev V.N.: Evaluation of documents semantic proximity based on latent-semantic analysis with automatic selection of rank values. *SPIIRAN proceedings*, 2017. no. 5(54), pp. 185-204 (2017)
13. Rehman Abdur, Barbi H., Saeed M., Feature Extraction for Classification of Text Documents. *International Conference on Communications and Information Technology (ICCIT 2012)*, Hammamet, Tunisia, 2012, pp. 234 - 239. (2012)
14. Budanitsky A. Hirst G.: Evaluating WordNet-based Measures of Lexical Semantic Relatedness *Computational Linguistics*. 2006. Vol. 32. pp. 13-47 (2006)
15. Bondarchuk D.V. Vector model of knowledge representation based on semantic proximity of terms]. *Bulletin of SUSU. Series: Computational Mathematics and Computer Science*. 2017. vol. 6 no. 3. pp. 73–83 (2017) doi: 10.14521/cmse170305.
16. Tsoumakas G., Katakis I.: Multi-label classification: an overview. *International Journal of Data Warehousing & Mining*. 2007. vol. 3(3). pp. 1–13 (2007)
17. Rubin T.N., Chambers A., Smyth P., Steyvers M.: Statistical topic models for multilabel document classification. *Machine Learning*. 2012. vol. 88. no. 1–2. pp. 157–208 (2012)
18. Erpev A.S.: Automatic classification of text documents. *Mathematical Structures and Modeling*. 2010, vol. 21, pp. 65-81 (2010)
19. Zyuz'kov V. M.: *Mathematical logic and theory of algorithms*. Tomsk, El Content (2015)
20. Willett P. The Porter Stemming Algorithm: Then and Now *Program: Electronic Library and Information Systems*. 2006. vol. 4, no. 4. pp. 219-223 (2006)
21. Golub T.V., Tyahunova M.YU.: The method of Ukrainian language stitemming for the classification of documents based on Porter's algorithm. *Scientific papers of the Donetsk National Technical University. Series: Informatics, Cybernetics and Computing* 2017, no. 1, pp. 59 – 63 (2017)
22. Oliynyk YU. O., Katyushchenko D. O.: Analysis of the methods of determining the text documents signs weight. *Scientific Review*, 2018, 3(46), pp. 112 – 123 (2018).

Discrete Signals with Special Correlation Properties

Alexandr Kuznetsov ¹[0000-0003-2331-6326], Diana Kovalchuk ¹[0000-0002-4499-3732],

Kateryna Kuznetsova ¹[0000-0002-5605-9293], Mykola Pastukhov ³[0000-0002-7784-4725],

Oleksii Smirnov ²[0000-0001-9543-874X], Dmytro Prokopovych-Tkachenko ⁴[0000-0002-6590-3898]

¹V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, 61022, Ukraine,
kuznetsov@karazin.ua, dianakovalhyk@ukr.net

²Central Ukrainian National Technical University, avenue University, 8, Kropivnitskiy, 25006,
Ukraine, dr.smirnova@gmail.com

³National Technical University «KPI», 2, Kyrpychova str., 61002, Kharkiv, Ukraine,
denart66@gmail.com

⁴University of Customs and Finance, st. Volodymyr Vernadsky, 2/4, Dnipro, 49000, Ukraine,
omega2@email.dp.ua

Abstract. The methods of synthesis of discrete signals are analyzed with the special cross-correlation properties: m- sequences, signals of Legendre, Barker, Paley-Plotkin, Gold, small and large set of Kasami. Comparative researches of properties of the formed discrete signals are conducted. Separate direction develops in development of methods of forming of discrete signals, that is based on the use of algebraic and structural properties of circular shifts of group codes. It is shown that offered approach allows forming the great numbers of discrete sequences, ensemble and cross-correlation properties of that are set by the properties of the corresponding group controlled from distance, structural and cyclic.

Keywords. Discrete signals, cross-correlation and ensemble properties, group codes, circular shifts.

1 Introduction

The effective functioning of digital communication networks with providing plural access on the technology of code channel separation directly depends on ensemble, cross-correlation and structural properties of the formed discrete signals [1-3]. It is special it is important at the construction of perspective digital communication of fifth (5G) and sixth (6G) generation networks [4].

The perspective direction of researches is the development of methods of synthesis of discrete signals with the special correlation properties [5-30]. The values of lateral ejection of the function of correlation for such signals are determined by strict analytical correlations and directly related to structural and group properties of ensembles of discrete sequences [5-9]. In particular, such signals include [5–30]: m-sequences, Legendre, Barker, Paley-Plotkin, Gold signals, a small and large number of Kasami

and many others. At the same time, the main disadvantage of such methods is the small power of the ensembles of the formed sequences. In this sense, the most promising methods are the synthesis of large ensembles of discrete signals with a multi-level auto- and cross-correlation function [16].

In this paper, the analysis and comparative studies of methods of synthesis of discrete signals with special correlation properties are carried out. Theoretically, synthesis methods based on the cross-section of circular shifts of the group code are justified, which allow forming a set of sequences with predetermined distance properties and algebraically construct large ensembles of discrete signals with a multilevel function of auto- and cross-correlation.

2 Analysis of the known methods of synthesis of discrete signals with the special correlation properties

The most development to date was got by the methods of synthesis of discrete signals, based on the use of recurrent transformations and corresponding linear and nonlinear recurrent sequences [5-30]. Procedures of forming of such sequences easily will be realized with the use of the simplest switch charts with shift registers.

Linear recurrent sequences are formed with the use of shift registers with linear feedback (LFSR) and at the corresponding choice of the function of feedback allow to provide the maximal period of the formed sequences [5-9]. In literature, such signals got the name of linear recurrent sequences of a maximal period (MLRS) or m-sequences. On their basis, many other classes of signals are formed: sequences of Legendre, Paley-Plotkin, a Barker signal and many other.

The results of a comparative analysis of some methods for synthesizing signals with special correlation properties are given in Table 1. The comparison was made according to the following indicators: the period length n and the power M of discrete signals, the maximum value of the lateral lobe module of the correlation function ρ .

The analysis showed that the lateral ejections of the correlation function of the considered signals take finite, previously known values, which allows them to be used at various stages of digital communication, including for channel synchronization and in radio-location. The main disadvantage is the small capacity of the ensembles of the formed sequences. For example, the number of MLRS, Legendre, Paley-Plotkin and other sequences is determined by the number of irreducible polynomials (Euler function), which determine the rule for the formation of sequences. Improved ensemble properties are possessed by Gold's sequences, small and large sets of Kasami sequences. The power of the ensembles of such signals is significantly increased. The lateral lobes of the correlation function ρ for these sequences are also increased, but with an increase in the length of the sequences n the loss in the correlation properties is insignificant. Therefore, the construction of large ensembles of discrete signals is a promising direction for further research.

Table 1. Ensemble and correlation characteristics of some discrete signals with special properties

Signal class	n	M	ρ
m-sequences	$2^m - 1,$ $m \in \mathbb{Z}^+$	$\varphi(2^m - 1) / 2m$	$\rho = -\frac{1}{2^m - 1}$
Legendre Sequences	$n \leq (p^m - 1),$ $p, m \in \mathbb{Z}^+$	$M \leq \varphi(p^m - 1) / 2m$	$\rho = \frac{p^{m-1}}{n}$
Legendre sequences with $m = 2$, Barker signals	$p + 1,$ $p = 3, 5, 7, 11, 13$	$\frac{p+1}{8} \frac{\varphi(p^2 - 1)}{\varphi(p - 1)}$	$\rho = 1 / p$
Legendre Sequences, $m = 2$	$p + 1, p \in \mathbb{Z}^+,$ $p \neq 3, 5, 7, 11, 13$	$\frac{p+1}{8} \frac{\varphi(p^2 - 1)}{\varphi(p - 1)}$	$\rho = 1 / \sqrt{p}$
Paley-Plotkin signals	$p, p \in \mathbb{Z}^+$	p	$\rho = 1 / p$
Gold signals	$2^m - 1,$ $m = 2p + 1, p \in \mathbb{Z}^+$	$2^m + 1$	$\rho = \frac{1 + 2^{(m+1)/2}}{2^m - 1}$
Gold signals	$2^m - 1,$ $m = 2p, p \in \mathbb{Z}^+$	$2^m + 1$	$\rho = \frac{1 + 2^{(m+2)/2}}{2^m - 1}$
Small set of Kasami	$2^m - 1,$ $m = 2p, p \in \mathbb{Z}^+$	$2^{m/2}$	$\rho = \frac{1 + 2^{m/2}}{2^m - 1}$
Large set of Kasami	$2^m - 1,$ $m = 2p, p \in \mathbb{Z}^+$	$2^{m/2} (2^{m/2} + 1)$	$\rho = \frac{1 + 2^{(m+2)/2}}{2^m - 1}$

The most important, in this sense, are methods based on the use of algebraic and structural properties of group codes. Thus, in [16, 29, 30], it was shown that the sub-orthogonal discrete signals, the three-level Gold signals are a special case of n-level discrete sequences formed by the section of cyclic orbits of a group binary code, and can be analytically formalized using the mathematical apparatus of the theory finite fields and, in particular, the theory of rings of polynomials.

3 Algebraic and structural properties of circular shifts of group codes

The proposed approach to the formation of discrete signals with a multilevel correlation function is based on the use of the algebraic and structural properties of cyclic orbits of group codes over finite fields, as well as the procedure for selecting the corresponding discrete sequences [8, 16]. Consider the algebraic structure of a finite field and the cyclic orbits contained in it. We will research the algebraic and structural properties of cyclic orbits of group codes to form discrete sequences with special properties.

We fix a finite field $GF(q)$, consider the vector space $GF^n(q)$ as a set of n -sequences of elements from $GF(q)$ with component-wise addition and multiplication by a scalar. A linear (n, k, d) code V is a subspace $GF^k(q)$ in space $GF^n(q)$, i.e. nonempty set of n -sequences (code words) over $GF(q)$, k is the dimension of a linear subspace, d is the minimum code distance (the minimum weight of a nonzero code word). A cyclic code is a special case of a subspace that has the additional property of cyclicity. Each vector from $GF^n(q)$ can be represented by a polynomial in a formal variable x of degree not higher than $n-1$. The components of the vector are identified with the coefficients of the polynomial. The set of polynomials has the structure of a vector space, identical to the structure of the space $GF^n(q)$, as well as the structure of the ring of polynomials $GF(q)[x]/(x^n-1)$. In the ring of polynomials, the multiplication over polynomials is defined: $p_1(x) \cdot p_2(x) = R_{x^n-1}[p_1(x) \cdot p_2(x)]$, where $R_b[a]$ is the remainder of the polynomial a divided by the polynomial b . The cyclic shift on $\tau \in \{0, \dots, n-1\}$ elements in terms of polynomial algebra can be written as:

$$x^\tau \cdot p(x) = R_{x^n-1}[x^\tau \cdot p(x)]. \quad (1)$$

If the code words (n, k, d) of a code over $GF(q)$ are given in the form of polynomials, then code V is a subset of the ring $GF(q)[x]/(x^n-1)$. The code V is cyclic if, along with the code word $C(x)$ it also contains the polynomial $x \cdot C(x)$. The only nonzero given polynomial $g(x)$ of the smallest degree $r = n-k$ uniquely defines (n, k, d) the cyclic code over $GF(q)$ and is denoted by the generating polynomial, moreover $g(x) = \prod_i (x - \beta^i)$, where $\beta^i \in GF(q^m)$. It is connected with the check polynomial $h(x)$ by the relation $g(x) \cdot h(x) = x^n - 1$, or, equivalently, $R_{x^n-1}[g(x) \cdot h(x)] = 0$.

Consider the structure of a finite field $GF(q^m)$, as a set of polynomials of degree $\leq m$ with coefficients from $GF(q)$, i.e. polynomial ring structure $GF(q)[x]/(x^m-1)$. This ring with modulo irreducible polynomial operations is an extended Galois field $GF(q^m)$. Such a field consists of a set of classes of conjugate elements $\alpha^{i(q^s)}$, $s = 0, 1, \dots, m_i - 1$, where m_i is the smallest natural number, such that equality holds [8]:

$$iq^{m_i} = (i) \bmod (q^m - 1).$$

The algebraic structure of a finite Galois field is given in table 2. Each class of conjugate elements specifies (through the roots) the minimal polynomial $f_i(x)$. The product of all minimal polynomials $f_i(x)$ of a finite field $GF(q^m)$ defines the polynomial $(x^{q^m-1} - 1)$, i.e. we have the equality:

$$(x^{q^m-1} - 1) = \prod_{\forall i \in \{0, \dots, q^m-1\}} f_i(x) = \prod_{\forall i \in \{0, \dots, q^m-1\}} (x - \alpha^i),$$

where α is a primitive element of the field $GF(q^m)$, whence follows:

$$g(x) = LCM \left(\prod_j f_j(x) \right) = LCM \left(\prod_j \prod_{s=0}^{m_j} (x - \alpha^{j(q^s)}) \right),$$

$$h(x) = \frac{x^n - 1}{g(x)} = LCM \left(\prod_{i \neq j} f_i(x) \right) = LCM \left(\prod_{i \neq j} \prod_{s=0}^{m_i} (x - \alpha^{i(q^s)}) \right),$$

where LCM is least common multiples.

Table 2. Classes of adjoint elements and minimal polynomials

Adjoint elements					Minimal polynomials
α^0					$f_0(x) = (x - \alpha^0)$
α^1	α^q	α^{q^2}	...	α^{q^m}	$f_1(x) = f_q(x) = f_{q^2}(x) = \dots = f_{q^m}(x) =$ $= (x - \alpha^1) \cdot (x - \alpha^q) \cdot (x - \alpha^{q^2}) \cdot (x - \alpha^{q^m})$
...	...	...	...	...	...
α^i	α^{iq}	α^{iq^2}	...	α^{iq^m}	$f_i(x) = f_{iq}(x) = f_{iq^2}(x) = \dots = f_{iq^m}(x) =$ $= (x - \alpha^i) \cdot (x - \alpha^{iq}) \cdot (x - \alpha^{iq^2}) \cdot (x - \alpha^{iq^m})$
...	...	...	...	...	...

Let us consider the structure of the group (n, k, d) code V over $GF(q)$ from the point of view of the cyclic properties of the sequences that form it. We will use the

concept of a *cyclic orbit* V_ξ is a set of sequences with elements from $GF(q)$, equivalent to each other with respect to the cyclic shift operation, i.e. many such:

$$C_i = (c_0^i, c_1^i, \dots, c_{n-1}^i), \quad c_v^i \in GF(q) \quad \text{and} \quad C_j = (c_0^j, c_1^j, \dots, c_{n-1}^j), \quad c_v^j \in GF(q),$$

that equality holds:

$$(c_0^i, c_1^i, \dots, c_{n-1}^i) = (c_{(\tau) \bmod(n)}^j, c_{(\tau+1) \bmod(n)}^j, \dots, c_{(\tau+n-1) \bmod(n)}^j), \quad (2)$$

for any $\tau \in \{0, \dots, n-1\}$.

Expression (1) using (2) is expressed in terms of polynomial algebra:

$$p_i(x) = x^\tau \cdot p_j(x) = R_{x^n-1} \left[x^\tau \cdot p_j(x) \right],$$

$$p_i(x) = c_0^i + c_1^i x + \dots + c_{n-1}^i x^{n-1}, \quad p_j(x) = c_0^j + c_1^j x + \dots + c_{n-1}^j x^{n-1}.$$

Consider the set $W = GF^n(q)$ of all n -sequences with elements from $GF(q)$, which form the so-called "full code". The structure of a set is equivalent to a vector space $GF(q)$ with componential addition and multiplication by a scalar.

We divide the entire set W into subsets of orbits $V_0, V_1, \dots, V_L$, each of which contains a set of sequences equivalent to each other in relation to the cyclic shift operation. Thus, we obtain the decomposition of the vector space $GF(q)$ into sets of non-intersecting orbits (Fig. 1). In Fig. 1 $S_{i,j} \in GF^n(q), i=0, \dots, L, j=1, \dots, Z_i$ is schematically denoted by n -sequences as elements of the set W . All $S_{i,j}$ are grouped on the basis of equivalence in relation to the cyclic shift operation. Each group is a set V_i , all elements of the set V_i form i orbit of a set W . It is also obvious that each sequence $S_{i,j}$ can be represented in the form (2) as some sequence C_L in the notations introduced above. The power of the set V_i (the number of elements of the i orbit of the set W) is equal to Z_i . Obviously, for a set W formed by n -sequences with elements from $GF^n(q)$, the power of the set V_i for an arbitrary $i=0, \dots, L$ does not exceed n , i.e. $Z_i \leq n, i=0, \dots, L$, the number of nonzero orbits L is bounded below by the following value:

$$L \geq \frac{q^n - 1}{n}. \quad (3)$$

Under the zero is understood the orbit V_0 , consisting of one zero sequence S_{0,z_0} (a sequence consisting of only zero elements $GF(q)$), the corresponding number of sequences of the orbit $Z_0 = 1$.

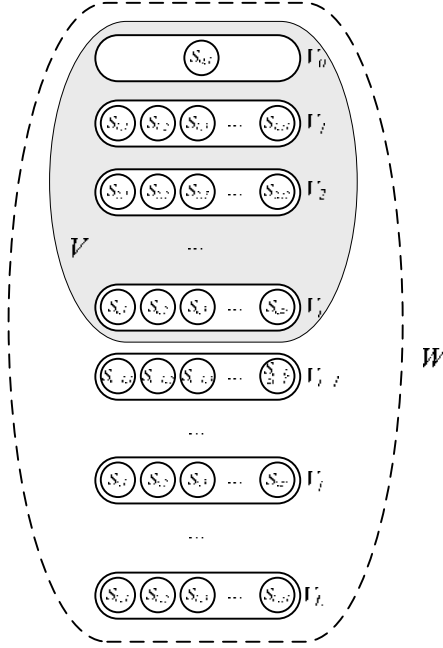


Fig. 1. A diagram of decomposition of vector space $GF^n(q)$ on the sets of not-intersecting circular shifts $V_i, i = 0, \dots, L$

Analyzing (2) we come to the conclusion that each orbit contains n - sequences with a fixed weight (the cyclic shift operation does not change the weight of the sequence). At the same time, the code V does not contain sequences of weight less than d . Strictly speaking, the weights of the sequences that form the code V , are determined solely by the weight spectrum of code V .

Suppose that the considered code V has a weight spectrum of the form:

$$A(w), \quad w=0, \dots, n,$$

where $A(w)$ is the number of code words of the V code with weight w .

It is obvious that for (n, k, d) code V the weight spectrum has the form:

$$\begin{cases} A(w) = 1, & w = 0; \\ A(w) = 0, & w = 1, \dots, d-1; \\ A(w), & w = d, \dots, n, \end{cases} \quad (4)$$

i.e. nonzero components of the weight spectrum $A(w) \neq 0$ (except for one zero sequence $A(0)=1$) are concentrated in the weight range $W=d, \dots, n$.

Taking into account the above considerations, the diagram for decomposing the vector space $GF^n(q)$ into sets of disjoint cyclic orbits $V_i, i = 0, \dots, L$ is represented as a diagram in Fig. 2.

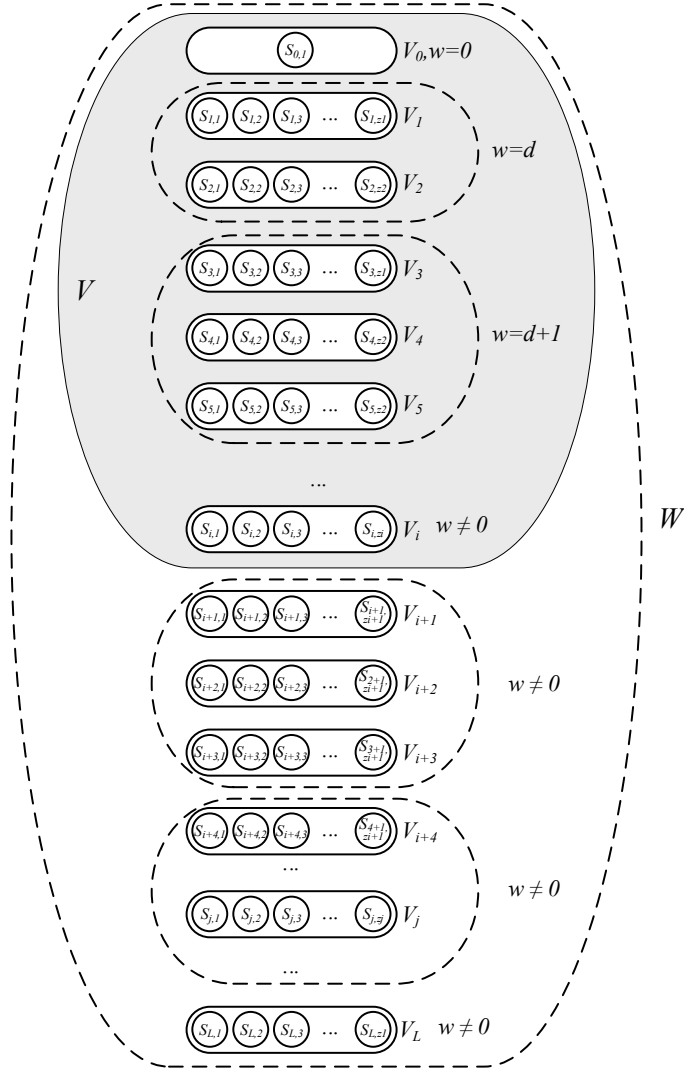


Fig. 2. The diagram of decomposition of vector space $GF^n(q)$ on the sets of not-intersecting circular shifts $V_i, i = 0, \dots, L$

Obviously, the «complete code» W is represented as a union of a finite number of disjoint orbit-sequences of fixed weight, equivalent to each other with respect to the cyclic shift operation. The code V as a subset of the space $GF^n(q)$ is the union of a finite number of orbits, and the weights of the sequences of the orbits are determined exclusively by the weight spectrum, i.e. relevant $w \neq 0$.

The number of orbits of a fixed weight is also determined by the weight spectrum of the code, i.e. the number of sequences in these orbits. Thus, the number of orbits of a weight $w \neq 0$ of a code V with a weight spectrum (4) is bounded below by an expression (by analogy with (3)):

$$L(w) \geq \frac{A(w)}{n}. \quad (5)$$

So, the group code is uniquely defined by the leaders (representatives) of its cyclic orbits.

4 Conclusions

The efficiency of functioning of digital communication systems with the provision of multiple access on the technology of code division of channels directly depends on ensemble, correlation and structural properties of used discrete signals. Promising in this sense are methods for the synthesis of discrete signals with special correlation properties, the magnitudes of lateral ejections of the correlation function of which are determined by strict analytical relations and are directly related to the structural properties of ensembles of discrete sequences.

The analysis of methods of synthesis of discrete signals with special correlation properties has shown that application of formed sequences allows providing the set level of noise immunity of communication. The lateral ejections of function of correlation of discrete signals with the special properties take on final beforehand known values, that allows to use them on the different stages of digital communication. At the same time, the main disadvantage of such methods is the low power of the ensembles of formed sequences. So, for example, the number of MLRS, sequences of Legendre, Paley-Plotkin and other is determined by the number of irreducible polynomials that define the rule for the formation of sequences. The improved ensemble properties are possessed by sequences Golda, small and large set of sequences of Kasami. Their construction is based on the use of the developed mathematical vehicle of the theory of the finite fields and, in particular, the theory of polynomial rings, which allows us to connect the correlation properties of the formed sequences with the group and structural properties of signal ensembles. The most perspective in this sense are the methods of synthesis based on the cross-section of circular shifts of the group codes.

Research of cyclic properties of group codes and presentation of them through the combination of circular shifts allowed to ground the rule of forming of discrete signals as leaders of corresponding circular shifts. Due to the cross-section of circular shifts, it is possible to form a set of sequences, the ensemble and correlation proper-

ties of which are determined by the distance, structural and cyclic properties of the corresponding group codes. The further researches can be focused on using in some different other areas [31-34].

References

1. Popoff, A.: *Fundamentals of Signal Processing in Metric Spaces with Lattice Properties: Algebraic Approach*. CRC Press (2017)
2. Benedetto, S., Biglieri, E.: *Principles of Digital Transmission: With Wireless Applications*, Springer (2002)
3. Miao, G., Zander, J., Sung, K.W., Slimane, B.: *Fundamentals of Mobile Data Networks*. Cambridge University Press (2016)
4. David, K., Berndt, H.: 6G Vision and Requirements: Is There Any Need for Beyond 5G? *IEEE Vehicular Technology Magazine*. (2018) doi:10.1109/mvt.2018.2848498
5. Barker, R.H.: *Group Synchronizing of Binary Digital Systems*. Communication Theory. London: Butterworth (1953)
6. Zierler, N.: Legendre sequences. Armed Services Technical Information Agency Arlington Hall Station, Massachusetts Institute of Technology Lincoln Laboratory, Group Report 34-71 (1958)
7. Plotkin, M.: Binary codes with specified minimum distance. *IRE Transactions on Information Theory*, 1960, vol. 6, pp. 445–450. (1960) doi:10.1109/TIT.1960.1057584
8. MacWilliams, F.J., Sloane, N.J.A.: *The Theory of Error-Correcting Codes*. North Holland Publishing Co. (1977)
9. Mather, F., Sloane, N.J.A.: *A Handbook of Integer Sequences*, Academic Press (1973)
10. Borwein, P., Mossinghoff, M.J.: Barker sequences and flat polynomials. In James McKee; Chris Smyth. *Number Theory and Polynomials*. LMS Lecture Notes. 352. Cambridge University Press (2008)
11. Naumenko, N.I., Stasev, Y.V., Kuznetsov, A.A.: Methods of synthesis of signals with prescribed properties. *Cybernetics and Systems Analysis*, 2007, vol. 43(3), pp. 321–326. (2007) doi:10.1007/s10559-007-0052-8
12. Stasev, Y.V., Kuznetsov, A.A., Nosik, A.M.: Formation of pseudorandom sequences with improved autocorrelation properties. *Cybernetics and Systems Analysis*, 2007, vol. 43(1), pp. 1–11. (2007) doi:10.1007/s10559-007-0021-2
13. Leung, K.H., Schmidt, B.: The Field Descent Method. *Designs, Codes and Cryptography*, 2005, vol. 36(2), pp. 171–188. (2005) doi:10.1007/s10623-004-1703-7
14. Hedayat, A.S., Sloane, N.J.A., Stufken, J.: *Orthogonal Arrays. Theory and Applications*. Springer-Verlag New York (1999)
15. Sloane, N.J.A., Plouffe, S.: *The Encyclopedia of Integer Sequences*. Academic Press, San Diego (1995)
16. Karpenko, O., Kuznetsov, A., Sai, V., Stasev, Y.: Discrete Signals with Multi-Level Correlation Function. *Telecommunications and Radio Engineering*, 2012, vol. 71(1), pp. 91–98. (2012) doi:10.1615/TelecomRadEng.v71.i1.100
17. Amadei, M., Manzoli, U., Merani, M.L.: On the assignment of Walsh and quasi-orthogonal codes in a multicarrier DS-CDMA system with multiple classes of users. *Global Telecommunications Conference. GLOBECOM '02. IEEE*, Taipei, Taiwan, 2002, vol. 1, pp. 841–845. (2002) doi:10.1109/GLOCOM.2002.1188196
18. Arambepola, B., Partington, K.C.: Walsh-Hadamard transform for complex-valued signals. *Electronics Letters*, 1992, vol. 28(3), pp. 259–261. (1992) doi:10.1049/el:19920160

19. Zulfikar, Abbasi, S.A., Alamoud, A.R.M.: A Novel Complete Set of Walsh and Inverse Walsh Transforms for Signal Processing. In: 2011 International Conference on Communication Systems and Network Technologies, Katra, Jammu, 2011, pp. 504–509. (2011) doi:10.1109/CSNT.2011.108
20. Gold, R.: Optimal binary sequences for spread spectrum multiplexing (Corresp.). IEEE Transactions on Information Theory, 1967, vol. 13(4), pp. 619–621. (1967) doi:10.1109/TIT.1967.1054048
21. Wojuola, O.B.: Cross-correlation index and multiple-access performance of Gold codes in a spread-spectrum system. In: 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon-si Gangwon-do, Korea (South), 2018, pp. 764–768. (2018) doi:10.23919/ICACT.2018.8323914
22. Gouda, M., El-Hennawy, A., Mohamed, A.E.: Detection of Gold Codes Using Higher-Order Statistics. In: 2011 First International Conference on Informatics and Computational Intelligence, Bandung, 2011, pp. 361–364. (2011) doi:10.1109/ICI.2011.65
23. Pereira, J.S., da Silva, A.J.H.: M-ary mutually orthogonal complementary gold codes. In: 2009 17th European Signal Processing Conference, Glasgow, 2009, pp. 1636–1640. (2009)
24. Huda, T.M.N., Islam, S.F.: Correlation analysis of the gold codes and Walsh codes in CDMA technology. In: 2009 First Asian Himalayas International Conference on Internet, Kathmandu, 2009, pp. 1–4. (2009) doi:10.1109/AHICI.2009.5340345
25. Kasami, T., Lin, S., Peterson, W.W.: Some Results on Cyclic Codes Which Are Invariant under the Affine Group and Their Applications. INFORMATION AND CONTROL, 1968, vol. 11(5-6), pp. 475–496. (1968) doi:10.1016/S0019-9958(67)90691-2
26. Chandra, A., Chattopadhyay, S.: Small Set Orthogonal Kasami codes for CDMA system. In: 2009 4th International Conference on Computers and Devices for Communication (CODEC), Kolkata, 2009, pp. 1–4. (2009)
27. Tsai, Y., Li, X.: Kasami Code-Shift-Keying Modulation for Ultra-Wideband Communication Systems. IEEE Transactions on Communications, 2007, vol. 55(6), pp. 1242–1252. (2007) doi:10.1109/TCOMM.2007.898859
28. Choi, J.H., Chung, H.K., Lee, H., Cha, J., Lee, H.: Code-Division Multiplexing Based MIMO Channel Sounder with Loosely Synchronous Codes and Kasami Codes. In: IEEE Vehicular Technology Conference, Montreal, Que., 2006, pp. 1–5. (2006) doi:10.1109/VTCF.2006.33
29. Kuznetsov, A.A., Smirnov, A.A., Sai, V.N.: Formation of discrete signals with a multilevel correlation function (in Russian). Processing Systems Information, Kharkiv, 2011, vol. 5(95), pp. 50–60. (2011)
30. Stasev, Y., Kuznetsov, A., Sai, V., Karpenko, O.: Discrete Signals with Multi-Level Correlation Function. Telecommunications and Radio Engineering, 2010, vol. 71(1), pp. 91–98. (2010) doi:10.1615/TelecomRadEng.v71.i1.100
31. Kavun, S.: Conceptual fundamentals of a theory of mathematical interpretation. Int. J. Computing Science and Mathematics, 2015, vol. 6(2), pp. 107–121. (2015) doi:10.1504/IJCSM.2015.069459
32. Kavun, S.: Indicative-geometric method for estimation of any business entity. Int. J. Data Analysis Techniques and Strategies, 2016, vol. 8(2), pp. 87–107. (2016) doi:10.1504/IJDATS.2016.077486
33. Zamula, A., Kavun, S.: Complex systems modeling with intelligent control elements. Int. J. Model. Simul. Sci. Comput., 2017, vol. 08(01). (2017) doi:10.1142/S179396231750009X

34. Kavun, S., Zamula, A., Mikheev, I.: Calculation of expense for local computer networks. In: Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 2017 4th International, Kharkiv, Ukraine, 2017, pp. 146–151. (2017) doi:10.1109/INFOCOMMST.2017.8246369

Automation Methods for Processing Medical Images Based on the Application of Grids

Liudmyla Vasylieva¹[0000-0002-9277-1560], Oleksandr Tarasov²[0000-0002-0493-1529]

Donbass State Engineering Academy, Akademichna str., 72, Kramatorsk, 84313, Ukraine

¹vasileva.dgma@gmail.com

²alexandrtar50@gmail.com

Abstract. The problem of reducing the complexity of the process of automated recognition and obtaining a quantitative assessment of the graphic objects parameters in medical images are considered. A method for applying a grid to an image is proposed. It makes possible to provide acceptable accuracy of calculations, a high level of generalization of data, as well as reduce the complexity of calculations. The software that implements the proposed methods is developed. The developed software is used in the analysis of metallographic images of materials for medical application.

Keywords: microstructure, image, processing automation, cluster, recognition

1 Introduction

Many devices for determining the medical parameters of a person represent information in graphical form. In particular, these are the results of studies using microscopes, the results of X-ray analysis, MRI, ultrasound, etc. In addition, medical instruments, implants, materials have a certain internal structure and surface structure, which must be controlled. Quality control in medicine plays an important role in ensuring the reliability of manufacturing prostheses, implants, medical instruments. Therefore, to obtain objective quantitative data in the processing of any information presented in graphical form, automation of image processing is required. Determining the quantitative characteristics of image elements is necessary for quick and correct interpretation of results. However, processing and analysis of medical, metallographic images is a very nontrivial task due to the complex shape and mutual arrangement of the elements, as well as the quality of the images provided by modern equipment.

Currently, effective methods of digital image processing have been developed [1, 2], including for the analysis of microstructures, surface coatings of products [3], etc. The complete process of image processing in analyzing the results of medical diagnostics and quality control of products is a complex multi-step time-consuming procedure and includes a number of basic steps: pre-processing and image restoration, segmentation, filtering, normalization of selected objects, recognition and, often, comparison with reference objects. Multistep is due to the fact that various processing tasks are closely related and are performed sequentially, therefore the quality of the

solution of one of them affects the choice of the method of solving the other problems [4]. Automation of image processing is the main problem in obtaining quantitative information about the results of the study.

The aim of the study is to reduce the complexity of the process of automated recognition and obtaining a quantitative assessment of the parameters of graphic objects in medical images.

2 Problem statement

To implement automated image processing for medical purposes, it is necessary to solve the following tasks:

- the analysis of methods, models and information technologies for recognition of graphic objects in images, microstructures of implants and instruments;
- the development of a method for reducing the amount of information analyzed through the use of a specified dimensions grid, which is superimposed on the image;
- the development of an algorithm for recognizing objects in an image using software implementation for automating successive stages of recognizing objects in an image, including clustering methods, binarization methods, building concave contours, and applying these methods to processing medical images;
- the experimental obtaining of quantitative characteristics of recognized images and their elements on the basis of the developed methods, algorithms and software.

3 Literature review

Currently, the development of computing technology has significantly increased the resolution of microscopes, in addition, it has become possible to share a microscope with a computer, so it is possible to create software systems for analyzing and processing metallographic images. To obtain image characteristics by means of modern information technologies, a number of algorithms are used, which can be divided into three categories: algorithms based on comparison with templates; algorithms that use the methods of decision-making theory; algorithms using neural networks, etc. Each of these groups of algorithms has its own strengths and weaknesses, different areas of application. Methods using patterns are inflexible, but because of their simplicity and low resource intensity, they often are used in solving particular recognition problems. The method of neural networks, which uses the principles of artificial intelligence, is more versatile, but requires a significant amount of input data and computational resources necessary for the organization of the learning process; its disadvantage is poor predictability of recognition results. Algorithms based on methods of decision theory, make the learning process simple and provide an opportunity to achieve a high degree of predictability of recognition results.

Among the existing solutions in the field of analysis and processing of metallographic images are the following specialized software systems. Image Expert Pro 3

allows you to get a wide range of geometrical parameters of the elements of the microstructure. The resulting characteristics are available both for each type of microparticles separately, and in the form of their statistical sampling [5]. The Fuzzy Clustering and Data Analysis Toolbox is a Matlab-based software package that includes three main categories of functions: clustering algorithms (K-means, K-medoid, FCMclust, GKclust and GGclust), analysis functions (Dunn, Alternative Dunn, Xie and Beni's, Partition Index), as well as data visualization functions (Sammon's method) [6]. Altami Studio allows real-time image capture from various devices (microscope or camera), has the ability to pre-process bitmap images (rotate, crop, change brightness, gamma, contrast). Altami Studio provides the ability to generate analysis reports in a user-friendly format [7]. The authors have developed algorithmic software for a specialized software package that implements the use of clustering methods for processing metallographic images [8].

It should be noted that there is no universal mathematical apparatus that allows you to create a common formalized approach to the construction of systems for analyzing metallographic images. Therefore, a systematic approach is needed to solve practical problems of image processing, and the development of interactive tools to automate the process of obtaining quantitative information in metallographic studies. Existing software systems for solving problems of processing and analysis of metallographic images use various methods of cluster analysis. The principle of the K-means method is to build k clusters located at the greatest distances from each other. The number of required clusters is specified manually by the user [9]. The EM algorithm searches for clusters based on hidden characteristics that the algorithm determines based on known data about each object. The main disadvantage of the algorithm can be called probable quasi-optimal solutions for its work [10]. The C-means method is a modification of K-means, but taking into account fuzzy clustering, i.e. objects are not strictly belong to any cluster, but have a certain value of belonging to each cluster [11]. Kohonen neural networks are a two-layer neural network in which the neurons of the output layer are cluster elements corresponding to the number of clusters into which the given objects are divided. One of the drawbacks of the method is the need to pre-tune and train the neural network [12].

4 The method of "overlaying the grid" in image processing

To solve the set tasks, the authors analyzed the characteristics of metallographic images, identified the problems of their processing during process automation, associated with the need to analyze a large number of pixels in the presence of actual and process noise in the images. The authors propose a method and an algorithm for extracting image regions using cluster analysis [13] and grids, which will make it possible to abandon the processing of each image pixel.

Image analysis showed the redundancy of information on the image associated, for example, with the presence of the background on which the structure elements are located. The essence of the proposal is to reduce the amount of information analyzed, to use a grid of specified dimensions, which is superimposed on the image, and to

analyze the image directly by grid pixels. The grid is built on the basis of the grid dimension entered by the user, which can be changed to ensure the best fit to the size of the elements selected in the image. Thus, the accuracy of determining the size of objects is regulated.

The next processing step is the selection of the desired points (pixels) in the image using a binarization algorithm. The binarization method implemented in this paper consists in comparing pixels in the RGB color model. Each pixel of the grid and the corresponding pixel on the image with a user-specified tolerance are compared. As a result of binarization, selected image pixels are obtained for further analysis.

Pseudocode of the proposed algorithm:

```
if (grid.R >= image.R - sens && grid.R <= image.R + sens)
if (grid.R >= image.R - sens && grid.R <= image.R + sens)
if (grid.R >= image.R - sens && grid.R <= image.R + sens)
return true;
```

where grid.R, grid.G, grid.B – a specified color in RGB format, image.R, image.G, image.B – RGB image color, sens – tolerance color parameters.

The algorithm was tested on a set of test color and grayscale images. To illustrate the operation, the result of the image binarization of the submicrocrystalline titanium BT 1-0 structure, obtained using transmission electron microscopy [14] and a color image obtained using an optical microscope is shown in Fig. 1. The search was performed according to the color of the object and the background color, with the use of the different grids sizes. An HSL color scheme has been added to analyze the image in grayscale. Each pixel was converted to HSL using standard formulas. Image analysis was performed within the grid.

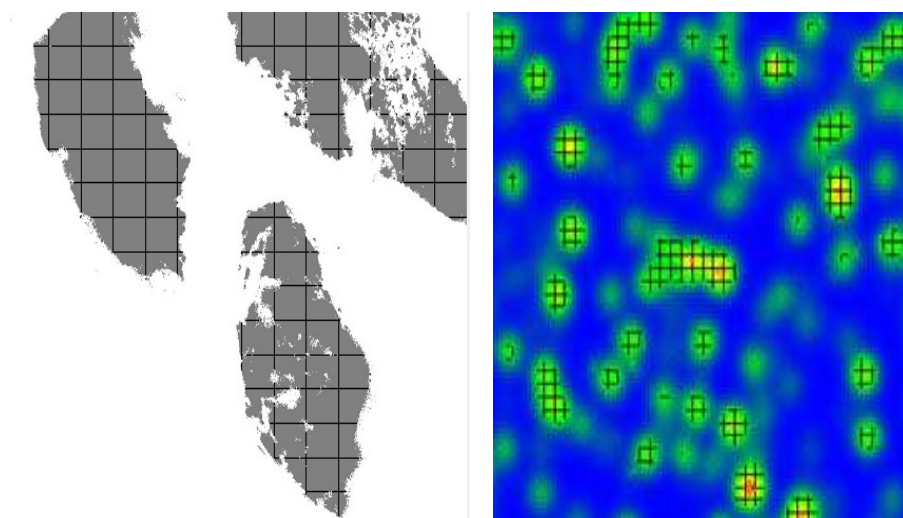
The results of the algorithm are obtained in the form of data arrays:

$$O^k = \{x_i, y_i, R_i, G_i, B_i, H_i, S_i, L_i, i = 1, 2, 3, \dots\}^k, k = 1, 2, 3, \dots,$$

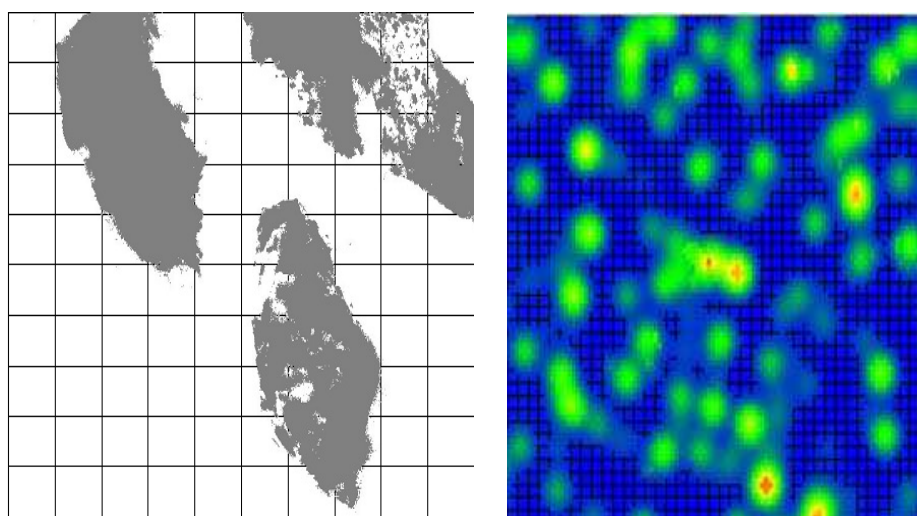
where k – the number of separate objects in the image; x_i, y_i – screen coordinates of the grid pixel; $R_i, G_i, B_i, H_i, S_i, L_i$ – RGB and HSL data for this pixel.

A simple contour can be constructed by connecting the “extreme” points of the grid superimposed on an object with the help of line segments. Analysis of the results of applying the algorithm on a set of test images showed that the amount of information processed is reduced by 4 – 25 times. This depends on the parameters of the elements in the image under study.

An important problem is the study of the effect of grid size on the quality of image processing and the exclusion of small elements of the original image. It is proposed to take the initial grid cell size no more than $d_x/5 \times d_y/5$, where d_x, d_y – the dimensions of the minimum picture element along the respective axes are. At this stage of work, the user for the whole image sets the grid size. The planned direction of the improvement of the algorithm is the splitting of the original image into separate zones.



a



b

Fig. 1. Selection of areas on the image using the “grid overlay” method with search by object color (a) and background color (b)

The resulting image points on the grid were used for the subsequent clustering algorithm. The algorithm was developed on the basis of the following requirements: the method should automatically determine the number of clusters based on

the data provided; each analysis method must give the same result with the same input data. To find the clusters, you need to specify the value of the Euclidean distance (*dist*) between the points inside the cluster. In this way, we can isolate the desired regions with different values of *dist* to control the results of cluster allocation. The use of the grid in the considered algorithm allows us to significantly speed up the process of extracting the desired image elements.

The next task of analyzing the selected image elements associated with obtaining quantitative characteristics is to determine their orientation relative to the horizontal axis. To do this, build a segment, the extreme points of which will be the most distant grid points p_1, p_2 in the cluster, then the angle of the selected image areas is determined. To determine the length, the square of the Euclidean distance $d^2 = (x.p_2 - x.p_1)^2 + (y.p_2 - y.p_1)^2$ is used, which optimizes the execution time of the analysis. The result of applying the algorithm analyzing the test image is shown in Fig. 2.

The next step to identify the elements in the image is to find the contour of each selected separate area. The problem of finding the convex contour of a set of points is well known and studied in detail. To solve it, there are many algorithms with different efficiency. The quick hull algorithm uses the idea of quick sort. The complexity of this algorithm in the average case is $O(N \cdot \log(N))$, where n is the number of starting points [15, 16]. Another well-known algorithm is Graham's three-step algorithm [17]. In the case of its implementation using quick sort, we obtain the total complexity of the algorithm $O(N \cdot \log(N))$. This also applies to the multi-step Jarvis algorithm [18]. The Divide-and-Conquer algorithm is a popular method for solving problems in computer science, it consists in dividing a problem into several parts, recursively looking for solutions in each of them and further integrating solutions to these problems into a common, complex solution. The obvious advantage of this approach is the possibility of paralleling the search for solutions to each individual task. A feature of the algorithm is that the algorithm for constructing contours itself can be performed using other, more optimal algorithms, with small volumes of initial points [19].

To solve the problem, an algorithm for constructing a concave contour was used. The idea of the algorithm is to first calculate the convex contour using any known algorithm, and then turn the convex contour into a concave [20]. The requirements for the algorithm are as follows: the algorithm must be stable; should be able to handle any set of points in size; the method should ensure the continuity of the contours, the absence of intersections between the contours and ensure that all points of the cluster fall into the selected contour. The speed of the algorithm of this part of the overall process is less important, given the basic tasks. In addition, from the user's point of view, the algorithm should be easy to use, allow control of the image processing process, the user is not required to enter a large number of additional parameters. A concave contour of a set of points can be defined as a shape that minimizes the area of a figure bounded by a contour, but allows any angle between adjacent edges of the contour.

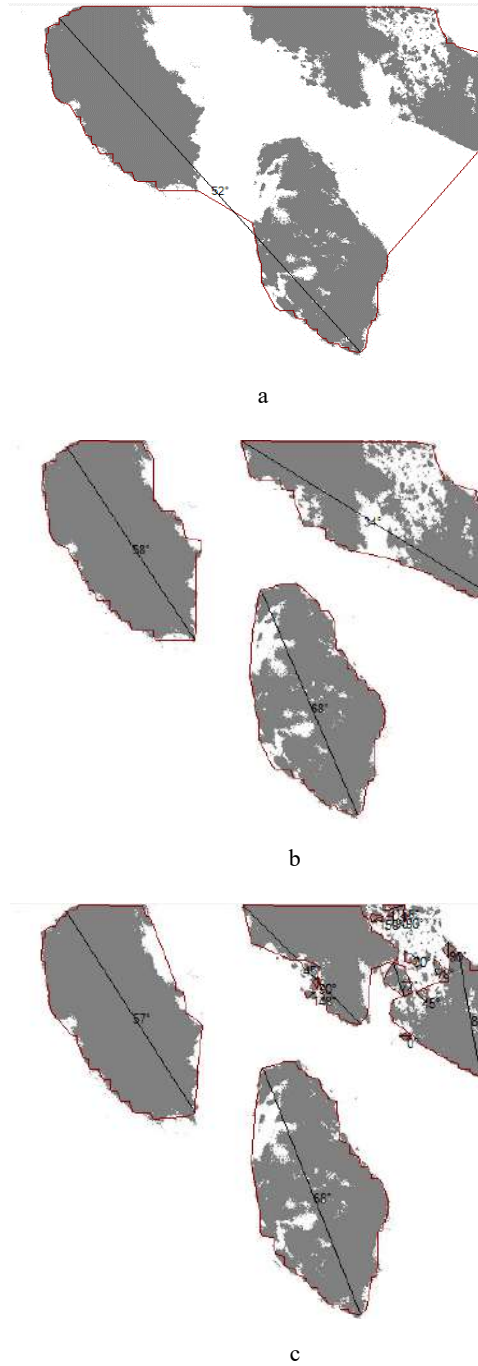


Fig. 2. Determination of clusters and the angle of inclination of the cluster to the horizontal axis
at $dist = 70$ (a), $dist = 10$ (b), $dist = 2$ (c)

In most cases, there is no need to find the minimum area, as this can distort the figure. In our implementation of the concave contour method, we need a contour that defines and describes the general shape of a set of points. It is necessary that the contour is sufficiently smooth and show as much detail as possible.

The developed algorithm for calculating the concave contour first calculates the convex contour. The algorithm used in this case for calculating convex contours is the Divide and Conquer algorithm. Then the convex contour opens iteratively to create a concave contour. The first step of the concave contour construction algorithm is to add all the edges of the convex contour to the list, where they are sorted by length. Then the longest edge in the list is selected and removed.

The next step is to search through all the points that are most suitable for choosing the best candidate for adding a new part of the contour between the endpoints of the remote edge. The search criteria for points is that the largest angle from the new point to the end points of the old contour must be less than the cosine of the user-entered angle between the new edges (concavity coefficient K). This ensures that the point is not outside the contour. After each iteration, a check is made to see if the new edges of the contour intersect with the existing edges.

In the case of $K = 1$, we obtain an actually convex contour, the constant $K = 0$ gives a concave smooth contour, and the constant $K = -1$ gives a broken concave contour (Fig. 3). The ability to select the angle between new edges and a region-defined grid size to search for new points is an opportunity for the user to balance between a smoother or sharper contour.

5 Conclusion

The analysis of the possibilities and shortcomings of existing medical image processing methods showed that in order to reduce the complexity of processing, it is necessary to automate the entire sequence of analysis steps with expanding the process control capabilities at each stage and using methods to reduce the amount of information processed.

An algorithm for identifying image element search areas and an algorithm for forming the boundaries of found clusters with concave sections with simultaneous application of the grid overlay method to reduce the amount of information processed was proposed and implemented.

The object-oriented software that automates the stages of image processing is developed. The final result of the analysis is the areas highlighted in the image with concave contours and certain quantitative characteristics of their form.

The application of the grid overlay method made it possible to abandon the processing of each pixel of the image and reduce the processing time of the original data. The analysis of the results of the algorithm on the test image of the submicrocrystalline structure of titanium BT 1-0 showed that, depending on the problem being solved, it is possible to reduce the amount of information processed by 4-25 times.

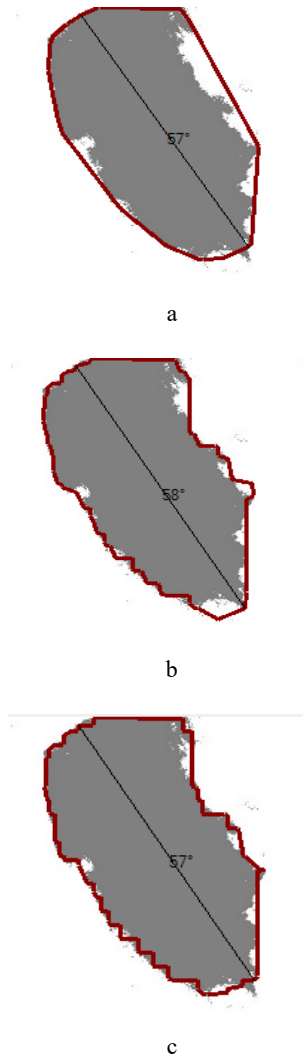


Fig. 3. Examples of the construction of a convex contour (a), a smooth concave contour (b), a broken concave contour (c)

The prospects for further research are to test the proposed methods and algorithms on a wider set of applied problems, to study the dependence of the accuracy of working methods on the type of image.

References

1. Gonzalez, Rafael C., and Richard E. Woods. "Digital image processing." (2002).
2. Pratt, W. K. Introduction to digital image processing. CRC Press. (2013).
3. Panteleev, V. G., Egorova, O. V. and Klykova, E. I., (2005), *Computer microscopy*, [Komp'yuternaja mikroskopija], Technosphere, Moscow.
4. Jaroslavskij, L. P., (1987), *Digital signal processing in optics and holography: introduction to digital optics*, [Cifrovaja obrabotka signalov v optike i golografii: vvedenie v cifrovuju optiku], Radio i svjaz', Moscow, 296 p.
5. NEXSYS ImageExpert Pro 3. http://www.nexsys.ru/nexsys_iepro3x.htm.
6. Balasko, Balazs, Janos Abonyi, and Balazs Feil. "Fuzzy clustering and data analysis toolbox." *Department of Process Engineering, University of Veszprem, Veszprem* (2005).
7. Altami Studio. <http://alelso.ru/projects/altami-studio/>.
8. Vasilyeva, L., Tarasov, A. and Getman, I. (2016), "Development of algorithmic support and model of the software complex of image processing", [Razrabotka algoritmicheskogo obespecheniya i modeli programmnogo kompleksa obrabotki izobrazhenij], *Proceedings of the tenth international scientific-practical conference «Internet-Education-Science» (IES-2016)*, Vinnytsia, pp. 214–215, available at: <http://ir.lib.vntu.edu.ua/handle/123456789/13371>
9. Fasulo, D. (1999). *An analysis of recent work on clustering algorithms* (No. 01-03, p. 02). Technical report.
10. Oreshkov. V. (2012). EM-masshtabiruyemyy algoritm klasterizatsii. <https://basegroup.ru/community/articles/em>.
11. Leonenkov, A.V. (2003), *Fuzzy modeling in the MATLAB and fuzzyTECH environment* [Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH], BHV-Peterburg, Peterburg, p. 385-388.
12. Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1, No. 10). New York: Springer series in statistics.
13. Tarasov, O., Vasilyeva, L., & Efremov, M. (2017). Automation of processing of microstructures of metals based on contour and texture analysis of images. *Naukovi Praci Donec'kogo Nacional'nogo Tehnichnogo Universitetu. Seria, Informatika, Kibernetika i Obcisluval'na Tehnika*, 2(25), 109–117. doi:10.31474/1996-1588-2017-2-25-109-117
14. Tarasov A.F. et al (2015). *Osobennosti ispolzovaniya protsessa reversivnogo sdviga dlya polucheniya submikrokristallicheskikh obyemnykh zagotovok* [Features of using the reverse shift process for obtaining submicrocrystalline bulk billets]
15. Barber, C. B., Dobkin, D. P., Dobkin, D. P., & Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4), 469-483.
16. Wenger, R. (1997). Randomized quickhull. *Algorithmica*, 17(3), 322-329. <https://doi.org/10.1007/BF02523195>
17. The Graham Scan, Convex Hull Algorithm. <http://www.algomation.com/algorithm/graham-scan-convex-hull>.
18. Jarvis, R. A. (1973). On the identification of the convex hull of a finite set of points in the plane. *Information processing letters*, 2(1), 18-21. [https://doi.org/10.1016/0020-0190\(73\)90020-3](https://doi.org/10.1016/0020-0190(73)90020-3)
19. Bentley, J. L. (1980). Multidimensional divide-and-conquer. *Communications of the ACM*, 23(4), 214-229.
20. Lee, D.T. & Schachter, B.J. International Journal of Computer and Information Sciences (1980) 9: 219. <https://doi.org/10.1007/BF00977785>

Two-stage method for adaptive binarization of raster engineering drawings

Vera Molchanova^[0000-0001-5646-1410], Dmitry Mironenko^[0000-0001-5646-1410]

Pryazovskyi State Technical University, str. Universytets'ka 7, Mariupol 87555, Ukraine

molchanova_vs@i.ua, mironenko_ds@ukr.net

Abstract. The paper describes the proposed original model and the developed two-stage method for adaptive binarization of raster engineering drawings. The method allows binarization of images whose brightness histograms are not bimodal. At the same time, thin contour lines on resulting images are saved. The method is adaptive to brightness of image background. Comparative analysis of testing results of the developed method and some well-known methods for binarization raster images confirms that the proposed solutions are effective. The method was tested on a representative set of 40 different quality raster engineering drawings.

Keywords: drawing, binarization, brightness histogram, threshold, contour, background, local area.

1 Introduction

The task of raster-vector transformation of engineering drawings images requires ensuring homotopy original and result images at all stages of image processing. The binarization stage solves the problem of dividing a bitmap to a contour and a background. Binary image processing requires less computational resources and from point of view of theoretical complexity it is simpler than gray-scale image processing. This explains binarization expediency.

Literary analysis has shown that there are many threshold methods for image processing [1-5]. They describe binarization of generalized images, but they don't pay respect to the specifics of brightness distribution and contrast in engineering drawings.

The GMT (Global Multi-stage Thresholding) [6] methodology is the most promising binarization technology for this subject domain. It provides images multistage binarization. The researches for three-section brightness histograms [7, 8] is also interesting for further development.

2 Formal problem statement

The bitmap drawing represents a two-dimensional matrix of raster points $R^{(*)}$ (1) .

$$R^{(*)} = \{r_{ij}\}_{i=0, j=0}^{M-1, N-1} \mid r_{ij} \in (C = \{c_k\}_0^{K-1}) \quad (1)$$

where r_{ij} is the point (i,j) of raster drawing, C is the K -colors color palette, M, N is the width and height of the image.

The color palette of a color image ($R^{(CF)}$) consists of K colors $\{c_k\}_0^{K-1}$. The color palette of a grayscale image ($R^{(M)}$) consists of 256 tints of gray color. The color palette of a binary images $R^{(B)}$ is bicolored $\{c_k\}_0^1 = \{0, 1\}$. Here white color coded as «0», it is a background color. And black color coded as «1», it is a contour-line color.

The purpose of a color image binarization $R^{(CF)}$ is its homotopic mapping to a binary image $\varphi: R^{(CF)} \xrightarrow{\varepsilon} R^{(B)}$ with a permissible error ε . The transform colored points $\{r_{ij}\}_{i=0, j=0}^{M-1, N-1}$ to bicolored performs in accordance with the classification method to a background and a contour.

As a binarization quality criterion we selected the minimal Hamming distance between binarized $R^{(B)}$ and reference images $R^{(0)}$ (2) [9, 10].

$$\rho(R^{(B)}, R^{(0)}) = \sum_{i=1}^M \sum_{j=1}^N (r_{ij}^{(B)} \wedge r_{ij}^{(0)}) \leq \varepsilon \rightarrow 0 \quad (2)$$

The image quality metrics (MSE, PSNR, CNR, SSIM) perform as measures of this distance.

3 Literature review

A generalized n -stages scheme for raster images binarization was proposed in [5–7]. Images of engineering drawings are a typical example of contour-line images. The qualitative binarization result for contour-line images can be obtained even with $n = 2$ [6]. Therefore, we will consider only two-stage binarization.

Let's imagine that a dark contrast contour and an uneven chromatic image background there is in a color drawing image (for example, it is a blueprint or tracing paper). Brightness histograms of such images are usually three-sectioned. (Fig. 1). As a rule, in such histograms the most points have brightness which there is at the vicinity of the two prevailing values $r_{\max}^{(B)}$ (for points which belong to contour) and $r_{\max}^{(T)}$ (for points which belong to background).

In this case, if the shadow zone is limited by a bottom global binarization threshold $T^{(B)}$, the points of this zone with brightness $r_{ij} < T^{(B)}$ will certainly belong to a contour. If the light zone is limited by a top global binarization threshold $T^{(T)}$, the points of this zone with brightness $r_{ij} > T^{(T)}$ will certainly belong to a background. Points with brightness $T^{(B)} \leq r_{ij} \leq T^{(T)}$, to the uncertainty interval $[T^{(B)}; T^{(T)}]$. They need to extra

classification for a contour and a background based on the brightness estimates of their local neighborhoods

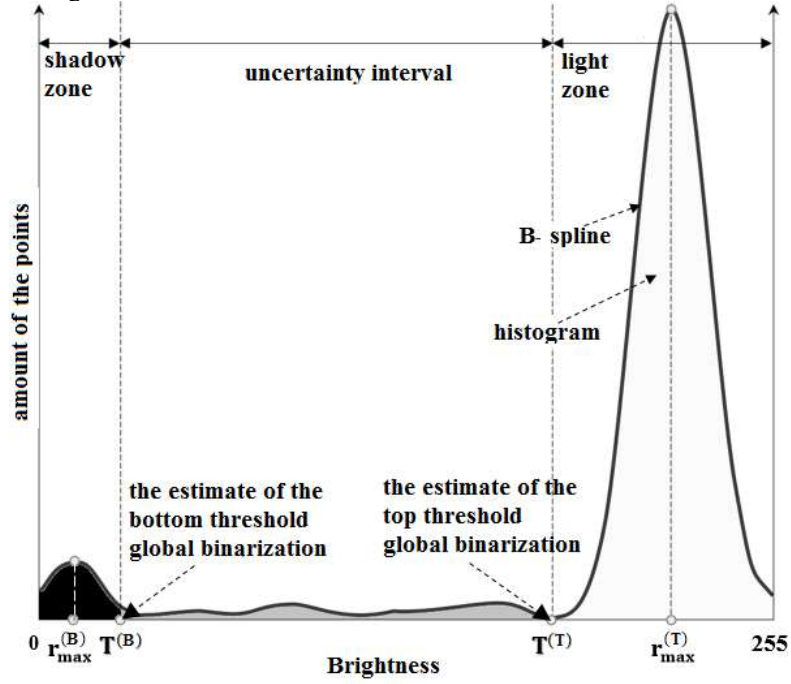


Fig. 1. Partitioning of the brightness histogram for contour-line image in gray-scale

In [7, 8] a two-stage method for calculating the estimates of global thresholds is proposed for three-sectioned brightness histograms (fig. 1).

At the first stage, the lower and upper binarization thresholds are estimated. For it, we calculate minimum values of brightness for the points which belonging to the shadow or to the light zone on a brightness histogram. The solution of this task is proposed in [6], where the method QIR are described. This method allows to find the extremum of a quadratic polynomial, which approximates the brightness histogram on an information interval.

At the second stage, a final threshold value is selected as an average between the lower and the upper global thresholds. But this approach is unacceptable here. This is due to the presence of many local extremes in the brightness histograms of images with a non-uniform background (for example, drawings on tracing paper and blue-prints). In addition, this approach doesn't take the mutual influence of the shadow zone and light zone when estimates of global thresholds are calculating. This leads to incomplete removal of the chromatic background from the images and to appearance of binary noise near the contour.

4 The proposed two-stage method for contour-line images binarization

In view of the foregoing, we proposed an adaptive two-stage binarization method for images with three-sectioned brightness histograms. This method takes into consideration the brightness characteristics of contour-line images with a non-uniform background (tracing paper, blueprint). It includes:

- estimation of global binarization thresholds and search points in image those belong to the contour or background definitely;
- the points in an uncertainty interval are classified as a contour and a background based on the local binarization thresholds estimates.

The proposed method is adaptive to the image brightness characteristics for different carrier types (drawing paper, tracing paper, blueprint) on account of injection an additional adaptive coefficient α to the formulas for a global and a local binarization thresholds calculation.

5 The method for global adaptive binarization thresholds estimation

The proposed method for calculating two global binarization thresholds assumes a statistical analysis of the image points brightness distribution. Further, the obtained result is adjusted depending by the drawing carrier type (drawing paper, tracing paper, blue). It includes next steps:

1. The color drawing image is converted to a grayscale image. A brightness histogram $H(n, r_{ij})$ is calculated for it [5] and approximated by cubical B-spline. This allows you to get rid of minor fluctuations in a brightness histogram of an image [11].
2. Brightness distribution of an image is analysed by sizes of shadow zone and light zone in the range of its brightness. The brightness characteristics at a left (marked as «B») and a right (marked as «T») parts of the histogram are proposed to analyze separately (3-6). It allows you to eliminate the mutual influence of a shadow and a light zones when the upper and the lower binarization thresholds are calculating.

$$\lim_{i \rightarrow 0} h_i = L_{\min}^{(B)} \mid h_i = 0, \quad \lim_{i \rightarrow 127} h_i = L_{\max}^{(B)} \mid h_i = 0, \quad \bar{L}^{(B)} = \frac{1}{n^{(B)}} \cdot \sum_{i=0}^{127} i \cdot h_i \quad (3)$$

$$\beta^{(B)} = \frac{\left(\bar{L}^{(B)} - L_{\min}^{(B)} \right)}{\left(L_{\max}^{(B)} - L_{\min}^{(B)} \right)}, \quad \gamma^{(B)} = 1 - \beta^{(B)} \quad (4)$$

$$\lim_{i \rightarrow 128} h_i = L_{\min}^{(T)} \mid h_i = 0, \quad \lim_{i \rightarrow 255} h_i = L_{\max}^{(T)} \mid h_i = 0, \quad \bar{L}^{(T)} = \frac{1}{n^{(T)}} \cdot \sum_{i=128}^{255} i \cdot h_i \quad (5)$$

$$\beta^{(T)} = \frac{\left(\bar{L}^{(T)} - L_{\min}^{(T)}\right)}{\left(L_{\max}^{(T)} - L_{\min}^{(T)}\right)}, \quad \gamma^{(T)} = 1 - \beta^{(T)} \quad (6)$$

where h_i is the bin of colour «i» in the brightness histogram of an image, $n^{(B)}$, $n^{(T)}$, is the amount of points of an image which are belonged to a left or to a right part of histogram respectively, $[L_{\min}^{(B)}; L_{\max}^{(B)}]$ и $[L_{\min}^{(T)}; L_{\max}^{(T)}]$ is the ranges of brightness, $\bar{L}^{(B)}$, $\bar{L}^{(T)}$ is the average values of brightness in a left and a right part of brightness histogram respectively, $\beta^{(B)}$, $\beta^{(T)}$ is the estimates of relative sizes of a shadow zone, $\gamma^{(B)}$, $\gamma^{(T)}$ is the estimates of relative sizes of a light zone.

1. The estimates of the lower ($T^{(B)}$) and the upper ($T^{(T)}$) binarization thresholds can be calculated using the formulas (7-8).

$$T^{(B)} = \alpha^{(B)} \cdot \left(\beta^{(B)} \cdot \bar{L}^{(B)} + \gamma^{(B)} \cdot L_{\min}^{(B)} \right), \quad \alpha^{(B)} = 1 - k \cdot \beta^{(B)}, \quad (7)$$

$$T^{(T)} = \alpha^{(T)} \cdot \left(\beta^{(T)} \cdot \bar{L}^{(T)} + \gamma^{(T)} \cdot L_{\min}^{(T)} \right), \quad \alpha^{(T)} = 1 - k \cdot \beta^{(T)}, \quad (8)$$

where $\alpha^{(B)}$, $\alpha^{(T)}$, k is the tuning coefficients.

2. The method sensitivity to an image background tonality on three equiprobable carriers (blueprint, tracing paper, drawing paper) we can change using the tuning coefficients, $\alpha^{(B)}$, $\alpha^{(T)}$. Their values should be selected from the range [0,95; 1.00].

As a result, points of a drawing image $\{r_{ij}\}$ are classified to:

- Points of a binary contour ($r_{ij}^{(b)} = 0$). Their brightness are less then the lower global binarization threshold (i.e. $r_{ij} < T^{(B)}$);
- Points of a binary background of the image ($r_{ij}^{(b)} = 1$). Their brightness are more than the higher global binarization threshold (i.e. $r_{ij} > T^{(T)}$).

The rest of an image points are remained in an interval of uncertainty.

6 The method for local adaptive binarization thresholds estimation

As noted above, in brightness histograms of the images of drawings, a shadow zone less than a light zone. The bimodality of those histograms is weakly expressed. However, a bimodality at local areas of an image is more distinct. Therefore, the points of the drawing, the brightness of which is in the interval of uncertainty, are divided into points of the background and contour by the local threshold $\{t_m = f(\omega_m, L_{\omega_m})\}$. The formula, which used for a local binarization threshold calculating, takes to account influence of the uneven distribution L_{ω_m} of the background brightness and the type of image carrier. It allows you to vary the size d_ω for each local area. A local binarization includes some steps:

1. We build local areas ω_m around each point r_{ij} , brightness of which belongs to uncertainty interval. We choose the points are equidistant and centrated relatively the point r_{ij} . The brightness histograms of those areas are quasi-bimodal (Fig. 2).

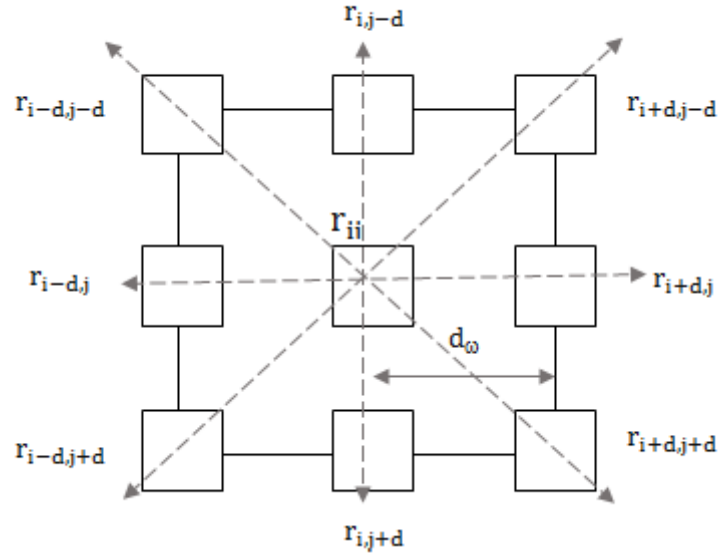


Fig. 2. The local ω_m area for the point r_{ij} with the radius d_{ω}

In accordance with the Gibbs principle [11], it suffices to consider only points of an image which are located along the base vectors. Their brightness has the greatest effect on a result of binarization the point r_{ij} .

It experimentally established [12], the thin contour lines can be binarized successful when $d_{\omega} \in [1; 9]$. They are usually deformed when $d_{\omega} > 9$. This is due to the peculiarities of a smoothed raster contour lines construction by Bretzenheim [5]. In accordance to it, the lines are depicted not only in black, some points of them are grey. Their brightness value can overcome the local binarization threshold t_m .

The initial value d_{ω} is proposed to calculate according to the expressions (9-10).

$$d_{\omega} = 9 \cdot \left\lceil \frac{r_{ij} - L_{\max}}{L_{\max} - L_{\min}} + 1 \right\rceil \quad (9)$$

$$\lim_{i \rightarrow 0} h_i = L_{\min} \mid h_i = 0, \quad \lim_{i \rightarrow 255} h_i = L_{\max} \mid h_i = 0, \quad \bar{L} = \frac{1}{n} \cdot \sum_{i=0}^{255} i \cdot h_i \quad (10)$$

So, the value of d_{ω} is inversely proportional to brightness of an analyzed point. It allows you to save thin contour lines after binarization.

At the next, we calculate a minimal $(L_{\min}^{(\omega)})$, a maximal $(L_{\max}^{(\omega)})$ and an average $(\bar{L}^{(\omega)})$ estimates of points brightness inside the local ω_m area.

$$L_{\min}^{(\omega)} = \min \begin{pmatrix} \min(r_{k+i,l}) & \forall k \in [-d_\omega, d_\omega], l = 0, \\ \min(r_{k,l+j}) & \forall l \in [-d_\omega, d_\omega], k = 0, \\ \min(r_{k+i,l+j}) & \forall k, l \in [-d_\omega, d_\omega] \end{pmatrix} \quad (10)$$

$$L_{\max}^{(\omega)} = \max \begin{pmatrix} \max(r_{k+i,l}) & \forall k \in [-d_\omega, d_\omega], l = 0, \\ \max(r_{k,l+j}) & \forall l \in [-d_\omega, d_\omega], k = 0, \\ \max(r_{k+i,l+j}) & \forall k, l \in [-d_\omega, d_\omega] \end{pmatrix} \quad (11)$$

$$\bar{L}^{(\omega)} = \frac{1}{9 \cdot d_\omega} \left(\sum_{k=-d_\omega}^{d_\omega} r_{k+i,l} + \sum_{k=-d_\omega}^{d_\omega} r_{k,l+j} + \sum_{k=-d_\omega}^{d_\omega} r_{k+i,l+j} \right) \quad (12)$$

After that, we estimate the sizes of a shadow zone $(\beta^{(\omega)})$ and of a light zone $(\gamma^{(\omega)})$ in the local ω_m - area of the point r_{ij} (14).

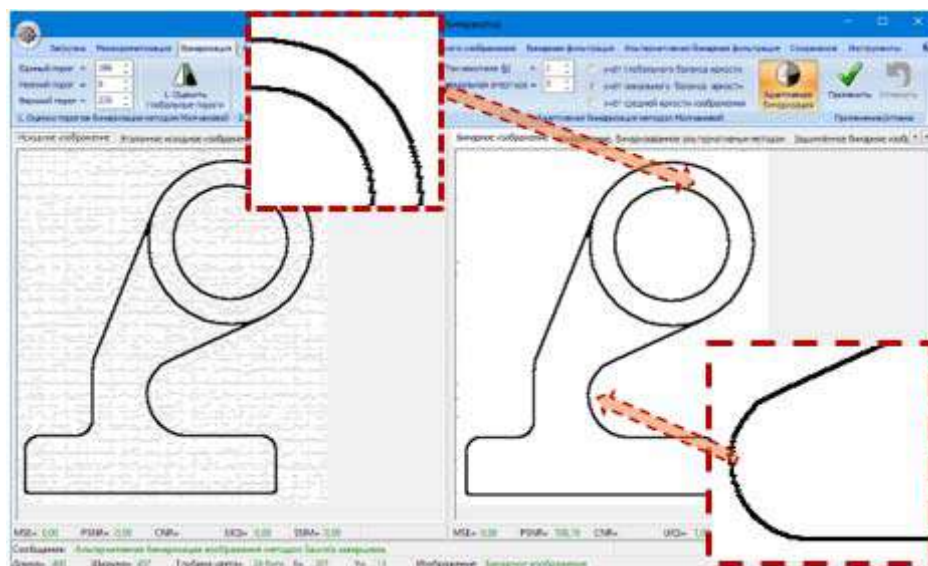
$$\beta^{(\omega)} = \frac{\bar{L}^{(\omega)} - L_{\min}^{(\omega)}}{L_{\max}^{(\omega)} - L_{\min}^{(\omega)}}, \quad \gamma^{(\omega)} = \frac{L_{\max}^{(\omega)} - \bar{L}^{(\omega)}}{L_{\max}^{(\omega)} - L_{\min}^{(\omega)}} = 1 - \beta^{(\omega)} \quad (13)$$

The local binarization threshold can be calculated by the formula (15).

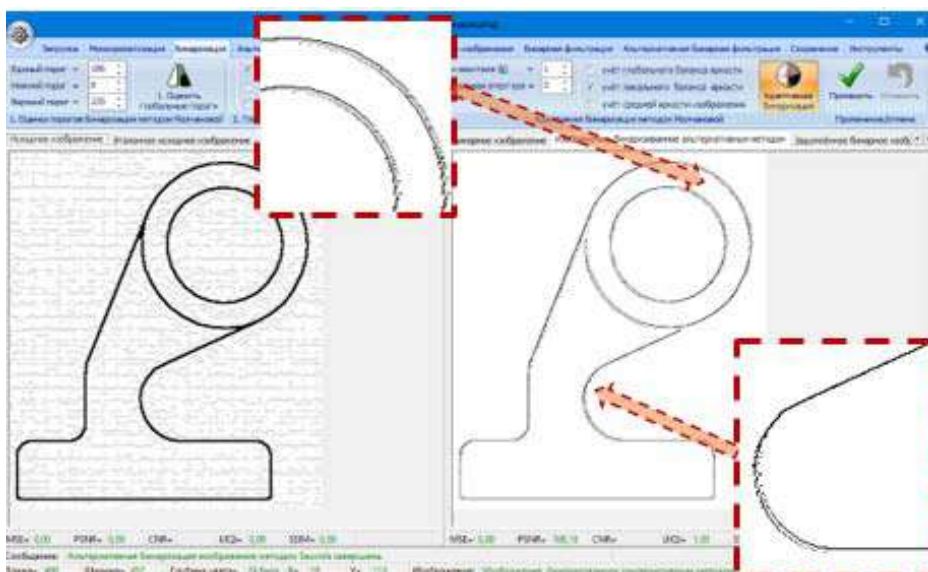
$$t_m = \alpha \left(\beta^{(\omega)} \cdot \bar{L}^{(\omega)} + \gamma^{(\omega)} \cdot L_{\min}^{(\omega)} \right), \quad \alpha = 1 - k \cdot \beta \quad (14)$$

So, the proposed method for local binarization threshold calculating is adaptive to brightness of the central point. This is ensured by the choice availability of an optimal diameter for local area of each point, which brightness value is in uncertainty zone. In addition, this method takes into account background tonality effect to the binarization result for different types of drawing carriers. So, the result binary image contains minimum number of artifacts.

We have developed an information system that shows testing results of the developed method and of some most popular alternative binarization methods. As you can see the proposed method allows you to save contour lines integrity. In this case, all artifacts near the contour are removed (Fig. 3, a). Alternative binarization methods leads to contour line doubling (Fig. 3, b).



a)



b)

Fig. 3. The result of the color image binarization using the developed method

7 Experiments and results

Analysis of the testing results of the method proposed in this work confirms that binarization quality is sufficient to preserve the topological features of the contours. We used a random sample of 40 color images of various sizes to test the method. Such sample guarantees representative results with a 95% probability when $n > 23$ [15] (16).

$$n = \frac{p \cdot (100 - p) \cdot t^2}{\delta^2} = \frac{15 \cdot 85 \cdot 2^2}{15^2} = 23 \quad (15)$$

where n is the minimum size of sample with an unknown value of the general concurrence, p is the average error (15%), t is the confidence coefficient, it shows a probability that the error will not exceed the a limit value (when $t = 2$ a probability of an accurate prediction is 95%), δ is the boundary error ($\pm 15\%$).

The reliability and efficiency analysis of the developed method is based on a results comparison, which we got using this method and some well-known alternative methods (Bernsen, Niblack, Otsu, Sauvola, Phansalkar). These methods usually are used for solving the similar tasks, they're generally accepted and effective.

We've chosen several general accepted metrics for results quantification: PRECISION, RECALL, F-MEASURE, ACCURACY [16]. For each image from the test sample, we repeated an experiment for 5 times, and then we averaged the results. The relative improvement of values of basic metrics for the proposed and several alternative methods is summarized in the table 1.

Table 1. The relative improvement of values of basic metrics for the proposed and several alternative methods

The basic metric	The alternative method					Average by methods
	Bernsen	Niblack	Otsu	Sauvola	Phansalkar	
$\Delta \text{Recall}^{\text{AM}}, \%$	1,70	0,07	-0,08	8,17	0,22	2,01
$\Delta \text{Precision}^{\text{AM}}, \%$	11,02	81,67	44,42	3,57	0,45	28,22
$\Delta F - \text{measure}^{\text{AM}}, \%$	7,58	71,03	36,60	8,05	0,45	24,74
$\Delta \text{Accuracy}^{\text{AM}}, \%$	2,11	49,34	19,45	2,15	0,01	14,61
Average performance ($\Delta M_{\text{avg}}^{\text{AM}}, \%$)	5,60	50,53	25,10	5,48	0,28	17,40

So, using the developed method to reach relative values improvement for the basic metrics of image quality by 17.4%. This allows us to make a positive conclusion about its reliability and effectiveness.

8 Conclusion

It has been established that the homotopic mapping of a color drawing image to a binary drawing image can be done as converting the brightness of its points to a two-color palette by classifying it as a background and a contour. The proposed method takes to account influence of brightness characteristics of images, which made on different carriers and the need to preserve the topological features of the contour (for example, small primitives and thin lines).

The proposed method allows us to significantly increase the quality quantitative indicators of the converting color images of drawings to binary ones. It includes:

- the method for global adaptive binarization thresholds estimation;
- the method for local adaptive binarization thresholds estimation.

The first method is based on the forms analysis and a light and a shadow zones ratio in the brightness histograms. This allows us to take to account the global brightness heterogeneity of the drawing images.

The second method is based on the balance of brightness analysis for areas with quasi-bimodal brightness histograms. This allows us to take to account the local brightness features of the drawing images by adaptation to different types of carriers.

References

1. Otsu, N. Y.: A threshold selection method from gray-level histograms. *IEEE transactions on systems, man and cybernetics*. 9(1): 62-66 (1979) doi:10.1109/TSMC.1979.4310076
2. Ablamejko, S. V.: *An Introduction to Interpretation of Graphic Images*. SPIE Press, Washington (1997)
3. Nixon, M. S.: *Feature Extraction and Image Processing for Computer Vision*. Academic Press, New York (2013)
4. Garg, N. K.: Binarization Techniques used for Grey Scale Images. *Binarization Techniques used for Grey Scale Images. International Journal of Computer Applications*. 71(1):8-11 (2013) doi:10.5120/12320-8533
5. Gonzalez, R. S Woods, R. E.: *Digital image processing*. Second edition. Prentice Hall, New Jersey (2006).
6. Solihin, Ya Leedham, C.G.: Integral Ratio: A New Class of Global Thresholding Techniques for Handwriting Images. *IEEE Transactions on pattern analysis and machine intelligence*, 21(8): 761-768 (1999) doi: 10.1109/34.784289
7. Dou, J Zhang, W.: A Fast Thresholding Technique in Image Binarization for Embedded System. *Indonesian Journal of Electrical Engineering and Computer Science*. 12(1):592-598 (2014) doi: 10.11591/telkomnika.v12i1.3359
8. Muñoz, I. Rubio-Celorio, M.: Computer image analysis as a tool for classifying marbling: A case study in dry-cured ham. *Journal of Food Engineering*. 166:148-155 (2015) doi: 10.1016/j.jfoodeng.2015.06.004
9. Blahut, R. *Theory and Practice of Error Control Codes*. Addison-Wesley Press, Boston (1983)
10. Chandler, D. M.: Seven Challenges in Image Quality Assessment: Past, Present, and Future Research. *I SRN Signal Processing*. 2013(8):1-55 (2013) doi: 10.1155/2013/905685

11. Myasnikov, V.V. Popov, S.P. Sergeyev, V.V. Soifer, V.A. Computer Image Processing, Part I: Basic concepts and theory. Edited by Victor A. Soifer. - VDM Verlag, Saarbrücken (2009)
12. Molchanova, V. S.: The adaptive threshold binarization method of raster images of technical drawings. Radio Electronics, Computer Science, Control. 33(2): 62-70 (2015) doi: 10.15588/1607-3274-2015-2-8
13. Molchanova, V. S.: Development of a hybrid adaptive noise suppression method in a raster image. Eastern-European Journal of Enterprise Technologies. 74 (4): 35-43 (2015) doi: 10.15587/1729-4061.2015.47415
14. Parker, J. R. Algorithms for Image Processing and Computer Vision. Second edition. Wiley Publishing, Choboken (2011)
15. Lindgren, B. W.: Statistical Theory. MacMillan Publishing Company, London (1976)
16. Wang, Zh, Bovik. C. A.: A Universal Image Quality Index. IEEE Signal Processing Letter. 9(3): 81-84 (2002) doi: 10.1109/97.995823

Quadratic Optimization Problem on Permutation Set with Simulation of Applied Tasks

Liudmyla Koliechkina^{1[0000-0002-4079-1201]}, Alla Nahirna^{2[0000-0002-2863-8706]},

Olena Dvirna^{3[0000-0002-0750-6958]}

¹University of Lodz, Uniwersytecka Str. 3, 90-137 Lodz, Poland
ludaplt1971@gmail.com

²Kyiv International University, Lvivska str., 49, Kyiv, 03179, Ukraine
naghirnaalla@ukr.net

³Poltava University of Economics and Trade, Koval str., 3, Poltava, 36000, Ukraine
lenadvirna@gmail.com

Abstract. The article discusses the formulation of an optimization problem with a quadratic target function and additional constraints on the permutation set, which can be a model of many applied problems. An algorithm for solving an optimization problem with a quadratic target function and additional constraints on permutations is proposed.

During the implementation of the method the first reference plan is found and additional restrictions for it are checked at the first stage. Thus, in the beginning of the algorithm, the number of considered solutions decreases. This makes it possible at the first stage to reduce the number of possible solutions and narrow the area of the problem study. An example of solving a theoretical problem using this method, demonstrating its effectiveness, is proposed.

Such task can be used to modeling various technological processes. The reason for this is the optimization of mathematical models and algorithms for the proposed models.

Keywords: optimization problems, combinatorial set of permutations, model of optimization problems, quadratic target function, optimal solutions.

1 Introduction

Computer simulation of technological processes consists in optimization of the study object by a mathematical model, and further model study with the help of implemented computational algorithms on personal computers. The computer simulation process, as a single process of building and researching a model, is used for research, analysis, design and optimization of technological objects and systems [1]. Today, computer modeling approaches combines both theory and practice. Working with a

model that represents a research object gives you the opportunity to explore its properties and behavior in all situations.

Computer modeling is the process of building a real study object's model and setting up computational experiments on this model in order to understand and evaluate various strategies based on the use of algorithms that ensure the functioning of this study object. Thus, the process of computer simulation includes the model design, and its application to solve the problem of optimization and design of technological processes in production. All these problems are extremely complex and include numerous elements, variables, parameters, constraints, etc. Such problems can be modeled by combinatorial optimization tasks [1-8].

The combinatorial optimization problems study comprises a fairly wide range of mathematical models associated with the need to solve various important practical problems of optimal planning, management and design [1, 2, 5, 7, 9, 10-12, 20-29]. In this regard, many papers have recently appeared which investigate the combinatorial optimization problems and propose approaches to their solution [8, 11-34, 36-38].

The paper presents a model of optimization problems on a combinatorial set of permutations, which is a model of many applied problems. It is proposed the algorithm for solving an optimization problem with a quadratic target function and additional restrictions on permutations.

During the method's implementation the first reference plan is located and additional restrictions are checked, which reduces the number of considered solutions at the beginning of the algorithm

It is offered an example of solving a theoretical problem by the given method and demonstrating its efficiency.

2 Literature review

Computer models have become a common tool for mathematical modeling and are used in various areas of electronics, engineering, industry, and so on. Computer models are used to obtain new knowledge about the object or for an approximate assessment of the system's behavior that are too complex for analytical research.

Computer modeling is one of the effective methods for studying complex systems. The construction of a computer model is based on abstracting from the specific nature of the phenomena or the original object under study and consists of two stages – first, the creation of a qualitative, and then a quantitative model. The more significant properties will be identified and transferred to the computer model – the more approximate it will be to the real model, the greater the capabilities that the system using this model will have. Today the works of many scientists are devoted to this research area [1-5, 8-10, 21].

An important computer simulation area is analytical and simulation modeling. In analytical modeling, mathematical (abstract) models of a real object are studied in the algebraic equations form, as well as providing for the implementation of an unambiguous computational procedure leading to their exact solution. In simulation, mathematical models are studied in the algorithms form, which reproduce the

functioning of the system under study by sequentially performing numerous elementary operations. Very relevant today for simulation and analytical modeling are discrete models, in particular, combinatorial optimization, their study is the subject of a great number of papers [1–17], [22–31].

A significant contribution to the development of modern discrete optimization theory, the development and implementation of its methods in the study and solving of important applied problems have been made over the course of more than forty years under the scientific guidance of Academician I.V. Sergienko scientists of the Institute of Cybernetics named by V.M. Glushkov National Academy of Sciences of Ukraine. They obtained important results, which became the basis for the theory development and the creation of new lines in discrete optimization.

At present, intensive studies of the stability problems of vector discrete optimization problems are carried out at the Institute of Cybernetics named by V.M. Glushkov of the National Academy of Sciences of Ukraine (I.V. Sergienko, T.T. Lebedev, N.V. Semenova, T.I. Sergienko) [16, 17], Belarusian State University (V.O. Emelyichev, D.P. Podkopayev, Yu.V. Nikulin, etc.) [6, 7], the Joint Institute of Computer Sciences of the National Academy of Sciences of Belarus (Yu.N. Sotskov), the Computational Center of the Russian Academy of Sciences (E.M. Gordeyev, V.K. Leontiev), Omsk Branch of the Institute of System Studies of the Polish Academy of Sciences (M. Libura), a number of universities of the Federal Republic of Germany (E. Girlich), the Netherlands (ES van der Poort, G. Sierkma, APM Wagelmans, JAA van der Veen), in the uniquely Colorado US (H.J. Greenberg).

Belarusian scientists under the scientific guidance of Professor V.O. Yemelycheva successfully develops a constructive approach to the problem of the vector discrete optimization problems correctness associated with the reception of quantitative characteristics of the stability of specified problem statements, a mathematical apparatus for studying the stability of multicriteria discrete problems with different types, vector criteria functions, principles of optimality, as well as generalized and parameterized principles of optimality are developed.

It is important to study the relation between different classes of point configurations triangulation (regular, weakly regular, deployed, symmetric, political, etc.), in particular, the study of the structure of a partially ordered set constructed on the family of triangulation classes considered in relation to the inclusion relation. Recently, algorithms for solving convex polyhedral admissibility problems are actively explored [11, 18, 19]. One of the interesting computational approaches is the reduction of the admissibility problem for a polyhedron to the projection problem on a normal cone generated by a dual system of inequalities [18, 29, 31], which is sufficiently close to the projection problem for a binary polyhedron. Today, in the research area of various classes of combinatorial models, the new methods development for their solution, great attention is paid to methods based on the use of structural properties of combinatorial sets. The properties of combinatorial sets study is closely related to the theory of polyhedral and graphs. The use of information about the structure of the convex shell of admissible multivariate solutions, which is the basis for many methods, is one of the most successful approaches to solving combinatorial optimization problems for today. But when solving such tasks there are

problems related to the complexity of mathematical models, large volume of information, etc.

Applied problems simulated by extreme discrete tasks often have a high dimensionality, so they are quite complex from a computational point of view. The main task is to determine the value of an argument belonging to a certain combinatorial configuration for which the target function acquires the global optimum. So it is necessary to develop the most effective algorithm, which is based on the specific properties of the combinatorial configurations.

For the extreme problems of combinatorial optimization, polynomial algorithms for finding an optimal solution based on the properties of the input data structure have been developed, but there are few such work compared to methods based on partial overview of the options. One of the approaches to solving such problem is to carry out research and analysis of the combinatorial configurations properties, in which the target function, which reflects the combinatorial nature of the tasks, is determined. Analysis and study of combinatorial configurations as a target function argument, setting the change in the values of the target function, depending on the elements ordering of the selected combinatorial configuration and the structure of the input data specificity, does not pay sufficient attention in the literature. But it should be noted that the study of the certain tasks properties in order to identify their characteristic properties and their use for solving the problem, gives the possibility of constructing new approaches and the development of known methods.

Hence, one of the important problems in the discrete optimization area is the detection of the properties of combinatorial configurations in extreme problems, the use of which would allow to establish the regularity of the change in the values of the target functions, depending on the argument ordering and on the specificity and structure of the combinatorial configurations sets.

Today, significant results have been obtained in the area of research of combinatorial models' various classes and the development of new methods for their solution. The following foreign scientists made a fundamental contribution to the development of discrete, in particular, combinatorial optimization: M. Gary, S. Berge, D. Johnson, H. Papadimitriou, P. Pardalos, K. Staiglich, R. Stanley, F. Harari, V.A. Emelichev, V.M. Sachkov [1-8, 10, 11, 12].

In turn, the many Ukrainian scientists' works are devoted to the various classes of combinatorial optimization problems' study: L.F. Gulyanitsky, P.I. Stetsyuka, I.V. Sergienko, N.S. Shor, Yu.G. Stoyan, S.V. Yakovlev, A.O. Yemetsa, V.O. Perepelitsy and many others.

In particular, in [18–21], the authors describe the convex extensions theory and its applications in combinatorial optimization problems. Combinatorial models' applications in practical problems of geometric design are presented in the works of L.F. Gulyanitsky, I.V. Sergienko, Yu.G. Stoyan, S.V. Yakovlev, N.S. Shor [22-27, 38].

Quadratic optimization on the permutation set is reflected in [28–31].

The permutation set's representation as the intersection of a permutation polyhedron and a hypersphere is interesting, as well as optimization methods on permutation configurations using the intersection described in [32].

The development of an integrated approach to the analysis of the properties of combinatorial optimization supplies covers a wide range of studies of combinatorial functions, combinatorial polyhedral, combinatorial configurations as an argument of the target function. The results give the opportunity to improve existing methods for solving such problems and develop new methods for optimal solutions searching. The problems based on the properties of combinatorial configurations are actual problems of combinatorial optimization. Of particular importance in this aspect is the consideration of extreme problems in combinatorial configurations using graph theory.

The research of tasks in graphs deals with such scientists as F. Harari, O. Ore, I.V. Sergienko, V. O. Yemelichev, A. O. Zikov, V. O. Perepelytsya, R. I. Tyshkevich and others. Despite quite large achievements in the area of discrete optimization, in the process of modeling, there are extreme problems classes for which a number of issues have not yet been investigated. Principal difficulties that arise during modeling are also related to various types of uncertainty. These include: the availability of many criteria for evaluating the quality of solutions, interval setting of task parameters, etc. In these conditions, classical methods are not sufficiently suitable for solving problems. As you know, most combinatorial optimization tasks can be reduced to integer programming tasks, but this is not always justified, since it eliminates the possibility of taking into account the combinatorial properties of task solutions.

This work is a continuation of research in the extreme discrete problem's area, in particular, combinatorial optimization, as well as optimization problems under the conditions of multicriteria, which arise in the study of many theoretical and applied problems.

3 Formal problem statement

We consider the permutation as an ordered sample of elements $a = (a_{i_1}, a_{i_2}, \dots, a_{i_k})$, where $a_{i_j} \in A$, $\forall i_j \in N_n$, $\forall j \in N_n$ $i_s \neq i_t$ if $s \neq t \forall s \in N_n$, $\forall t \in N_n$ from some multi-set $A = \{a_1, a_2, \dots, a_q\}$, which is characterized by the base $S(A) = \{e_1, e_2, \dots, e_k\}$, where $e_j \in R^1$, $\forall j \in N_k$ and the multiplicity of the elements $k(e_j) = r_j$, $j \in N_k$, $r_1 + r_2 + \dots + r_k = q$ $j \in N_k$, according to [14, 17].

A set of permutations with repetitions of n real numbers, among which k different, is called the general permutation set and is denoted as $P_{nk}(A)$. This is the set of ordered n -samples from the multiset A under the condition $n = q > k$.

If we have $n = k = q$ set of permutations without repetition, we denote it as P_n . Obviously, $P_n(A) = P_{nn}(A)$. In cases where the form of the set of permutations is not indicated, it will be possible to write down these sets as $P(A)$.

It is known [14] that the convex hull of the set of permutations is a permutation polyhedron $\Pi = \text{conv } P(A)$, which vertex set $P(A)$ is equal to the set of permutations: $\text{vert } \Pi(A) = P(A)$.

Without loss of generality, we order the elements of the multiset A in non-descending order: $a_1 \leq a_2 \leq \dots \leq a_q$ and the elements of its foundation—in descending order: $e_1 > e_2 > \dots > e_k$.

Then the convex hull of a common set of $P(A)$ permutations is a generic polyhedron $\Pi(A) = \text{conv } P(A)$, that is described by a well-known system of linear inequalities:

$$\begin{cases} \sum_{j=1}^n x_j \leq \sum_{j=1}^n a_j, \\ \sum_{j=1}^i x_{\alpha_j} \leq \sum_{j=1}^i a_j, \end{cases} \quad (1)$$

$\alpha_j \in N_n, \alpha_j \neq \alpha_t, \forall j \neq t, \forall j, t \in N_i, \forall i \in N_n$ and $P(A) = \text{vert } \Pi(A)$.

Consider the optimization problem:

$$Z(\Phi, P(A)) : \max \{\Phi(a) \mid a \in P(A)\} \quad (2)$$

$$D = \{x \in R^n \mid Gx \leq (\geq) b\} \quad (3)$$

where $G \in R^{m \times n}$, $b \in R^m$

$$\Phi(a) = \sum_{j=1}^n c_j x_j^2 + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (4)$$

on a set of permutations $P_n(A)$.

Additional linear constraints form a multifaceted set $D \subset R^n$.

Then, to every point $a \in P_n(A)$ will match point $x \in X$, one that satisfies equality $F(x) = \Phi(a)$:

$$Z(F, X) : \max \{F(x) \mid x \in X\} \quad (5)$$

$$\text{where, } F(x) = \sum_{j=1}^n c_j x_j^2 + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (6)$$

and additional constraints:

$$D = \{x \in R^n \mid Gx \leq (\geq) b\} \quad (7)$$

where $G \in R^{m \times n}$, $b \in R^m$

X - nonempty set in R^n , which is defined as follows:

$$X = \text{vert } \Pi(A), \quad \Pi = \text{conv } P(A) \quad (8)$$

It is natural to assume that the maximum of a quadratic function will be one of the vertices of the permutation polyhedron, and the vertices of the graph $G(P_n)$ will match to all points of the set of permutations $P_n(A)$.

The adjacency of the vertices of a permutation polyhedron is determined by a one-time transposition of two vertex elements. The number of transpositions in the graph is determined by the formula [35]:

$$C_n^2 = \frac{n(n-1)}{2} \quad (9)$$

4 The algorithm for solving an optimization problem with a quadratic target function and additional constraints on the set is rearranged

The algorithm for finding the problem optimal solution consists of four steps, which in a few steps make it possible to obtain an optimal solution.

STEP 1. Finding the first support solution.

Consider the first transposition of the target function:

$$x_1 \leftrightarrow x_2, (i = 2, \dots, n):$$

$$f_{12} = (x_1 - x_2)(a_1x_1 + a_1x_2 + \dots + a_nx_n)$$

We form the necessary conditions for finding the first solution:

$$\begin{cases} x_1 > x_2, \\ x_i \rightarrow \max, (\max a_i). \end{cases} \quad (10)$$

Variants of transpositions:

$$x_1 \leftrightarrow x_2, (i = 2, \dots, n):$$

$$f_{12} = (x_1 - x_2)(a_1x_1 + a_1x_2 + \dots + a_nx_n)$$

$$x_1 \leftrightarrow x_3 : f_{13} = (x_1 - x_3)(a_{13}x_1 + a_{23}x_2 + \dots + a_{n3}x_n)$$

... ..

$$x_1 \leftrightarrow x_n : f_{1n} = (x_1 - x_n)(a_{1n}x_1 + a_{2n}x_2 + \dots + a_{nn}x_n) \quad (11)$$

$$x_2 \leftrightarrow x_i, (i = 3, \dots, n);$$

• • • • •

• • • • •

The first solution will be: $(x_1, x_2, \dots, x_n)$. It should be noted that there may be several first solutions.

$$\Delta g_i = \Delta g_i^2 - \Delta g_i^1 = \left(x_i^{g_i^2} * c_j + x_j^{g_i^2} * c_i \right) - \left(x_j^{g_i^1} * c_j + x_i^{g_i^1} * c_i \right) \quad (14)$$

All increments satisfy conditions (12), then the found point is the optimal solution. Otherwise, we return to the beginning of the step 4.

4.2. If condition (13) is not fulfilled, we return to the beginning of step 4.

It should be noted that conditions (12) are sufficient for finding the optimal solution, and the fulfillment of inequality (13) is necessary for finding the optimal solution.

5 Example

Find the maximum value of the target function:

$$\max F(x) = \frac{1}{2}(2x_1 - x_2 - x_3 + x_4)^2 + (x_2 + x_3)^2 - \frac{1}{2}(x_3 - x_2 + 2x_4)^2 + 2x_4^2$$

on the permutations set A_4 , where $A = (1,2,3,4)$, with the following restrictions:

$$\begin{cases} g_1 = 5x_1 - 2x_2 + 3x_3 + 4x_4 \geq 15, \\ g_2 = -3x_1 + 6x_2 + 8x_3 - x_4 \leq 31. \end{cases}$$

Solving.

For formula (1), the number of possible transpositions is: $C_4^2 = 6$.

Consider the first transposition $x_1 \leftrightarrow x_2$, presenting the target function as a product:

$$f_{12} = (x_1 - x_2)(x_1 + x_2 - 6x_3 + x_4),$$

accordingly, when searching for the first solution, it is necessary to consider the following conditions:

$$\begin{cases} x_1 > x_2, \\ x_3 \rightarrow \min. \end{cases}$$

Then the first solution may be the following points of the permutations set: (4,3,1,2), (4,2,1,3), (3,2,1,4).

Consider the first point (4,3,1,2):

$$\begin{cases} g_1(4,3,1,2) = 25 \geq 15, \\ g_2(4,3,1,2) = 12 \leq 31. \end{cases}$$

Additional restrictions are enforced. Then the initial conditions for improving the first support solution will be:

$$\max f_1(4,3,1,2) = 40,$$

$$\begin{cases} \Delta g_1^1(4,3,1,2) \geq -10, \\ \Delta g_2^1(4,3,1,2) \leq 19. \end{cases}$$

When considering point 2, it is necessary to fulfill the condition Δf_i - increases:

$$\Delta f_2(4,2,1,3) = (x_2 - x_4)(-4x_1 + 0.5x_2 + 0.5x_4 + 7x_3) = 6,5.$$

Consequently, the target function increases by 6.5 units, so there is a need to check the increment of additional restrictions:

$$\begin{cases} \Delta g_1^1(4,2,1,3) = 6 \geq -10, \\ \Delta g_2^1(4,2,1,3) = -7 \leq 19. \end{cases}$$

Constraints are met, so the support solution can be improved:

$$\max f_2(4,2,1,3) = f_1(4,3,1,2) + \Delta f_2(4,2,1,3) = 46,5.$$

$$\begin{cases} \Delta g_1^1(4,2,1,3) \geq -16, \\ \Delta g_2^1(4,2,1,3) \leq 26. \end{cases}$$

Consider the point (3,2,1,4), then you need to transpose $x_1 \leftrightarrow x_4$ regarding point 2 (4,2,1,3) :

$$\Delta f_{14}(3,2,1,4) = \frac{1}{2}(x_1 - x_4)(3x_1 - 6x_2 + 2x_3 + 3x_4) = \frac{1}{2}(3 - 4)(9 - 12 + 2 + 12) = -5,5.$$

Since the target function decreases by 5.5 units, there is no point in considering this point of the set of permutations.

Consider ascending, point (4,2,3,1), respectively, transposition, $x_3 \leftrightarrow x_4$ regarding point 2 (4,2,1,3) :

$$\Delta f_{34}(3,2,1,4) = \frac{1}{2}(x_3 - x_4)(6x_2 - 8x_1 + x_3 + x_4) = \frac{1}{2}(3 - 1)(12 - 32 + 3 + 1) = -16.$$

The target function decreases by 16 units, respectively, this point is not considered.

Therefore, point 2 (4,2,1,3) is optimal and cannot be improved.

$$\max f_2(4,2,1,3) = 46,5.$$

6 Conclusion

The article represents an optimization problem model with a quadratic target function and additional constraints on the combinatorial set of permutations as a model of many applied problems. An algorithm for solving the optimization problem is proposed and a numerical example is demonstrated. It should be noted that the problem are very complex from a computational point of view. Therefore, their solutions re-

quire a lot of time and resources. This method significantly simplifies the procedure for finding the optimal solution of an optimization problem with a quadratic target function and additional constraints on the combinatorial set of permutations, since the inequality of restrictions increments allows you to immediately determine whether the point of the permutation set is a support solution or not. There is not a necessity to do complex calculations of all constraints and the target function; it suffices to find the increment of constraint and functions in the case of an improvement the solution.

The further development of this study is going to be aimed at realizing and adapting the formulated method on other combinatorial constructions, as well as developing new methods for solving combinatorial optimization problems, taking into account the input data.

References

1. Korte, B., Vygen, J.: Combinatorial Optimization: Theory and Algorithms. Springer, New York (2018)
2. Pardalos, P.M., Du, D.-Z., Graham, R.L. (Eds.): Handbook of combinatorial optimization. Springer, New York (2013)
3. Bernhard, K., Jens, V.: Combinatorial Optimization, Theory and Algorithms. Springer-Verlag, Berlin (2012)
4. Karin, R., Saoub, A.: Tour through Graph Theory. Chapman and Hall/CRC, New York (2017)
5. Berge, C.: Principes de combinatoire. Dunod, Paris (1968)
6. Yemelichev, V.A., Kovalev, M.M., Kravtsov, M.K.: Polytopes, graphs and optimisation. Cambridge University Press, Cambridge (1984)
7. Sachkov, V.N.: Combinatorial methods of discrete mathematics. Nauka, Moscow (1975)
8. Hulianytskyi, L., Riasna, I.: Formalization and classification of combinatorial optimization problems. Springer Optimization and its Applications. 130: pp. 239-250 (2017)
9. Yakovlev, S.V., Grebennik, I.V.: Localization of solutions of some problems of nonlinear integer optimization. Cybernetics and Systems Analysis. 29(5): pp. 727-734 (1993)
10. Schrijver, A.: Combinatorial Optimization: Polyhedra and Efficiency. Springer Science & Business Media, Berlin (2012)
11. Balakrishnan, R., Sridharan, S.: Foundations of Discrete Mathematics with Algorithms and Programming. Chapman and Hall/CRC, New York (2018)
12. Bona, M.: Combinatorics of Permutations, Second Edition. Chapman and Hall/CRC, New York (2012)
13. Shor, N.Z., Stetsyuk, P.I.: Modified r-algorithm to find the global minimum of polynomial functions Cybernetics and Systems Analysis. 33(4): pp.482-497 (1997)
14. Emets, O.A., Koliechkina, L. N.: Solution of optimization problems with fractional-linear objective functions and additional linear constraints on permutations. Cybernetics and Systems Analysis. 40(3): pp. 329-339 (2004)
15. Donec, G. A., Kolechkina, L. M.: Method of ordering the values of a linear function on a set of permutations. Cybernetics and Systems Analysis. 45(2): 204-213 (2009)
16. Sergienko, I.V., Volkovich, O.V., Roshchin, V.A.: Models and methods of solution of quadratic integer programming problems. Cybernetics. 23(3): 289-305 (1987)

17. Semenova, N.V., Kolechkina, L.N., Nagornaya, A.N.: On Approach to Solving Vector Problems with Fractionally Linear Functions of the Criteria on the Combinatorial Set of Arrangements. *Journal of Automation and Information Sciences*. 42(1): pp. 67-80 (2010)
18. Yakovlev, S.V.: The theory of convex continuations of functions on vertices of convex polygons. *Computational Mathematics and Mathematical Physics*. 34: pp. 959-965 (1994)
19. Yakovlev, S.: Convex extensions in combinatorial optimization and their applications. *Springer Optimization and its Applications*. 130: pp. 567-584 (2017)
20. Yakovlev, S.V.: Bounds on the minimum of convex functions on Euclidean combinatorial sets. *Cybernetics*. 25(3): pp. 385-391 (1989)
21. Pichugina, O.S., Yakovlev, S.V.: Continuous representations and functional extensions in combinatorial optimization. *Cybernetics and Systems Analysis*. 52(6): pp. 921-930 (2016)
22. Yakovlev, S.V.: The method of artificial space dilation in problems of optimal packing of geometric objects. *Cybernetics and Systems Analysis*. 53(5): pp. 725-732 (2017)
23. Stoyan, Y.G., Yakovlev, S.V.: Configuration space of geometric objects. *Cybernetics and Systems Analysis*. 54(5): 716-726 (2018)
24. Yakovlev, S.V.: On some classes of spatial configurations of geometric objects and their formalization. *Journal of Autom and Information Sciences* . 50(9): 38-50 (2018)
25. Stoyan, Yu.G., Sokolovskii, V.Z., Yakovlev, S.V.: Method of balancing rotating discretely distributed masses. *Energomashinostroenie*. 2: pp. 4-5 (1982)
26. Stoyan, Y., Romanova, T., Pankratov, A., Kovalenko, A., Stetsyuk, P.: Balance layout problems: mathematical modeling and nonlinear optimization. *Springer Optimization and its Applications*. vol. 114, pp. 369-400 (2016)
27. Gulianitsky, L.F. Sergienko, I.V.: Meta-evolutionary method of deformed polyhedron in combinatorial optimization. *Cybernetics and system analysis*. 44(6): pp. 70-79 (2007)
28. Stoyan, Y.G., Yakovlev, S.V., Parshin, O.V.: Quadratic optimization on combinatorial sets in R^n . *Cybernetics and Systems Analysis*. 27(4): pp. 561–567 (1991)
29. Yakovlev, S.V., Valuiskaya, O.A.: Optimization of linear functions at the vertices of a permutation polyhedron with additional linear constraints. *Ukrainian Mathematical Journal*. 53(9): pp. 1535-1545 (2001)
30. Pichugina, O.S., Yakovlev, S.V.: Functional and analytic representations of the general permutations. *Eastern-European Journal of Enterprise Technologies*. 1(4): pp. 27-38 (2016)
31. Pichugina, O., Yakovlev, S.: Convex extensions and continuous functional representations in optimization, with their applications. *Journal of Coupled Systems and Multiscale Dynamics*. 4(2): pp. 129–152 (2016)
32. Yakovlev, S.V., Pichugina, O.S.: Properties of combinatorial optimization problems over polyhedral-spherical sets. *Cybernetics and Systems Analysis*. 54(1): pp. 99-109 (2018)
33. Kolichekina, L. N., Dvirna, O. A., Nagornaya, A. N.: Modified Coordinate Method to Solve Multicriteria Optimization Problems on Combinatorial Configurations. *Cybernetics and Systems Analysis*. 59(4): pp. 620-626 (2014)
34. Kolichekina, L. N., Dvirna, O. A.: Solving Extremum Problems with Linear Fractional Objective Functions on the Combinatorial Configuration of Permutations Under Multicritheriality. *Cybernetics and Systems Analysis*. 53(4): pp. 590-599 (2017)
35. Donec, G. A., Kolechkina, L. M.: Construction of hamiltonian paths in graphs of permutation polyhedra. *Cybernetics and Systems Analysis*. 46(1): pp.7-13 (2010)
36. Kolechkina, L. M., Dvirna, O. A.: Solving extremum problems with linear fractional objective functions on the combinatorial configuration of permutations under multicritheriality. *Cybernetics and Systems Analysis*. 53(4): pp. 590-599 (2017)

37. Kolehckina, L. M., Dvirna, O. A., Nagornaya, A. N.: Modified coordinate method to solve multicriteria optimization problems on combinatorial configurations. *Cybernetics and Systems Analysis*. 50(4): pp. 620-626 (2014)
38. Kolehckina, L., Pichugina, O.: Multiobjective Optimization on Permutations with Applications, In *DEStech Transactions on Computer Science and Engineering*, 2018 IX International Conference on Optimization and Applications (OPTIMA 2018) (Supplementary Volume), pp. 61-75 (2018)

The sampling method preserving interclass boundaries

Dmytro Kavrin¹[0000-0002-8952-4067], Sergey Subbotin²[0000-0001-5814-8268],

^{1,2}Zaporizhzhia National Technical University, Zhukovsky str., 64,
Zaporizhzhia, 69063, Ukraine

¹kavrin@gmail.com, ²subbotin@zntu.edu.ua

Abstract. The sample dimensionality reduction problem for classification is addressed. The sampling method with the preservation of the most significant instances near the interclass boundaries is proposed. It calculates the interclass distances in the sample used to build hyperspheres, removes redundant instances inside the hyperspheres, and creates a subsample from the set of hypersphere centers. The experiments to study the proposed method properties are conducted. They allow recommending the proposed method for use in practice as a significant to reduce the complexity and ensure acceptable accuracy of classification models.

Keywords: class, classification, interclass boundaries, cluster, diagnosis, instance, metric, sample.

1 Introduction

The decision making automation in the problems of technical and biomedical diagnostics requires the construction of classification model that usually created on a set of observations (instances, precedents) characterized by a set of features [1]. In applied problems often we need to operate with a large-size datasets. This leads to a significant expense of time for data processing and requires big computer memory resources. Therefore, the urgent problem is to reduce the data sample dimensionality.

There are two main approaches to solve a data dimensionality reduction problem: a feature selection and a sample selection [2].

The approach of feature selection is the most widely used in practice [3]. It enumerates various combinations of features and determines their suitability for a model building based on a specific criterion for feature combination. Among all the best feature combinations the resulting combination the one is selected as a result that contains the smallest feature number. However, if the feature set is not initially redundant, this approach may lead to the loss of important information.

Another approach is a sample (instance) selection using different methods to reduce the number of instances [4–7]. The most common practice is a random extraction of smaller subsamples from the big original sample. But random methods in most cases cannot guarantee that the resulting sample of a small size will reflect the main properties of the original sample, especially near the interclass boundaries.

The purpose of this paper is to develop a sample selection method, which allows minimizing the size of original samples and at the same time to preserve the most important instances located near the interclass boundaries.

2 Formal problem statement

Let we have an original data sample $X = \langle x, y \rangle$, where $x = \{x^s\}$, $y = \{y^s\}$, $s = 1, 2, \dots, S$, where S is a number of instances, y^s is a class of s -th instance, $x^s = \{x_j^s\}$ is a s -th instance inputs, $j=1, 2, \dots, N$, x_j^s is a value of j -th input feature x_j for s -th instance, N is a number of input features.

Then the problem of sample selection may be presented in the form: find $X' = \langle x', y' \rangle$: $x' \in \{x^s\}$, $y' = \{y^s \mid x^s \in x'\}$, $S' \leq S$, $f(\langle x', y' \rangle, \langle x, y \rangle) \rightarrow \text{opt}$, where $\langle x', y' \rangle$ is a selected subsample, S' is a number of instances in a selected subsample, $f()$ is a criterion describing quality of sample selection, opt is an acceptable value of f criterion. As a rule, for the problems of approximation the model quality criterion is determined as a function of the model error.

3 Literature review

The sample dimensionality reduction methods by extracting smaller subsamples from the original samples can be divided into two main categories: probabilistic and deterministic methods [4–20].

Probabilistic methods involve randomly selecting of each instance (group of instances) from the original sample with a known non-zero probability that can be accurately determined. The probabilistic sampling [4–18] includes the following methods:

- simple random sampling method [10], which allows randomly selecting a given number of instances from the original sample. Moreover, all instances of the original sample have the same probability of being selected;
- systematic selection method [11], which arranges the original sample in a certain form and splits it into consecutive groups of instances. Then an object with a given sequence number from each group is selected and included in the subsample being formed;
- stratified selection method [12], which divides the original sample into non-intersecting homogeneous subsets (strata), including instances of all types. Then random or systematic selection methods are applied to each subset;
- probability proportional to size sampling method [13–14], which is applied when there is additional information about the classes and their size, and the probability of selecting each instance of the original sample will be proportional to the size of the class to which it belongs;

- cluster sampling method [15–18], which divides instances of the original sample into clusters. Then, from each cluster, a subset of the instances for the subsample is randomly selected.

The advantages of probabilistic methods [5, 10] are their relative simplicity and the possibility of estimating the sampling error. The disadvantages of probabilistic methods are that they do not guarantee that a subsample will display the properties of the original sample well or will not be redundant and will not artificially simplify the task.

Deterministic sampling methods [5, 19] involve the selection of instances based on assumptions about their informativeness, which forms the selection criteria. In this case, the sample data contains instances that may not be selected or the probability of their selection cannot be accurately determined. Therefore, the theory developed for probabilistic samples is not applicable to such samples. The deterministic sampling includes the following methods:

- convenience sampling [5, 20], which forms a non-representative sample of the instances most easily available for study;
- quota sampling [5], which divides the original sample into disjoint subgroups with different properties, after which instances are selected from each subgroup based on a given proportion and on the researcher's preferences;
- purposive sampling [20], which extracts instances from the original sample in accordance with the researcher's opinion about their relevance to the study.

A major problem with these methods [19–20] is the impossibility of estimating the error of the formed samples. The advantage of deterministic methods is their ability to identify the most significant instances for building a diagnostic model of precedents, which can also be used to initialize recognition patterns and speed up the learning process.

The original data samples used in solving the problem of building a diagnostic model by precedents can be very large or have redundant data. Model operation using such samples may require significant computational and time resources. Extracting smaller subsamples from the original data is an effective and natural solution to the big data problem when building a diagnostic model by precedents. It is important to preserve significant instances of the original sample to obtain a representative subsample of a smaller volume. Therefore, when constructing diagnostic models, based on cluster analysis deterministic methods are the most relevant, since they make it possible to identify the most significant instances. As a rule, when solving problems of technical and biomedical diagnosis there is information about classes, this greatly simplifies the task of clustering, which can be performed more accurately, especially near the class boundaries. However, in the common case, cluster analysis solves the clustering problem for unlabeled data and does not allow selecting important instances located near interclass boundaries.

Therefore, to create representative subsample of a small size with well-defined class boundaries, it is necessary to develop a method of sampling from the original data, which will reduce the computational load and preserve the topological represen-

tativeness of the original sample in the feature space by keeping significant instances at the interclass boundaries.

4 The sampling method preserving interclass boundaries

To preserve the topological representativeness of the original sample in the feature space by keeping significant instances at the interclass boundaries a deterministic method of labeled datasets reduction is proposed. It separates the original labeled sample by the hyperspheres of different radii in the feature space, depending on the distances between instances of different classes. The reduction is performed sequentially for each class (primary class) relative to all other classes. On each method's iteration, the most perspective instance of primary class is selected. This is calculated using the Gromov–Hausdorff distance between the primary class and the united set of other classes:

$$D_h = \arg \min_{p=1,2,\dots,S_p} \{d(x^p, x^u)\}, \quad (1)$$

where x^p is an instance of the primary class and the most perspective instance, x^u is an instance of the united set of other classes.

Then a hypersphere is formed with a center at the most perspective point and a radius equal to the distance to the nearest instance of the joint set of other classes in the feature space. All instances of primary class inside the hypersphere are excluding from further consideration. The procedure is repeated until all instances of the primary class will be excluded from the consideration. The reduced subsample is formed from the centers of the resulting hyperspheres. The construction of hyperspheres is performed separately for each class of the original sample. Thus, the method adapts to the distribution of data in the sample, automatically adjusting the number of instances in the reduced subsample. Due to the adaptability of the hypersphere radii, closer to the class boundary more hyperspheres of a smaller radius are formed. In other words, in a reduced sample, the density of instances at the class boundary will be higher than far from the boundaries, which allows for more accurate determination of the class boundaries.

The proposed method is based on the hypothesis of compactness of classes. In the ideal case, when classes do not intersect (compact), the proposed method greatly reduces the original sample, including by removing meaningful instances at interclass boundaries, thereby skewing interclass boundaries. In this case, in order to form a more representative subsample, it is proposed to regulate the number of hyperspheres by the fractional change in their radii. With a decrease in the hypersphere radii, the number of clusters increases, and the most of the clusters accumulate at the class boundaries. Thus, in the formed subsample, it is possible to regulate the number of instances, especially at the class boundary, increasing its representativeness. Formally, this method can be written as follows:

1. The initialization stage. Set the initial data sample $X = \langle x, y \rangle$ and initialize the reduced sample $X' = \langle x', y' \rangle = \emptyset$.
2. The stage of sample splitting into classes. Split the original sample $X = \langle x, y \rangle$ into K separate subsamples $X(k)$ for the instances of each class: $X(k) = \bigcup_{s=1}^S \{x^s \mid y^s = k\}$, where $k = 1, \dots, K$. Determine the volume (number of instances) of each k -th subsample S_k .
3. The stage of subsample reduction. Set $k = 0$. Then while $k < K$ perform in a cycle: $k = k + 1$, set primary subsample $P = X(k)$, merge all other subsamples into one virtual subsample $U = \bigcup_{s=1}^S \{x^s \mid y^s \neq k\}$, further while $P \neq \emptyset$ perform in a cycle:
 - calculate the Gromov–Hausdorff distance between the primary subsample P and the virtual subsamples U :

$$d(x^p, x^u) = \arg \min_{p=1,2,\dots,S_p} \left\{ \sqrt{\sum_{j=1}^N (x_j^p - x_j^u)^2} \right\}; \quad (2)$$

- set the instance x^p as the most perspective instance of the primary subsample P and join it to the subsample X' : $X' = X' \cup x^p$;
 - determine the radius of the hypersphere centered at x^p as $r = \lambda d(x^p, x^u)$, where λ is a fraction of distance d , $0 \leq \lambda \leq 1$;
 - join the most perspective instance x^p to the subsample X' : $X' = X' \cup x^p$;
 - remove from the subsample P all instances that are included in the hypersphere of radius r : $P = P \setminus P^r$, where P^r is a set of instances remote from the center x^p of the hypersphere at a distance $r < d(x^a, x^s)$.
4. The stage of model building. Construct the recognition model using reduced sample X' .

The proposed method allows automating the process of reducing the size of the original sample, which contains information about the classes, for solving the problems of technical and biomedical diagnosis.

The disadvantage of the method is the necessity of calculating and memory storing of the pairwise distances between instances of opposite classes. Therefore, if the size of the original sample is large enough and does not allow all pairwise distances to be simultaneously loaded into the computer's memory, or the data are received dynamically, it is possible to process the original sample in packets.

To provide the correct work of the method each packet should be represented by all classes. Applying the described method to the first data packet, we obtain the initial distribution of the extracted sample. The instances of the next data packet are com-

bined with the instances of the sample obtained in the previous step. The combined sample is processed again by the sampling method with preserves interclass boundaries. The procedure is repeated until all the data of the original sample has been processed, or until a stop has been made in the specified phase of the data stream, if the data is dynamic. Formally, this method can be written as follows:

1. The initialization stage. Set the initial data sample $X = \langle x, y \rangle$ and initialize the reduced sample $X' = \langle x', y' \rangle = \emptyset$.
2. The stage of packets initialization. Determine the number of packets in the sample $P = \text{round}(S/Q)$, where Q is a number of instances in the packet, specified by the user, round is an argument rounding function to the nearest integer. Split the original sample X into P packets $X(\rho)$, $\rho = 1, \dots, P$.
3. The stage of packet processing. Set $\rho = 0$. While $\rho < P$, perform in a cycle: set $\rho = \rho + 1$, join packet $X(\rho)$ with reducing sample X' : $X(\rho) = X' \cup X(\rho)$, then process the packet $X(\rho)$ using the sampling method with preserving class boundaries described above using the adaptive reduction method and set $X' = X(\rho)$.
4. The model building stage. Construct the recognition model from a reduced sample X' .

The proposed batch sampling method allows to process of very large samples in a batch of a given size. The data batch processing combined with the sampling method preserving interclass boundaries allows to obtain representative reduced samples from large data samples, or dynamic data sets, without requiring significant computational resources.

5 Experiments

To study the properties of the proposed methods they were implemented as computer software. The set of experiments were conducted to solve the practical problems of classification using the developed software implementing proposed methods.

Considering the possible variability of relative performances of method across datasets, the results were obtained based on two datasets of various size and dimension, characterized in Table 1.

Table 1. Datasets used for validate the method

Dataset	Number of instances	Number of features	Number of classes
Breast Cancer Wisconsin (Diagnostic)	569	30	2
Activity recognition with healthy older people using a batteryless wearable sensor	9101	8	5

The Breast Cancer Wisconsin Diagnostic dataset (BSW) requires to predict whether the cancer is benign or malignant. Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image [21].

The activity recognition with healthy older people using a batteryless wearable sensor (RHOP) dataset contains sequential motion data from 14 healthy older people aged 66 to 86 years old using a batteryless, wearable sensor on top of their clothing for the recognition of activities in clinical environments [22].

In each problem the data were normalized to map feature values into the interval $[0, 1]$:

$$x_j^s = \frac{x_j^s - x_j^{\min}}{x_j^{\max} - x_j^{\min}}, \quad (3)$$

where x_j^s is a j -th feature value of s -th instance of a sample, $x_j^{\min}$ is a minimum value of j -th feature, $x_j^{\max}$ is a maximum value of j -th feature.

Further, each initial data set was divided into test and training samples by the stratification method [12] in the ratio of 25% and 75%, respectively. This made it possible to compare the work of models built on the original sample and reduced sample by the proposed method.

The classification model was built using a two-layer feed-forward neural network with 10 neurons in a hidden layer trained by the Levenberg-Marquardt method and Error backpropagation technique [23].

For the comparison of models built on different training samples the mean square error of the test sample classification was calculated:

$$MSE = \frac{\sum_{s=1}^S (y^s - f(x^s))^2}{S}, \quad (4)$$

To provide a reliable result in our study ten simulations of model building for each training sample were performed. After model testing their MSE values were averaged.

The results of the proposed method depend on the density of the instance distribution near the interclass boundaries and the class compactness. So, for a sample with heavily mixed classes, the size of the reduced subsample may remain almost unchanged, or it may change slightly. If the classes are well separable then the size of the reduced subsample will be small enough and the interclass boundaries will be poorly defined.

Therefore, the key issue of applying the proposed method is the correct selection of the λ coefficient. Selection of λ coefficient allows to change the radii of hyperspheres, thus adjusting the number of instances near the interclass boundaries and the representativeness of the reduced sample. In the experiments, a training set was created for different values of λ coefficient (0.25, 0.5, 0.75, 1). Then, the MSE dependences on the λ coefficient were studied.

6 Results and Discussion

The conducted experiments showed that the proposed method works well in automatic mode due to its adaptability, which is an important factor in solving technical and biomedical diagnostics problems.

The results of the proposed method are presented in Table 2.

Table 2. Results of experiments

λ	BSW		RHOP	
	S'	MSE	S'	MSE
0	427	0.026194	6902	0.003484
0.25	427	0.027152	623	0.010707
0.5	372	0.027807	329	0.007960
0.75	228	0.027599	201	0.012432
1	127	0.035558	136	0.026053

As it can be seen from the Table 2, the results for the different data sets are also different. This is due to the different distribution of instances at interclass boundaries.

In the case of BSW, the original size of the training sample was 427 instances. From Table 2 it can be seen that MSE has changed slightly for all λ values, while the sample size has reduced almost 4 times with $\lambda = 1$. Thus, using the proposed method for the BSW dataset, it was possible to significantly reduce the size and save the most important instances on the interclass boundaries.

The original size of the training sample of the RHOP dataset was 6902 instances. It can be seen that for given λ values, the sample size has significantly reduced, and the value of MSE has increased. This may mean that the classes are very well separated and the density of the instances at the interclass boundaries is low, so the use of the developed method led to a deficiency of instances at the interclass boundaries, respectively to a decrease in the model's efficiency. Therefore, in this case, it is possible to recommend to set the minimum λ values ($0 < \lambda \leq 0.1$).

The proposed method makes possible to find a compromise between the size and representativeness of the reduced samples, depending on the tasks by changing the proportion of the hypersphere radii using the λ coefficient.

The disadvantage of the developed method is that it is computationally expensive, especially for large datasets.

Therefore, there is a need to apply approaches to reduce the method computational load. For example, in the case of large samples, it is possible to use the method in an ensemble with feature selection methods [24].

It is also possible to reduce the computational load by eliminating the stage of determining the most promising instance from the calculations and it's replace by a random instance of the class. However, this approach requires further study and identification of the necessary restrictions. The next approach could be parallelization of the computational load using multiprocessor systems.

7 Conclusion

The problem of reducing the labeled large data samples for diagnostic model building by precedents is addressed in the paper.

The scientific novelty of study results consists in the fact that a new sampling method preserving interclass boundaries has been developed. It makes possible to significantly reduce the size of the original labeled sample, retaining the most significant instances near the interclass boundaries and removing less informative instances located inside the classes. Thus, the proposed method allows in the automatic mode to solve the sample reduction problem adapting to the data distribution in the labeled sample.

The practical significance of the obtained results is that proposed method is implemented as software, which provides possibility of batch processing of large data samples, or samples formed from data streams (dynamically incoming data). This software has been experimentally studied at solving the problems of real datasets sample selection. The experiments were confirmed the efficiency of the developed software and of implemented method. The results of the conducted experiments allow to recommend the use of the developed method and its software implementation for solving the problems of technical and biomedical diagnosis.

The prospects for further research may be concerned on study the proposed method on a wider class of practical problems. Also the study of the possibilities of the proposed method in ensembles with the methods of feature selection for large datasets seems to be appropriate. The development of implementations of the proposed method for multiprocessor systems operating in a parallel mode also is important for many practical problems.

References

1. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2016)
2. García, S., Luengo, J., Herrera, F.: Data Preprocessing in Data Mining. Springer, Switzerland (2016)
3. Khalid, S., Khalil, T., Nasreen, S.: A survey of feature selection and feature extraction techniques in machine learning. In: 2014 Science and Information Conference, pp. 372-378. IEEE Press, London (2014). doi: 10.1109/SAI.2014.6918213
4. Thompson, S.K.: Sampling. John Wiley & Sons, Hoboken (2012)
5. Lavrakas, P.J.: Encyclopedia of survey research methods. Sage Publications, Thousand Oaks (2008)
6. Cochran, W.G.: Sampling Techniques. John Wiley & Sons, New York (1977)
7. Chaudhuri, A., Stenger, H.: Survey sampling theory and method. Chapman & Hall, New York (2005)
8. Subbotin, S.A.: The sample properties evaluation for pattern recognition and intelligent diagnosis. In: The 10th International Conference on Digital Technologies 2014, pp. 321-332. IEEE Press, Zilina (2014). doi: 10.1109/DT.2014.6868734
9. Subbotin, S.A.: The training set quality measures for neural network learning. In: Optical Memory and Neural Networks (Information Optics), vol. 19(2), pp. 126-139. Allerton Press (2010). doi: 10.3103/S1060992X10020037

10. Tille, Y., Wilhelm, M.: Probability Sampling Designs: Principles for Choice of Design and Balancing. In: Statistical Science, vol. 32(2), pp. 176–189. Project Euclid (2017). doi:10.1214/16-STS606
11. Kalton, G.: Systematic Sampling. In: Wiley StatsRef: Statistics Reference Online. John Wiley & Sons (2017). doi: 10.1002/9781118445112.stat03380.pub2
12. Parsons, V.L.: Stratified Sampling. In: Wiley StatsRef: Statistics Reference Online. John Wiley & Sons (2017). doi: 10.1002/9781118445112.stat05999.pub2
13. Skinner, C.J.: Probability Proportional to Size (PPS) Sampling. In: Wiley StatsRef: Statistics Reference Online. John Wiley & Sons (2016). doi: 10.1002/9781118445112.stat03346.pub2
14. Subbotin, S.A.: Methods of sampling based on exhaustive and evolutionary search. In: Automatic Control and Computer Sciences, vol. 47(3), pp. 113–121. Allerton Press (2013). doi: 10.3103/S0146411613030073
15. Nelson, G.F.: Cluster Sampling: A Pervasive, Yet Little Recognized Survey Design in Fisheries Research. In: Transactions of the American Fisheries Society, vol. 143(4), pp. 926–938. John Wiley & Sons (2014). doi: 10.1080/00028487.2014.901252
16. Ly, T., Cockburn, M., Langholz, B.: Cost-efficient case-control cluster sampling designs for population-based epidemiological studies. In: Spatial and Spatio-temporal Epidemiology, vol. 26, pp. 95–105. Elsevier (2018). doi: 10.1016/j.sste.2018.05.002
17. Subbotin, S., Oliinyk, A.: The Sample and Instance Selection for Data Dimensionality Reduction. In: Szewczyk R., Kaliczyńska M. (eds.) Recent Advances in Systems, Control and Information Technology. SCIT 2016. Advances in Intelligent Systems and Computing, vol. 543, pp. 97–103. Springer, Cham (2017). doi: 10.1007/978-3-319-48923-0_13
18. Subbotin, S.: The Instance and Feature Selection for Neural Network Based Diagnosis of Chronic Obstructive Bronchitis. In: Bris R., Majernik J., Panczer K., Zaitseva E. (eds.) Applications of Computational Intelligence in Biomedical Technology. Studies in Computational Intelligence, vol. 606, pp. 215–228. Springer, Cham (2016). doi: 10.1007/978-3-319-19147-8_13
19. Elliott, M.R., Valliant, R.: Inference for Nonprobability Samples. In: Statistical Science, vol. 32(2), pp. 249–264. Project Euclid (2017). doi: 10.1214/16-STS598
20. Etikan, I., Musa, S.A., Alkassim, R.S.: Comparison of Convenience Sampling and Purposive Sampling. In: American Journal of Theoretical and Applied Statistics, vol. 5(1), pp. 1–4. Science PG, New York (2016). doi: 10.11648/j.ajtas.20160501.11
21. UCI machine learning repository. [https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)), last accessed 2019/01/10
22. UCI machine learning repository. <https://archive.ics.uci.edu/ml/datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor>, last accessed 2019/01/10
23. Ravindran, A., Ragsdell, K.M., Reklaitis, G.V.: Engineering optimization: methods and applications. John Wiley & Sons, New Jersey (2018)
24. Subbotin, S.: Quasi-Relief Method of Informative Features Selection for Classification. In: 2018 IEEE 13th International Scientific and Technical Conference on CSIT, pp. 318–321. IEEE Press, Lviv, Ukraine (2018). doi: 10.1109/STC-CSIT.2018.8526627

Information Model of Evaluation and Output Rating of Start-up Projects Development Teams

Volodymyr Polishchuk¹ [0000-0003-4586-1333], Miroslav Kelemen² [0000-0001-7459-927X]

¹ Uzhhorod National University, Narodna Square, 3, Uzhhorod, 88000, Ukraine
volodymyr.polishchuk@uzhnu.edu.ua

² Technical university of Kosice, Rempova str., 7, Kosice, 04121, Slovak republic
miroslav.kelemen@tuke.sk

Abstract. The problem of constructing an informational model of evaluation and output of the start-up team rating is considered. This model is based on neuro-fuzzy network when there are expert fuzzy data on the teams of developers. As the success of a start-up implementation depends on the quality of the team of developers, then the development of such a model will increase the degree of validity of the decision to finance the start-up projects.

Keywords: start-up team, assessment, rating, neuro-fuzzy model

1 Introduction

For any project, there are people who implement it. Even for a very good start-up project, with a very high score and prospects for success, successful commercialization depends, to a greater extent, on a qualitative form of developers, ready to bring the product to the market and successfully decide on its sale. Therefore, investors in start-ups like to say that they primarily invest in a team.

The urgency of the work consists of the developed informational model of evaluation and withdrawal of the start-up team rating using the neuro-fuzzy network when there are expert and sometimes fuzzy data on the team of developers. The development of such a model will allow increasing the degree of validity of financing start-up projects since the success of the start-up implementation directly depends on the qualitative composition of the team of developers.

2 Formal problem statement

Let's formulate the task of evaluating and eliminating the ranking of teams of developers of start-up projects as follows. Let the set of teams of developers be set $X = (x^1, x^2, \dots, x^n)$, which should be evaluated according to many indicators (criteria) $K = (K_{11}, K_{12}, \dots, K_{34})$, organize according to a certain rule and draw a linguistic rating $Y = \{y_1, y_2, \dots, y_5\}$ by command.

Each criterion for evaluating the team of start-up project developers is evaluated expertly using one of the terms, the next term-set of linguistic variables $L=\{H; HC; C; B\}$, where: H – “Low-level indicator”; HC – “Indicator below average”; C – “Average level of the indicator”; B – “High level of the indicator”. Also, for every assessment, the expert puts “confidence factor” d in assigning it an assessment, from the interval $[0; 1]$.

3 Literature review

Analysing scientific sources, we see that there is a need to systematize tools and develop algorithms assessment teams of developers’ start-up projects. Still not produced a holistic concept, definition of the rating of the teams of developers for the successful implementation of the start-up of the projects taking into account the subjective aspects of evaluation. Thus, the problems of project start-up evaluation are raised in the work [1], where the fuzzy set is used and the existing group of criteria "authors of the idea" is used, but not enough attention is paid to the analysis of the teamwork on the project. In the work [2] shows a cognitive star rating model that can be used only as an auxiliary tool for improving decision-making accuracy by venture funds. In [3] offers a fuzzy management model that can help select and filter applications for grants. On the one hand, this model considers business ideas, and on the other - the person of the entrepreneur. The approach is based on linguistic variables, which reveals subjectivity.

Fuzzy exclusion systems can use human expertise and perform fuzzy output to obtain initial estimates [4-5]. Formation of rules and related membership functions very much depends on a priori knowledge of the system under consideration. Therefore, there is no universal way of transforming the experimental knowledge of human experts into the knowledge base of the fuzzy output system. Therefore, there is also a need to develop teaching methods for obtaining an initial assessment with the required level of accuracy. In addition, the mechanism of training neural networks does not rely on human expertise, but through a homogeneous structure of neural networks [6-8] it is difficult to extract structured knowledge. Therefore, for the task of evaluating and withdrawing the rating of the team of developers of the start-up projects, it is necessary to develop its own neuro-fuzzy network, working with fuzzy expert input signals and based on the knowledge base displays adequate results [9-10].

Selected theoretical framework within the Simulation and modelling of Security issues is in the work of Fuchs et al. [11] focused on the simulation of dangerous substances outflows into the environment because of traffic accidents by dangerous substances transport, in the study of Dvorak et al. [12] on the enhancing of security on critical accident locations using telematics support, in the work of Balatka et al. [13] on the exposure of the environment and surface water by dangerous liquid - the slop outflow model, or the modelling and evaluation of risks in Soušek et al. [14], or Madarász [15] on the situational Management Methodology and its Application.

Consequently, there are no special models for evaluating and withdrawing the ratings of developers implementing the start-up projects.

4 Neuro-fuzzy model for outputting the ranking of start-up project teams

We describe a neuro-fuzzy model of teams' start-up evaluation, based on input linguistic terms. Input signals are presented in the form of linguistic terms and coefficients of expert confidence in their assignment.

Let the input of the neuro-fuzzy network provide expert data start-up teams (alternatives) $X = (x^1, x^2, \dots, x^n)$ by the set of criteria $K = (K_{11}, K_{12}, \dots, K_{34})$. The criteria are divided into three groups, and the second group has two subgroups of criteria. For each criterion, we obtain a linguistic variable $L = \{H; HC; C; B\}$ and "confidence factor" d in the assignment expert assessment [16]. For example, if the answer is not the one that corresponds to the developer team, then the metric d corrects the accuracy of the answer.

Then let's look at the object of the species $Y = f(x^1, x^2, \dots, x^n)$ for which the connection "input x^k – output Y " can be submitted in the form of an expert matrix U , Table 1.:

Table 1. Expert matrix U

Name criterion	Input signals						
	x^1		x^2		...	x^n	
	L	d	L	d		L	d
K_{11}	L_{11}^1	d_{11}^1	L_{11}^2	d_{11}^2	...	L_{11}^n	d_{11}^n
K_{12}	L_{12}^1	d_{12}^1	L_{12}^2	d_{12}^2	...	L_{12}^n	d_{12}^n
K_{21}	L_{21}^1	d_{21}^1	L_{21}^2	d_{21}^2	...	L_{21}^n	d_{21}^n
...	...	...	...	...	...	...	...
K_{25}	L_{25}^1	d_{25}^1	L_{25}^2	d_{25}^2	...	L_{25}^n	d_{25}^n
K_{31}	L_{31}^1	d_{31}^1	L_{31}^2	d_{31}^2	...	L_{31}^n	d_{31}^n
...	...	...	...	...	...	...	...
K_{34}	L_{34}^1	d_{34}^1	L_{34}^2	d_{34}^2	...	L_{34}^n	d_{34}^n
Exit by alternatives	y_g		y_g		...	y_g	

The set of fuzzy knowledge base rules of production follows:

If [$K_{11} = (L_{11}^k; d_{11}^k)$ (with weight α_{11}) and $K_{12} = (L_{12}^k; d_{12}^k)$ (with weight α_{12})] (with weight α_1) **also** [$K_{21} = (L_{21}^k; d_{21}^k)$ (with weight α_{21}) and ...and $K_{23} = (L_{23}^k; d_{23}^k)$ (with weight α_{23})] (with weight β_{21})] and [$K_{24} = (L_{24}^k; d_{24}^k)$ (with weight α_{24}) and $K_{25} = (L_{25}^k; d_{25}^k)$ (with weight α_{25})] (with weight β_{22})]]

(with weight α_2) **also** $[K_{31}=(L_{31}^k;d_{31}^k)$ (with weight α_{31}) and ...and $K_{34}=(L_{34}^k;d_{34}^k)$ (with weight α_{34})] (with weight α_3) **then** $Y=y_g, \quad k=\overline{1,n};$
 $g=\overline{1,5}$.

Where $K_{ij}, i=\overline{1,3}; j=\overline{1,5}$ – criterion of evaluation of the i -th group, j – serial number of the rule in the group; L_{ij} – variable with term-set L for the j -th group indicator i ; d_{ij} – “confidence factor” expert on assigning a variable L_{ij} ; $(L_{ij}^k;d_{ij}^k)$ – grouped input data received from κ -th start-up team by K_{ij} criterion; $\alpha_{11}, \alpha_{12}, \alpha_{21}, \dots, \alpha_{25}, \alpha_{31}, \dots, \alpha_{34}$ – synaptic weight criteria from the interval $[1;b]$; $\beta_{21}, \beta_{22} \in [1;b]$ – synaptic weight for subgroups of the criteria of the second group; $\alpha_1, \alpha_2, \alpha_3$ – synaptic weight groups of criteria according to the interval $[1;b]$; $Y = \{y_1, y_2, y_3, y_4, y_5\}$ – linguistic interpretation of the rankings of the teams of developers of the start-up.

Getting an aggregated rating of the start-up team rating can be presented in the form of a four-layer neuro-fuzzy network of type integrated neuro-fuzzy systems (similar to Mamdani neurofuzzy approximator), Figure 1.

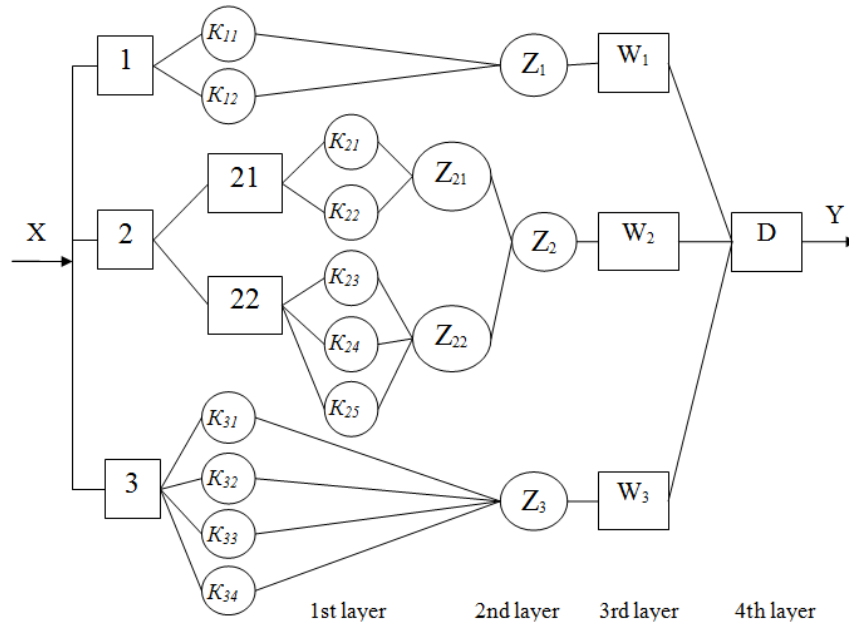


Fig. 1. The structure of the neuro-fuzzy network

Next, consider in more detail what happens on each layer of the neuro-fuzzy network.

1st layer

In the neurons of the first layer, the fuzzification operation is performed, that is, for each input value $(L_{ij}^k; d_{ij}^k)$ the value of the membership function is brought into conformity $\mu(O_{ij}^k)$. Therefore, at the first level, it is necessary to build membership rules in order to get a normalized estimate of the input data.

Let the term-set of linguistic variables $L=\{H; HC; C; B\}$ represent on a certain numerical interval $[a_1; a_5]$, where $H \in [a_1; a_2]$, $HC \in [a_2; a_3]$, $C \in [a_3; a_4]$, $B \in [a_4; a_5]$. The value of breakdowns may be determined in the learning process of a neuro-fuzzy network using real data from teams of developers of start-up projects.

Calculate criterion estimates $O_{ij}^k, k = \overline{1, n}$ with the help of a characteristic function:

$$O_{ij}^k = \begin{cases} a_2 \cdot d_{ij}^k, & \text{if } L_{ij}^k \in H; \\ a_3 \cdot d_{ij}^k, & \text{if } L_{ij}^k \in HC; \\ a_4 \cdot d_{ij}^k, & \text{if } L_{ij}^k \in C; \\ a_5 \cdot d_{ij}^k, & \text{if } L_{ij}^k \in B. \end{cases} \quad (1)$$

This will make it possible to adjust the assessment regarding the expert's confidence in its assignment, or how close is the answer to the questions of the team of developers to the truth. Without diminishing generality, for example, we introduce the membership rule to help S-similar membership function [17-18]:

$$\mu(O_{ij}^k) = \begin{cases} 0, & O_{ij}^k \leq a_1; \\ 2 \left(\frac{O_{ij}^k - a_1}{a_5 - a_1} \right)^2, & a_1 < O_{ij}^k \leq \frac{a_1 + a_5}{2}; \\ 1 - 2 \left(\frac{a_5 - O_{ij}^k}{a_5 - a_1} \right)^2, & \frac{a_1 + a_5}{2} < O_{ij}^k < a_5; \\ 1, & O_{ij}^k \geq a_5. \end{cases} \quad (2)$$

Constructed in this way, the membership function says that the resulting value $\mu(O_{ij}^k)$ will go to 1, in case if the high estimation of the project by the criterion and the sufficiently high confidence of the expert on its assignment. Therefore, of course, S-similar membership function best suited for this task.

Thus, we turn from experts' evaluations of teams of developers of start-up projects and expert confidence in their assignment to normalized comparable data [19].

For example, if we take the interval value $[a_1; a_5] = [0; 10]$, then the membership function (2) will have the form:

$$\mu(O_{ij}^k) = \begin{cases} 0, & O_{ij}^k \leq 0; \\ 0,02 \cdot (O_{ij}^k)^2, & 1 < O_{ij}^k \leq 5; \\ 1 - 0,02 \cdot (10 - O_{ij}^k)^2, & 5 < O_{ij}^k < 10; \\ 1, & O_{ij}^k \geq 10. \end{cases} \quad (3)$$

The membership function constructed in this way has the following content, if the answer to the question corresponds to the high value of the term – B and "confidence factor" expert is low, at level 0,2, then the value of the membership function is obtained accordingly $\mu(O_{ij}^k)$ will be low – 0,08.

2nd layer

On the second layer, the calculation of functions of postsynaptic potential is grouped according to the criteria of evaluation. The second layer contains the number of neurons that corresponds to the number of groups of criteria.

Let the person who makes the decision set the synaptic weights $\alpha_{11}, \alpha_{12}, \alpha_{21}, \dots, \alpha_{25}, \alpha_{31}, \dots, \alpha_{34}$, from the interval $[1; b]$ for each criterion and set the synaptic weight of the rules for the subgroups of the second group of criteria β_{21}, β_{22} from the interval $[1; b]$. We calculate the functions of postsynaptic potential as follows:

$$Z_1^k = \frac{1}{\alpha_{11} + \alpha_{12}} \cdot (\mu(O_{11}^k) \cdot \alpha_{11} + \mu(O_{12}^k) \cdot \alpha_{12}), \quad k = \overline{1, n}, \quad (4)$$

$$Z_2^k = \frac{1}{\beta_{21} + \beta_{22}} \cdot (Z_{21}^k \cdot \beta_{21} + Z_{22}^k \cdot \beta_{22}), \quad (5)$$

$$\text{where } Z_{21}^k = \frac{1}{\alpha_{21} + \alpha_{22}} \cdot (\mu(O_{21}^k) \cdot \alpha_{21} + \mu(O_{22}^k) \cdot \alpha_{22}), \quad (6)$$

$$Z_{22}^k = \frac{1}{\alpha_{23} + \alpha_{24} + \alpha_{25}} \cdot (\mu(O_{23}^k) \cdot \alpha_{23} + \mu(O_{24}^k) \cdot \alpha_{24} + \mu(O_{25}^k) \cdot \alpha_{25}), \quad (7)$$

$$Z_3^k = \frac{1}{\alpha_{31} + \alpha_{32} + \alpha_{33} + \alpha_{34}} \cdot \left(\mu(O_{31}^k) \cdot \alpha_{31} + \mu(O_{32}^k) \cdot \alpha_{32} + \mu(O_{33}^k) \cdot \alpha_{33} + \mu(O_{34}^k) \cdot \alpha_{34} \right), \quad k = \overline{1, n}. \quad (8)$$

Output neurons of the second layer Z_1, Z_2, Z_3 will be normalized since the calculations use the relative importance of the synaptic scales of the criteria.

3rd layer

On the third layer, the second layer of neurons is corrected in relation to the importance of one or the other group of evaluation criteria. In this case, for each group of

criteria, person who makes the decision has his own considerations regarding the synaptic weights $\alpha_1, \alpha_2, \alpha_3$ respectively, from some interval $[1; b]$. We compute the functions of the postsynaptic potential of the third layer of neurons in the following way:

$$W_1^k = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_3} \cdot Z_1^k, \quad (9)$$

$$W_2^k = \frac{\alpha_2}{\alpha_1 + \alpha_2 + \alpha_3} \cdot Z_2^k, \quad (10)$$

$$W_3^k = \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} \cdot Z_3^k, \quad k = \overline{1, n}. \quad (11)$$

Similarly, the output neurons of the third layer W_1, W_2, W_3 will be normalized.

4th layer

On the fourth layer, we will be defuzzification the data. To do this, use the following activation function in the output neuron:

$$Z^k = W_1^k + W_2^k + W_3^k, \quad k = \overline{1, n}. \quad (12)$$

As a result of the training of the neuro-fuzzy network, the rankings of teams of start-up design teams for comparing the aggregated score are determined Z with output variable $Y = \{y_1, y_2, y_3, y_4, y_5\}$ as follows: $Z \in (0, 87; 1] - y_1$; $Z \in (0, 67; 0, 87] - y_2$; $Z \in (0, 37; 0, 67] - y_3$; $Z \in (0, 21; 0, 37] - y_4$; $Z \in [0; 0, 21] - y_5$.

5 Training a neuro-fuzzy network

We offer the method of forming the knowledge base by generating new production rules that do not contradict the rules from the knowledge base of the system, based on the analysis of experimental data about the teams of developers [9].

Let's have a sample S value pairs $\langle x^s, Z^s \rangle, s = \overline{1, S}$. Method of the formation knowledge base of the start-up team developer is next.

Stage 1. With $m, (m < S)$ arbitrary values $\langle x^s, Z^s \rangle$, the initial knowledge base of the model, which is represented by a matrix with strings, is composed $\langle x^s, Z^s \rangle = \langle K_{11}^s, K_{12}^s, \dots, K_{34}^s, Z^s \rangle$. This representation is equivalent to the formulated set of production rules, the fuzzy knowledge base described above.

Stage 2. Next, for each new experimental point $\langle x^*, Z^* \rangle$ we calculate the predicted

value by the centroid method [8]:

$$Z_{new}^* = \frac{\sum_{s=1}^m Z^s \mu(\|x^s - x^*\|)}{\sum_{s=1}^m \mu(\|x^s - x^*\|)}. \quad (13)$$

Where μ – function of exponential form: $\mu(\|x^s - x^*\|) = \exp(-\lambda \sum_{h=1}^{hs} |x_h^s - x_h^*|)$, λ – function parameter (considered predefined), hs – number of rules.

Stage 3. If $|Z^* - Z_{new}^*| > \varepsilon$, where ε – a constant is given that determines the error of the approximation, then the knowledge base is replenished by expanding the matrix U , in the opposite case, the matrix U remains unchanged.

Stage 4. The rule of stop is checked. In this variant, the construction of the model is considered complete if, in accordance with steps 2 and 3, all are selected S experimental points, otherwise we go to stage 2.

It was accomplished training of the neuro-fuzzy network on a training set of data from a university team of developers (a total of 23 teams) taken from Incubator of Uzhhorod National University. Verified correctness of work the neuro-fuzzy network based on test data of successful start-up projects and their developers. Based on the training of the neuro-fuzzy network, the rankings of teams of developers of start-up projects are set. The described teaching method corresponds to the simplified method of fuzzy logic output but differs that the knowledge base is not fixed but is complemented by the arrival of experimental data. The contradiction of the new production rule is guaranteed by the procedure for updating the knowledge base [20].

6 Informational model for assessment teams of developers' start-up projects

Consider which characteristics are typical for an effective team? For this purpose, we propose, for example, the following set of criteria for evaluating the start-up team of developers divided into three groups. Evaluation criteria are presented in the form of a questionnaire, where each team chooses the answer that comes closest to them.

The first set of criteria is stability and team cohesion. For this group we offer the following indicators and options for answers:

K_{11} – The length of work in the project, measured in months of work on the project: 1. from 0 to 3 months; 2. from 3 to 6 months; 3. from 6 to 12 months; 4. more than 12 months.

K_{12} – The stability of the team is determined by the change of leaders and team members:

1. Completely new team members and part of the leaders;

2. The slight change in the number of team members;
3. The composition of the team is unchanged, as all members and leaders meet the requirements of professionalism;
4. The initial membership of the team is unchanged, but there was an expansion of the members and team leaders for the highest competence of the project.

The second group of criteria is professional competence and team experience. For this group, we propose to divide into two subgroups: professional competencies of leaders and professional competencies of team members.

The professional competence of leaders.

K_{21} – Successful work experience on topics or close to it:

1. Experience is absent as this project is the first one;
2. Availability of the first experience on the subject and receiving a small income;
3. A successful innovative project on the subject has been implemented;
4. Leaders have implemented a successful project on topics or close to it.

K_{22} – Successful management experience:

1. Management experience is absent as this project is the first;
2. Management experience is available but insignificant;
3. Middle managers are available;
4. Available high-level managers.

K_{23} – Education leaders:

1. Technical or managerial education is absent;
2. Graduated from college or university student in the technical or managerial direction;
3. Completed higher technical or managerial education;
4. Available degree from at least one of the leaders.

Assessment of professional competence of team members.

K_{24} – Successful experience in large or similar projects:

1. Work experience is absent as this project is the first one;
2. Work experience available but in small projects;
3. Available experience in large projects but not in all team members;
4. All team members have experience in large or successful projects.

K_{25} – Professional education of team members:

1. Team members do not have special education to implement the project;
2. Some team members have a special education to implement the project;
3. Most team members have a special education to implement the project;
4. All team members have a special education to implement the project.

The third group of criteria is the professional activity of the team.

K_{31} – Team participation in professional project conferences, investment sessions or profile events:

1. There is no involvement of professional project activities;
2. There is a single activity;
3. Available activity;
4. Existing and systematic activity of advanced training.

K_{32} – Publications in the media or professional online sources for the project:

1. No posts;
2. Available information about the project and the team, but mainly in social networks;
3. There is no single information about the project and the team;
4. Available and systematic activity of publishing and popularizing the project.

K_{33} – The presence of team ties in social networks and messengers:

1. No links; 2. There are insignificant, isolated links; 3. A wide range of mutual friends in various social networks;
4. Great activity with a large number of subscribers.

K_{34} – The presence of communications with advisers in social networks:

1. No links; 2. There are insignificant, isolated links; 3. Available links;
4. Wide circle of friends.

So, “Low-level indicator” will be considered as the first answer to the question, and the last answer, respectively, is “High level of the indicator”.

Scale of the output variable $Y = \{y_1, y_2, y_3, y_4, y_5\}$ we (&) offer the following:

y_1 = “The rating of the project start-up team is high”. The highest level of start-up team rating. Very low expectations regarding the risks of non-fulfilment of project development obligations. Very high ability to respond and solve current or strategic problems of project realization in a timely manner.

y_2 = “The rating of the project start-up team is higher than the average”. High ranking team start-up. Low expectations of non-fulfilment of project development obligations. Ability to react in a timely manner and solve current or strategic problems of project implementation. However, negative changes in circumstances and economic conditions are likely to reduce this ability.

y_3 = “The rating of the project start-up team is average”. Speculative level of start-up team rating. There is a possibility of development of project risks or the risk of conflicts in the middle of the team, especially as a result of negative economic changes that may occur over time.

y_4 = “The rating of the project start-up team is low”. The rating says that realizing the project in time is not a real opportunity. The ability to fulfil the project obligations of the team entirely depends on the favourable business and economic conditions.

y_5 = “The rating of the project start-up team is very low”. Very high risks of non-fulfilment of project development obligations. Formed start-up team is not able to work on a project.

7 General algorithm for obtaining a rating assessments and ranking start-up command

1st step. For the considered teams, developers of start-up projects $X = (x^1, x^2, \dots, x^n)$ conduct an expert survey and get the input data separately for each team.

2nd step. Person who makes the decision sets his own wishes for the synaptic scales of the criteria – $\alpha_{11}, \alpha_{12}, \alpha_{21}, \dots, \alpha_{25}, \alpha_{31}, \dots, \alpha_{34} \in [1; b]$, synaptic weights of sub-groups for the second group of criteria – $\beta_{21}, \beta_{22} \in [1; b]$ and synaptic scales of the criteria groups – $\alpha_1, \alpha_2, \alpha_3 \in [1; b]$.

3rd step. We make fuzzification of the input signals $(L_{ij}^k; d_{ij}^k)$ in neurons of the first layer, according to (1)-(2), and we obtain the value of the membership function $\mu(O_{ij}^k)$.

4th step. We calculate the output of the neuron with the following activation function (14).

$$Z^k = \frac{1}{\alpha_1 + \alpha_2 + \alpha_3} \cdot \left(\alpha_1 \cdot \left(\frac{1}{\alpha_{11} + \alpha_{12}} \cdot (\mu(O_{11}^k) \cdot \alpha_{11} + \mu(O_{12}^k) \cdot \alpha_{12}) \right) + \right. \\ \left. + \alpha_2 \cdot \left[\frac{1}{\beta_{21} + \beta_{22}} \cdot \left(\left(\frac{1}{\alpha_{21} + \alpha_{22}} \cdot (\mu(O_{21}^k) \cdot \alpha_{21} + \mu(O_{22}^k) \cdot \alpha_{22}) \right) \cdot \beta_{21} + \right. \right. \right. \\ \left. \left. \left. + \frac{1}{\alpha_{23} + \alpha_{24} + \alpha_{25}} \cdot \left(\mu(O_{23}^k) \cdot \alpha_{23} + \mu(O_{24}^k) \cdot \alpha_{24} + \mu(O_{25}^k) \cdot \alpha_{25} \right) \right) \cdot \beta_{22} \right] + \right. \\ \left. + \alpha_3 \cdot \left(\frac{1}{\alpha_{31} + \alpha_{32} + \alpha_{33} + \alpha_{34}} \cdot (\mu(O_{31}^k) \cdot \alpha_{31} + \mu(O_{32}^k) \cdot \alpha_{32} + \mu(O_{33}^k) \cdot \alpha_{33} + \mu(O_{34}^k) \cdot \alpha_{34}) \right) \right] \quad (14)$$

5th step. Defuzzification of data and ranking of start-up teams $Y = \{y_1, y_2, y_3, y_4, y_5\}$.

6th step. Ranking of teams of developers. Based on quantities $Z^k(x^k), k = \overline{1, n}$ we build a ranking line of developers of start-up projects:

$$Z = (Z^1, Z^2, \dots, Z^n). \quad (15)$$

8 Experiments and results

Let the venture fund get 5 start-ups of transport projects submitted by teams of developers – $X = (x^1, x^2, \dots, x^5)$, which should be evaluated, bring the rating of the successful implementation of the project by the team and build their ranking line. All

considered developers of the start-up projects are real, and data is taken from the university incubators (Uzhhorod National University and Technical university of Kosice). We evaluate according to the developed neuro-fuzzy model and constructed general algorithm.

1st step. The teams were expertly evaluated and received the following input scores. Table 2.

Table 2. Input signals by evaluation criteria

Name criteria	x^1		x^2		x^3		x^4		x^5	
	L	d	L	d	L	d	L	d	L	d
K_{11}	B	0,8	C	0,7	H	0,9	B	0,9	H	0,6
K_{12}	C	0,9	C	0,6	HC	0,8	C	0,8	HC	0,7
K_{21}	B	0,7	H	0,5	C	0,8	B	0,7	B	0,5
K_{22}	C	0,8	HC	0,8	C	0,7	C	0,9	HC	0,8
K_{23}	HC	0,6	HC	0,6	C	0,8	B	0,9	HC	0,6
K_{24}	C	0,5	C	0,8	C	0,6	C	0,8	C	0,9
K_{25}	C	0,7	HC	0,9	C	0,7	HC	0,7	C	0,8
K_{31}	HC	0,8	HC	0,8	B	0,8	HC	0,9	C	0,9
K_{32}	B	0,9	H	0,8	B	0,6	B	0,9	H	0,8
K_{33}	B	0,9	B	0,9	B	0,7	B	0,6	H	0,7
K_{34}	C	0,8	HC	0,8	C	0,8	HC	0,8	C	0,9

2nd step. Person who makes the decision sets his own wishes for the synaptic scales of the criteria (8; 9; 8; 10; 9; 10; 7; 8; 6; 7; 9) $\in [1;10]$, synaptic weights of subgroups for the second group of criteria – (10; 8) $\in [1;10]$ and synaptic scales of the criteria groups – (10; 9; 8) $\in [1;10]$.

3rd step. We perform fuzzification of the input signals in the neurons of the first layer. To do this, we define the membership function on a numerical interval $[0;10]$, where $H \in [0;2]$, $HC \in [2;5]$, $C \in [5;8]$, $B \in [8;10]$. We use formula (3) and get the value of the membership function $\mu(O_{ij}^k)$, the result will be written in Table 3.

4th step. We calculate the output of the neuron by the activation function (14): $Z^1 = 0,7383$; $Z^2 = 0,4238$; $Z^3 = 0,5223$; $Z^4 = 0,7381$; $Z^5 = 0,3613$.

5th step. Defuzzification of data and ranking of start-up teams:

“team rating x^1 – higher than the average”; “team rating x^2 – average”; “team rating x^3 – average”; “team rating x^4 – higher than the average”; “team rating x^5 – low”.

Table 3. Fuzzification of input signals

Name criteria	x^1	x^2	x^3	x^4	x^5
K_{11}	0,920	0,613	0,065	0,980	0,029
K_{12}	0,843	0,461	0,320	0,741	0,245
K_{21}	0,820	0,020	0,741	0,820	0,500
K_{22}	0,741	0,320	0,613	0,843	0,320
K_{23}	0,180	0,180	0,741	0,980	0,180
K_{24}	0,320	0,741	0,461	0,741	0,843
K_{25}	0,613	0,405	0,613	0,245	0,741
K_{31}	0,320	0,320	0,920	0,405	0,843
K_{32}	0,980	0,051	0,680	0,980	0,051
K_{33}	0,980	0,980	0,820	0,680	0,039
K_{34}	0,741	0,320	0,741	0,320	0,843

6th step. Based on the initial estimates, we build a ranking line-up of start-up project developers: $Z = (x^1, x^4, x^3, x^2, x^5)$. We conclude that the best team of developers submitted the start-up of the project – x^1 with a rating higher than the average.

9 Discussion of results

The informational neuro-fuzzy model of the output of the start-up team's ranking has been constructed with a number of advantages, namely: raises the objectivity of expert assessments in the evaluation of teams of developers, using incoming linguistic variables and “confidence factor” expert opinion on their assignment; based on a neuro-fuzzy network that has the ability to change the settings of synaptic weights: criteria and groups of criteria for evaluating teams of developers; when receiving experimental data, we can conduct neuro-fuzzy network training by completing the knowledge base and adjusting the rankings of teams of developers of start-up projects.

The disadvantages of this approach can be attributed to the fact that the acquired membership function in the neuro-fuzzy network corresponds to the stage of rough debugging. Therefore, the process of debugging a neuro-fuzzy network, which depends on the partition of the interval $[a_1; a_5]$ possible if there, is a sample of reliable experimental data. In addition, the learning process of the neuro-fuzzy network requires a large amount of real reliable data from the teams of developers and the results of the successful implementation of the start-up projects.

10 Conclusion

The research of the actual task of the development of an information model of assessment and output of the start-up team rating was conducted using neuro-fuzzy network. To do this, the following tasks were solved. For the first time, a four-layer neuro-fuzzy model was developed to obtain a resultant estimate. The production rules of the fuzzy knowledge base are formulated. The model does not require much computation, reveals the subjectivity of expert opinions and displays the rating of teams of developers. The approach to training developed by the neuro-fuzzy network of team start-up evaluation and the method of forming the knowledge base by generating new production rules are given. A general 5-step algorithm for constructing a rating and rank-starter commands is described. For the information model, for the first time, a set of 11 criteria for evaluating start-up project teams has been formed, classified them into 3 groups and presents the input data in the form of 4 linguistic terms and the expert confidence coefficient for their assignment. For the first time, there are 5 levels of developer team rankings. The research has been tested and the results of the verification have been verified on the real data of five teams of start-up project developers.

The developed neuro-fuzzy informational model will be a useful tool for substantiating the choice of teams by investors for the implementation of their projects. Further study of the problems we see in approbation of the developed model on a large sample to increase the knowledge base and the accuracy of the evaluation.

Acknowledgments. This work was carried out within the project "The technological aspects of defining the level of security of project finance in the fight against financial fraud in the FINANCIAL AND TRANSPORT sectors" funded National Scholarship Program of the Slovak Republic.

References

1. Polishchuk, V., Malyar, M., Sharkadi, M., Liakh, I.: Model of start-ups assessment under conditions of information uncertainty. *EEJET* 3/4 (81), 43-49 (2016). doi:10.15587/1729-4061.2016.71222
2. Csaszar, F.: Strategic and cognitive criteria for the selection of startups. *Original Research Article Technovation* 26, 151-161 (2006)
3. Mendialdua, J.C.: Using fuzzy logic in selecting people and ideas to participate in public programs of support to business start-ups. *Cuadernos de Gestion* 14(2), 73-98 (2014)
4. Zade, L.: *Ponyatiye lingvisticheskoy peremennoy i yego primeneniye k prinyatiyu priblizhennykh resheniy*. Mir, Moskva (1976)
5. Rotshteyn, O.P.: *Intelektualni tekhnolohiyi identyfikatsiyi: nechitki mnozhyny, henetychni alhorytmy, neyronni merezhi*. UNIVERSUM, Vinnytsya (1999)
6. Snytyuk, V. YE.: *Prohnozuvannya. Modeli. Metody. Alhorytmy*. Maklout, Kyiv (2008)

7. Subbotin, S. O.: Podannya ta obrobka znan u systemakh shtuchnoho intelektu ta pidtrymky pryynyattya rishen. ZNTU, Zaporizhzhya (2008)
8. Subbotin, S.O., Oliynyk, A.O., Oliynyk, O.O.: Intelektualnyy analiz danykh: navchalnyy posibnyk. ZNTU, Zaporizhzhya (2011)
9. Shin, Y. C., Xu, C.: Intelligent systems: modeling, optimization, and control. Boca Raton, CRC Press (2009)
10. Kruse, R., Borgelt, C., Klawonn, F. et. al.: Computational intelligence: a methodological introduction. London, Springer-Verlag (2013). doi: 10.1007/978-1-4471-5013-8_1
11. Fuchs, P., Novak, P., Saska, T., Smida, J., Dvorak, Z., Kelemen, M., Sousek, R.: Simulation of dangerous substances outflows into the environment because of traffic accidents by dangerous substances transport. In: 14th world multi-conference on systemics, cybernetics and informatics, WMSCI 2010 proceedings, June 29 - July 2 2010, Orlando, Florida, USA: Sánchez, M., vol. 1, pp. 204-207 (2010)
12. Dvorak, Z., Cekerevac, Z., Kelemen, M., Sousek, R.: Enhancing of security on critical accident locations using telematics support. In: International conference on society and information technologies, ICSIT 2010 proceedings, 6-9 April 2010, Orlando, Florida, USA: Sánchez, M., pp. 414-417 (2010)
13. Balatka, M., Fuchs, P., Kamenicky, J., Sousek, R., Kelemen, M.: Exposure of the environment and surface water by dangerous liquid - the slop outflow model. In: 15th World Multi-Conference on Systemics, Cybernetics and Informatics, Proceedings, 19-22 July 2011, Orlando, USA: Sánchez, M., vol. 3, pp. 280-284 (2011)
14. Soušek, R., Šustr, M., Fuchs, P., Endridzalová, E., Novák, M., Müllerová, J.: Evaluation of risks in air transport. In: 22nd World Multi-Conference on Systemics, Cybernetics and Informatics, WMSCI 2018 Conference Proceedings, 8-11 July 2018, Orlando, USA: WMSCI, NagibCallaos et al, vol. 3, pp. 149-153 (2018)
15. Madarász, L.: Situational Management Methodology and its Application. ELFA TUKE (2003)
16. Polishchuk, V.V., Malyar, M.M., Sharkadi, M.M.: Model informatsiynoyi tekhnolohiyi otsynyuvannya ryzyku finansuvannya proektiv. Radioelektronika, informatyka, upravlinnya 2017/2, 44-52 (2017). doi: 10.15588/1607-3274-2017-2-5
17. Malyar, M.M.: Modeli i metody bahatokryterialnoho obmezhenno-ratsionalnoho vyboru. RA SHARK (2016)
18. Zadeh, L.: Fuzzy Sets. Information and Control №8, 338-353 (1965)
19. Gaber, M. M.: Scientific Data Mining and Knowledge Discovery – Principles and Foundations. Springer, New York (2010). doi: 10.1007/978-3-642-02788-8
20. Hu, Z., Bodyanskiy, Ye.V., Kulishova, N.Ye., Tyshchenko, O. K.: A Multidimensional Extended Neo-Fuzzy Neuron for Facial Expression Recognition. International Journal of Intelligent Systems and Applications. Volume 9, No. 9, 29-36 (2017)

Simulation of ACS for magnetic susceptibility measurements in ECAD based on time domain functions

Olga Vasylenko¹[0000-0001-6535-3462], Vitalii Reva²[0000-0002-3265-1735],
Gennadii Snizhnoi³[0000-0003-1452-0544]

^{1,2,3}Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063,
Ukraine

¹traven03@yahoo.com

²vitalii.reva@gmail.com

³snov@zntu.edu.ua

Abstract. The task of ACS analyzing in ECAD in the large-scale signal mode was considered. New models based on continuous time domain functions allow us to solve the problem of transient analysis in non-stationary nonlinear ACS. This approach does not require a transition to the frequency domain by calculating the transfer functions for each block, while the model retains its nonlinearity. It was found that such modeling in ACS in ESAV is quasi-causal indeed. According to the chosen approach, dynamic nonlinear macromodels for subsystems of ACS for determine magnetic susceptibility were developed: for the actuator/stepper motor, PWM-controller, Choke with movable core, etc.

Keywords: Magnetic Susceptibility Measurement System, ACS, CAE, ECAD, quasi-causal modeling, macromodels, time-domain simulation

1 Introduction

For measure of magnetic susceptibility, which has a high informative value when conducting research on the structural state of austenitic materials, various installations are used, for example, magnetometric scales. The block diagram of the automated system for magnetic susceptibility measuring is given in [1]. This information-measuring system with microcontroller control is a multi-domain Automated Control System (ACS).

Traditionally, for the analysis of such ACS is used approach, in which for the each block it was necessary to obtain transfer functions by converting the equations of the time domain to the equations of the frequency domain [2]. Such transition is possible when the several simplifications and initial conditions are met: the system must be stabilized at a certain operating point (the transition process is neglected), the static characteristics are linearized, which is possible for the Small Signal Mode. Thus, the ranges of adequacy of such models are limited to the calculation of frequency characteristics, for example, hodographs, which are often used to analyze the stability of systems, etc.

However, it is quite often necessary to calculate the time of transients, response delays, it is also necessary to take into account the nonlinearity of characteristics to predict the behavior of systems in real time and to write and debug programs for generating control signals by microcontroller.

2 Formal problem statement

In ACS the microcontroller performs measurements, calculations and control, so, the main requirements for the designed model are data representation in the time domain and the ability to determine the nonlinearity of parameters in full scale variation of variables. This approach does not require calculation the transfer functions for each block. The model must retain its nonlinearity.

Consider an example of building a dynamic non-linear model of an ACS adapted to the capabilities of the Electronic Computer Aided Design (ECAD) mathematical processor Micro Cap v.11 (MC 11) [3]. Based on the new developed models, it is necessary to obtain the optimal variant of the automated system for measuring magnetic susceptibility.

3 Literature review

Simulation of ACS as multi-domain dynamical systems, for example, mechatronic, is one of the most powerful tools of system research [4]. Simulation is an extremely important tool for every control engineer who is doing practical control system design in industry. For arbitrarily nonlinear plants, there is often no alternative to designing controllers by means of trial and error, using computer simulation [5]. Once the dynamic behavior is understood, changes to the system can be made or control systems can be designed to make the system perform as desired.

Simulation of ACS is used in the design of microcontroller systems at the stage of Control Law Design [6,7]. The result of the simulation is the evaluation of integral criteria: accuracy, stability, performance. There are two main approaches to quality assessment. The first uses information about the time parameters of the system (the transition function). The second uses information about some frequency properties of the system: bandwidth; relative height of the resonant peak, etc. Frequency quality criteria are used when known or it is possible to determine experimentally the frequency characteristics of the ACS. The type of transitional process is not considered at the same time. Frequency criteria can be used to assess the stability and performance of control systems. Some Computer-aided Engineering (CAE) programs allow automatically determining the quality parameters of a designed ACS, using its various characteristics.

There are different approaches to the study of ACS and various levels of model abstraction, and, accordingly, different software (CAD/CAE). There are many tools from CAE and CAD market for automated mechatronic systems simulation [8,9]. Some of them are well-known and popular (MATLAB \ Simulink [10,11], Maple \ MapleSim [12], VisSim [13]), others have not been widely distributed (SimApp [14],

20-sim [15] etc). Some programs are universal and can be used to simulate any technical systems (20-sim, Dymola [17]). Others specializing in a particular subject area (Modelica [18], Powersim [19], CVODES [20]). Simulink is integrated with MATLAB and data can be easily transferred between the programs [21].

The capabilities of many packages are largely overlapping and approaches to solving the same tasks are roughly the same.

The computer simulation model is built visually through block diagrams. Usually for the study of ACS, it is necessary to obtain transfer functions for its subsystems by converting the functions of time and transition to the frequency domain. Conversion (e.g., Laplace transform) allows to obtain simple model as algebraic equations with complex coefficients. However, the area of the adequacy of such models is small, because for an adequate transition to the frequency domain, it is necessary to linearize the models and make an assumption that the processes are stationary. Transients in nonlinear systems with this approach cannot be modeled. As a rule, in CAE programs, blocks are built on the basis of transfer functions, however, SimApp [14] allows for dynamic modeling in the time and frequency domains.

Analysis of ECAD programs [8] showed that they also provide the ability to simulate ACS using the flowchart method. Despite the similarity of models at the stage of modeling in ECAD and CAE, the simulation process is fundamentally different. In CAE, a causal algorithm is used, in ECAD – an a-causal algorithm [4]. The a-causal algorithm makes it possible to simulate systems, composed of blocks with complex frequency functions, blocks with analytical functions of time (programmable, including), elements of circuit diagrams, both analog and digital. Since most of the designed automated system for measuring magnetic susceptibility is an electronic system, the choice of the ECAD for its preliminary analysis is quite logical.

MC 11 contains many features that help you grow the model sophistication to the level needed for realistic results [3]. Non-linear elements and custom inputs bring reality to the models. The library has a standard PID controller model. The graphemes used in the classical theory of automatic control (adders, integrators, and so on) are also used in the construction of the ACS model. This makes it possible to investigate the behavior of multi-domain systems, in which blocks of continuous and discrete physical nature are interconnected by summers, dividers, multipliers etc. Blocks with Transfer Functions may also be part of the model.

4 Conceptual model of information measuring system

The optimality of the designed system was checked by mathematical modeling, which was carried out in two stages: the construction of the model according to the selected criteria (modeling) and the implementation of the model experiment (simulation). The purpose of the simulation is to determine the quality of the system (stability, dynamic characteristics, sensitivity) and choose the directions of its parametric (structural) optimization.

Information-measuring system is an ACS (magnetometric scales with the registration of the zero position on the basis of determining the frequency changes of LC

generator). ACS is designed to compensate for the displacement of the rod caused by the magnet's impact on the magnetic material. The resulting scheme of ACS is shown in fig. 1.

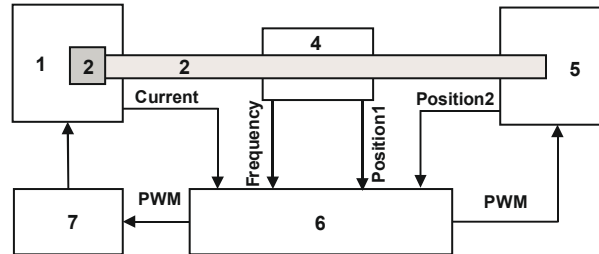


Fig. 1. Structural diagram of the ACS model

The total model of ACS is built visually through block diagrams and consists of subsystems: 1 – Experimental subsystem / Power Magnet; 2 – tested sample; 3 – the movable rod; 4 – Measurement subsystem on Colpitts variable frequency oscillator; 5 – Compensating Subsystem (implemented on an inductive choke); 6 – Microprocessor Control System; 7 – SMPS (Switch Mode Power Supply) with DAC (Digital to Analog Converter).

Microcontroller allowed increasing the accuracy and stability of automated measurement of magnetic susceptibility. ACS measuring and fixing the current of a power magnet (*CURRENT*), resulting displacement of the rod (*POSITION1* – displacement of the rod (ejection/retraction of the tested sample from the Power Magnet); *POSITION2* – result displacement (ejection/retraction from Compensating Subsystem)), frequency difference (*FREQUENCY*), duty circle of pulse width modulation signal (*PWM*-signal) for the compensating choke (*PWM*) and for SPMS, which generated signal for Power Magnet.

The rod, to which the ampule with the test materials attached, is pushed out or retracted by the magnetic field in the Experimental Subsystem, accordingly, its position changes relative to the coil in Measuring Subsystem. The subsystem for measuring the displacement of the rod is realized on the composition of the alternator generator and the electronic frequency meter, which transmits the data about frequency of oscillations changing relative to the base frequency f_0 to the microprocessor system, which, based on these data, determines the Duty Cycle (D.C.) of PWM output voltage for the Compensation System to return the rod in start position.

Operations of the measuring and compensating subsystems are based on the variation of the chokes' inductance, for which a specific choke model was developed. During the development of the system, the simulation based on a-causal block modeling principle was applied.

5 Measurement Subsystem's modeling

The biasing sensor is actually implemented on the Colpitts oscillator with an inductance, which changes its value by changing the core's depth of penetration into the

coil [23,24]. Parameters of the Colpitts oscillator are calculated for the reference frequency $f_0=300$ kHz based on the formula:

$$f_0 = \frac{1}{2\pi \sqrt{L_0 \cdot \frac{C_1 \cdot C_2}{C_1 + C_2}}} \quad (1)$$

where L_0 – initial inductance (solenoid without core); C_1, C_2 – capacitance of C_1, C_2 capacitors from circuit.

All parameters have been optimized by the Powell algorithm in MC11 [22]. The Goal Function is the maximum of amplitude of output signal.

The inductance of the generator's coil depends on the current position of the core (Δl) and the magnetic field parameters (μ_{VAR}):

$$L_{VAR} = L_0 + \mu_0 \cdot \mu_{VAR} \cdot k^2 \cdot \Delta l \cdot S, \quad (2)$$

$$k = \frac{N}{l},$$

where μ_0 – absolute magnetic permeability; μ_{VAR} – varying magnetic permeability; L_0 – initial value, calculated by (1) on 300 kHz; N – number of turns; l, S – length and area of the solenoid cross, respectively; Δl – variation of the rod's penetration depth.

To take into account the change in the magnetic permeability μ_{VAR} from the intensity of the magnetic field H [23,24], the following approximation of the Stoletov's curve is used:

$$\mu_{VAR} = \mu_0 + \mu_{max} \cdot \sqrt{H} \cdot e^{-0.01 \cdot H}, \quad (3)$$

where μ_{max} – maximum magnetic permeability on the Stoletov's curve for the selected material.

The scheme for modeling the Stoletov's curve and nonlinear inductance is shown in fig. 2. Basic parameters of approximation predefined. The field intensity changes (node with the name H) is emulated by the generator $V1$. The subsystem of L_{VAR} (determining the inductance of the measuring oscillator by the formula (2)) is consist of: generator $X1$, which defines a step change in the depth of penetration of the rod into the coil (node with the name $DELTA\_L$); resistors $R2, R3$; inductance L_{VAR} , through which the pulsating current flows from the generator $V2$ and resistor R_{mu} , which variate its resistivity by the formula (3). The results of L_{VAR} simulation in the full range of field intensity variation are shown in fig. 3.

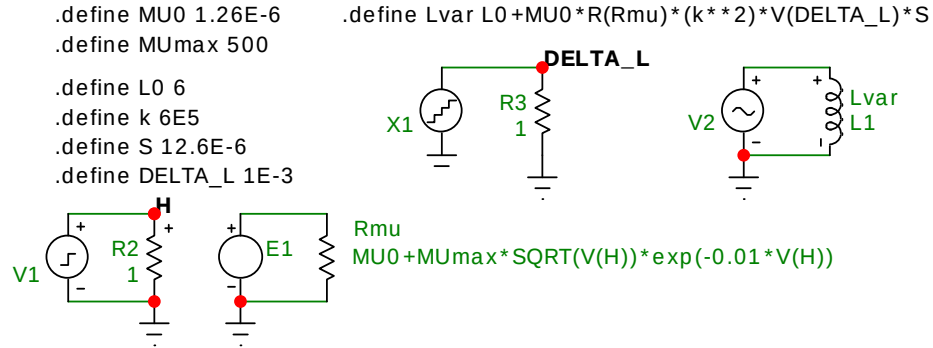


Fig. 2. Nonlinear inductance Modeling

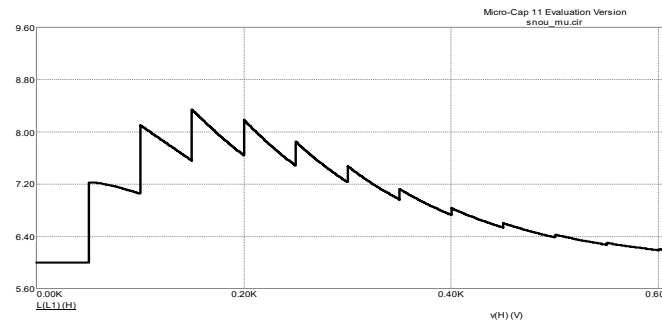


Fig. 3. Simulation of nonlinear inductance

But we must note that with increasing frequency inductance sensor sensitivity nonlinearly decreases (fig.4).

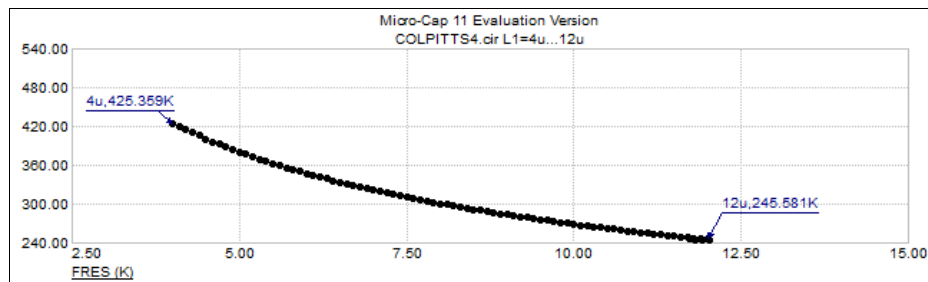


Fig. 4. Frequency vs. Inductance Curve

The dependence of the resonant frequency FRES on the value of the inductance is approximated by the results of a multivariate analysis in the following way:

$$FRES = 465 \cdot L(LCOIL)^{-0.2}, \quad (4)$$

where $L(LCOIL)$ – variable choke inductance.

The characteristics obtained by simulation are shown in fig. 5.

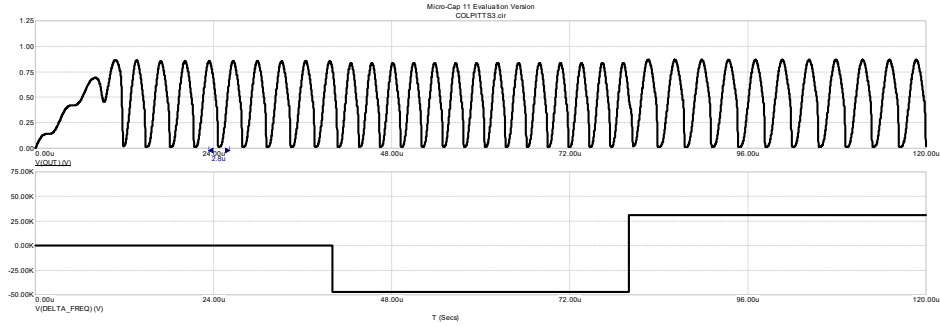


Fig. 5. Transient Analysis of Measurement subsystem

The upper graph shows the output voltage oscillogram, the frequency of which varied by the variation of the inductance: from the base frequency at 300 kHz initially increased by 48kHz, then decreased by 28kHz, as illustrated by the lower graph $DELTA\_FREQ$ vs T (Time). The initial frequency is set after the full charge of resonance circuit capacitance.

Inductance's changing leads to the frequency change in the oscillation circuit. Current value of measured frequency is compared with the reference frequency 300 kHz. The difference between the measured and the base frequency $DELTA\_FREQ$ is transmitted to the voltage source and then sent to the subsystem, which generates certain pulse duration of the PWM-signal.

According to the simulation results, the program for the microcontroller has been adjusted.

6 Compensation Subsystem modeling

Model of the Compensating Subsystem (nonlinear solenoid with moving core) is based on the relay model [3] with several modifications, includes a flux circuit and derives a magnetizing force from the flux. Choke current depends on the initial inductance and the core's position (variable $POSITION2$). Reactive component given in the functional current source by the empirical formula:

$$I_L = \frac{1}{LCOIL} \cdot (V(FLUX)) \cdot (1.1 - V(POSITION2) + (1.1 - V(POSITION2))^2) \quad (5)$$

where $LCOIL$ – initial choke inductance;

The $FLUX$ is determined by the magnetizing $FORCE$, by the formula (given in functional source E3):

$$FORCE = K_{FORCE} \cdot \left(K_{FLUX} \cdot \frac{V(FLUX)^2}{AREA} \right), \quad (6)$$

where K_{FORCE} , K_{FLUX} – coefficients are given by operators `.define`; $V(FLUX)$ – flux emulation, received as the voltage of the “FLUX” node; $AREA$ – the cross-sectional area of the rod is given by the operator `.define`.

By Faraday's law of induction, any change in flux through a circuit induces an electromotive force (EMF) or voltage in the circuit, proportional to the rate of change of flux. Any alteration to a circuit which increases the flux (total magnetic field) through the circuit produced by a given current increases the inductance. The force produced by a magnetic field [25], so for there to be a force on a piece of iron (core) then a displacement of the rod must result in an alteration to the field energy:

$$FORCE = \frac{i}{2} \cdot \left(\frac{dL}{dx} \right), \quad (7)$$

where i is the coil current and x is displacement of the core.

This result is proved in textbooks such as Hammond, and also Smith [23,24]. Unfortunately, it might be tricky to calculate how the inductance changes unless the system you have is particularly simple to analyze.

There is another way to determine this FORCE. Force affects to the rod by pushing it, or pulling it at the speed $VELOCITY$, which is determine through the second law of Newton. The current position of the rod (variable $POSITION2$) is determined by the integration of $VELOCITY$.

7 Total ACS modelling and simulation

The simplified model of ACS is shown on fig.6.

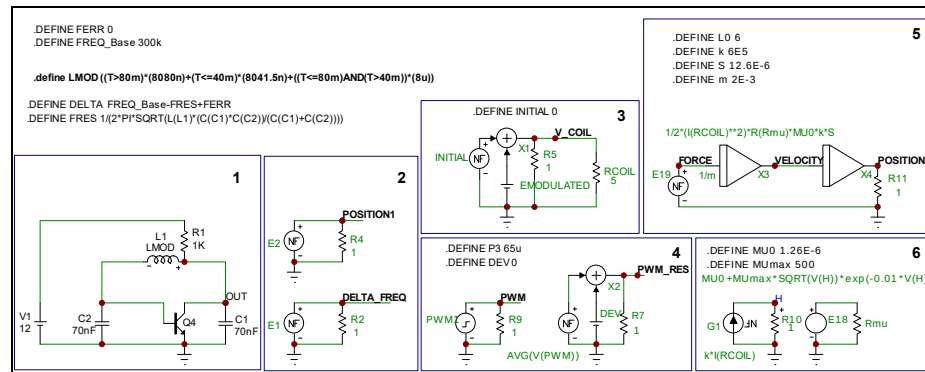


Fig. 6. Model of automated measurement system for Micro-Cap 11

Total model consist of macros: 1. Measurement subsystem on Colpitts oscillator; 2. Finding the frequency difference – *DELTA_FREQ*; displacement of the rod – *POSITION1*; 3. Calculation of Current/voltage for compensating system – *V_COIL*; 4. Emulation of PWM signal – *PWM_RES*; 5. Finding the result displacement in compensating system – *POSITION2*; 6. Stoletov's curve fitting.

PWM controller model [26] in which compensating signal produced as a constant continuous signal with variation of its amplitude (staircase signal). This assumption is acceptable provided if the output signal is smoothing effectively by filter (Choke's inductance). Since the PWM signal regulates the average value of the voltage by changing the duty cycle, it is possible to directly use the variation of the mean/average value of PWM signal. The output PWM signal is control the current of the choke with inductance 10 mH. The compensation signal stops changing when the values of *POSITION1* and *POSITION2* become equal.

Verification during the simulation was carried out by comparing the founded values of the rod movement (*POSITION1*), which should be the same in the measuring and compensating subsystems. The displacement of the rod in the measuring system (*POSITION1* – lower graph) and its response in the compensation system (*POSITION2* – upper graph) is shown on fig.7.

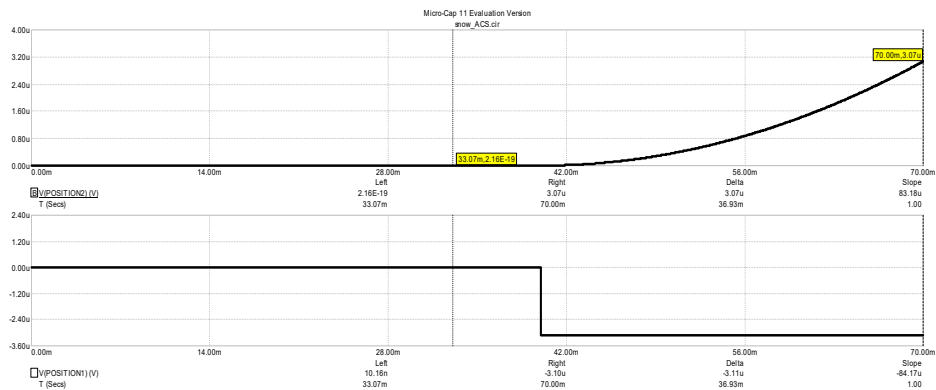


Fig. 7. Displacement of the rod in the measuring system (*POSITION1*) and response in the compensation system (*POSITION2*)

The variation of the inductance in the range 8040 ± 40 nH has changes the frequency (± 800 Hz), the displacement of the rod (± 3.2 μ m), and the voltage on the compensation circuit (± 3.2 mV).

A-causal models work together with causal models very well due to the development of the simulation SPICE algorithm [3,27]. This makes it possible to analyze the system composed of models of different levels of abstraction. For example, the a-causal model of the Colpitts generator and the causal macromodel of a PWM-controller.

8 ACS implementation

This system is implemented on the ATmega8 microcontroller family. It has 23 inputs, six of which can receive an analog signal through the built-in ten-bit analog-to-digital converter (ADC). This controller is capable of generating a PWM signal by hardware, as it has three built-in PWM controllers. One PWM signal is used for stepped pumping of a Power Magnet, and the other is for regulating the current in a Compensating Subsystem.

The PWM controller is tuned by setting the control bits in the corresponding registers. The PWM controller operates in Fast PWM mode, in this mode the counter counts from zero to 255, after an overflow is reached, it is reset to zero and the counting starts again. The timer has a special OCR comparison register, the reset also occurs when the value in the counting register is compared with the value written to the comparison register. The D.C. factor of the signal is determined with an accuracy of 2-8 by writing a number from 0 to 255 in the OCR register. With a clock frequency of 8 MHz, PWM frequency will be equal to 31.25kHz. If frequency dividers from 8 to 1024 is used, the frequency will decrease proportionally, but the accuracy of the interval setting will not improve.

The current measurement of the Power Magnet is made using a Hall sensor. At the output, it produces a voltage proportional to the current flowing. The signal enters the controller via an ADC. In the Atmega8 controller, the ADC has the following characteristics: resolution of 10 bits; the time of conversion of one indication from 13 to 250 μ s depending on the measurement bit depth, as well as on the clock frequency of the controller generator; start support by interrupts, which allows you to automatically process data and save them to the correct register, for processing them in the main cycle.

The change in the state of the system and the processing of information occur after system initialization and after calculating the frequency of the Colpitts generator. The signal from the generator after converting it to a rectangular view is fed to the input of the controller. To determine the frequency, two interrupt vectors are used: the timer overflow interrupt and the external pulse interrupt. An interrupt from an external pulse increases the register value by one. The higher the timers interrupt frequency, the faster the controller response, but the frequency measurement accuracy is lower. However, there is a limit value of the frequency at which the controller does not have time to make calculations. The minimum frequency of the timer 244.14 Hz, which is quite suitable for this ACS, has been determined.

The speed of the measurements is determined mainly by the transition process in the compensating coil, which is about ten times longer than the period of the microcontroller timer (46 ms vs 4.1 ms). Measurement delay is also associated with transitions to a stationary state in a power coil. The system is stable, oscillations of transient processes are not revealed either during modeling, or in «field» experiments.

9 Conclusion

A new way of modeling ACS is based on dynamic models for its subcircuits. This will allow you to remain inside the time domain for a full study of the dynamics even for non-stationary systems. Different levels of abstraction Macromodels for all subsystems of the Automated Magnetic Susceptibility Measurement System have been developed. Analysis of static and dynamic characteristics in the environment of the ECAD program Micro-Cap 11 has been conducted.

The resulting model is dynamic, nonlinear, quasi-causal [4,8]. Causality eliminates unwanted feedback and sets a certain sequence of calculations. In fact, this model is composition of proportional-integration blocks and analog schemes. Its time-behave and nonlinearities have been evaluated during the Transient Analysis. The delay of operation of the compensation system is found ($4\text{ }\mu\text{s}\pm 10\%$). Directions of its parametric optimization and improvement on criteria of accuracy and versatility are found.

The relative error of the model is not more than 20%, so it can be recommended for the study of such systems.

After the model experiment, in which the parameters of the system has been optimized and refined on the basis of physical features of the real hardware, a prototype of compensation system was designed [1]. So, the adequacy of the models has been confirmed not only by comparison of the simulation results (*POSITIONS1* vs *POSITIONS2*), but with the results of experiments on real ACS [11].

New approach, chosen for measurement system modeling lies within the multidomain modeling paradigm, which allows us to investigate a wider spectrum of characteristics of nonlinear dynamic systems, in contrast to the standard approach adopted in the theory of automatic regulation, which focuses on the analysis of stationary processes in linearized systems. The developed macromodels are distinguished by minimalism and simplicity and therefore, may be recommended for the study of ACS on the different levels of abstraction in ECAD and in other similar mathematical processors.

References.

1. Snizhnoi, G. V., Zhavzharov, Ye. L.: Automated equipment for determining the magnetic susceptibility of steels and alloys. Visnyk NTUU "KPI". Ser. Radiotekhnika. Radioaparatorobuduvannya. 49, 136-141 (2012).
2. Otter, M., Cellier, F.E.: Software for Modeling and Simulating Control Systems (1996) http://www.engr.arizona.edu/~cellier/publications_oo.html
3. Micro-Cap 11, Electronic Circuit Analysis Program Reference Manual, Spectrum Software <http://www.spectrum-soft.com/manual.shtm>
4. Vasylenko, O.V., Petrenko, Ya.Y.: Povyshenye kachestva modelyrovaniya dynamicheskikh system vyborom opy-malnykh alhorytmov symuliatsyy. Radio Electronics, Computer Science, Control. 4., 11-18 (2016).
5. Safety Analyzes of Mechatronics Systems: a Case Study Benjamin. <https://hal.archives-ouvertes.fr/hal-01651282/document>

6. Stanford University Lecture.
https://web.stanford.edu/class/archive/ee/ee392m/ee392m.1056/Lecture9_ModelSim.pdf
7. Luiz, S., Adrielle, C., Ricardo, O., Gilberto, A.: Processor-in-the-Loop Simulations Applied to the Design and Evaluation of a Satellite Attitude Control. Computational and Numerical Simulations.(2014).
doi: 10.5772/57219
8. Vasylenko, O.V.: Analiz prohram dlia modeliuvannia mekhatronnykh system. Radio Electronics, Computer Science, Control . 3, 80-88 (2015).
9. Mrozek, Z.: Computer aided design of mechatronic systems. Applied Mathematics and Computer Science. 13, 101-113(2003).
10. Simulink Basics Tutorial
http://ctms.engin.umich.edu/CTMS/index.php?aux=Basics_Simulink
11. Haugen, F.: Functions for Control, System Identification and Simulation in LabVIEW. Drammen, Norway (2004).
<http://www.techteach.no/presentations/nidays04/>
12. MapleSim: Modeling & Simulation Software
https://www.maplesoft.com/contact/webforms/google/maplesim/simulation_software.asp
13. VisSim is Now Altair Embed
<http://www.vissim.com/products/vissim.html>
14. SimApp – Product Description.
<https://www.simapp.com/simulation-software-description.php>
15. Kleijn, C., Groothuis, M. A., Differ, H.G.: Reference manual 20-sim 4.6. Enschede : Controllab Products B.V., 1099 (2015). ISBN 978-90-79499-19-9
16. Simscape Overview
<http://www.mathworks.com/products/simscape/index.html>
17. Dynamic Modeling Laboratory. Dymola User's Manual, Version 5.3a
<https://www.inf.ethz.ch/personal/cellier/Lect/MMPS/Refs/Dymola5Manual>
18. PSIM. Design a controller in minutes, not hours
<https://powersimtech.com/products/smartctrl/>
19. Fritzson, P.: Principles of object-oriented modeling and simulation with Modelica 2.1. Piscataway, NJ : IEEE Press (2006). ISBN 978-0471471639
20. Description of CVODES
https://computation.llnl.gov/casc/sundials/description/description.html#descr_cvodes
21. Control Tutorials for MATLAB and Simulink (CTMS).
http://ctms.engin.umich.edu/CTMS/index.php?aux=Basics_Simulink#65
22. Vasylenko, O.V., Petrenko, Ya.I.: Vybir metodu optymizatsii system avtomatychnoho keruvannia v systemakh avtomatyzovanoho inzhynirynhu. Enerhetyka i avtomatyka. 1, 75-89 (2017).
23. Haus, H. A., James, R. M.: Electromagnetic fields and energy. Englewood Cliffs, N.J.: Prentice Hall, (1989). ISBN: 9780132490207.
24. Knuteson, B. and Hudson, E. and Stephans, G. and Belcher, J. and Joannopoulos, J. and Feld, M. and Dourmashkin, P.: 8.02T Electricity and Magnetism. Lecture notes. Massachusetts Institute of Technology
<https://ocw.mit.edu/courses/physics/8-02t-electricity-and-magnetism-spring-2005/lecture-notes/>
25. The force produced by a magnetic field

<http://info.ee.surrey.ac.uk/Workshop/advice/coils/force.html>

26. Vasylenko, O.V., Snizhnoi, G.V.: PWM controller's models for investigation ACS in SPICE-family ECAD programs. Radio Electronics, Computer Science, Control. 1, 64-71 (2018).
27. Kendall, R.: Modular macromodeling techniques for Spice simulators (2002).
<https://www.edn.com/design/analog/4348119/Modular-macromodeling-techniques-for-Spice-simulators>

Using Computer Modeling and Artificial Neural Networks for Ensuring the Vibration Reliability of Rotors

Ivan Pavlenko¹[0000-0002-6136-1040], Calin Neamtu²[0000-0003-0899-0451], Anton Verbovyi¹,
Jan Pitel³[0000-0003-1942-0438], Vitalii Ivanov¹[0000-0003-0595-2660] and Grigore Pop²

¹ Sumy State University, 2 Rymskogo-Korsakova St., 40007 Sumy, Ukraine

² Technical University of Cluj-Napoca, 28 Memorandum St., 400114 Cluj-Napoca, Romania

³ Technical University of Kosice, 1 Bayerova St., 080 01 Presov, Slovakia
i.pavlenko@omdm.sumdu.edu.ua

Abstract. The paper is devoted to the development of the refined numerical and mathematical models of rotor dynamics for high-performance turbomachines. These models consider the nonlinear regression dependencies between the radial stiffness of bearing supports and rotor speed of the shaft. As a result of the comprehensive application of the proposed approach, the algorithms of identification of the stiffness parameters of the bearing supports were proposed based on the critical frequencies of the rotor system. Additionally, the methodology for estimating the system of initial imbalances by the related displacements of the rotor axis in the correction and calculation cross-sections was improved. The proposed approaches based on the comprehensive application of the CAE software and computational intelligence systems allow conducting the modal and harmonic analysis and realizing the virtual balancing with the significant decrease of the preparatory and machine times without the loss of relative accuracy. Additionally, the developed mathematical models of free and forced oscillations of rotor systems were implemented as a program code of the operating files “Critical frequencies of the rotor” and “Forces oscillations of the rotors” within the computer algebra system that allows improving the dynamic balancing procedure in order to estimate the system of initial imbalances. The high accuracy of the proposed approach is confirmed by the verification of dynamic deflections of the rotor axis by the system of residual imbalances in accordance with the norms of vibration reliability.

Keywords: Critical Frequencies, Forced Oscillations, Virtual Balancing, CAE Software, Computational Intelligence Systems.

1 Introduction

The permanent acceleration of the development in the field of energy technologies is due to the use of up-to-date highly-intensive equipment, particularly multistage centrifugal machines (i. e. pumps and turbocompressors). The consequent increase in of their technical characteristics leads to increasing complexities in terms of vibration

reliability. In this case, features in designing of the power equipment require a more detailed analysis of the dynamic state for pump and compressor units as the most energy-intensive ones. Due to this trend, the use of existing mathematical and numerical models is practically inappropriate, since the relatively high rotational speeds that caused significant nonlinear effects. Due to the abovementioned, there is a necessity to create the refined and reliable models of rotor dynamics, which will be allow carrying out the modal and harmonic analysis, as well as evaluating the impact of the operating parameters on the stiffness characteristics of the bearing supports in order to qualitatively realizing the rotor balancing procedure.

The abovementioned problem is solved by the determination of critical frequencies and related mode shapes of rotors, as well as by conducting the previous virtual balancing procedure. Thus, the paper is aimed at developing computational and analytical approaches both for ensuring vibrational reliability of rotary systems both for estimating the parameters of the dynamic state of turbomachines and for realizing the virtual balancing procedure.

To achieve the stated aim, the ANSYS software, as well as working files “Critical frequencies of the rotor” and “Forced oscillations of the rotor” are used. These two files were developed at the Department of General Mechanics and Machine Dynamics of Sumy State University and implemented into the MathCAD software.

2 Literature Review

The abovementioned statement of the problem can be successfully solved after the precise analysis of the recent researches in the fields of computer modeling of rotor dynamics. Particularly, the research paper [1] deals with the estimation of the reliability of the rotors’ automatic axial-balancing devices in multistage centrifugal pumps. As a result, the methodology for numerical calculations was proposed. The research work [2] presents the methods of static and dynamic analysis of the closing rotor balancing device of the multistage centrifugal pumps. However, the proposed approaches are based on the criterion of the permissible amplitudes of axial oscillations only.

The phenomenon of rotor’s instability for the case of tilting pad bearings is described in the paper [3]. As a result of numerical simulation of rotor dynamics considering stiffness of bearing supports, spectral analysis was carried out, and related phase paths of rotor’s motion were presented, as well as features of rotor dynamics in a range of unstable rotational speeds were analyzed. Consequently, ways for designing the radial sliding bearing supports with the hydrostatically suspended pads for ensuring the dynamic stability of rotors are proposed in the paper [4], as well as the technology of improving the quality of contact surfaces for optimizing the operating characteristics of bearing supports is presented in the research work [5]. However, the abovementioned approaches do not allow estimating parameters of bearing stiffness.

A comprehensive methodology for identification of load categories in rotor systems based on vibration analysis is proposed in the paper [6]. It can be successfully

applied for the fault diagnosis of pumps and compressors. Additionally, ways of the faults detection in gas turbines' rotors using vibration analysis under varying conditions are completely described in the paper [7]. Moreover, enhancement of performance for steam turbines using artificial intelligent systems and electric circuit design is realized within the research [8]. As a result, a good agreement between the data of numerical simulation and a physical experiment was achieved that allowed predicting the operating parameters with a relatively small error.

The reliability analysis on resonance for low-pressure compressor rotor blade based on least squares support vector machine with leave-one-out cross-validation is presented in the paper [9]. Additionally, the approach for detection and modeling of vibration behavior of a gas turbine based on artificial neural networks is used in the research works [10, 11], as well as the methodology for the detection of propagating cracks in rotors using neural networks is used in the paper [12]. Moreover, ways of using artificial neural networks and genetic algorithms for bearing fault detection are presented in the paper [13]. However, the abovementioned approaches do not allow estimating parameters of bearing supports by means of the implemented computational intelligence systems.

Due to the abovementioned, a comprehensive approach of combined using the numerical simulation and computational intelligence systems should be developed both for identification of the related mathematical models and ensuring the vibration reliability of rotor systems.

3 Research Methodology

3.1 Eigenfrequencies and Critical Frequencies of the Rotor

The objects of the study are free and forced oscillations of the flexible rotor of the centrifugal compressor 225GC2-135/12-50M1245 (Fig. 1) at the 1643 MW Navoiy Thermal Power Plant (Yangi-Ariq, Uzbekistan) with the following technical parameters: flow rate $25 \text{ m}^3/\text{s}$; inlet pressure 1.20 MPa; outlet pressure 4.95 MPa; pressure ratio 4.13; polytropic efficiency 70 %; operating frequency 934 rad/s; nominal power 7.33 MWt; inlet temperature 301 K; temperature difference 156 °C; compression ratio 0.98.



Fig. 1. A rotor of the multistage centrifugal compressor 225GC2-135/12-50M1245.

The calculation of eigenfrequencies provides the determination of potentially dangerous rotor speeds under its rotation at the operating frequency, which is equal to 934 rad/s. Therefore, the bearing stiffness for the common calculations was adopted constant and equal to $c_w = 2.94 \cdot 10^8$ N/m at the operation rotor speed.

In fact, the calculation of critical frequencies needs to consider that bearing supports have different stiffness at different rotor speed. The working file “Critical frequencies of the rotor” allows directly determining the nonlinear dependencies of the stiffness characteristics of the rotor’s bearing supports on the rotational speed as the following approximation [14]:

$$c(\omega) = c_0 + \alpha\omega + \beta\omega^2 \quad (1)$$

with the unknown coefficients c_0 , α , and β evaluated on the basis of the linear regression procedure by minimizing the following error function:

$$R(c_0, \alpha, \beta) = \sum_{k=1}^n (c_0 + \alpha\omega_k + \beta\omega_k^2 - c_k)^2 \rightarrow \min, \quad (2)$$

where c_k , ω_k – k -th experimental values of bearing stiffness and frequency, respectively; n – total number of experimental dataset.

In this case, the system of equations for determining the unknown parameters takes the form:

$$\begin{cases} \frac{\partial R}{\partial c_0} = 2 \sum_{k=1}^n (c_0 + \alpha\omega_k + \beta\omega_k^2 - c_k) = 0; \\ \frac{\partial R}{\partial \alpha} = 2 \sum_{k=1}^n (c_0 + \alpha\omega_k + \beta\omega_k^2 - c_k) \omega_k = 0; \\ \frac{\partial R}{\partial \beta} = 2 \sum_{k=1}^n (c_0 + \alpha\omega_k + \beta\omega_k^2 - c_k) \omega_k^2 = 0, \end{cases} \quad (3)$$

which can be rewritten as a matrix equation:

$$[\Omega]\{X\} = \{C\}, \quad (4)$$

where $[\Omega]$, $\{C\}$, and $\{X\}$ are the following matrix 3×3 and the column vectors 3×1 :

$$\Omega_{i,j} = \sum_{k=1}^n \omega_k^{i+j-2}; \quad C_i = \sum_{k=1}^n c_k \omega_k^{i-1}; \quad \{X\} = \{c_0, \alpha, \beta\}^T. \quad (5)$$

The solution of the matrix equation (4) is

$$\{X\} = [\Omega]^{-1} \{C\}. \quad (6)$$

In the case of the compressor 225GC2-135/12-50M1245, the evaluated values of the coefficients are $c_0 = 1.77 \cdot 10^8$ N/m, $\alpha = 1.30 \cdot 10^5$ N·s/m, and $\beta = -4.97$ N·s²/m.

The operating file “Critical frequencies of the rotor” is based on the calculation of the 2D-beam model of rotor dynamics (Fig. 1 a). Consequently, the input data is a description of each i -th finite element for the following characteristics: density ρ_i , kg/m; Young’s modulus E , Pa; length L_i , m; external and internal diameters D_i and d_i , correspondently, m; single mass m_i , kg; polar moment of inertia I_i , kg·m²; radial stiffness of the bearing support $c_i(\omega)$, N/m according to the dependence (1).

The ANSYS software allows simulating the rotor dynamics based on the finite element method both for 2D-beam and 3D finite elements (Fig. 2 b). The first one is not of interest because the obtained results will not be fundamentally different from the previously calculated using the working file “Critical frequencies of the rotor”.

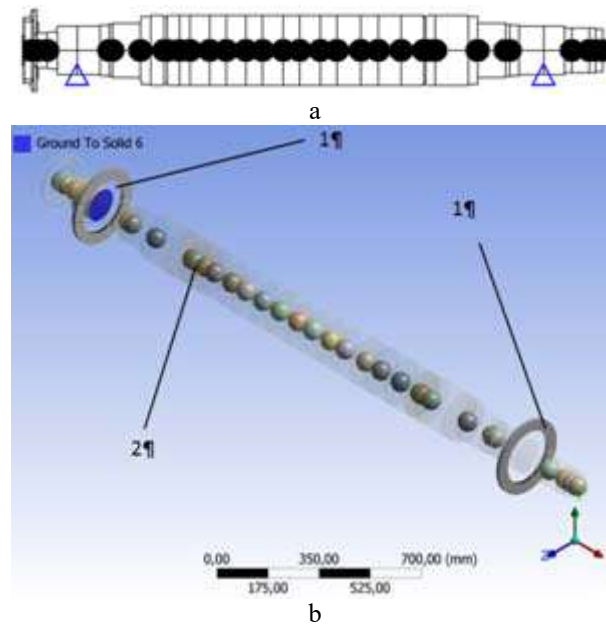


Fig. 2. The finite element models for simulating free oscillations using the working file “Critical frequencies of the rotor” (a) and ANSYS software (b): 1 – bearings; 2 – single masses.

It should be noted that capabilities of the ANSYS software do not allow calculating the critical frequencies directly, as it possible using the working file “Critical frequencies of the rotor”. Therefore, the research in ANSYS is based only on building and further analyzing the Campbell diagram.

3.2 Forced Oscillations of the Rotor

The working file “Forced oscillations of the rotor” allows calculating the forced oscillations of a rotor under the given rotation speed and the influence of a certain system of imbalances. As a result, amplitudes of the rotor’s deflections and mode shapes are determined an operating frequency.

The following data is required to solve the problem of forced oscillations: density ρ_i , kg/m; Young's modulus E , Pa; length L_i , m; external and internal diameters D_i and d_i , correspondently, m; single mass m_i , kg; polar moment of inertia I_i , kg·m²; operating frequency $\omega_w = 934$ rad/s; radial stiffness of the bearing support at the operating frequency $c_i(\omega_w)$, N/m, determined by the equation (1); the system of imbalances D_i , kg·m.

Theoretically, imbalances of a flexible rotor are characterized by a certain spatial curve as a hodograph of the continuous system of imbalance vectors perpendicular to the rotor axis. Since this approach is rather time-consuming, subsequent calculations can be limited by a finite number of local imbalances located in the cross-sections of impellers and other massive parts.

The harmonic analysis is based on the 3D finite element model and realized using the ANSYS software. In this case, there is a need to set parameters for imbalances distribution. Fig. 3 presents the calculation models based on the 2D-beam and 3D finite element models.

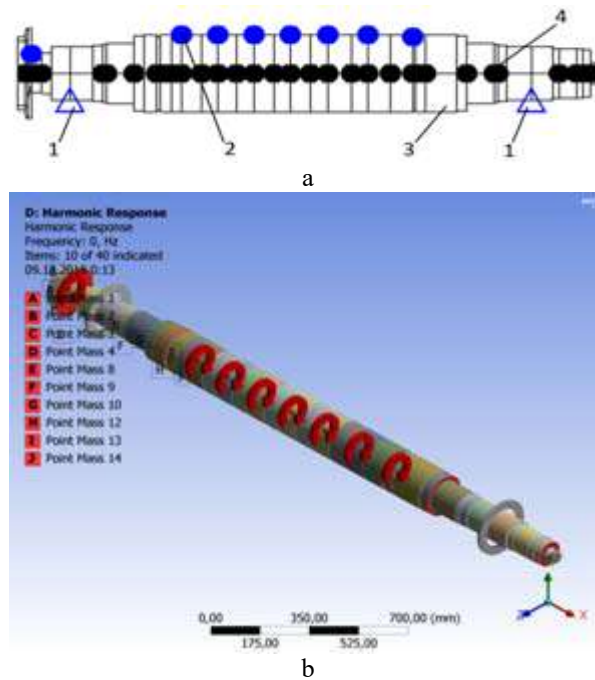


Fig. 3. The finite element models for simulating forced oscillations using the working file “Forced oscillations of the rotor” (a) and ANSYS software (b): 1 – bearings; 2 – single masses; 3 – beam finite elements; 4 – local imbalances.

3.3 Virtual Balancing of the Rotor

For realizing the virtual balancing procedure, initially, the amplitudes of the rotor's oscillations are measured based on the numerical simulation in 8 correction planes (7 impellers and 1 drive unit), in which balancing masses will be installed.

Traditionally, the balancing procedure based on the calculation model assumes a linear relationship between the imbalances and related movements. In this case, the experimental results allow calculating the coefficients of influence, as well as estimating the imbalances, which should be located in the related cross-sections and cause complex displacements $W_{a,i}$. For this purpose, it is necessary to conduct 9 measurements of the amplitude in the correction planes. The first measurement is in its original state, and the other one – with installing a trial imbalance $\Delta D = 0.01 \text{ kg}\cdot\text{m}$ in each of the 8 correction planes. The dependence of the amplitude for the rotor displacements on the system of imbalances is as follows [15]:

$$\{Y\}_0 = [W]\{D\}, \quad (7)$$

where $\{Y_0\}$ – column vector of amplitudes, m; $[W]$ – a matrix of the influence coefficients, kg; $\{D\}$ – column vector of the estimated imbalances, $\text{kg}\cdot\text{m}$.

After introducing the i -th vector of trial imbalances as a set of imbalances located in the i -th correction plane, the following equation can be written for the i -th start ($i = 1, 2, \dots, 8$) of the rotor system:

$$\{Y\}_i = [W](\{D\} + \{\Delta D\}), \quad (8)$$

where $\{\Delta D\}$ – column vector of trial imbalances.

Subtracting (7) from (8) allows determining the matrix of influence coefficients, elements of which are equal to

$$W_{a,i} = \frac{Y_{a,i} - Y_{0a}}{\Delta D}, \quad (9)$$

where $a = \{1, 2, \dots, k\}$ – row number of the matrix $[W]$; $k = 8$ – total number of the correction planes.

The determined matrix of the influence coefficients $[W]$ and the column vector of amplitudes $\{Y\}_0$ allow directly estimating the imbalances in the correction planes. Thus, from the formula (7), it can be obtained:

$$\{D\} = [W]^{-1} \{Y\}_0. \quad (10)$$

In the general case of rectangular matrix $[W]$, this equation should be rewritten in the following form:

$$\{D\} = ([W]^T [W])^{-1} [W]^T \{Y\}_0. \quad (11)$$

4 Results and Discussion

4.1 Results of the Modal and Harmonic Analysis

For the bearing stiffness calculated by the formula (1) in a range of rotation speeds from 0 to 3000 rad/s with a step 250 rad/s including the operating rotor speed 934 rad/s, the eigenfrequencies were determined. As a result, the Campbell diagram was built (Fig. 4) with the X-axis as an operating frequency, and Y-axis as an eigenfrequency.

Fig. 5 shows the following calculation results of the rotor: critical frequencies and their mode shapes. The corresponding numerical values are summarized in Table 1. The results of calculating forced oscillations as rotor displacements are shown in Fig. 6.

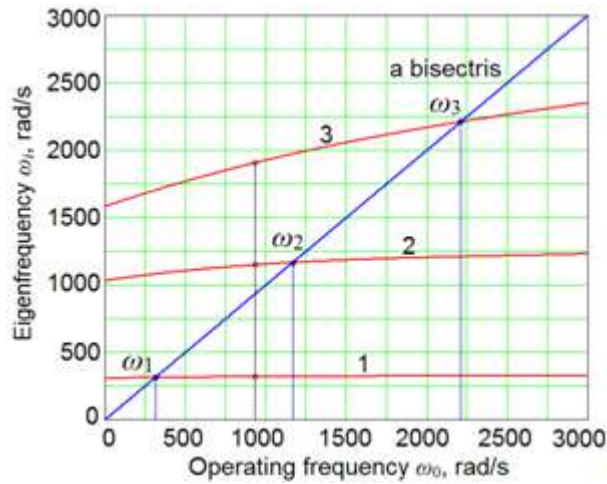


Fig. 4. The Campbell diagram.

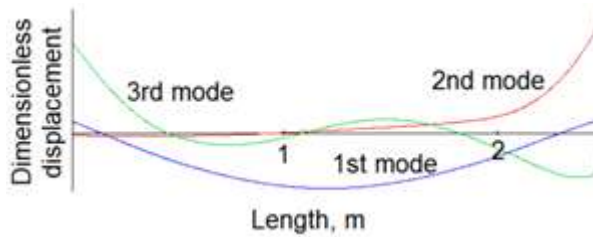


Fig. 5. The mode shapes of the rotor.

Table 1. Eigenfrequencies and critical frequencies, rad/s.

Mode shape number	Eigenfrequency			Critical frequency		
	MathCAD	ANSYS	Relative error, %	MathCAD	ANSYS	Relative error, %
1	318	319	0.3	312	313	0.3
2	1148	1125	2.0	1164	1140	2.1
3	1906	1867	2.0	2210	2150	2.7

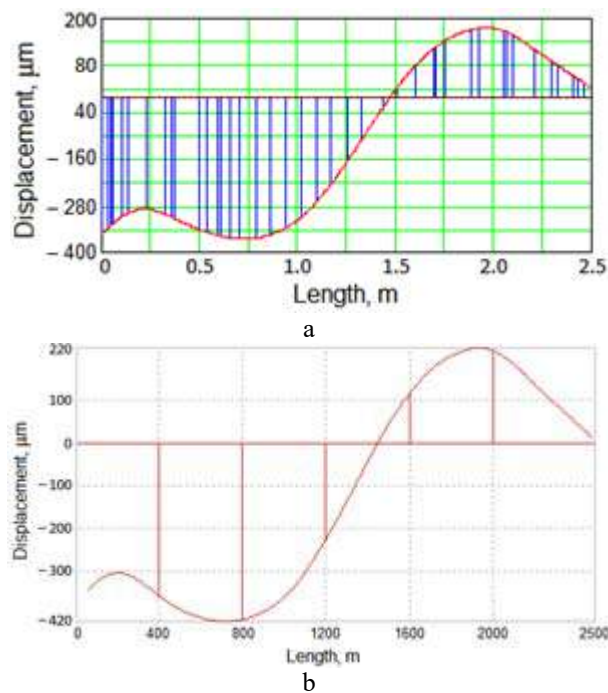


Fig. 6. The mode shapes of the unbalanced rotor at the operating frequency using the working file “Forced oscillations of the rotor” (a) and ANSYS software (b).

4.2 Evaluation of the System of Imbalances

The results of the virtual balancing procedure are given in Table. 2. The related shapes of the rotor under the residual imbalances after the balancing procedure is presented in Fig. 7.

Thus, as a result of the virtual balancing procedure, the deflections of the rotor’s axis under the system of residual imbalanced meet the established standards in vibration reliability [16].

Table 2. The results of the rotor balancing procedure.

Correction plane number i	Initial displacement Y_{i1} , m	Balancing imbalance D_i , kg·m
1	$-3.36 \cdot 10^{-4}$	0.038
2	$-3.63 \cdot 10^{-4}$	0.010
3	$-3.63 \cdot 10^{-4}$	0.038
4	$-3.37 \cdot 10^{-4}$	0.007
5	$-2.69 \cdot 10^{-4}$	0.069
6	$-1.64 \cdot 10^{-4}$	0.018
7	$-3.15 \cdot 10^{-5}$	0.009
8	$7.90 \cdot 10^{-5}$	0.080

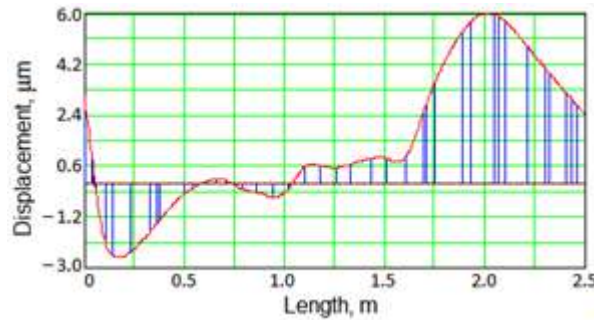


Fig. 7. Displacements of the rotor's axis after the balancing procedure.

4.3 General Approach for the Application of Artificial Neural Networks

The linearity of the traditional mathematical models is the main disadvantage of the abovementioned procedures for determining the stiffness of bearing supports, critical frequencies of the rotor, and dynamic deflections of the axis. Additionally, the general approach does not regulate the general rule for the numerical value of a trial imbalance. Moreover, conducting a series of calculations in an amount not less than $(k + 1)$ is obligatorily to determine the system of imbalances. However, considering the nonlinear dependencies between the stiffness parameters of the bearing supports and critical frequencies of the rotor, as well as between the axis deflections and a system of residual imbalances is a highly-complicated problem, solving of which will allow offering a refined virtual balancing procedure.

In the case of significant nonlinearities of mathematical models, using the computational intelligence systems were realized by the row of researchers. Particularly, in the paper [17], a method of wavelet packet combining neural network fusion is implemented In order to improve the accuracy of fault diagnosis. A neural network monitoring system is used in the research [18] for the frequency vibration prediction in a gas turbine. Additionally, the paper [19] deals with the early detection of damages in rolling bearings of an asynchronous machine. As a result, vibration

simulations were proved that neural networks can be effective agents in the diagnosis of motor bearing fault through vibration signature. Finally, in the paper [20], the motor bearing fault diagnosis was realized using deep convolution neural networks with the related analysis of the vibration signal.

Therefore, a new approach is proposed based on the use of artificial neural networks for the direct identification of the system of imbalances, avoiding the procedure of determining the matrix of the influence coefficients. The scientific and methodological approaches for the identification of mathematical models of mechanical systems, in general, are presented in the research [21].

The fundamental approach of using an artificial neural network for identification of bearing stiffness characteristics in rotor dynamics analysis was previously used for the case of solving the direct and inverse problems of rotor dynamics [22].

The design scheme of the comprehensive application of computer modeling systems and artificial neural networks for the dynamic analysis of rotor systems is presented in Fig. 8. Non-variable input parameters are colored by black, and variable input parameters – by red, as well as output parameters – by blue.

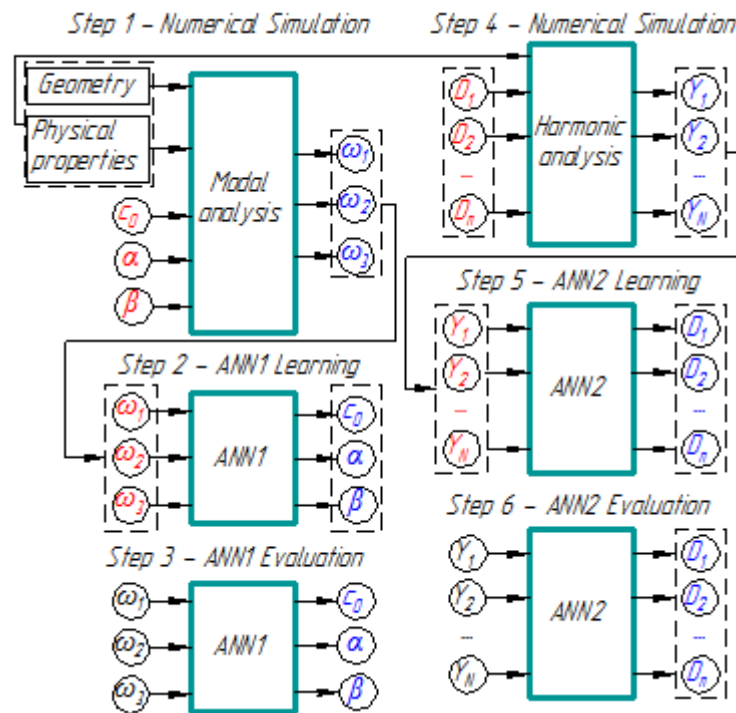


Fig. 8. The design scheme of the comprehensive implementation of numerical simulation and artificial neural networks for the dynamic analysis of the rotor.

At the first stage, modal analysis is carried out using either ANSYS software or working file “Critical frequencies of the rotor”. Variation of the stiffness parameters

c_0 , a , and β for the bearing supports in a given range with the corresponding step allows obtaining an array of 3 critical frequencies ω_1 , ω_2 , and ω_3 of the rotor. This data is used at the second stage as the input data for training the artificial neural network ANN1 created, particularly, using the Visual Gene Developer software. The process of estimating the stiffness coefficients is realized at the third stage and based on the evaluation of the output parameters of ANN1 by the experimental values of critical frequencies measured on the balancing machine “Schenck” at PJSC “Sumy Machine-Building Science-and-Production Association” (Sumy, Ukraine).

At the fourth stage, harmonic analysis is conducted using either ANSYS software or working file “Forced oscillations of the rotor”. Variation of imbalances $D_1, D_2, \dots, D_n$ (n – total number of the correction planes) in a given range with the corresponding step allows obtaining an array of dynamic deflections $Y_1, Y_2, \dots, Y_N$ (N – total number of the planes of measurement). This data is used at the fifth stage as the source data for training the artificial neural network ANN2. The process of estimating unknown imbalances is realized at the sixth stage. It is based on the calculation of the output parameters of ANN2 by the experimental values of dynamic deflections in the planes of measurement.

5 Conclusions

In the research paper, the refined mathematical model of rotor dynamics for the multistage centrifugal compressor is developed using the 2D-beam finite element models. This model considers the dependence of the radial stiffness of bearing supports on the rotational speed. It was justified on the example of the compressor 225GC2-135/12-50M1245. The reliability of the mathematical model is confirmed by the permissible difference between the obtained results in a range of 0.3–2.7 %, as well as by observing the theorem on the mutual arrangement in the spectrum of eigenfrequencies and critical frequencies.

The obtained results indicate that the use of the working file “Critical frequencies of the rotor” implemented into the system of the computer algebra is more expedient to solve the applied problems in the field of rotor dynamics since it allows arbitrary creating the dependencies of the stiffness characteristics of bearing supports on the rotation speed. Moreover, unlike the modal analysis by the ANSYS software, the calculation using the proposed approach avoids building the Campbell diagram, as well as leads to significant increase the preparatory and machine times without loss of the calculation accuracy.

A model of forced oscillations of the rotor under the influence of the system of local imbalances is also developed. The corresponding algorithm for calculating the vibration state of the rotor is implemented using the working file “Forced oscillations of the rotor” with further improvement through the virtual balancing procedure. The accuracy of this procedure is confirmed by the verification of dynamic deflections of the rotor’s axis under the system of residual imbalances in terms of vibration reliability.

Additionally, the paper proposes the use of a general approach, which allows comprehensively applying the computational means and artificial intelligence systems. Particularly, on the example of combining the ANSYS and Visual Gene Developer software, the algorithm of identification of the stiffness parameters for bearing supports by three critical frequencies is proposed. This approach also allows evaluating the system of imbalances by the corresponding displacements of the rotor.

The obtained results will be useful for researchers in the field of computational means for numerical modeling of rotor dynamics, as well as for specialists in the implementation of the calculation intelligence systems for solving applied highly-complicated nonlinear problems in the field of rotor dynamics.

6 Acknowledgments

The research was realized between the closed collaboration between Sumy State University (Ukraine), Technical University of Cluj-Napoca (Romania), and Technical University of Kosice (Slovakia). Particularly, numerical simulations using the working files “Critical frequencies of the rotor” and “Forced oscillations of the rotor” were realized at Sumy State University. The CAE software was provided within the internship program “Ensuring Vibration Reliability and Numerical Simulation Methods for Dynamic Analysis of Centrifugal Machines’ Rotors” at the Technical University of Cluj-Napoca (Romania). Computational intelligence systems were implemented at the Faculty of Manufacturing Technologies with a seat in Presov at Technical University of Kosice (Slovakia) within the research project “Identification of Parameters for Technological Equipment using Artificial Neural Networks” supported by the National Scholarship Programme of the Slovak Republic.

References

1. Pavlenko, I., Trojanowska, J., Gusak, O., Ivanov, V., Pitel, J., Pavlenko, V.: Estimation of the reliability of automatic axial-balancing devices for multistage centrifugal pumps. *Periodica Polytechnica Mechanical Engineering* 63(1), 52–56 (2019), doi: 10.3311/PPme.12801.
2. Pavlenko, I.: Static and dynamic analysis of the closing rotor balancing device of the multistage centrifugal pump. *Applied Mechanics and Materials* 630, 248–254 (2014), doi: 10.4028/www.scientific.net/AMM.630.248.
3. Gadyaka, V., Leikykh, D., Simonovskiy, V.: Phenomena of stability loss of rotor rotation at tilting pad bearings. *Procedia Engineering* 39, 244–253 (2012), doi: 10.1016/j.proeng.2012.07.031.
4. Martsinkovsky, V., Yurko, V., Tarelnik, V., Filonenko, Y.: Designing radial sliding bearing equipped with hydrostatically suspended pads. *Procedia Engineering* 39, 157–167 (2012), doi: 10.1016/j.proeng.2012.07.020.
5. Tarel’nik, V. B., Paustovskii, A. V., Tkachenko, Y. G., Martsinkovskii, V. S., Konoplyanchenko, E. V., Antoshevskii, K.: Electric-spark coatings on a steel

base and contact surface for optimizing the working characteristics of babbitt friction bearings. *Surface Engineering and Applied Electrochemistry* 53(3), 285–294 (2017), doi: 10.3103/S1068375517030140.

6. Zhang, K., Yang, Z.: Identification of load categories in rotor system based on vibration analysis. *Sensors* 17(7), 1676 (2017), doi: 10.3390/s17071676.
7. Djaidir, B., Hafaifa, A., Kouzou, A.: Faults detection in gas turbine rotor using vibration analysis under varying conditions. *Journal of Theoretical and Applied Mechanics* 55(2), 393–406 (2017), doi: 10.15632/jtam-pl.55.2.393.
8. Salim, H., Faisal, K., Jawad, R.: Enhancement of performance for steam turbine in thermal power plants using artificial neural network and electric circuit design. *Applied Computational Intelligence and Soft Computing* 2018, 8042498 (2018), doi: 10.1155/2018/8042498.
9. Gao, H., Bai, G.: Reliability analysis on resonance for low-pressure compressor rotor blade based on least squares support vector machine with leave-one-out cross-validation. *Advances in Mechanical Engineering* 7(4), (2015), doi: 10.1177/1687814015578351.
10. Mohamed, B., Ahmed, H., Mouloud, G., Qi, C. X.: Detection and modeling vibrational behavior of a gas turbine based on dynamic neural networks approach. *Journal of Mechanical Engineering – Strojnický Casopis* 68(3), 143–166 (2018), doi: doi.org/10.2478/scjme-2018-0032.
11. Ogbonnaya, E. A., Adigio, E. M., Ugwu, H. U., Anumiri, M. C.: Advanced gas turbine rotor shaft fault diagnosis using artificial neural network. *International Journal of Engineering and Technology Innovation* 3(1), 58–69 (2013).
12. Adewusi, S. A.: Detection of propagating cracks in rotors using neural networks. *Pipe and Component Analysis and Diagnosis* 447, 71–78 (2002).
13. Samanta, B., Al-Balushi, K. R., Al-Araimi, S. A.: Bearing fault detection using artificial neural networks and genetic algorithm. *EURASIP Journal on Applied Signal Processing*, Hindawi Publishing Corporation 3, 366–377 (2004).
14. Pavlenko, I. V., Simonovskiy, V. I., Demianenko, M. M.: Dynamic analysis of centrifugal machines rotors supported on ball bearings by combined application of 3D and beam finite element models. *IOP Conference Series: Materials Science and Engineering* 233(1), 012053 (2017), doi: 10.1088/1757-899X/233/1/012053.
15. Yashchenko, A. S., Rudenko, A. A., Simonovskiy, V. I., Kozlov, O. M.: Effect of bearing housings on centrifugal pump rotor dynamics. *IOP Conference Series: Materials Science and Engineering* 233(1), 012054 (2017), doi: 10.1088/1757-899X/233/1/012054.
16. Kim, Y. W., Jeong, W. B.: Reliability evaluation technique of compressor using pressure pulsation and vibration signals. *Journal of Physics: Conference Series* 1075, 012076 (2018), doi: 10.1088/1742-6596/1075/1/012076.
17. Ding, F., Wang, Z., Qin, F.: Two kinds of neural network fusion of aero-engine rotor vibration signal fault diagnosis. In: *4th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering*, pp. 1546–1552 (2015).
18. Ben Rahmoune, M., Hafaifa, A., Guemana, M.: Neural network monitoring system used for the frequency vibration prediction in gas turbine. In: *3rd International Conference on Control, Engineering and Information Technology*, pp. 15418537 (2015), doi: 10.1109/CEIT.2015.7233185.

19. Tanoh, A., Konan, D. K., Koffi, M., Yeo, Z., Kouacou, M. A., Koffi, B. K., N'guessan, K. R.: A neural network application for diagnosis of the asynchronous machine. *Journal of Applied Sciences* 8, 3528–3531 (2008), doi: 10.3923/jas.2008.3528.3531.
20. Manjurul, M. M., Kim, I.-M.: Motor bearing fault diagnosis using deep convolutional neural networks with 2D analysis of vibration signal. *Lecture Notes in Computer Science* 10832, 144–155 (2018), doi: 10.1007/978-3-319-89656-4_12.
21. Pavlenko, I., Trojanowska, J., Ivanov, V., Liaposhchenko, O.: Scientific and methodological approach for the identification of mathematical models of mechanical systems by using artificial neural networks. In: 3rd Conference on Innovation, Engineering and Entrepreneurship, Regional HELIX 2018, *Lecture Notes in Electrical Engineering* 505, 299–306 (2019), doi: 10.1007/978-3-319-91334-6_41.
22. Pavlenko, I., Simonovskiy, V., Ivanov, V., Zajac, J., Pitel, J.: Application of artificial neural network for identification of bearing stiffness characteristics in rotor dynamics analysis. In: *Advances in Design, Simulation and Manufacturing, DSMIE 2018, Lecture Notes in Mechanical Engineering*, pp. 325–335 (2019), doi: 10.1007/978-3-319-93587-4_34.

Computer Modeling of Viscous Fluid Flow Based on the Regularized Lattice Boltzmann Model

Artem Ostapenko ^[0000-0003-3368-8203]

Pryazovskyi State Technical University, Universitetska str., 7, Mariupol, 87555, Ukraine

ostapt5@gmail.com

Abstract. The problem of the laminar fluid flow simulations with the lattice Boltzmann method at moderate and at large Reynolds numbers up to 20000 is considered. To avoid the fluctuations that appear in the numerical solutions with the LBM the regularization method that based on the median filtration is proposed. The regularization method was tested on the classical problem of the flow over the circular cylinder. The fluid flow modeling over the Nasa 0012 profile at different angles of attack is carried out at the Reynolds number 1000. The flow patterns, the drag and lift coefficients of the NASA profile studied in detail.

Keywords: viscous fluid, Boltzmann equation, regularization, circular cylinder, Nasa 0012 profile, Reynolds numbers

1 Introduction

Methods of the computational fluid dynamics (CFD) are classified according to the different approaches in continuous medium describing. These approaches are the Euler and the Lagrange ones [1]. Besides them there are hybrid methods that combine the advantages of the Euler and Lagrange approaches. The first hybrid CFD method for the non-stationary flow calculating was developed in 1955 in the USA by Harlow on the basis of the Euler equation. It was called Particles in Cell method (PiC) [2]. The computational domain divides by a stationary Euler grid into cells in which particles are located. Fluid describes by a discrete model, as a collection of particles of fixed mass, moving along the cells of a grid. The method of particles in cells is convenient to study the dynamics of multicomponent flows and flows with free boundaries. However, it has two significant disadvantages. The first one follows from the fluid discrete model: the fluctuations appearing which is the irregularity of the movement of particles. The second disadvantage is the high memory demands and the low speed of calculations.

In 1965 as a development of the PiC method was proposed the method of large particles, developed by Davidov U. M. and Bilatserkovskiy O. M [1]. The computational domain divides by an Eulerian grid which cells are considered as large particles. The motion of such large particles is modeled on the basis of the non-stationary

Euler model. Thus, the method occupies an intermediate position between the Harlow's PiC method and the classical finite-difference methods.

In 1973 was appeared another approach in the fluid flow modeling based on the cellular automata - Lattice Gas Cellular Automata (LGA) [3]. In this method particles can move to the next cells and collide in them according to the laws of gas dynamics (Boltzmann equation). The disadvantage of this method is the stochastic noise that appears when densities, impulse and velocities are calculating. As a means of sealing such noise the integer number of particles replaces by the particles distribution function. The evolution of the distribution function is described by the Boltzmann kinetic equation. This approach was called the lattice Boltzmann method (LBM) [3-6].

The LB method becomes widespread. A lot of scientists in Europe and America are engaged in its development [7-10]. The popularity of the method is due to its advantages:

- Simplicity in description and programming, since all stages of the method are the linear equations that solves by the explicit schemes [3, 6];
- the simplicity in the boundary conditions settings, due to the fact that the method is based on the kinetic approach in the continuous medium describing [6, 11];
- the ability to simulate flows in different difficult domains;
- ease in external forces adding [12];
- the possibility of using parallel computing technologies [13, 14];
- solving a wide class of tasks.

However, despite the growing popularity, there are still the following problems:

1. Significant calculation time, which increases with Reynolds number increasing [15];
2. Conditional stability [3, 15, 16].

These problems make it difficult to obtain numerical solutions for the flows at moderate Reynolds numbers $Re \sim 100$ and prevent simulations at large numbers $Re > 100$. For a partial solution of these problems at small Reynolds numbers $Re < 100$ schemes with several relaxation parameters [17, 18] or implicit schemes [19] are used. However, the above problems are still not solved to the end and determine the relevance of the article as well as its scientific and practical significance.

2 Regularized lattice Boltzmann model

According to the lattice Boltzmann method, the computational domain divides into a grid which cells are treated as large particles [3-6]. These large particles can move to one of the neighbors cells and collide in them according to the Boltzmann's kinetic equation. Such large particles are describing statistically using the particle distribution function [20].

The mathematical formulation of the problem consists of the Boltzmann equation (1), in which the particles collision integral is replaced by the Bhatnagar-Gross-Kruck approximation (2) [21]:

$$\partial_t f + \vec{v} \partial_x f + \frac{\vec{F}}{m} \partial_v f = \int |\vec{v}_1 - \vec{v}_2| \left(f_1' f_2' - f_1 f_2 \right) d\sigma d\vec{v}_2 \quad (1)$$

$$I_{coll} = \int |\vec{v}_1 - \vec{v}_2| \left(f_1' f_2' - f_1 f_2 \right) d\sigma d\vec{v}_2 \approx \frac{f^{eq} - f}{\tau} \quad (2)$$

Numerical model is a system of the lattice equations (3), obtained by sampling the mathematical model (1) - (2) with the absence of the external forces $\vec{F} \equiv 0$.

$$\underbrace{f_k(\vec{r} + \vec{V}_k \Delta t, t + \Delta t)}_{stream} = \underbrace{f_k(\vec{r}, t) - \frac{1}{\tau} \left[f_k(\vec{r}, t) - f_k^{eq}(\vec{r}, t) \right]}_{collision} \quad (3)$$

where $f_k(\vec{r}, t)$ – discrete particle distribution function;

$\vec{V}_k$ – discrete series of particles possible velocities;

Δt – time step;

τ – dimensionless relaxation parameter;

$f_k^{eq}(\vec{r}, t)$ – discrete approximation of the Maxwell-Boltzmann local equilibrium distribution function.

The justification of using the lattice Boltzmann method to model viscous flows is given in [22]. It was shown that using the Chapman-Enskog expansion on the equations (3) the Navier-Stokes equation and the continuity equation for a uncompression isothermal fluid (4) can be obtained as well as the formulas that associate macroscopic and lattice parameters (5) - (8).

$$\begin{aligned} \frac{\partial u}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} &= -\nabla p + \nu \Delta \vec{u} + o(\Delta t); \\ \nabla \cdot \vec{u} &= 0 + o(\Delta t) + o(M_p) \end{aligned} \quad (4)$$

$$\begin{array}{ll} \text{Kinematic viscosity} & \\ \text{of the liquid} & \nu = c_s^2 \Delta t (\tau - 0.5) \end{array} \quad (5)$$

$$\begin{array}{ll} \text{Lattice speed of the} & \\ \text{sound} & c_s = \frac{c}{\sqrt{3}} \end{array} \quad (6)$$

$$\begin{array}{ll} \text{Lattice speed} & \\ & c = \frac{d}{\Delta t} \end{array} \quad (7)$$

where d – cell's size

Lattice Mach number

$$M_p = \frac{U_{\max}}{c_s} \quad (8)$$

where $U_{\max}$ – maximum velocity value in the domain

The stability of the LBM numerical scheme was investigated analytically in [3, 5, 16, 22]. The instability of numerical solutions may be caused by one of the following factors:

- lattice Mach number increasing. The Boltzmann kinetic equation approximates the Navier-Stokes equation only at small Mach numbers $M_p < 0.3$ [22].
- relaxation parameter decreasing. The "safe" value (according to [5]) is $\tau=1$. The relaxation parameter decrease τ causes instability: the particles accumulate in some cells, which results in pulsations in the velocity field (fig. 1);
- speed increasing. The numerical model assumes modeling flows only at low speeds $U_{\max} < c_s$.
- Reynolds number increasing. As the Reynolds number increases, the flow becomes turbulent (according to [23-24]).

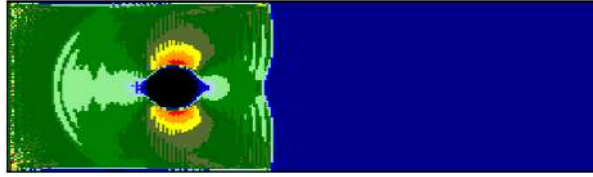


Fig. 1. Fluctuations in the velocity field

Macroscopic fluid parameters: density, velocity and pressure are defined as moments of the distribution function according to the formulas:

$$\rho(\vec{r}, t) = \sum_{k=0}^8 f_k(\vec{r}, t), \quad \vec{u}(\vec{r}, t) = \frac{1}{\rho(\vec{r}, t)} \sum_{k=0}^8 \vec{V}_k f_k(\vec{r}, t), \quad p(\vec{r}, t) = c_s^2 \rho(\vec{r}, t) \quad (9)$$

The possible particles moving directions $\{\vec{e}_k\}_{k=1}^9$ determines the lattice of the model. The lattice is defined in such a way as to ensure the isotropy of the corresponding tensors of the 4th and 6th orders and to provide the momentum moment conservation law. Numerous scheme for solving two-dimensional tasks on the basis of lattice model with nine possible directions of particle movement (D2Q9) was used.

To simulate viscous fluid flow at large Reynolds numbers with the LBM a smoothing scheme based on the A. L. Chudov [25] idea was constructed. The smoothing scheme is based on the function value correction at a space point in accordance with the neighboring values. The basis of such correction is the median filtration (10) with the smoothing condition (11):

$$u_x(i, j) = \text{med}(u_x(i, j-2), u_x(i, j-1), u_x(i, j), u_x(i, j+1), u_x(i, j+2)) \quad (10)$$

$$u_x(i, j) > u_x(i, j \pm 1) \vee u_x(i, j) < u_x(i, j \pm 1) \quad (11)$$

The advantage of such filter is the nonlinearity with the following properties [26]:

- after smoothing, the sharp boundaries of the solution areas are preserving;
- non-correlated or weakly corrected obstacles are suppressing;
- fluctuations are reducing.

Smoothing using (10-11) equations schematically depicted in fig. 2.

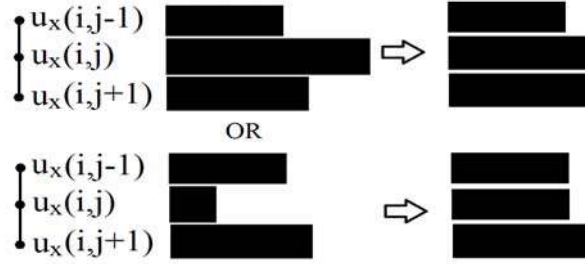


Fig. 2. Velocity smoothing

3 Verification

To verify the proposed method a series of numerical experiments of numerical solutions regularizing was conducted. Considered the flow pattern at $Re=500$. Fig. 3a illustrates velocity diagram got with the LBM algorithm at the moment $t=2.5$. The fig. 3a shows pulsations, which arise mainly on the boundary of the region. Further solution is unstable. Figure 3b illustrates velocity diagram of smoothed solution.

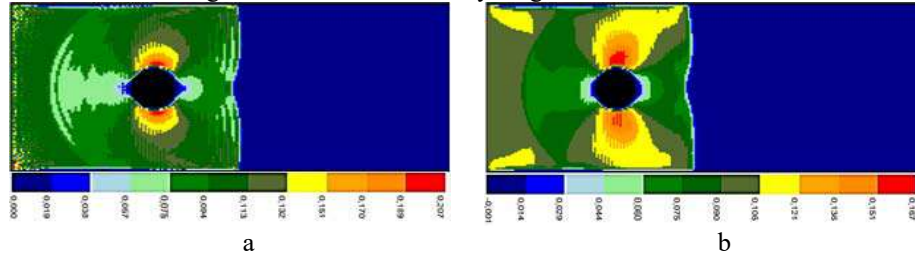


Fig. 3. Velocity diagram at $Re = 500$ at time $t = 2.5$ a) without smoothing b) smoothed solution

The regularization method has allowed to smooth the fluctuations and obtain a stable solution over large time intervals, such as the velocity diagram (fig. 4 a) and the streamlines (fig. 4 b).

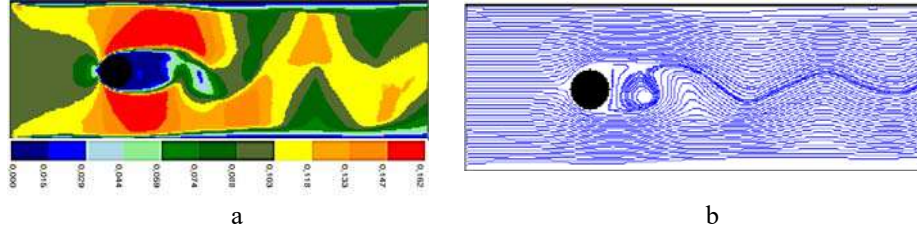


Fig. 4. Simulation with the LBM with smoothing at $Re = 500$ at time $t = 50$ a) velocity diagram b) streamlines

The resulting solution was compared with the solution of a similar problem that was obtained in the Comsol Multiphysics package with the finite element method (fig. 5). Figures 4 and 5 shows the similarity of the obtained patterns and the streamlines.

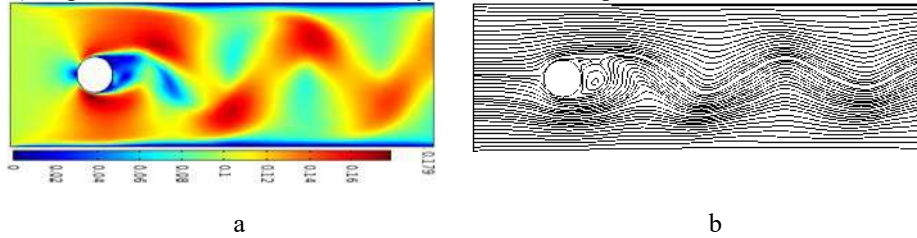


Fig. 5. Simulation in the Comsol Multiphysics package at $Re = 500$ at time $t = 50$ a) velocity diagram b) streamlines

A more detailed comparison was made in vertical sections near the streamlined body $x=0.8, 1.0$ (fig. 6). The graphs of the velocity distribution, obtained by the LBM method with smoothing and the finite elements method (FEM), were constructed.

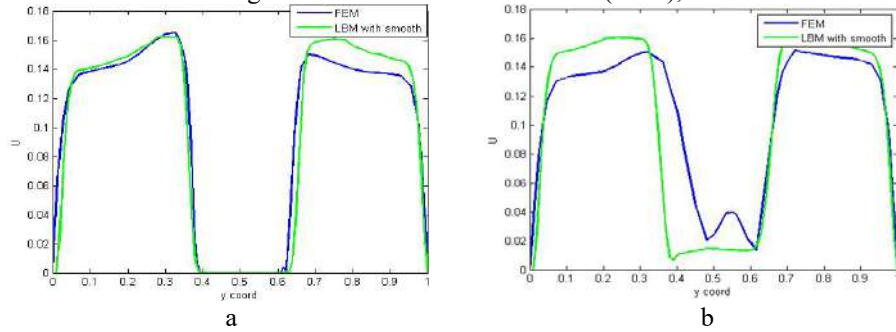


Fig. 6. Velocity distribution at the section a) $x=0.8$ b) $x=1.0$

Consequently, the solutions of the test problem obtained by the lattice Boltzmann method with smoothing showed a good correspondence with the results of other numerical experiments at the Reynolds number $Re = 500$.

Increase the Reynolds number up to $Re = 10,000$. The structuring of such flow disappears and the turbulent is appearing [23, 24]. A photo of such flow is shown in fig. 7 [27].

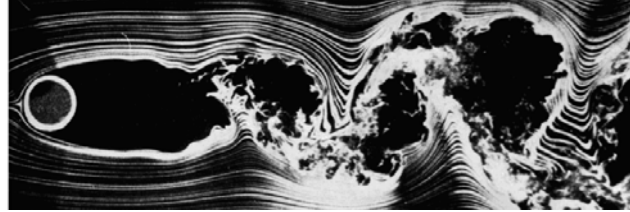


Fig. 7. Photo of the fluid flow at $Re = 10\,000$

The conducted numerical experiments give the following results. Fig. 8 illustrates the simulation results got with the LBM with smoothing.

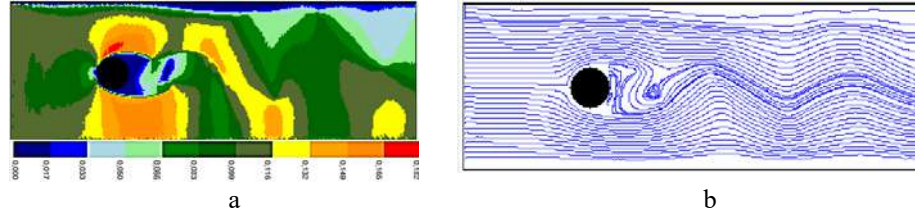


Fig. 8. Simulation with the LBM with smoothing at $Re = 10000$ at time $t = 100$ a) velocity diagram b) streamlines

Fig. 9 illustrates the results of flow simulation with the FEM method.

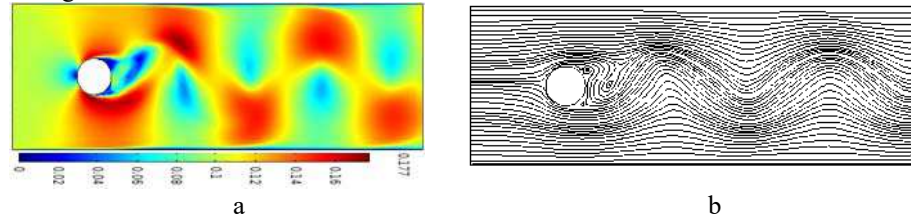


Fig. 9. Simulation in the Comsol Multiphysics package at $Re = 10000$ at time $t = 96$ a) velocity diagram b) streamlines

As the Reynolds number increases, the periodicity of vortex formation increases: vortices descend from the edges of the cylinder at about $t = 2$. Such a rapid formation of vortices led to a divergence of numerical solutions by the FEM and LBM methods at about $t = 2$ (figures 8 and 9).

The results of the fluid flow modeling at Reynolds number $Re = 20000$, are depicted in fig. 10 with LBM method and fig.11 with the FEM method.

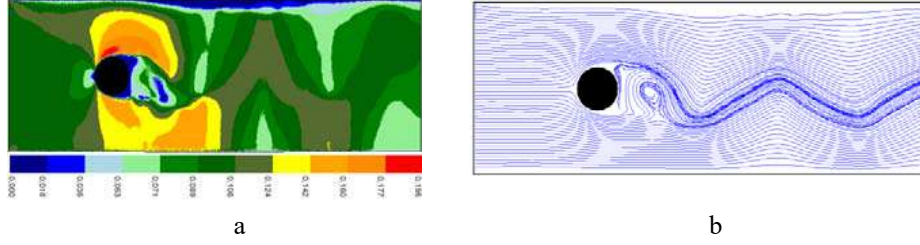


Fig. 10. Simulation with the LBM with smoothing at $Re = 20000$ at time $t = 50$ a) velocity diagram b) streamlines

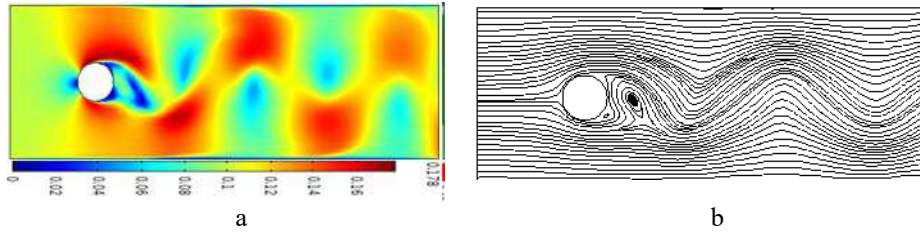


Fig. 11. Simulation in the Comsol Multiphysics package at $Re = 20000$ at time $t = 50$ a) velocity diagram b) streamlines

The results of numerical experiments for the viscous fluid flow modeling with the lattice Boltzmann method and the finite element method in the Comsol package at Reynolds numbers $500 < Re < 20000$ have been studied in detail. The comparison shows the good correspondence of the results, namely the process of the boundary layer separation and the formation of vortices over the cylinder. The structure of the streamlines also coincides.

4 Experiments and results

The fluid flow modeling over the Nasa 0012 profile at different attack angles was carried out at the Reynolds number $Re=1000$. Figure 12 illustrates the flow scheme: the velocity vector V_{in} , angle of attack α , drag and lift coefficients $\vec{F}_{drag}, \vec{F}_{lift}$.

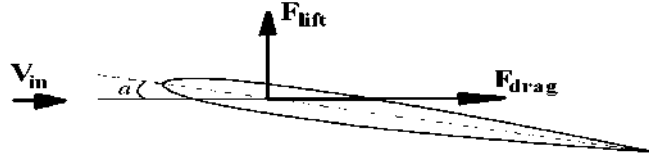


Fig. 12. Scheme of the flow over the NASA 0012 profile

The influence of the angle of attack on the flow pattern and on the drag and lift coefficients was studied. The flow pattern with the zero of the angle of attack is depicted in fig. 13.

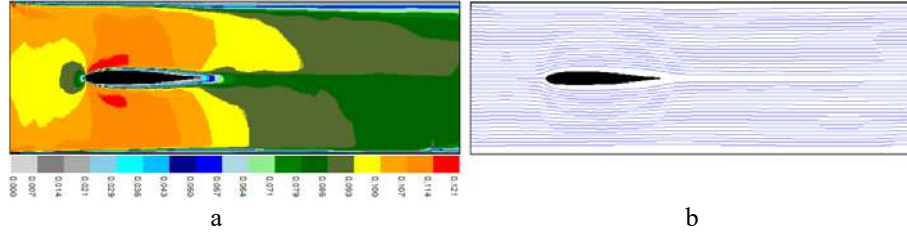


Fig. 13. Flow over the Nasa 0012 profile with the absent angle a) velocity diagram b) streamlines

As can be seen from the fig. 12, the flow around the profile is symmetric. Was calculated the average values of the drag C_d and lift C_l coefficients at different times up to $t=100$. The appropriate graphs are depicted in fig. 14.

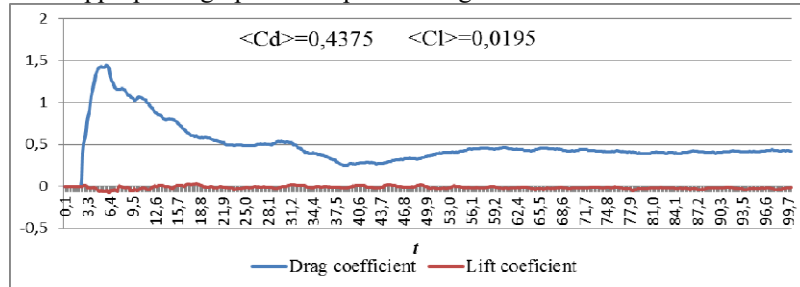


Fig. 14. Graphs of the drag and lift coefficients average values at different times

The change of the flow pattern by the angle of attack increasing $\alpha=10^\circ, 20^\circ, 30^\circ$ was considered. The average values of the hydrodynamic coefficients shown in fig. 15.

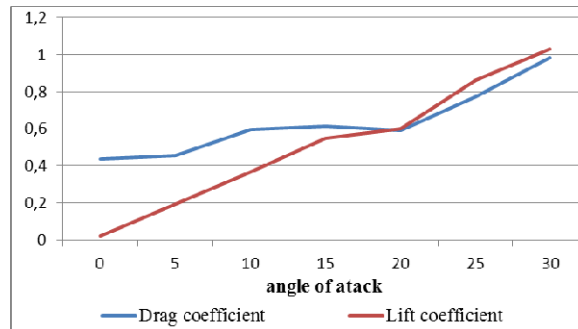


Fig. 15. Graphs of the dependence of the hydrodynamic coefficients on the angle of attack

The angle of attack increasing cause the drag and lift coefficients increasing too as can be seen from the fig. 15. The numerical simulations with the angles of the Nasa 0012 profile more than 30° are unstable with the regularized lattice Boltzmann model. Fig. 16 illustrates the velocity diagrams and streamlines of the flow over the profile with the different angles of attack $\alpha=10^\circ, 20^\circ, 30^\circ$.

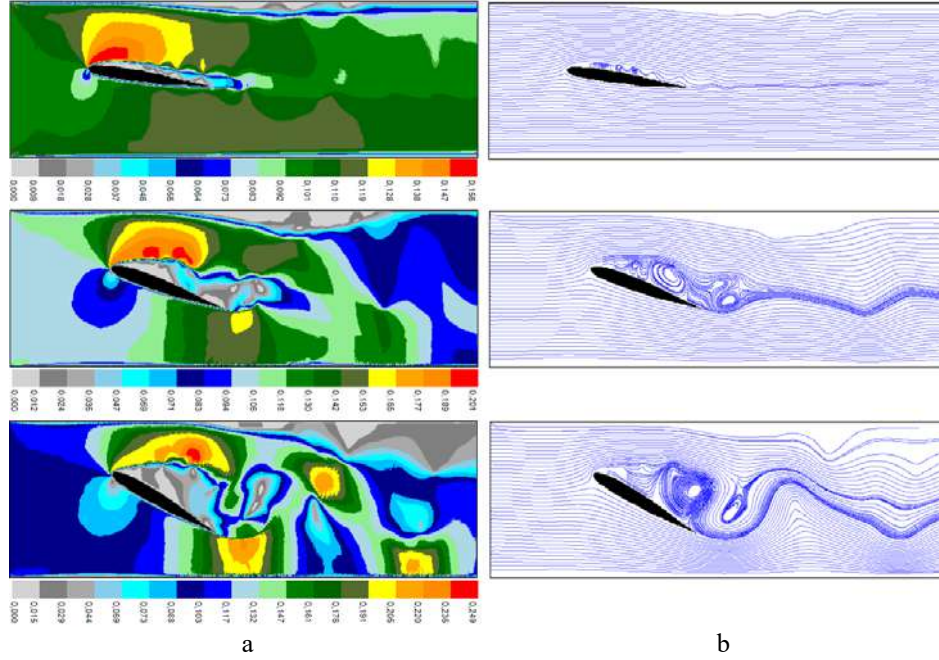


Fig. 15. Flow pattern over the Nasa 0012 profile with the different angels of attack $\alpha=10^\circ$, 20° , 30° *a*) velocity diagram *b*) streamlines

The conducted numerical experiments indicate the growth of the hydrodynamic coefficients and the formation of vortices over the profile with the angle of attack increasing.

5 Conclusion

The problem of the laminar fluid flow simulations with the lattice Boltzmann method at moderate and at large Reynolds numbers up to 20000 was considered. Noticed, that LBM is a new approach in CFD that used the Boltzmann equation to simulate fluid behavior. Although this method becomes more popular there are some problems that make fluid flow modeling at moderate and large Reynolds numbers impossible. They are the significant calculation time, which increases with Reynolds number increasing and conditional stability. The last one may causes fluctuations which are the irregularity of the movement of particles. To avoid the fluctuations the regularization method that based on the median filtration was proposed.

The regularization method was tested on the classical problem of the flow over the circular cylinder. The comparison of the numerical experiments with the lattice Boltzmann method and the finite element method in Comsol Multiphysics package showed the good correspondence of the results, namely the process of the boundary layer separation and the formation of vortices over the cylinder. The structure of the

streamlines also coincides. So, the regularization method has allowed to smooth the fluctuations and to obtain a stable solution. This allows to use the LBM for studying the laminar patterns at large Reynolds numbers.

The fluid flow modeling over the Nasa 0012 profile at different angles of attack was carried out at the Reynolds number $Re=1000$. Declared the vortex formation at the angles more than 7° . Size of the vortex is increasing with the angle increasing as well as the hydrodynamics coefficients. The numerical simulations with the angles of the Nasa 0012 profile more than 30° are unstable even with the regularized lattice Boltzmann model so demands more detail study.

References

1. Bilatserkovskiy, O.M.: Numerical modeling in continuum mechanics. Fizmatlit, Moscow (1994)
2. Harlow, F.X.: Numerical method of particles in cells for problems of hydrodynamics. Mir, Moscow (1967)
3. Wolf-Gladrow, D.: Lattice-Gas Cellular Automata and Lattice Boltzmann Models - An Introduction. Alfred Wegener Institute for Polar and Marine, Bremerhaven (2005)
4. Succi, S., Benzi R., Higuera, F.: The Lattice Boltzmann Equation: A New Tool for Computational Fluid-Dynamics. *Physica D: Nonlinear Phenomena*. 47, 219-230 (1991)
5. Sucop, M.: Lattice Boltzmann Modeling. An Introduction for Geoscientists and Engineers. Springer, Miami (2006)
6. Succi, S.: The Lattice Boltzmann Equation for Fluid Dynamics and Beyond. University Press, Oxford (2001)
7. Grazyna, K.: The numerical solution of the transient heat conduction problem using the lattice Boltzmann method. *Scientific Research of the Institute of Mathematic and Computer Science*. 11, 23-30 (2006)
8. Leclaire, S., Pellerin, N., Reggio, M., Trepanier, J.-Y.: Enhanced equilibrium distribution functions for simulating immiscible multiphase flows with variable density ratios in a class of lattice Boltzmann models. *International Journal of Multiphase Flow*. 57, 159-168 (2013)
9. Favier, J., Revell, A., Pinelli, A.: A lattice Boltzmann-immersed boundary method to simulate the fluid interaction with moving and slender flexible objects. *Journal of Computational Physics*. 261, 145-161 (2014)
10. Anderl, D., Bogner, S., Rauh, C., Rude, U., Delgado, A.: Free surface lattice Boltzmann with enhanced bubble model. *Computers and Mathematics with Applications*. 67(2), 331-339 (2014)
11. Latt, J., Chopard, B., Malaspinas, O., Deville, M., Michler, A.: Straight velocity boundaries in the lattice Boltzmann method. *Physical Review*. 77, 1-17 (2008)
12. Kupershtokh, A.L.: Accounting for the action of volume forces in the lattice Boltzmann equations. *Bulletin of the Novosibirsk State University. Series "Mathematics, mechanics, computer science"*. 6(2), 75-96 (2004)

13. Kupershtokh A.L.: 3D modeling with the lattice Boltzmann Method on GPU. Numerical methods and programming. 13, 130-138 (2012)
14. Agrawal, T., Lin, C.: Implementation of an incompressible lattice Boltzmann model on GPU to simulate Poiseuille flow. Proceeding of the Forties National Conference on Fluid Mechanics and Fluid Power. 40, 991-995 (2013)
15. Bulanchuk, G., Bulanchuk, O., Ostapenko, A.: Stability investigation of the two-dimensional nine-vectors model of the lattice Boltzmann method for fluid flows in a square cavity. Bulletin of Kharkov national University named after V.N.Karazin. Series: Mathematical modeling. Information technology. Automated control system. 28, 113-125 (2015)
16. Aslan, E.: Investigation of the Lattice Boltzmann SRT and MRT Stability for Lid Driven Cavity Flow. International Journal of Materials, Mechanics and Manufacturing. 2(4), 317-324 (2014)
17. Rettinger, C.: Fluid Flow Simulation using the Lattice Boltzmann Method with multiple relaxation times. Friedrich-Alexander Universuty Erlanger-Nuremberg, Erlanger (2013)
18. Zaharov, A.M., Senin, D.S., Grachev, E.A.: Lattice Boltzmann modeling with multirelaxationals times. Numerical methods and programming. 15, 644-657 (2014)
19. Krivovichev, G.V.: On the computation of viscous fluid flows by the lattice Boltzmann method. Computer Reserch and Modeling. 5(2), 165-178 (2013)
20. Liapin, I.I.: Introduction to the kinetic equations theory. UGTU, Ekaterinburg (2003)
21. Mussa, M.: Numerical Simulation of Lid-Driven Cavity Flow Using the Lattice Boltzmann Method. Applied Mathematics. 13, 236-240 (2008)
22. He, X.: Lattice Boltzmann Model for the Incompressible Navier – Stokes Equation. Journal of statistical physics. 88, 927–944 (1997)
23. Schlichting, H.: Boundary layer theory. Nauka, Moskow (1974)
24. Bilatserkovskiy, O.M., Bilatserkovskiy, S.O., Gushin, V.A.: Numerical modeling of the unsteady periodic flow of the viscous fluid past a cylinder. Computational mathematics and mathematical physics. 24(8), 1207-1216 (1984)
25. Paskonov, V.M., Polezhaev, V.I., Chudov, L.A.: Numerical simulation of heat and mass transfer process. Nauka, Moscow (1984)
26. Micek, J., Kapitulok, J.: Median Filter. Journal of Information, Control and Management Systems. 1(2), 51-56 (2003)
27. Van Dyke, M.: An Album of Fluid Mechanics. The Parabolic Press, Stanford (1982)

Computer Modeling of the Obtaining Nanostructures Process under the Action of Laser Radiation on Steel

Gennadiy Kostyuk¹[0000-0002-7584-2240], Viktor Popov²[0000-0001-9189-6882],
Kateryna Kostyk³[0000-0003-4139-9970]

¹National Aerospace University named by N.Ye. Zhukovsky «KhAI», Chkalov str., 17,
Kharkiv, 61070 Ukraine
g.kostyuk@khai.edu

²Joint Stock Company “FED”, Sumska str., 132, Kharkiv, 61023 Ukraine
vvpopov123@gmail.com

³National Technical University «Kharkiv Polytechnic Institute», Kyrpychova str., 2, Kharkiv,
61002 Ukraine
eklitus@gmail.com

Abstract. The conditions for the formation of nanostructures in the surface layer of steels with different carbon content (steel 20, 40, 45, 40Cr, U8 and U12), which determine the required temperature (500-1500 K) and the rate of increase (more than 10^7 K/s). The zones of formation of nanostructures depending on the heat flux density on the time of action of the ionizing radiation are determined. It is shown that it is necessary to take into account the rate of temperature rise and the probability of thermoelastic destruction due to the action of temperature stresses. In low speed temperature rise formed micro and submicrostructure that has been confirmed experimentally.

Keywords: nanostructure, submicrostructure, technological parameters, steel, pulse laser radiation

1 Introduction

Nanostructured surface layers can significantly improve the performance of parts by increasing the microhardness of the surface, reducing the modulus of elasticity with increasing yield strength and toughness, which allows you to create on the details of the surface layers to increase the life of the parts working under shock loads and guarantee their long-term strength.

As our reviews presented in [1-5] showed, despite the fact that nanostructures were experimentally obtained under the action of ionizing radiation, the theoretical consideration of the possibility of obtaining nanostructures was not considered. All this is due to the fact that the criteria for obtaining nanostructures (NS) were not formulated, which are formed only in the temperature range 500-1500 K at a temperature rise rate of more than 10^7 K/s, and are intensified by the action of non-stationary temperature stresses of the order 10^8 - 10^{10} Pa. In addition, the time of action of temperatures should

be such that the process of grain size growth under the prolonged action of temperature is not realized. Then these criteria can be expanded to the following: time of cooling to temperatures close to 500 K was not more than $e \cdot \tau_u$ (τ_u – the duration of the radiation pulse) that will ensure the stability of the formation of the NS.

2 Effect of laser radiation on structural materials

2.1 Features of the description of the heat source under the action of the laser on opaque materials

For technological purposes, a focused source of coherent radiation is used, the heat flux density of which is distributed in the focal plane as follows:

$$q(r) = q_0 \left[\frac{2I_1(Br)}{(Br)^2} \right]^2 \quad (1)$$

where $I_1(Br)$ is a Bessel function of the first kind of the first order;

$$B = I_1 D / \lambda F . \quad (2)$$

Where D is the diameter of the lens; λ – wavelength radiation; F – focal length; q_0 – intensity in the center of the spot, calculated by the formula:

$$q_0 = \frac{I_1 D^2}{4\lambda^2 F^2} P_0 \quad (3)$$

where P_0 is the radiation power absorbed by the material.

The heat flux density can be obtained from the product of a time-dependent function and a surface coordinate function:

$$q_n(r\tau) = A\varphi(\tau)q(r) . \quad (4)$$

Where A is the absorption capacity of the processed material, which generally depends on both the state of the surface (degree of processing, roughness) and its temperature; $q(r)$ – the spatial distribution of the radiation power calculated by the formula (1) taking into account (2) and (3); $\varphi(\tau)$ – describes the time structure of the pulse; for example, for a laser pulse operating in a beam mode with ordered generation, the function can be represented as:

$$\varphi(\tau) = (1 - \cos \omega\tau) . \quad (5)$$

For envelope of lumps the expression is true:

$$\varphi(\tau) = \tau^n \exp(br^m), \quad (6)$$

where n and m are some numbers (integer or fractional).

Modulation of the q-switched laser allows to monoimpulse radiation duration $\approx 10^{-8}$ s, is a temporary structure which can be described by a function that is close to triangular and the slope front may be different from the rear.

The light flux falling on the surface of the material is partially reflected, and the rest of it passes inside the body volume and is absorbed. Inside and on the surface of the body is the heat source distributed in space and time.

The density of the absorbed heat flux for almost all technological applications of the laser varies within the volume of the material according to the Booger law:

$$q_V(z) = q_{V_0}(1-R)e^{-\alpha z}, \quad (7)$$

where $q_V(z)$, q_{V_0} are, respectively, the volume densities of the heat flux of radiation at a distance z and on the surface, W/cm³,

$$q_{V_0} = \frac{q_n(r, \tau)}{0.1\delta}. \quad (8)$$

Where $\delta \approx 10^{-4} \dots 10^{-5}$ m – layer, where the light flux is absorbed in the interaction with the conduction electrons; R and α are respectively the reflectivity and light absorption coefficient.

2.2 Model of the interaction of the light beam fluxes and construction materials

Subject to the volume and source of radiation, the amount of probability density is calculated according to the expression (7) with (8), and the surface heat source with heat density of current according to the expression (4) with (5), (6), solve the equation.

Heat balance in the unit volume of the part

The heat balance is represented in the expression:

$$\begin{aligned} & C[T]\gamma[T] \frac{dT(x, y, z, t)}{dt} + C[T]\gamma[T] \frac{\partial T(x, y, z, t)}{\partial y} V_n + C[T]\gamma[T] \tau_p \frac{d^2 T(x, y, z, t)}{dt^2} = \\ & = \nabla \lambda [T] \nabla T(x, y, z, t) + C[T]\gamma[T] V_{tm} \frac{\partial T(x, y, z, t)}{\partial x} - AL_{III} \gamma[T] \frac{dV_{nl}}{dt} + \\ & + B \frac{q_n(r, \tau)}{0.1\delta} \pm D \frac{dW(x, y, z, t)}{dt} \pm m_a C_a [T_a] \frac{dn_a}{dt} (T_a - T(x, y, z, t)) \pm \\ & \pm P_{T.X.P} (n_A, n_B, T, t_e) \frac{dn_{A(B)}}{dt} L_{T.X.P} + q_v(z), \end{aligned} \quad (9)$$

where $C[T]$ and $\gamma[T]$ is the specific heat and density of the target material corresponding to a temperature T ; τ_p is the relaxation time temperature by one Kelvin; V_n is the velocity of the plasma flow of laser radiation or the target relative to it; $L_{\Pi\Pi}$ and $L_{T.X.P}$ are specific heat of fusion and thermochemical reaction; $V_\phi[T]$ – the displacement speed of the evaporation front; V_{ni} – the volume of molten metal; $W(x, y, z, t)$ – energy of deformation of a unit volume target; m_a – mass of the diffusing atom; $Ca[Ta]$ is the heat capacity of the diffusing material at a temperature T_a ; $P_{T.X}(n_A, n_B, T, t_{B3})$ – the probability of thermochemical reactions that depend on concentration of the reagents; n_A and n_B , are the temperature T and the time of interaction; t_{B3} ; n_A and n_B are the concentration of the reactants that determine the possibility of a reaction.

The amount change of heat in a unit volume (the first term in the left part of the equation) is realized by moving the laser radiation flow along the treated surface or moving the target relative to the laser radiation flow at the rate of V_n (the second term); thermophysical processes: effects on the heat transfer of the final rate of heat propagation (the third term), thermal conductivity(the first term on the right), displacement of the evaporation front(the second term), melting (the third term); collision processes: volumetric heat source due to the action of laser radiation (the fourth term), thermoelastic, thermoplastic and thermo-fatigue processes that determine the energy of deformation of the material of elementary volume (the fifth term); diffusion processes that determine the heat transfer of the diffusing material (the sixth term); thermochemical processes associated with the implementation of chemical reactions between the part material and the coating material or between the components of alloys and composite materials, the volumetric heat source due to the action of the light-beam flow.

When evaporating from a layer of molten metal (Poole-Frenkel conduction mechanism [6, 7]), the evaporation rate is determined by the formula

$$V_\phi = V_0 \exp\left(\frac{T^*}{T(0, x, y, z, t)}\right) \quad (10)$$

where V_0 and T^* are critical evaporation rates and surface temperatures [8].

To determine these values, we use the approximation of the graphs of the dependence of V_0 and T^* on the heat flux density q . For a good approximation, a quadratic function is suitable, the coefficients of which can be determined by the program of approximation of the function with a power basis by the least squares' method. On the charts, the values of q and V_0 are indicated on a logarithmic scale, so use the following substitutions:

$$V_0 = 10^{z-3}; \quad q = 10^{s+8}; \quad \Rightarrow \quad S = \lg q - 8 \quad (11)$$

The value of the heat flux density q can be determined by the formula (2).

V_0 and T^* dependences on q through auxiliary variables have the following form:

$$\begin{aligned} Z &= (nS^2 + kS + p) \cdot 10^{-1}; \\ T^* &= (mS^2 + rS + f) \cdot 10^{-2}. \end{aligned} \quad (12)$$

The coefficients in the formulas (18) for different materials are given in Table. 1.

Table 1. Coefficients n, k, p, m, r, f in the formula (18)

Material	n	k	p	m	r	f
Al	1.619	3.0952	8.8571	1.4048	2.5833	3.2321
W	1.7143	10.2857	9.9997	3.2857	3.4428	1.8032
Fe	1.751	3.2502	10.5021	3.2849	-1.9286	37.942
Mo	0.8928	4.7321	10.9731	4.5357	-1.6752	53.652
Cu	-0.0858	1.1319	1.0605	0.8451	3.9713	19.116

The third term on the right of the formula (9) takes into account the change in the amount of heat during melting and is calculated for the volume of the V_{nl} material in which the calculated temperature exceeds the melting point. The specific heat of fusion L_{nl} can be calculated according to the formula:

$$L_{nl} = nT_{nl}f(T_{nl}) \cdot 4186.8 \quad (13)$$

where n is the number of atoms in the molecule; $f(T_{nl})$ is a function of the dependence of L_{nl} on T_{nl} , which can be approximated as a linear dependence of the form:

$$f(T_{nl}) = (1.57T_{nl} + 1428) \cdot 10^{-3} \quad (14)$$

For example, for iron $f(T_{nl}) = 3.5$ cal/g. The specific melting heat L_{nl} in the formula (12) is measured in calories per gram.

The fourth term of the formula (9) calculates the change in the amount of heat due to the action of laser radiation as a volumetric heat source.

The fifth term of the formula (9) takes into account the energy expended on the deformation of the body during the action of the heat source t_u , and the energy returned to the material during stress relaxation (for a time greater than t_u and less than $t_u + \tau_p$).

The deformation energy of a single volume is determined by the formula:

$$W = G \left[\begin{aligned} &\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz}^2 + 2(\varepsilon_{xy}^2 + \varepsilon_{yz}^2 + \varepsilon_{zx}^2) + \\ &+ \frac{\mu}{1-\mu} \ell^2 - \frac{2(1+\mu)\alpha_1}{1-2\mu} \ell (T(x, y, z, t) - T) \end{aligned} \right] \quad (15)$$

where $2\varepsilon_{ik} = 2\varepsilon_{ki} = \frac{\partial u_k}{\partial i} + \frac{\partial u_i}{\partial k}$ ($k, i = x, y, z$); $\ell = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$; $\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}$ – e – elongation; $\varepsilon_{xy}, \varepsilon_{yz}, \varepsilon_{zx}$ – shifts relative to the corresponding axes; u_x, u_y, u_z – displacements relative to the corresponding axes; α_1 – coefficient of linear expansion of the target material; μ is the Poisson's ratio (the ratio of lateral strain to longitudinal value

μ concluded between 0.1 and 0.5), G is the shear modulus (modulus of the second kind) (for the iron $G = 3,5 \dots 10,3 \cdot 10^{10}$ N/m², $\mu = 0.23 \dots 0.31$); T_H – the initial temperature.

Deformations of shifts of ε_{ik} cannot be set arbitrarily, they are connected by differential relations-compatibility conditions [9].

$$\begin{aligned} \frac{\partial^2 \varepsilon_{xx}}{\partial y^2} + \frac{\partial^2 \varepsilon_{yy}}{\partial x^2} &= 2 \frac{\partial^2 \varepsilon_{xy}}{\partial x \partial y}; \quad \frac{\partial^2 \varepsilon_{xx}}{\partial y \partial z} = \frac{\partial}{\partial x} \left[-\frac{\partial \varepsilon_{yz}}{\partial x} + \frac{\partial \varepsilon_{zx}}{\partial y} + \frac{\partial \varepsilon_{xy}}{\partial z} \right]; \\ \frac{\partial^2 \varepsilon_{yy}}{\partial z^2} + \frac{\partial^2 \varepsilon_{zz}}{\partial y^2} &= 2 \frac{\partial^2 \varepsilon_{yz}}{\partial z \partial y}; \quad \frac{\partial^2 \varepsilon_{yy}}{\partial x \partial z} = \frac{\partial}{\partial y} \left[-\frac{\partial \varepsilon_{zx}}{\partial y} + \frac{\partial \varepsilon_{xy}}{\partial z} + \frac{\partial \varepsilon_{yz}}{\partial x} \right]; \\ \frac{\partial^2 \varepsilon_{xx}}{\partial z^2} + \frac{\partial^2 \varepsilon_{zz}}{\partial x^2} &= 2 \frac{\partial^2 \varepsilon_{zx}}{\partial x \partial z}; \quad \frac{\partial^2 \varepsilon_{zz}}{\partial y \partial x} = \frac{\partial}{\partial z} \left[-\frac{\partial \varepsilon_{yx}}{\partial z} + \frac{\partial \varepsilon_{yz}}{\partial x} + \frac{\partial \varepsilon_{xz}}{\partial y} \right]. \end{aligned} \quad (16)$$

These conditions verify the correct definition of elongation and shear, and their adjustment-the input of additional stresses. To determine the elongation ε_{xx} , ε_{yy} , ε_{zz} and shifts ε_{xy} , ε_{yz} , ε_{zx} use the expression of the thermoelastic displacement potential Φ :

$$\Delta \Phi - \frac{1-2\mu}{2(1-\mu)} \frac{\gamma}{G} \frac{\partial^2 \Phi}{\partial t^2} = \frac{1+\mu}{1-\mu} \alpha_1 [T(x, y, z, t) - T_H] \quad (17)$$

where γ is the density of the part material.

Given that the solution of the problem is carried out in a mobile coordinate system and when entering a stationary or close to it, the second term in the left part of the equation (17) becomes insignificant, we obtain:

$$\Delta \Phi = \frac{1+\mu}{1-\mu} \alpha_1 [T(x, y, z, t) - T_H] \quad (18)$$

The magnitude of the thermoelastic potential of displacements and the known ratios to find the magnitude of the elongation and changes [10-11].

$$\varepsilon_{ik} = \frac{\partial^2 \Phi}{\partial i \partial k}, \quad (i, k = x, y, z) \quad (19)$$

The values of temperature stresses are determined by the expression

$$\sigma_{ik} = 2G \left(\frac{\partial^2 \Phi}{\partial i \partial k} - \Delta \Phi \cdot \delta_{ik} \right), \quad (20)$$

where δ_{ik} is subject to the conditions: $\delta_{ik} = 0$ at $i \neq k$ ($i, k = x, y, z$); $\delta_{ik} = 1$ at $i = k$.

In the seventh term of the formula (9), taking into account the heat transfer of the diffusing material, the mass of the diffusing atom is determined by the formula.

$$m_a = Mm_p \quad (21)$$

where M is the atomic weight of the applied material; m_p is the mass of the proton.

The change in the concentration of diffusing atoms per unit time can be determined by the expressions:

- at $t < t_k$:

$$\frac{dn_a}{dt} = \frac{j_a}{ez\sqrt{\alpha\tau}}$$

- at $t \geq t_k$:

$$\frac{dn_a}{dt} = \frac{j}{ezL_D}, \quad (22)$$

where j is the current density of the introduced atoms; e is the electron charge; z – the charge number of the applied material; L_D – the thickness of the part; t_k – time for which the part will warm up to the entire thickness, $t_k = \frac{L_D^2}{\alpha}$; α is the thermal diffusivity.

The diffusion coefficient is calculated:

$$K_{diff} = a_d d_s^2 V_0 \exp(-U / (kT)) \quad (23)$$

where a_d is a multiplier of the order of 0.1, determined by the type of crystal lattice; d_s is the distance between the nearest equivalent positions of vacancies in the crystal; V_0 is the value of the order of the frequency of atomic oscillations in the crystal ($10^{12} \dots 10^{14} \text{ s}^{-1}$); U is the potential barrier that must be overcome by vacancies when displaced to the neighboring position; k is the Boltzmann constant; T is the absolute temperature.

The eighth term of the formula (9) takes into account the influence of chemical processes on the heat balance in the target. In the interaction of laser radiation with substances consisting of polyatomic molecules, a whole cycle of chemical transformations is possible, such as the excitation of a molecule with its subsequent dissociation (decomposition) into active particles (ions); rearrangement of atoms in the structure of the molecule; movement of individual atoms from one part of the molecule configuration to another; accession to the excited molecule of another molecule; transfer of excitation energy from one molecule to another; capture electrons to form negative ions; recombination of an ion with an electron or an ion with a molecule. The concentration of chemical transformations in the first approximation is directly proportional to the density of the absorbed energy and the chemical yield (the average number of chemical transformations in the absorption of single energy). As a result of chemical

reactions between ions and radicals of the plasma-forming gas and atoms of the material, it is possible to remove the material in the form of volatile compounds, which is used in plasma chemical etching (for example, adding 10 % oxygen to argon can increase the rate of removal of the material due to chemical reactions by 10-15 times). The rate of plasma-chemical etching is 2...10 nm/s. In the processing of elements W, Te, Mo, Ta fluorine-containing gases are formed volatile fluoride, and Al in the processing of chlorine-containing gases – volatile chlorides. Thermo- and plasma-chemical reactions can also contribute to an increase in the mass and volume of the workpiece due to the formation of chemical compounds with the reaction gas. Laser irradiation in the atmosphere of a chemically active gas or in a mixture of inert and chemically active gases, such as O_2 , $Ar + N_2$, $Ar + N_2$, etc., is accompanied by the following processes that ensure the growth of the film of the chemical compound: the reaction between the chemically active ions A and the atoms of the target B, followed by the transfer of AB molecules to the substrate; reaction between atoms and ions of the chemically active gas A, A^+ , A^- and atoms of the coating; reaction between atoms and ions A, A^+ , A^- and sprayed particles B in the gas phase, followed by the deposition of molecules AB on the surface of the target.

In the case where a multicomponent substance AB (chemical compound, alloy) is exposed to laser irradiation, the ratio of surface concentrations in the established process will be as follows:

$$\frac{n_A(0)}{n_B(0)} = \frac{n_{AV}}{n_{BV}} \left(\frac{M_A}{M_B} \right)^{1/4} \quad (24)$$

where n_{AV} and n_{BV} are the concentrations of atoms A and B in the target volume.

The surface layer is enriched with a heavier component.

If the working substance continuously enters the surface and decomposes there under the action of irradiation, then in the simplest case the rate of change in the surface concentration of the molecules of the compound that have entered N_1 and have not entered N_2 will be determined by the equations:

$$\frac{dN_1}{dt} = \frac{N_0 \sigma_p J_e}{e} \quad (25)$$

$$\frac{dN_0}{dt} = n_0 - \frac{N_0}{\tau_0} - \frac{dN_1}{dt} \quad (26)$$

where σ_p is the reaction cross-section depending on the properties of the compound and the energy of laser radiation; n_0 is the number of molecules entering the unit surface area per unit time; τ_0 is the average time during which the unreacted molecules are on the surface before evaporation.

Having integrated the system of equations (20)–(26) in the case of a small flow n_0 , we obtain:

$$N_0(t) = n_0 \sigma_h \tau_0 J_\ell (e + \sigma_h \tau_0 J_\ell)^{-1} \quad (27)$$

Here the designation is introduced, $\frac{1}{\tau_1} = \frac{1}{\tau_0} + \frac{\sigma_p J_e}{e}$ and $N_0(t)$ is the initial concentration of molecules at the time of laser irradiation at $t = 0$. After a period of time $t \gg \tau_l$ establishes an equilibrium concentration of, equal to $n_0 \tau_1$, and for the reaction to proceed with a constant speed:

$$dN_1 / dt = n_0 \sigma_p \tau_0 J_e (e + \sigma_p \tau_0 J_\ell)^{-1} \quad (28)$$

In general, this rate depends on all the parameters of the process (E_e, J_ℓ, T, n_0) , but if the reaction is carried out at sufficiently low T and high J_ℓ , so that $\sigma_p \tau_0 J_\ell \gg 1$ the reaction rate is determined only by the rate of adsorption and for a small value of n_0 is not very high.

At high flux density n_0 , sufficient to create a thicker coating than a monolayer, the reaction rate is determined only by the irradiation regime and is independent of n_0 and T :

$$dN_1 / dt = N_0 \sigma_p J_\ell / e \text{ at } n_0 \tau_1 \gg N_0 \quad (29)$$

3 The heat transfer on the surface of the part

Heat flow on the target surface is created by the following factors:

- collision processes: heat released on the surface due to the action of laser radiation (the first term on the right), the heat flux withdrawn from the thermoelectrons (the second term), and secondary photons (the third term);
- thermophysical processes: removal of heat flow with evaporated material (fourth term), the material in the liquid phase, if the conditions for its release (fifth term), thermal radiation of the heated surface (sixth term) and condensed atoms, previously evaporated (seventh term);
- plasma chemical processes, realized by the reactions of the laser radiation flux with the evaporated material of the part or adsorbed gases (eighth term); this energy is transmitted by radiation.

The energy transfer is also carried out by the radiation of the laser radiation quantum flux (the last term):

$$-\lambda[T] \frac{\partial T(x, y, z, t)}{\partial x} = F_{\ell, r} - F_{m\beta} - F_{\beta\phi} - F_{ucl} - F_m - \sigma \varepsilon T^4(0, y, z, t) \pm \pm F_{\kappa o n o} + F_{n\kappa} + \sigma \varepsilon_c T_c^4, \quad (30)$$

where σ is the Stefan-Boltzmann constant; ε and ε_+ are the degree of blackness of the target surface and the medium; T_c is the temperature of the medium.

Each term is considered in more detail in [3].

Heat flux density due to the action of laser radiation:

$$F_{su} = \frac{I_i D^2}{4\pi\lambda^2 F^2} P_0 \quad (31)$$

The second term takes into account the removal of heat flow due to the emission of electrons by the heated surface of the target – thermo-electronic emission. The density of the heat flow discharged with electrons,

$$F_{T3} = J_s \phi(T) \quad (32)$$

where $\phi(T)$ is the work of the electron output, eV, at the surface temperature T ; J_s is the density of the emission current, which is determined by the Richardson equation:

$$J_s = (1 - \bar{r}) A \cdot T^2 \cdot \exp(-e\phi_0/kT) \quad (33)$$

Here $\bar{r}$ – the average electron energy reflection coefficient of the potential barrier at the boundary "solid-vacuum", it can reach several percent for pure metals; A – Richardson constant, $A = 4\pi m_e k^2 e/h^3 = 1,204 \cdot 10^6 \text{ A/m}^2(\text{K})^2$; ϕ_0 – the electron work function at $T = 273 \text{ K}$; k – is the Boltzmann constant.

For most metals ϕ lies in the range 4...5 eV.

4 Calculation results and discussion

Calculations of temperature fields under the action of laser radiation flow on the part are carried out on a computer by the finite element method. The process of computer simulation is implemented by dividing the equations by coordinates and computer calculations are carried out separately by coordinates, and then the coordinates of the corresponding temperature points are summed. The calculation of temperature stresses is also related to computer simulation, which allowed to divide each step into two half-steps. In the first half-step of the first step, calculations are carried out without taking into account computer modeling, that is, the temperature of the stresses is equal to zero. In the first half-step temperature is calculated on which the second half-step, taking into account the obtained stresses, and hence the deformation energy are adjusted by computer and the temperature stress values in the next step are already obtained temperatures are starting and in the first half-step of the second step (and subsequent) the previous temperature is the initial and the process is repeated by computer.

As a result of the calculations, the temperature fields in the zone of laser radiation on steels during heating and cooling were determined. Calculations were carried out for a wide range of heat flux densities and times of its action, but only those whose

maximum temperatures during heating and cooling are close to those necessary for obtaining nanostructures (500-1500 K) and the rate of temperature rise exceeds 10^7 K/s are discussed.

So in Fig. 1 the dependences of the maximum temperature in the spot ($r = 0.1$ mm) under the action of heat fluxes with peak density $q = 3 \cdot 10^{10}$ W/m² (1), $q = 2.5 \cdot 10^{10}$ W/m² (2); $q = 2 \cdot 10^{10}$ W/m² (3); $q = 1.5 \cdot 10^{10}$ W/m² (4) and $q = 10^{10}$ W/m² (5) with an action time of 10^{-7} s on steel 20 Fig. 1 (a) and steel 40Cr Fig. 1 (b). It can be seen that for the first two modes, the maximum temperatures on steel are 20 (Fig. 1 a) exceed the temperature of 1500 K, but the time for which they operate $2.1 \cdot 10^{-7}$ s for the first mode and $1.1 \cdot 10^{-7}$ s for the second during this time, the relaxation of the temperature field will not lead to a significant increase in the initial grain size so that not only in 3, 4 and 5 modes should be expected to obtain nanostructures but in the first and second modes. For heating and cooling modes, the rate of temperature rises and fall exceeds 10^7 K, which confirms the probability of nanostructures formation.

For steel 40Cr times in which the temperature exceeds 1500 K for the first mode is $1.7 \cdot 10^{-7}$, and 10^{-7} with, respectively, for the first and second mode is even less than for steel 20 and in this case, the effect of this temperature on grain growth is less significant. The results of similar calculations for steel 40 and 45 are shown in Fig. 2.

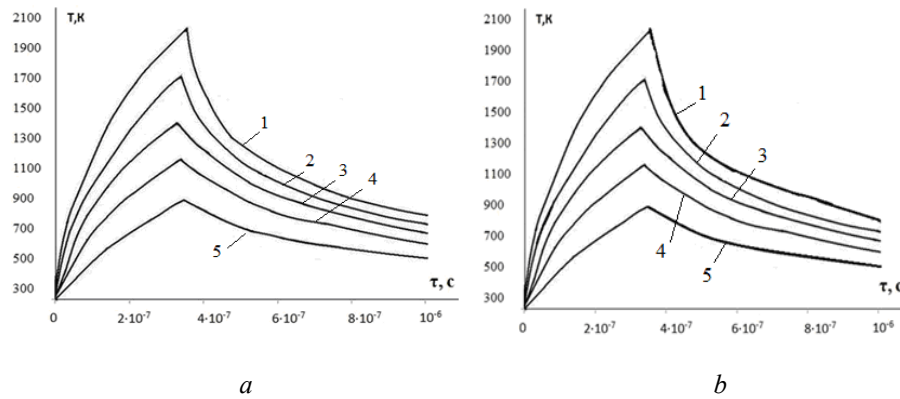


Fig. 1. Temperature at a depth of 1 μ m under the action of laser radiation, the peak heat flux density is $3 \cdot 10^{10}$ W/m² (1), $2.5 \cdot 10^{10}$ W / m² (2), $2 \cdot 10^{10}$ W/m² (3), $1.5 \cdot 10^{10}$ W/m² (4) and 10^{10} W/m² (5) for steel 20 (a) and steel 40Cr (b), operating at the initial time $3 \cdot 10^{-7}$ s

In this case, the values of the maximum temperatures are reduced and amount to values of the order of 1900 K, whereas they exceed 2000 K for steel 20, the nature of the change in the maximum temperature over time has been preserved, the temperature growth rate for these materials exceeds 10^7 K/s, which suggests that for these materials there is a real possibility of obtaining nanostructures.

For high-carbon steels U8 and U12 under the action of thermal $3 \cdot 10^{10}$ W/m² and $2.5 \cdot 10^{10}$ W/m² maximum temperatures also exceed 1500 K and are close to 2000 K, but the time of action of these temperatures does not exceed $2 \cdot 10^{-7}$ s, for these materi-

als for other modes up to the time of action 10^{-6} s temperature mode promotes the formation of nanostructures and the rate of temperature rise exceeds 10^7 K/s, this confirms that for these materials the probability of formation of nanostructures is high (Fig. 3).

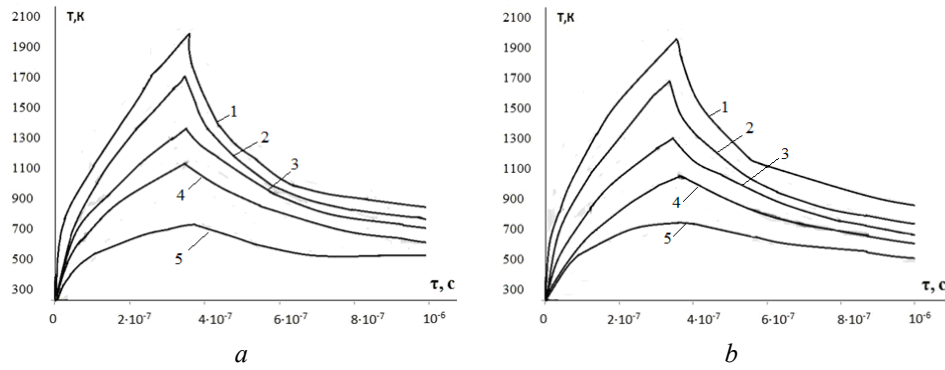


Fig. 2. Temperature at a depth of $1 \mu\text{m}$ under the action of laser radiation, the peak heat flux density is $3 \cdot 10^{10} \text{ W/m}^2$ (1), of $2.5 \cdot 10^{10} \text{ W/m}^2$ (2), $2 \cdot 10^{10} \text{ W/m}^2$ (3), and $1.5 \cdot 10^{10} \text{ W/m}^2$ (4) and 10^{10} W/m^2 (5) of the steel 40 (a) and steel 45 (b), the current at the initial moment of time $3 \cdot 10^{-7} \text{ s}$

When the temperature increases above 1500 K, the grain size increases and it is impossible to obtain nanostructures.

A wide range of steel grades was studied to obtain a common result and to find modes in which it is impossible to obtain nanostructures.

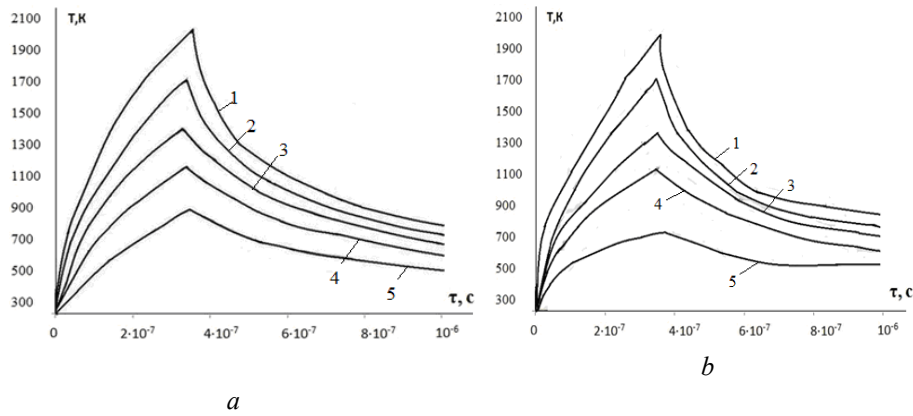


Fig. 3. Temperature at a depth of $1 \mu\text{m}$ under the action of laser radiation, the peak heat flux density is $3 \cdot 10^{10} \text{ W/m}^2$ (1), of $2.5 \cdot 10^{10} \text{ W/m}^2$ (2), $2 \cdot 10^{10} \text{ W/m}^2$ (3), and $1.5 \cdot 10^{10} \text{ W/m}^2$ (4) and 10^{10} W/m^2 (5) on steel U8 (a) U12 (b), the current at the initial moment of time $3 \cdot 10^{-7} \text{ s}$

To estimate the size of the zone of formation of the nanostructures it is necessary to have the radius of the zone of formation of nanostructures built for this space-time pattern of the temperature distribution along the radius and in time for example, steel 40Cr steel under the action of heat flux with density of $3 \cdot 10^{10} \text{ W/m}^2$ when the radius of the spot is 0.1 mm (Fig. 4).

It can be seen that the radius of the temperature decreases compared to the maximum at 200-250 K, which indicates the insignificance of the influence of the zone where the temperature exceeds the permissible (1500 K) on the nature of grain growth, which will be insignificant. All this once again confirms the possibility of obtaining nanostructures in the layer with a depth of about micrometers and a radius of more than 0.1 mm.

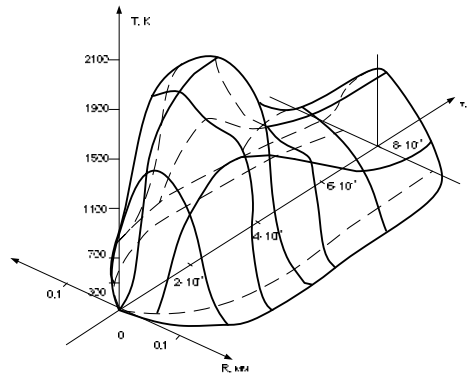


Fig. 4. Spatio-temporal pattern of temperature distribution over a radius in time under the action of laser radiation on steel 40Cr with a heat flux density of $3 \cdot 10^{10} \text{ W/m}^2$ at a spot radius of 0.1 mm

Practically the same or similar temperatures can be obtained at heat flux densities two orders of magnitude less than $q = 3 \cdot 10^8 \text{ W/m}^2$ and its action time $3 \cdot 10^{-3} \text{ s}$ (Fig. 5), but in this case, nanostructures are not formed, but micro and submicro clusters of 1-3 μm and 0.5-0.8 μm in size are formed [4]. all this confirms the thesis that to obtain nanostructures it is necessary to have temperature growth rates greater than 107 K/s, and in this case the maximum temperature growth rate reaches only $6 \cdot 10^6 \text{ K/s}$ and in this case only submicrostructures can be formed.

To select the technological parameters of laser radiation in the preparation of nanostructures on the basis of calculations of temperatures and rates of their increase, the dependences of the critical densities of heat fluxes $q_{\text{cr max}}$ and $q_{\text{cr min}}$ on the time of their action were constructed, at which nanostructures are formed on steel 40Cr (Fig. 6). The same method can be used to obtain dependencies for other steels.

It is seen that the zone of technological parameters for obtaining nanostructures is limited to direct $q_{\text{cr max}}$, $q_{\text{cr min}}$, the zone where the rate of temperature growth is insufficient and the zone where the high probability of thermoelastic destruction is shown to be possible to choose the technological parameters of laser radiation, the heat flux density and the time of its action, ensuring the production of nanostructures in the surface layer.

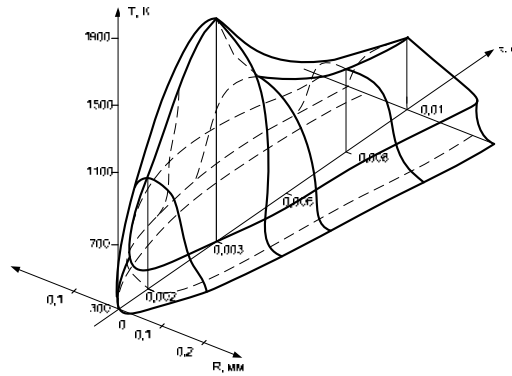


Fig. 5. Spatiotemporal pattern of temperature distribution over radius and time under the action of laser radiation with a heat flux density of $3 \cdot 10^8 \text{ W/m}^2$ at a spot radius of 0.1 mm

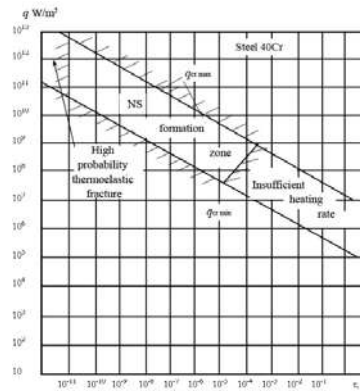


Fig. 6. Dependence of critical values of heat flux densities on the time of their action ensuring the production of nanostructures

5 Conclusions

The conditions for the formation of nanostructures in the surface layer of steels with different carbon content (steel 20, 40, 45, 40Cr, U8 and U12), which determine the required temperature (500-1500 K) and the rate of increase (more than 10^7 K/s). It was found that with an increase in temperature above 1500 K The grain size increases and it is impossible to obtain nanostructures.

The zones of formation of nanostructures depending on the heat flux density on the time of action of the ionizing radiation are determined.

It is shown that it is necessary to take into account the rate of temperature rise and the probability of thermoelastic destruction due to the action of temperature stresses. So with the low speed temperature rise formed micro and submicrostructure that has been confirmed experimentally.

The process of computer simulation is implemented by dividing the equations by coordinates and computer calculations are carried out separately by coordinates, and then the coordinates of the corresponding temperature points are summed. The calculation of temperature stresses is also related to computer simulation, which allowed to divide each step into two half-steps. In the first half-step of the first step, calculations are carried out without taking into account computer modeling, that is, the temperature of the stresses is equal to zero. In the first half-step temperature is calculated on which the second half-step, taking into account the obtained stresses, and hence the deformation energy are adjusted by computer and the temperature stress values in the next step are already obtained temperatures are starting and in the first half-step of the second step (and subsequent) the previous temperature is the initial and the process is repeated by computer. All this confirms that without computer modeling the process could not be realized.

References

1. Boles, M. A., Ling, D., Hyeon, T., and Talapin, D. V. The surface science of nanocrystals. *Nature materials*, vol. 15(2), p. 141 (2016).
2. Norhasri, M. M., Hamidah, M. S., and Fadzil, A. M. Applications of using nano material in concrete: A review. *Construction and Building Materials*, vol. 133, pp. 91-97 (2017).
3. Kostyuk G. I. *Nanotechnology: Theory, Experiment, Technology, Perspectives*. Academy of Sciences and Innovation, p. 648 (2012).
4. Kostyuk G. I. *Nanostructures and Nanocoatings: Perspectives and Reality*. University of Kharkov, p. 406 (2009).
5. Kostyuk G. I. *Scientific Foundations of the Creation of Modern Technologies*. University of Kharkov, p. 552 (2008).
6. Reshetnyak E. N. Synthesis of hardening nanostructured coatings. *Questions of Atomic Science and Technology*, Vol. 2, pp. 119-130 (2008).
7. Liu, Y., Wei, J., Liu, Y., Bai, X., Shi, P., Mao, S., Dkhil, B. Phase transition, leakage conduction mechanism evolution and enhanced ferroelectric properties in multiferroic Mn-doped BiFeO₃ thin films. *Journal of Materials Science: Materials in Electronics*, Vol. 27(3), pp. 3095-3102 (2016).
8. Jung, W. B., Cho, S. Y., Suh, B. L., Yoo, H. W., Jeon, H. J., Kim, J., and Jung, H. T. Polyelemental Nanolithography via Plasma Ion Bombardment: From Fabrication to Superior H₂ Sensing Application. *Advanced Materials*, 1805343 (2018).
9. Fortov, V. E., Ivlev, A. V., Khrapak, S. A., Khrapak, A. G., and Morfill, G. E. Complex (dusty) plasmas: Current status, open issues, perspectives. *Physics reports*, vol. 421(1-2), pp. 1-103 (2005).
10. Xiao, J., Liu, P., Wang, C. X., and Yang, G. W. External field-assisted laser ablation in liquid: An efficient strategy for nanocrystal synthesis and nanostructure assembly. *Progress in Materials Science*, vol. 87, pp. 140-220 (2017).
11. Lu, K. Stabilizing nanostructures in metals using grain and twin boundary architectures. *Nature Reviews Materials*, vol. 1(5), 16019 (2016).

Computer Modeling of Forming Surface by Sequential Wire Electrodischarge-Electrochemical Machining

Oleksandr Plakhotnyi^[0000-0002-5045-6132], Oleksii Timchenko^[0000-0001-8637-8749],
Vasyl Osypenko^[0000-0002-2166-7766],

Cherkasy State Technological University, Shevchenka blvd., 460, Cherkasy,
18006, Ukraine
o.plakhotny@chdtu.edu.ua, alec.timchenko@gmail.com,
osip5906@gmail.com

Abstract. The mathematical principles of computer modeling of the surface's formation process using the combined sequential technology of electrodischarge cutting and electrochemical machining by a wire electrode are developed. The data of computer modeling of complex surface formation processes using this technology, their analysis and the results of experimental verification are presented in this article. It is proved that the use of computer modeling methods allows to improve the surface forming accuracy and the quality of the parts' surface, which are manufactured by this modern promising technology.

Keywords: surface shape accuracy, wire electrical discharge machining, electrochemical machining, EDM/ECM combined process, wire-electrode deflection, ruled surface formation

1 Introduction

Wire electrical discharge machining (WEDM) has become one of the most popular processes for producing precise geometries in hard materials, such as those used in the tooling industry. Since it is recognized as a precision process, optimization of different aspects related to dimensional accuracy is a classic research topic [1].

For WEDM the problem is the reproduction of a given ruled surface by a non-rigid wire electrode, which is bent due to the forces caused by spark discharges. A particularly acute problem of surface forming error occurs when thick work-pieces are cut by a wire-cut EDM, the distance between the upper and lower guide is so wide that the wire is easily bent.

The researchers have proposed many effective strategies, including a hybrid computer-integrated system for the improvement of the corner cutting accuracy [2-4]. At the end of the processing of one corner's side plane using different methods, the wire electrode is straightened and continues the processing of the second side plane. Unfortunately, such a two-stage strategy is not applicable when continuously cutting in

circular interpolation mode along the circle arc. The bent wire electrode forms a barrel-shaped surface. The problem with barrel-shaped inaccuracy remains relevant, both for traditional WEDM and for high-speed WEDM, for workpieces with a thickness greater than 100 mm [5].

One of the latest cutting technologies, presented by Japanese researchers [6], is “High precision cutting for thick work-pieces”. The essence of the technology lies in the development of an optimum control of the discharge energy, paying attention to the wire bending. It can significantly reduce the wire bending and effectively improve the straightness of thick work-pieces. The straightness of 300 mm SKD11 work-piece is achieved to 7 μ m accuracy by cutting 3 times on the ROBOCUT machine. However, the reduction of the wire deflection in this case is achieved by decreasing the intensity of the discharges, and hence the loss of productivity and increasing of expensive processing time. Additionally, the entire heat affected zone of WEDM in the near-surface layers of a workpiece, which is usually detrimental, is not completely eliminated.

Significantly more promising is the use of the combined technology of sequential electrodischarge and electrochemical machining (ECM) with a wire electrode [7]. After roughing electrodischarge cutting, follow 1–2 passes with electrolyte feed to the processing area. The transverse forces do not act on the wire electrode, it remains straight and as a result of the anode dissolution, the surface of the part is aligned, the heat-affected zone is completely eliminated.

An actual task for the implementation of the proposed combined technology is the creation of mathematical methods for calculation support and modeling of the formation of ruled surfaces by a wire electrode. Computer simulation dramatically reduces the amount of experimental work and the accuracy degree can be acceptable for many applications.

The aim of the work is to develop and carry out testing of computer modeling methods to improve the accuracy of surface forming and surface quality of parts manufactured using the latest technology of combined electrodischarge and electrochemical machining with a wire electrode.

2 Experimental equipment, materials

Experimental studies of forming processes by electroerosive wire cutting and subsequent electrochemical machining with a wire electrode were carried out on an electrodischarge cutting machine SELD-02 (Ukraine) with computer numerical control (CNC) systems. The rough cutting mode of the generator GKI 300-200 A was used. Pulses' characteristics: the nominal pulse frequency 22 kHz, pulse duration 3 μ s, current amplitude at load 0.1 Ohm 170 A. The longitudinal tension of the wire varied within 6 - 8 N, the pressure of the working fluid was $5 \cdot 10^5$ Pa. A 0.20 mm Cobra Cut B (AGIE, Switzerland) wire electrode of a solid brass CuZn37 was used. Pulse electrochemical treatment was carried out by the same electrode in a 1 M NaCl aqueous solution in a potentiostatic regime. Workpieces 40, 68, 100 mm height made of tool stamping steel X12 Φ 1 (analogue DIN X155CrVMo12-1).

Shape deviations of the fabricated surface were studied using the Coordinate Measuring Machine 3D CMM AXIOM CNC ABERLINK. The limit of permissible error is $\pm (1.8 + 0.4L/100) \mu\text{m}$, where L , mm is the measured length. The software allows you to set the value of the deviations from the cylindrical shape of the obtained cloud of points. Surface microgeometry was studied using a TIME 3221 profilometer and DataView TIME3R Series software.

3 Methods

3.1 Mathematical principles of computer modeling of WEDM surface forming

For two-coordinate electroerosive cutting, the wire electrode is flexed due to the distributed transverse load caused by spark discharges along the entire height of the workpiece. The computational scheme for determining the shape of the wire electrode (Fig.1) is a special case of a mathematical model of the wire electrode shape in four-coordinate electroerosive machining given in [8]. The experimental-computational method for determining the transverse load $q(z)$ is presented there, depending on the cutting regimes and the material being processed from the class of tool steels and solid alloys.

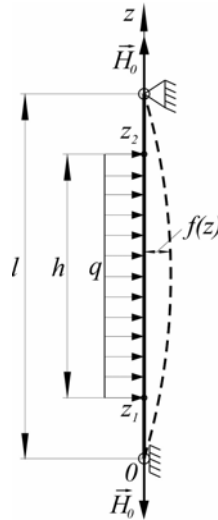


Fig. 1. The calculation scheme of wire electrode deflections: l – the distance between the nodes of the wire fixations; $h = z_2 - z_1$ – height of the machined part; $\vec{H}_0$ – longitudinal tension; $q = \text{const}$ – distributed transverse load

For our case, deflection magnitude of a wire electrode in an arbitrary section ($0 < z < l$) is described by the formula:

$$f(z) = \frac{q}{2H_0} \left[\frac{z}{l}(l-z_1)^2 - \frac{z}{l}(l-z_2)^2 - (z-z_1)^2 + (z-z_2)^2 \right], \quad (1)$$

where we state $(z-z_1) = 0$ at $z < z_1$ and $(z-z_2) = 0$ at $z < z_2$.

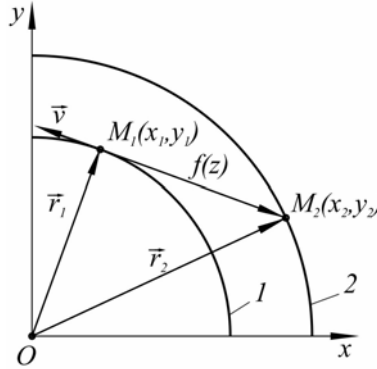


Fig. 2. Movement scheme of the points of the wire electrode: 1 – given movement contour of electrode fixation's points; 2 – trajectory of the point of the electrode with a deflection $f(z)$

To determine the shape of the surface, which is formed by a deflected wire electrode while moving along the arc of a circle in space, consider the scheme of motion of individual points of the electrode (Fig. 2). Let the wire fixation nodes (point $M_1(x_1, y_1)$) moving along the arc of a given radius r_1 (program trajectory). Then point $M_2(x_2, y_2)$, which is at a height z of a wire electrode with a corresponding deflection $f(z)$, describes another arc of the radius r_2 . Vector $\overrightarrow{M_1M_2}$ directed at a tangent to a circle at the point M_1 in the opposite direction of the velocity vector $\vec{v}$, then $\overrightarrow{M_1M_2}$ is perpendicular to the vector $\vec{r}_1$. Vector length $\overrightarrow{M_1M_2}$ is equal to $f(z)$. Then

$$r_2(z) = \sqrt{r_1^2 + f^2(z)}. \quad (2)$$

Or, in the form that is convenient for simulation, we obtain expressions for the coordinates of the point M_2 depending on the position in height z ($z_1 < z < z_2$):

$$x_2(z) = f(z) \cdot \frac{y_1}{r_1} + x_1, \quad y_2(z) = y_1 - f(z) \cdot \frac{x_1}{r_1}. \quad (3)$$

Thus, according to the established form of the wire electrode and taking into account the correction for half the cut width, which for WEDM consists of the wire electrode radius and the value of the interelectrode gap (IEG), the shape of the linear surface obtained as a result of the electrode movement along the arc section of the given trajectory is built.

3.2 Mathematical principles of computer modeling of wire ECM surface forming

The rate of anode dissolution, and hence the change of shape, depends on the density of the technological current on the surface of the part. For an electrochemical cell with the following configuration of electrodes, a cylindrical cathode – a cylindrical anode – the law of distribution must be established of the electric field strength in the IEG and the current density on the anode surface. The calculation scheme is presented in Fig.3. Using the method of conformal mappings of the complex variable function theory [9], we obtain the expressions for the intensity of the electric field:

$$E_x = A \left(\frac{x}{x^2 + (y-a)^2} - \frac{x}{x^2 + (y+a)^2} \right), \quad E_y = A \left(\frac{y+a}{x^2 + (y+a)^2} - \frac{y-a}{x^2 + (y-a)^2} \right),$$

$$E = \sqrt{E_x^2 + E_y^2}, \quad (4)$$

where $a = \sqrt{b^2 + 2bR}$, $b = \frac{\delta(\delta + 2r_0)}{2(\delta + r_0 + R)}$, $A = \frac{\varphi_1 - \varphi_0}{\ln p_1}$, $p_1 = \frac{a+b}{a-b} \cdot \frac{a+\delta-b}{a-\delta+b}$,

here R – radius of the cylindrical surface of the anode, r_0 – radius of the cathode, δ – value of the interelectrode gap, φ_0 , φ_1 – potentials on the cathode and the anode, respectively.

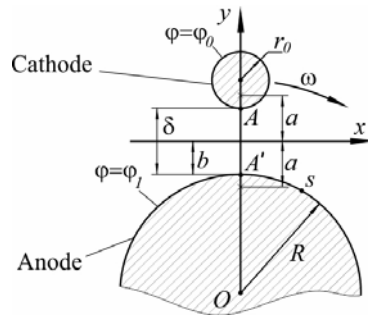


Fig. 3. Calculation scheme for determining the current density on an anode's surface for a wire ECM

For small IEGs, typical for wire ECM, a simplified formula can be used:

$$E \approx A \frac{2a}{s^2 + a^2}, \quad (5)$$

where s – arc coordinate, which is calculated from point A' (Fig.3).

The maximum value of the electric field strength:

$$E_{\max} = E \Big|_{s=0} = \frac{2(\varphi_1 - \varphi_0)}{a \ln p_1}. \quad (6)$$

The maximum value of current density at the anode's surface at the point A' :

$$j_{\max} = \frac{2\kappa(\varphi_1 - \varphi_0)}{a \ln p_1}, \quad (7)$$

where κ – specific electrolyte conductivity.

It is important to determine the thickness of the dissolved surface layer to simulate surface formation. To do this, consider the rotational motion of the wire cathode around the anode with the angular velocity ω (Fig. 3). The linear velocity of the cathode's movement in a given trajectory is determined by the expression:

$$v = \omega \cdot (R + \delta + r_0). \quad (8)$$

Then the intensity of the electric field on the anode's surface will take the form:

$$E = A \frac{2a}{(s - R \cdot \omega \cdot t)^2 + a^2}. \quad (9)$$

By Faraday's law, the depth of dissolution at the point A' during time dt :

$$d\Delta = K_v \eta j \Big|_{s=0} dt, \quad (10)$$

where K_v – coefficient of electrochemical dissolution of anode material, η – current ratio efficiency, j – current density.

After integrating (10) in the corresponding boundaries, we find the thickness of the dissolved layer in one processing run:

$$\int_0^{\Delta_m} d\Delta = 4K_v \eta \kappa A a \int_0^{t_k} \frac{dt}{(R \cdot \omega \cdot t)^2 + a^2}, \quad (11)$$

where $t_k = \pi/\omega$. Or

$$\Delta_m = 4K_v \eta \kappa A \frac{1}{R\omega} \arctg\left(\frac{\pi R}{a}\right). \quad (12)$$

For potentiostatic pulse regime of ECM at $\varphi_1 = 0$, $\varphi_0 = -U$, U – voltage, we finally get:

$$\Delta_m = 4K_v \eta \kappa \frac{U}{R \cdot \omega \cdot \ln(p_1)} \frac{t_i}{t_n} \arctg\left(\frac{\pi R}{a}\right), \quad (13)$$

where t_i – pulse on-time, t_n – period of pulses.

After the electroerosive cutting on the arc of radius R with the deformed wire electrode, the resulting surface will have a convexity in the middle of the workpiece along the z axis (Fig. 4). In the case of the subsequent ECM using nondeformed wire electrode, minimal and maximal IEGs will be δ_1 and δ_2 , respectively. From the formulas' analysis (4), (7), (13) it follows that the intensity of the electric field, the current density, and, as a result, the intensity of the anode dissolution increases with the decrease of IEG. Therefore, in the middle of the workpiece material removal will be bigger and the process will tend to leveling the surface in height. That is, when wire moves in a circle with radius $R + \delta + r_0$ there will be removed thicknesses Δ_{m1} , Δ_{m2} , considering $\Delta_{m1} > \Delta_{m2}$, as $\delta_1 < \delta_2$.

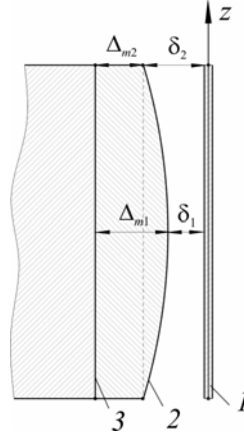


Fig. 4. Surface leveling using ECM after WEDM: 1 – wire electrode; 2 – distorted profile of the workpiece's surface; 3 – leveled generating line of workpiece's surface

A convex surface will become flat if equality is satisfied:

$$\delta_1 + \Delta_{m1} = \delta_2 + \Delta_{m2}. \quad (14)$$

We find the thickness of the dissolved layer in accordance with (13), then formula (14) will take form:

$$\delta_1 + \frac{A_1}{v} = \delta_2 + \frac{A_2}{v}, \quad (15)$$

$$\text{where } A_1 = 4K_v \eta \kappa \frac{U}{\ln(p_1)} \frac{t_i}{t_n} \operatorname{arctg}\left(\frac{\pi R}{a_1}\right), \quad A_2 = 4K_v \eta \kappa \frac{U}{\ln(p_2)} \frac{t_i}{t_n} \operatorname{arctg}\left(\frac{\pi R}{a_2}\right).$$

From (15) we find the wire cathode velocity in a given trajectory for leveling the surface in the process of anodic dissolution:

$$v = \frac{A_2 - A_1}{\delta_1 - \delta_2}. \quad (16)$$

4 Results and discussion

Based on the mathematical dependences obtained, a series of computational experiments on computer simulation of surface forming processes using the WEDM and ECM with a wire electrode with the experimental verification of individual results was performed using the Mathcad package. Computational experiments have shown that the wire electrode deflection is significantly increased with increasing distance between fixation points – electrode guides. The parameter l is included in (1) at the power 2. This explains the emergence of the unacceptable surface formation error problem in the processing of thick parts.

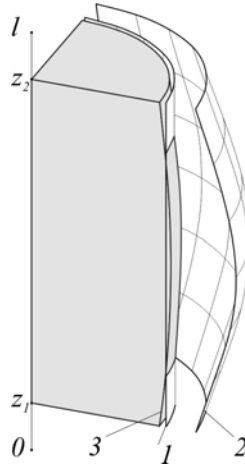


Fig. 5. The result of the surface formation by the combined technology: 1 – given cylindrical surface; 2 – convex surface after WEDM; 3 – the surface is formed as a result of ECM

Computer modeling of shaping of ruled surfaces (Fig. 5) using WEDM by deformed wire electrode during cutting in circular interpolation regime revealed a significant dependence of the error from the trajectory arc radius. The determinant factor is the relationship between the thickness of the workpiece and the magnitude of the arc radius – the generating line of the cylindrical surface.

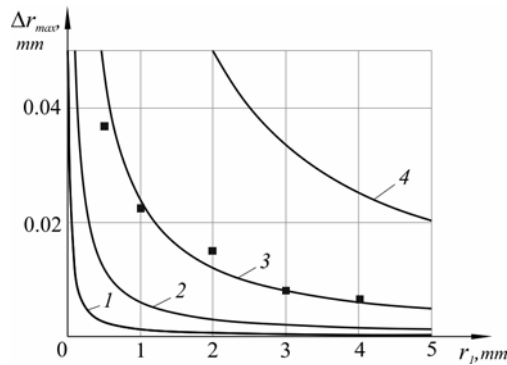


Fig. 6. Dependence of the maximum error of the surface forming of cylindrical surface Δr_{max} from the arc radius of program trajectory r_l for different thicknesses of the workpiece: 1 – 40 mm, 2 – 68 mm, 3 – 100 mm, 4 – 150 mm, ■ – experimentally measured error values

Fig. 6 shows the graphs of the dependence of the maximum errors magnitude of the cylindrical surfaces formation on the arc radius of the program trajectory, on which the wire electrode with a deflection moves in the process of WEDM. Using computer simulation, it was found that when cutting cylindrical parts from a 40 mm thick workpiece, the error of surface reproduction due to the wire electrode deflection does not exceed the requirements for the accuracy of cutting for most technological cases (curve 1, Fig.6). When cutting parts from a 68 mm thick workpiece, the forma-

tion error of the cylinders with a radius less than 0.5 mm becomes significant (curve 2). Curve 3 shows that surface formation error, due to the wire electrode deflection, will exceed 20 microns when cutting a cylindrical part with a radius less than 1 mm in diameter. With an increase in the thickness of the workpiece, the formation error will increase dramatically. The theoretically calculated curve 4 for the thickness of the workpiece of 150 mm, shows, that for radii of 5 mm and less, error becomes unacceptably large.

The following simulation is performed for the aforementioned workpiece materials and WEDM cutting regime. Experimental verification of computer simulation with the coordinate-measuring machine control of the convexity of the obtained cylindrical surfaces, confirmed the satisfactory accuracy of the model. The maximum measured difference between the results of the computational and physical experiment does not exceed 8% (Fig.6).

For other WEDM technological regimes, it is advisable to use the method of determining the power factors influencing the wire electrode described in [8] and to conduct a similar simulation according to the proposed method.

The task of the next stage of computer simulation of the cylindrical surface aligning process using the wire ECM technology is based on a reasonable choice of the amplitude-time parameters of the pulsed current and the velocity of the electrodes' movement along the trajectory. Theoretical calculations, simulations and particular experimental studies have shown that already at the beginning, anodic dissolution predominantly aligns the difference between the convex part in the middle of the height of the cylindrical surface and the less convex regions from above and below (Δ_{m1} and Δ_{m2} respectively, Fig. 4). Then, with the time of dissolution, the difference between Δ_{m1} and Δ_{m2} decreases, and more intense removal of the thickness of the entire surface layer becomes more predominant than the surface leveling in terms of the height of the workpiece. Indeed, from formulas (4), (13) it can be seen that the dissolution rate depends on the intensity of the electric field, which is higher on the humps. Over time, the humps become smaller and the unevenness of the distribution of tension is leveled. In practice, this means that it is not necessary to continue the process until the full alignment of the surface, but it is advisable to stop when achieving the required precision of reproduction of geometric parameters (Fig.5). This once again proves the importance of computer simulation in the design process of the combined technology. Its use allows to predict the magnitude of the surface formation error obtained after WEDM and, accordingly, the amplitude-time parameters of the pulsed current and the velocity of the electrode movement along the trajectory to reduce deviation level to the acceptable result using subsequent ECM with a wire electrode.

Increasing the intensity of the predominant dissolution of humps is possible by reducing the gap between the wire electrode and the surface of the part. The electric field strength concentration increase becomes more noticeable on the humps compared with the caverns. However, setting the IEG of less than 0.1 mm is not allowed due to disruption of a uniform electrolyte flow throughout the height of the part. The uneven and insufficient renewal of the electrolyte in the interelectrode space leads to

a decrease in the current efficiency (η in formulas (10-13), (15)), which significantly influences surface shaping and, accordingly, the accuracy of its modeling.

In addition to increasing accuracy, due to electrochemical machining with a wire electrode, surface quality is also improved. The heat affected zone, obtained due to spark discharge, is completely removed. Surface roughness is reduced. Experimentally obtained decrease in roughness from Ra 2.7 – 3.2 μm after WEDM to Ra 0.65 – 1.2 μm after the ECM stage. It should be noted that the possibilities of improving surface properties by finishing electrochemical machining, in particular, reducing the surface roughness, are not limited to the results obtained in this paper. By choosing the amplitude-time characteristics of pulses, the electrolyte composition, modern technologies of electrochemical surface polishing achieve even more significant results [10,11].

5 Conclusion

1. There are created mathematical principles of computer simulation of the process of surface forming using combined WEDM and ECM with a wire electrode.
2. The implemented complex of computational experiments allowed broadening the idea of complex surface forming processes using the technology of a combined sequential WEDM and ECM with a wire electrode. Experimental verification of computer simulation results confirmed the adequacy of the proposed mathematical models.
3. Usage of computer simulation technologies allowed to improve the shaping accuracy and surface quality of the parts manufactured with the combined technology WEDM and ECM with a wire electrode.
4. The obtained results are an element of the scientific basis for the further development of the modern promising technology.

References

1. Sanchez, J.A., Plaza, S., Lopez de Lacalle, L., Lamikiz, A.: Computer simulation of wire-EDM taper-cutting. *International Journal of Computer Integrated Manufacturing*, vol. 19(7), pp. 727-735 (2006) doi:10.1080/09511920600628855
2. Sanchez, J.A., Lopez de Lacalle, L.N., Lamikiz, A.: A computer-aided system for the optimization of the accuracy of the wire electro-discharge machining process. *International Journal of Computer Integrated Manufacturing*, vol. 17(5), pp. 413-420 (2004) doi:10.1080/09511920310001626590
3. Han, F., Zhang, J., Soichiro I.: Corner error simulation of rough cutting in wire EDM. *Precision engineering*, vol. 31(4), pp.331-336 (2007) doi:10.1016/j.precisioneng.2007.01.005
4. Chen, Z., Huang, Y., Zhang, Z., Li, H., Ming, W., Zhang, G.: An analysis and optimization of the geometrical inaccuracy in WEDM rough corner cutting.

- The International Journal of Advanced Manufacturing Technology, vol. 74(5-8), pp.917-929 (2014) doi:10.1007/s00170-014-6002-5
5. Deng, C., Liu, Z., Zhang, M., Ji, Y., Pan, H.: Minimizing drum-shaped inaccuracy in high-speed wire electrical discharge machining after multiple cuts. The International Journal of Advanced Manufacturing Technology, pp. 1-11 (2019) doi:10.1007/s00170-018-3068-5
 6. Takayama, Y., Makino, Y., Niu, Y., Uchida, H.: The Latest Technology of Wire-cut EDM. Procedia CIRP, vol. 42, pp. 623-626 (2016) doi:10.1016/j.procir.2016.02.259
 7. Osipenko, V.I., Plakhotny, A.P., Denisenko, A.Yu.: Improved methodology for calculating the processes of surface anodic dissolution of spark eroded re-cast layer at electrochemical machining with wire electrode. Pratsi Odeskoho politekhnichnoho universytetu, vol. 1(43), pp.55-60 (2014)
 8. Osypenko, V., Plakhotnyi, O.: Determination of force factors influencing surface forming in wire electrical discharge machining. Elektronnaya obrabotka materialov, vol. 54(6), pp.9-18 (2018) (in Russian) doi: 10.5281/zenodo.1968973. The English version of the journal: Surface Engineering and Applied Electrochemistry.
 9. Zhytnykov, V.P., Zaitsev, A.N.: Pulse precision electrochemical machining. Mashynostroenye, Moscow (2008) (in Russian).
 10. Taylor, E.J., Inman, M.: Electrochemical Surface Finishing. The Electrochemical Society Interface, vol. 23(3), pp. 57-61 (2014) doi:10.1149/2.F05143if
 11. Yang, G., Wang, B., Tawfiq, K., Wei, H., Zhou, S., Chen, G.: Electropolishing of surfaces: theory and applications. Surface Engineering, vol. 33(2), pp. 149-166 (2016) doi:10.1080/02670844.2016.1198452

Analysis of echo-pulse images of layered structures. The method of signal under space

Alexander Stepanenko^{1[0000-0002-0267-7440]}, Andrii Oliinyk^{2[0000-0002-6740-6078]},
Ievgen Fedorchenko^{3[0000-0003-1605-8066]}, Viktor Kuzmin^{4[0000-0003-4940-9395]},
Mariia Kuzmina^{5[0000-0003-2101-0630]} and Dmytro Goncharenko^{6[0000-0001-8973-3742]}

^{1,2,3,6} Dept. of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia
69063, Ukraine

^{4,5} Dept. of Social Work Zaporizhzhia National Technical University, Zaporizhzhia 69063,
Ukraine

alex@zntu.edu.ua, olejnikaa@gmail.com,
evg.fedorchenko@gmail.com, kuzmin2v@gmail.com, gdimk99@gmail.com

Abstract. A new approach to the analysis of pulse echo images is considered, which allows to increase the sensitivity of the analysis and provides the possibility of controlled compensation of the effect of multiple reflections. It is based on the use of the properties of eigenvectors of the correlation matrix of the spectral dependence of the signaling superposition that is being analyzed and the decomposition of the orthogonal space formed by the own vectors of the correlation matrix into two orthogonal subspaces: signaling and noise, that is, by searching for such a new coordinate system in which the informational harmonics of the spectral dependencies would be orthogonal within a given frequency interval. The results of experimental verification of the information capabilities of the method are presented. Probability of scientific positions is substantiated by experimental verification of developed models and methods, on examples of images with already known results (drilling of wells, etc.), which testifies the correspondence of theoretical and practical results.

Keywords: analysis of pulse, digital processing, mathematical model of the layered structure, analyzing pulse-echo images.

1 Introduction

The problem of processing echo-impulse images is very acute now, in order to improve their quality during visualization. It consists of analyzing superpositions of unknown pulse signals that make up these images. Methods of analysis, such as the inverse filtering method and the method of caps trial analysis, cannot be used in this area because of the number of constraints that they impose on incoming superpositions. Therefore, the development of such approaches and methods that are free of

existing restrictions and based on the use of a priori information, such as the method of signal subspace in the spectral region, is relevant.

The method of processing ultrasound images, based on the method of signal subspace, aimed at neutralizing the influence of multiple re-rebounds, is resistant to the influence of measuring and structural noise and does not require a priori knowledge of the model of the analyzed structure.

The area of application of the method is to improve the quality of images obtained using ultrasound devices or seismograms.

2 Literature Review

The book [1] contains interesting findings of some state-of-the-art research in the field of signal and image processing. It covers a wide range of signal processing applications involving filtering, encoding, classification, segmentation, clustering, feature extraction, denoising, watermarking, object recognition, reconstruction and fractal analysis. Various types of signals including image, video, speech, non-speech audio, handwritten text, geometric diagram, ECG and EMG signals, MRI, PET and CT scan images, THz signals, solar wind speed signals (SWS) and photoplethysmogram (PPG) signals have been dealt with. It demonstrates how new paradigms of intelligent computing like quantum computing can be applied to process and analyze signals in a most precise and effective manner.

Development of a large number of algorithms for a plurality of arrival directions of signals is simultaneously active in work [2]. Advantages: a large number of methods are considered, including the signal subspace method, a comparative analysis of methods is carried out, the effectiveness of these methods is discussed.

Work [3] provides an authoritative account of the art and science of reflection seismology. It offers a clear explanation of the methods by which artificially created seismic waves are employed in the economic discovery of oil and gas, as well as their growing use in the exploitation of coal reserves and other energy resources.

The book [4] dedicated to the digital processing of seismic data and their interpretation. Benefits: examines procedures for improving data quality; seismic velocity analysis; principles of structural constructions, to extract geologically significant information, information theory is used.

In the work [5] has been given of the surface-related multiple problems by making use of the so-called feedback model. From the resulting equations it has been concluded that the proposed solution does not require any properties of the subsurface. The surface-related multiple removal algorithms have been formulated in terms of a Neumann series and in terms of an iterative equation. The iterative formulation also has the advantage that it can be integrated easily with other multiple removal methods. An algorithm for the removal of internal multiples has been proposed as well.

A systematic review of many new (at that time) methods of spectral analysis of time series are given in the [6] work. Advantages: such methods are described, such as the classical method based on the use of period grams, the Blackman-Tukey method, the autoregression method (maximum entropy), the moving average method, the

maximum likelihood method, the autoregressive method - the moving average, the Prony method and the Pisarenko method; a summary table with a brief description of all the methods considered is given.

In the work [7] describes the methods of processing signals received on the sensor matrix to determine the location of the source. Mathematically, they represent forms of spectral estimation. Practical examples show Cramer-Rao accuracy for a variety of signals. Examples and comparisons with other methods of maximum probability and maximum entropy are included. An example is given of using MUSIC for time series to estimate multiple frequencies.

The work [8] devoted to digital spectral analysis using modern methods of spectral estimation. Such methods provide enhanced resolution with short sampling and the absence of false side lobes and, as a result, they can have a wide range of applications in radar, sonar, speech and music synthesis, image and seismogram processing, and many other areas. Advantages: programs that implement computational procedures for each of the considered methods are given.

The manual [9] a basic course on digital signal processing reviewed. Benefits: outlines the fundamentals of the theory of discrete signals and systems, discusses methods for spectral analysis and filtering of discrete signals, algorithms for the synthesis of discrete filters, the effect of quantization effects and finite computational accuracy on the operation of digital devices, describes modulation methods used to transmit digital information, discuss adaptive filters multi-speed signal processing.

In the work [10] an interference cancellation technique that combines a signal subspace approach with adaptive adjustment of the weight vector is derived. The weight vector is constrained to lie in a signal subspace computed by Eigen decomposition of the covariance matrix of the array outputs. The weight vector is rotated to maximize the output signal-to-interference ratio.

The article [11] reveals the meaning of the methods of the signal subspace, describes their features, application. Signal subspace methods are empirical linear methods for reducing dimension and noise level. The signal subspace is used in radio direction finding using MUSIC.

The objective of this paper [12] is threefold: to provide an extensive review of signal subspace speech enhancement; to derive an upper bound for the performance of these techniques; to present a comprehensive study of the potential of subspace filtering to increase the robustness of automatic speech recognizers against stationary additive noise distortions.

The book [13] outlines the elements of the classical theory of elastic waves in homogeneous and inhomogeneous media - ideal and real, isotropic and anisotropic, gives geological foundations and classification of seismic methods, analyzes the features of surface and linear hodographs of various waves, discusses the issues of modeling seismic fields. The chapters of the book study the techniques, methods and technology of modern fieldwork statistical theory, as well as the organization and economics of seismic exploration.

In work [14], various methods of reconstructing radio images reviewed, processing signals and images of the subsurface area and objects probed by ultra-wideband signals for their detection and interpretation during geophysical and archaeological re-

search are considered; considered the specificity of the propagation of UWB signals in environments with pronounced attenuation and dispersion; describes various approaches to solving inverse problems of subsurface sounding; various organizations of subsurface sounding radar software (RPG) and methods of their use by the operator are given; The main characteristics of modern RPZs, especially the construction of radars, as well as the results of their use for solving a number of scientific and engineering-physical tasks are considered.

The manual [15] outlines the theory and basic methods for processing geophysical data using spectral and correlation and regression analysis, linear filters, a statistical theory of weak signal detection, classification algorithms and pattern recognition methods for complex analysis of geophysical observations.

In [16], spectral, correlation, and cepstral methods for isolating signals from interference, used in the original field and numerical experiments, are considered. Among them: the selection and measurement of signal parameters when the object is located in the light; time compression of a narrowband signal by transforming its spectrum; elimination of signal distortion caused by multipath, using only a distorted signal (blind reverb); method for determining the delay of the pulse, fluctuating in form; wave propagation using M-sequences and others

In [17], the possibility of predicting the spectrum of a seismic signal based on a priori data on the velocity section of the medium being examined is considered. The condition of broadband and uniform spectrum of the signal at the source, linear absorption of the medium and statistical dependence between the decrement of absorption and seismic velocity is accepted.

In a work [18], a theory is presented to account for the structure of seismograms. It is demonstrated mathematically that a sharp seismic disturbance gives rise to a traveling wavelet of shape determined by the nature of the earth's absorption spectrum for elastic waves, and that a seismogram is composed of a succession of these wavelets, generally overlapping but sometimes in the clear.

The monograph [19] contains a presentation of the theory and practice of frequency filtering algorithms, in which the decomposition coefficients of a noisy signal or image on the Fourier basis and wavelet basis are subject to processing. At the same time, much attention is paid to the choice of the parameters of the algorithms, based on the condition of the minimum mean-square filtering error. Fragments of Mathcad math package documents that implement the constructed filtering algorithms are given.

The handbook [20] contains information about the physics and kinematics of seismic waves. The equipment, work methods, processing and interpretation of seismic data are described. Considered the use of seismic prospecting in solving geological problems, the organization and planning of work.

A resource [21] provides complete coverage of signal modeling, optimum filtering, ratios of the acoustic impedances, and adaptive filtering. Many worked examples illustrate particular algorithms and techniques in use in practical settings.

In the work [22], which is based on the analysis of methods for processing echo-pulse images, the existing approaches were found to be flawed and two new methods were proposed. The developed methods and the software system based on them make

it possible to increase the efficiency of visual analysis of medical and seismic ultrasound images. Disadvantages: similar software products for use in this subject area are not described.

The book [23] gives a systematic presentation of the results available in seismic surveys on linear transformations of seismic signals during their digital processing, as well as during the propagation of seismic waves in inhomogeneous media. The focus is on multi-channel and single-channel filtering of seismograms.

In the collection [24], the issues of construction and methods of application of seismic data processing algorithms for complex models of media and wave fields are considered. The results of the study of the processing procedures of the MOV and KMPV data, the continuations of the wave fields, the solution of direct dynamic and kinematic problems are presented. Advantages: great attention is paid to the efficiency and reliability of the algorithms, their stability in relation to various kinds of interference, including interference resulting from the computational structure of the algorithms

The book [25] describes the basics of processing and interpreting seismic data. Seismic exploration is one of the leading geophysical research methods with a variety of goals for the structure, structure and composition of rocks. Petroleum geology is the main and most effective field of application for seismic prospecting. The search and exploration of hydrocarbon deposits is currently difficult to imagine without the participation of seismic exploration.

The book [26] devoted to the use of the MATLAB matrix system in radio engineering calculations and in the simulation of electronic devices and systems. The newest versions of MATLAB with Simulink, Signal Processing Toolbox, Filter Design Toolbox, RF Toolbox and Blockset, Wavelet Toolbox, Control Systems, Sim Power Systems, etc. are described. The integration of MATLAB with modern digital radio measuring instruments and virtual oscillograms in semiactive radar is described.

The fundamental clinical guidelines are prepared by a team of leading ultrasound diagnostics specialists. The book [28] presents sections on ultrasound diagnostic systems, physical principles of ultrasound diagnostics.

In the work [29], methods of ultrasound imaging, widely used in modern medical diagnostic studies and non-destructive testing, are considered. For example, echopulse images of the biological tissue structure obtained in modern medical ultrasound diagnostics using the second harmonic are distinguished by an increased resolution due to the narrowing of the focal region of the second harmonic, lowering the side lobe level and suppressing reverberation noise compared to the fundamental frequency wave.

In the work [30], an algorithm was described for estimating arrival directions reviewed, delays, and Doppler shifts of the frequency of reflected signals in semi-active radar based on the signal subspace and the Esprit approach. Advantages: for semi-active radar, an algorithm for estimating arrival directions, delays and Doppler increments of the frequency of the reflected backlight signals received by a multi-element antenna array is described.

In the work [31], the accumulation effect was analyzed in estimating the correlation matrices formed in solving linear prediction problems using the least squares

method. By analyzing the dependence of the perturbations of the weight coefficients on the perturbations of the elements of the correlation matrix, it was determined that the most optimal approach to increasing the resolution of the autoregressive methods of spectral analysis is to reduce the conditional number of the correlation matrix of the signal subspace. Coherent methods of preliminary data processing and the formation of their matrices are proposed, which provide the best conditionality of the systems of linear prediction equations and a smaller order of the autoregressive model.

In the work [32], algorithms were considered for the joint detection and direction finding of a variety of radio emission sources in a wide viewing band, developed on the basis of methods with frequency selection and on the basis of methods of signal subspaces. Estimates of the azimuth and quadrature components of the signals are presented when overlapping the spectra of real signals.

In the work [34], the importance of solving the problem of determining the direction of arrival of signals from multiple sources is noted. Advantages: describes a new approach to solving the problem under consideration, it is proposed to simulate signal sensors with a linear combination of noisy control vectors, depending on the parameters characterizing the direction of arrival of signals, is described according to iterative procedure and results of numerical simulation.

In the work [35], progress has been made in the development and application of methods that use the signal subspace and make it possible to estimate the arrival directions of a set of simultaneously acting signals. Advantages: a rather simple algorithm is proposed that allows finding the envelopes maximum of the angular spectrum, its description is given and the resolution is estimated, recommendations are made on the effective use of the proposed algorithm.

The analysis of the subject domain suggests that the development of the signal subspace method is a rather topical task. The works in question have the following disadvantages:

- in some papers [1,4,5,7,10,11,13,16,18,20,23,26,28-30,32] some basic methods used in the field of signal processing are not described in the exploration of minerals and seismic data;
- a number of drawbacks of these methods are connected with the use of linear lattices and the growth of errors in the increase of the range of arrival angles [2];
- in work [3] stable methods of easing the regular and irregular interference in seismic exploration are not considered, while the task of suppressing multiple waves is not considered;
- some [6, 24] of the methods considered are currently irrelevant;
- absence [8] comparative table of the efficiency of using the considered methods;
- are not reflected [9] or are superficially affected, questions concerning adaptive filtering, multi-speed sampling, frequency-time analysis and non-linear digital signal processing, the bases of multidimensional signal processing, the important issue of sampling noise, the noise of arithmetic and problems, associated with minimizing these noise and maximizing the dynamic range.
- it is not described [12] how to increase the residual noise level, the risk of deleting important signal information for the internal interface of the detector is reduced.

Finally, it has also not been shown that the subspaces filtering advantageously differs from the well-known method of spectral subtraction.

- New features of display and effective methods of interactive processing of three-dimensional data of subsurface radar sounding are not described [14];
- The requirement [15] of the FCC to limit the lower bandwidth of the UWB signal does not seem appropriate. It may be that in terms of solving specialized telecommunication tasks in the field of telecommunications, this requirement has a definite meaning, but the concept of the U.S.P signal, in general, should not be limited to the scope of technology, which today is called ultra-broadband;
- not considered [17] the problem of noise immunity estimation of absorption measurements or Q-factors according to the data of the GSP;
- Similar software products for use in this subject area have not been described [19, 22, 31];
- no [21] is illuminated by modern software allowing to simulate and process signals, produce filtering, spectrum estimation.
- unfortunately [25], for a number of reasons, it is not possible to display various colored pictures in the textbook, which substantially limits the possibilities of demonstrating all the results of modern seismic technologies;
- the disadvantages [34] of the existing methods for solving this problem are connected either with insufficient resolution or with difficulties encountered in the direction of coherent sources;
- although [32] the solution of the problem of determining the position of the angles spectral peaks, in the theoretical consideration, is not a difficult problem; in the processing of real data, due to the large dispersion of the latter, certain difficulties arise.

Therefore, the development of the signal subspace method is relevant.

3 Research methodology and analysis of echo-pulse images of layered structures

Echo-pulse image analyze is the main source of information in relation to the problems of geophysical prospecting of hydrocarbon reserves. By virtue of great practical importance, this problem has always paid special attention. Most often, a mathematical model of a local earth structure is considered as an analogue of a layered structure, which allows us to represent a separate seismic trace in the form of

$$s(t) = a(t)\nu \sum_{i=1}^M r_i \delta(t - t_i) + n(t), \quad (1)$$

where $a(t)$ is the initial impulse of seismic excitation; r_i - reflection coefficients, characterized by the ratios of the acoustic impedance of the layer boundaries; t_i - time delays characterizing the acoustic thickness of the layers and their combinations; ν - convolution operation; $n(t)$ - measuring and structural noise. Within the model (1), the main indicator of the presence of hydrocarbons is a local change in amplitude (the so-called

“bright spot” model) [1], which, however, may contain false information due to the influence of interference effects caused by the influence of multiple reflections of seismic pulses between layers. This is expressed in the fact that in formula (1) the number of M observed pulses may absolutely not correspond to the number of significant layers of the L analyzed earth structure ($L \leq M$). In this connection, the development of methods for neutralizing the effect of multiple reflections is of particular importance from the point of view of increasing the sensitivity and reliability of visual analysis of echo-pulse images.

The most commonly used approach is based on direct fitting the parameters of the mathematical model of the layered structure to the measured seismic trace data. In this case, the known geometric model of the analyzed area is assumed, and the adjustment of parameters is carried out based on the use of a nonlinear least squares algorithm [2]. This approach is possible only if there is data from exploratory drilling (an extremely expensive operation), which predetermines its practical limitations. Another approach is based on the principles of backscattering theory and reduces to the summation of the Born backscattering series [3]. From a theoretical point of view, this method does not require a priori knowledge of the model of the geological structure, but by its very nature it belongs to the category of incorrect problems of mathematical physics, which, given the presence of significant measurement interference typical of tasks of exploratory geophysics, makes its practical application very problematic.

The purpose of this work is to analysis and development of the new method of neutralizing the influence of multiple reflections, which are resistant to the influence of measuring and structural noise, what will allow to increase the accuracy of determining the boundaries of deposits several times and do not require a priori knowledge of the parameters of the simulated geological structure, which are included in formula (1). As previously, it is assumed that the study area has a layered structure. The essence of the method is the transition to an information basis in which the useful (informative) and interfering subspaces are orthogonal, and the effect of multiple reflections is neutralized by the controlled choice of the dimension of the information subspace.

The basic idea of the proposed approach consists in the transition from the time domain of the initial measurements to the spectral domain of the Fourier transform, so

$$S(f) = \int_0^T s(t) \exp(-j2\pi ft) dt = A(f) \sum_{i=1}^M r_i \exp(-j2\pi t_i f) + N(f),$$

$$f_L \leq f \leq f_T, \quad (2)$$

where $S(f)$ is the spectral characteristic of the seismic trace; $A(f)$ - spectral characteristic of the probe pulse; $N(f)$ - spectral characteristic of noise. From consideration of expression (2) it can be seen that the time delays t_i in the function $S(f)$ are encoded by complex pulsations in the frequency domain (the period is inversely proportional to the value t_i), which makes it possible to use parametric adaptive spectral analysis methods [4] as applied to the tasks of compensating the effect of multiple reflections.

Regarding expression (2), one more remark needs to be made. Seismic measuring systems are low-frequency, their bandwidth is limited by the width of the spectral range $\Delta F = f_T - f_L$, where f_T and f_L , where and are the upper and lower frequencies of the range, respectively. Since the resolution in the time domain is limited by the ΔF magnitude, the use of parametric spectral analysis methods opens up an additional opportunity to increase the temporal resolution (also a very important practical task) of the analysis of pulse-echo images, but this issue will not be considered in this paper.

From a mathematical point of view, the method of suppressing multiple reflections in the spectral region, as noted above, is based on the idea of a signal subspace, first described in [5]. Such a choice is due to the high stability of the method to the influence of measuring noise and structural disturbances and, in particular, to the effect of changing the shape (and, consequently, the spectrum) of the probe pulse with depth.

Since the $S(f)$ function can be assumed to be sufficiently smooth, its influence can be ignored in the framework of the signal subspace method. All processing is carried out on a computer (i.e. in discrete form) and expression (2) can be represented as

$$S(f_k) = \sum_{i=1}^M r_i \exp(-j2\pi t_i f_k) + N(f_k), \quad k = 1, \dots, K, \quad (3)$$

where is the K number of frequencies in the spectral $S(f)$ characteristic. In the vector-matrix form of the expression (3) has the form

$$S = Hr + N, \quad (4)$$

where is the H matrix characterizing the set of complex exponentials responsible for the time delays t_i in the spectral Fourier domain.

It is assumed that the $N(f)$ statistical characteristics of the interference satisfy the conditions:

$$E\{N_i N_j\} = \sigma^2 \delta(i - j); \quad E\{HN\} = 0, \quad (5)$$

where E is the expectation operator.

The main idea of the signal subspace method is based on using the properties of the correlation matrix of the spectral characteristics of the analyzed seismic trace, so

$$R = E\{SS^T\} = Srr^T S^T + \sigma^2 I = SQS^T + \sigma^2 I, \quad (6)$$

where $Q = E\{rr^T\}$ is the correlation matrix characterizing the interrelation of complex harmonics in the spectral region; I - matrix unit.

As a new information orthogonal basis in the signal subspace method, we consider the linear vector space formed by the eigenvectors of the correlation matrix, which are found by solving the eigenvalue problem

$$R\Phi = \Lambda\Phi, \quad (7)$$

where $[\Phi_1 \Phi_2 \dots \Phi_K]$ is the matrix of eigenvectors; Λ - diagonal matrix of eigenvalues, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K$. By assumption, the H matrix is a superposition of complex exponentials, and their number is less than the number of samples of the spectral characteristics, so $K \geq M$. Consequently, the rank of the matrix R is equal N , which allows you to split the diagonal matrix of eigenvalues into two parts:

$$\lambda_i = \lambda_{hi} + \lambda_{ni}, i = 1, \dots, M; \lambda_i = \lambda_{ni}, i = M + 1, \dots, K, \quad (8)$$

where λ_{hi} and λ_{ni} are the eigenvalues corresponding to the SQS^T signal and σ^2 interference parts of the matrix R , respectively. Then, according to expression (8), the entire K dimensional space formed by the eigenvectors can be divided into two subspaces: M - the dimensional signal subspace formed by the first eigenvectors corresponding to the largest proper numbers, and $(K-M)$ dimensional disturbance subspace formed by the remaining eigenvectors and orthogonal to the first, so

$$\Phi = \left\{ \Phi_S^{(M)} | \Phi_N^{(K-M)} \right\}. \quad (9)$$

Taking into account the expression (7), the R matrix can be represented as

$$R = \sum_{i=1}^M \lambda_i \Phi_{si} \Phi_{si}^T + \sum_{i=M+1}^K \lambda_i \Phi_{ni} \Phi_{ni}^T = R_S + R_N, \quad (10)$$

those, the R correlation matrix can also be represented as a sum of two parts: the R_S signal and R_N interference. From expressions (6) - (8) it follows that

$$\Phi_N^T R \Phi_N = \sigma^2; \quad \Phi_N^T R \Phi_N = \Phi_N^T S Q S^T \Phi_N + \sigma^2 \Phi_N^T \Phi_N. \quad (11)$$

Since the eigenvectors are orthonormal, $\Phi_N^T \Phi_N = 1$ it follows from (11) that

$$\Phi_N^T S Q S^T \Phi_N = 0. \quad (12)$$

As by assumption $Q \neq 0$, it follows from expression (12) that

$$\Phi_N^T [s_1 s_2 \dots s_M] = 0. \quad (13)$$

Thus, from the consideration given it follows that, based on dividing the correlation matrix of the spectral characteristic of the recorded seismic trace into signal and interfering components, followed by projecting the spectral characteristic into a sub-domain formed by the eigenvectors of the interfering space, it is possible to control the process of extracting significant harmonics, and thus also time delays in the time domain, which allows for a controlled process of suppressing multiple reflections.

From a practical point of view, instead of the expression (13), focused on the selection of zeros, it is more convenient to use the expression.

$$\frac{I}{s_i^T R_N s_i} \Rightarrow \max; i = 1, \dots, M. \quad (14)$$

The features of seismogram analysis based on the expression (14) should include the following factors:

1. the analysis is carried out in the spectral region, but the result is obtained in the time domain, since the procedure of finding the spectrum from the spectrum;
2. the poles of expression (14) are visualized, and not the physical amplitudes as in the original $s(t)$ dependence;
3. since the actual number of significant layers is unknown, the selection of the reflections is carried out by varying the choice of values M in model (1);
4. from a mathematical point of view, the synthesis of a new seismogram based on expression (14) is a substantially non-linear process, which opens up additional possibilities for increasing the reliability and sensitivity of visual analysis of echo-pulse images.

4 Experimental results

The study of the informational capabilities of the new method was carried out using the example of a seismic image of a section of a gas condensate field, shown in fig.1a, the visual analysis of which is difficult due to the presence of numerous reflections that form a characteristic interference pattern. At the Fig.1b and Fig.1c are shown the results of applying the new method for the dimension of the signal subspace $M = 10$ and $M = 20$, respectively. Comparison of the results shows that the new method makes it possible to clearly distinguish the upper and lower boundaries of the gas-condensate collector. This result is especially clear when using pseudo color coding. The application of this approach to the data in Fig.1a did not give any positive results, whereas for the method of the signal subspace the result was excellent, which is apparently due to the essentially non-linear character of information display in this method.

The analysis of the results at the Fig.1 shows that the basic information stood out when the dimension of the signal subspace is 10. A further increase leads to the selection of subtle additional details (layers), but from the point of view of identifying the boundaries of the field, this no longer provides significantly new information.

Additional information about the features of the signal subspace method is provided by analyzing graphic image slices. So at the Fig.2a, a graph of the 100th column of the image in Fig.1a is shown, which clearly shows the effect of multiple reflections, which have a particularly negative effect on probing low-contrast areas with close values of impedances of the layer boundaries. Fig.2b presents the real part of the spectral characteristics of this path, which is the initial “raw material” for the signal

subspace method. Fig.3a and 3b show the display in the time domain of the results of the signal subspace method ($M = 10$ and $M = 20$, respectively) for the data presented at the Fig.2a and 2b.

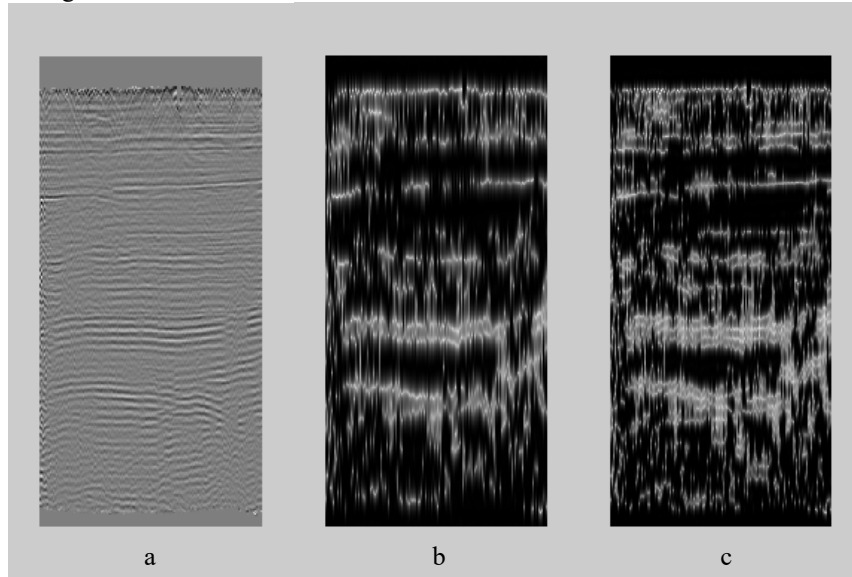


Fig.1. Echo-pulse images: a - the original image of the gas condensate field; b - the result of the signal subspace method (accounting for the 10 most significant signal components); c - the same, but for 20 signal components

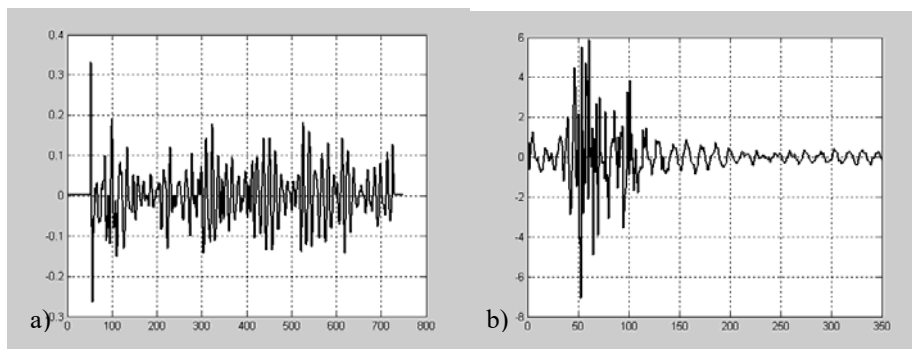


Fig.2. Graphic display of the results: a - the 100th column of the original image in Fig.1a; b - the real part of the spectral characteristics of the signal "a", which is the initial information for the method of the signal subspace.

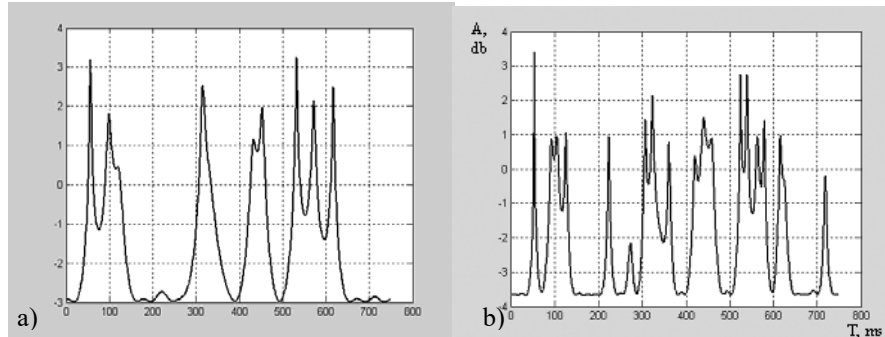


Fig.3. Results of the signal subspace method as applied to the data in Fig.2a: a is the dimension of the signal subspace $M = 10$; b - $M = 20$

Comparing the data at Fig.2 and Fig. 3, it follows that the signal subspace method allows one to effectively suppress multiple re-reflections by choosing the optimal value of the dimension of the signal subspace. The elimination of the effect of multiple reflections makes it possible to increase the reliability of visual analysis (compare Fig.2a and Fig.3a). An additional factor contributing to the informational significance of the method is, as noted above, the higher resolution of the analysis in the time domain, since the peaks in Fig.3 are sharper compared to the initial data at the Fig.2a. Comparing the data at Fig.3a, and Fig.3b clearly shows that increasing the dimension of the signal subspace leads to an increase in the sensitivity of the selection of image details, which can play an important role in the tasks of analyzing low-contrast seismic pulse-echo images.

5 Findings

- 1 The method of analyzing pulse-echo images of layered structures is a typical example of computer vision, because the visual analysis is based on the display of the poles of expression (14), and not on the physical amplitudes of the reflected signals.
- 2 A fundamentally important feature of the considered method of the signal subspace is the need to move from the time domain of the initial measurements to the spectral Fourier domain of the subsequent analysis.
- 3 The results of experimental studies show that the signal subspace method effectively neutralizes the effect of the negative effect of multiple reflections, which seriously complicates the visual analysis of pulse echo-images within the framework of traditional approaches.
- 4 Compensation of the effect of multiple reflections is carried out by optimizing the choice of the dimension of the signal subspace.
- 5 The new method has significant potential for further development and allows generalization to the field of other methods of echo-pulse diagnostics (ultrasound medical introspection, ultrasonic non-destructive testing).
- 6 The effectiveness of the developed method (technically achieved):

- - reduction of re-reflections on the synthesized image in comparison with the original echo-pulse image - up to 3x (vector subtraction method was used);
- - a decrease in noise - due to the optimal selection of the size of the signal subspace, which approximate the signal.

Acknowledgment

The work was performed as part of the project “Methods and means of decision-making for data processing in intellectual recognition systems” (number of state registration 0117U003920) of Zaporizhzhia National Technical University.

References

1. Boden M.: A guide to recurrent neural networks and backpropagation. Halmstad University, pp. 1–10 (2001).
2. Whitley D. Genetic Algorithms and Neural Networks. Genetic Algorithms in Engineering and Computer Science, pp. 203-216, (1995).
3. Van Tuc, N.: Approximation contexts in addressing graph data structures. University of Wollongong Thesis Collection, 30–55 (2015).
4. Erofeeva, V.A.: Review of the theory of data mining based on neural networks [Obzor teorii intellektualnogo analiza dannykh na baze neyronnykh setey. Stokhasticheskaya optimizatsiya v informatike], 11 (3), pp. 3–17 (2015).
5. Skillicorn, D.: Taxonomy for computer architectures. Computer (21), pp. 46-57 (1988). doi: 10.1109/2.86786.
6. Schmidhuber, J., Wierstra, D., Gagliolo, M., Gomez, F.: Training Recurrent Networks by Evolino. Neural computation. vol. 19(3), 757–779 pp. (2007). doi: 10.1162/neco.2007.19.3.757.
7. Parkinson's Disease Classification Data Set, <https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification>
8. Bland, J.M., Altman, D.G.: Statistics notes: measurement error. BMJ. 312 (7047): 1654 (1996). doi:10.1136/bmj.312.7047.1654
9. Viola, P., Jones M.J. Rapid Object Detection using a Boosted Cascade of Simple Features. Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 1–9 (2001).
10. Viola, P. Robust real-time face detection. International Journal of Computer Vision, 57, 2, 137–154 pp. (2004).
11. Peppas M. V., Bell D., Komar T., Xiao W. Urban traffic flow analysis based on deep learning car detection from CCTV image series. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 499–506 pp. (2018).
12. Ren S., He K., Girshick R., Sun J.. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39, 6, 1137–1149pp. (2017)
13. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., Berg A. SSD: Single Shot MultiBox Detector. Computer Vision ECCV, 21–37 (2016).
14. A CNN Vehicle Recognition Algorithm based on Reinforcement Learning Error and Error-prone Samples, <http://iopscience.iop.org/article/10.1088/1755-1315/153/3/032052/pdf>.

15. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Analysis of inrush currents of the unloaded transformer using the circuitfield modelling methods. *Eastern-European Journal of Enterprise Technologies*, vol. 3, № 5 (93), pp. 6-11 (2018). doi: 10.15587/1729-4061.2018.134248
16. Boden M.: A guide to recurrent neural networks and backpropagation. Halmstad University, pp. 1–10 (2001).
17. Yarymbash, D., Yarymbash, S., Kotsur, M., Divchuk, T.: Enhancing the effectiveness of calculation of parameters for short circuit of threephase transformers using field simulation methods. *Eastern-European Journal of Enterprise Technologies*, vol. 4, № 5 (94), pp. 22-28 (2018). doi: 10.15587/1729-4061.2018.140236.
18. Whitley D. Genetic Algorithms and Neural Networks. *Genetic Algorithms in Engineering and Computer Science*, pp. 203-216, (1995).
19. Yarymbash, D., Kotsur, M., Subbotin, S., Oliinyk, A.: A New Simulation Approach of the Electromagnetic Fields in Electrical Machines. *IEEE: The International Conference on Information and Digital Technologies*, July 5th - 7th, Zilina, Slovakia, 2017, Catalog Number CFP17CDT-USB, pp. 452-457, (2017). DOI: 10.1109/DT.2017.8024332
20. Oliinyk, A., Subbotin, S., Lovkin, V., Leoshchenko, S., Zaiko, T.: Development of the indicator set of the features informativeness estimation for recognition and diagnostic model synthesis. *14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET 2018)*, pp. 903–908, (2018). doi: 10.1109/TCSET.2018.8336342.
21. Leoshchenko, S., Oliinyk, A., Subbotin, S., Zaiko, T.: Methods of semantic proximity extraction between the lexical units in infocommunication systems. *4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T 2017)*, 7–13 pp. (2017). doi: 10.1109/INFOCOMMST.2017.8246137.
22. Leoshchenko, S., Oliinyk, A., Subbotin, S., Gorobii, N., Zaiko, T.: Synthesis of artificial neural networks using a modified genetic algorithm. *Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018)*, pp. 1-13 (2018). dblp key: conf/iddm/PerovaBSKR18
23. Stepanenko, O., Oliinyk, A., Deineha, L., Zaiko, T.: Development of the method for decomposition of superpositions of unknown pulsed signals using the secondorder adaptive spectral analysis. *EasternEuropean Journal of Enterprise Technologies*, vol. 92, issue 2/9, pp. 48–54, (2018). doi: 10.15587/1729-4061.2018.126578
24. Alsayaydeh, J.A., Shkaruplyo, V., Hamid, M.S., Skrupsky, S., Oliinyk, A.: Stratified Model of the Internet of Things Infrastructure, *Journal of Engineering and Applied Science*, vol. 13, issue 20, pp. 8634-8638, (2018). doi: 10.3923/jeasci.2018.8634.8638.
25. Shkaruplyo, V., Skrupsky, S., Oliinyk, A., Kolpakova T.: Development of stratified approach to software defined networks simulation. *EasternEuropean Journal of Enterprise Technologies*, vol. 89, issue 5/9, pp. 67–73 (2017). doi: 10.15587/1729-4061.2017.110142
26. Kolpakova, T., Oliinyk, A., Lovkin, V.: Improved method of group decision making in expert systems based on competitive agents selection. *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Institute of Electrical and Electronics Engineers, pp. 939–943, Kyiv (2017). doi: 10.1109/UKRCON.2017.8100388
27. Oliinyk A., Fedorchenko I., Stepanenko A., Rud M., Goncharenko D.: Evolutionary method for solving the traveling salesman problem. *Problems of Infocommunications. Science and Technology : 5th International Scientific-Practical Conference PICST2018*, Kharkiv, 9–12 October 2018, Kharkiv, Kharkiv National University of Radioelectronics, pp. 331-339 (2018). doi: 10.1109/INFOCOMMST.2018.8632033

28. Kotsur, M., Yarymbash, D., Kotsur, I., Bezverkhnia, Yu.: Speed Synchronization Methods of the Energy-Efficient Electric Drive System for Induction Motors. IEEE: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) 2018, pp. 304-307, Lviv-Slavske, Ukraine (2018). doi:10.1109/TCSET.2018.8336208
29. Van Tuc, N.: Approximation contexts in addressing graph data structures. University of Wollongong Thesis Collection, 30–55 (2015).
30. Schmidhuber, J., Wierstra, D., Gagliolo, M., Gomez, F.: Training Recurrent Networks by Evolino. *Neural computation*. vol. 19(3), 757–779 pp. (2007). doi: 10.1162/neco.2007.19.3.757.
31. Berstein, Y., et al.: Parametric nonlinear discrete optimization over well-described sets and matroid intersections. *Math. Program.*, vol. 124, no. 1–2, pp. 233–253 (2010).
32. Yemelichev, V.A., et al.: Polytopes, graphs and optimisation. Cambridge University Press, Cambridge (1984).
33. Pichugina, O.S., Yakovlev, S.V.: Functional and analytic representations of the general permutations. *Eastern-European Journal of Enterprise Technologies*, vol. 1, no. 4, pp. 27-38 (2016) doi:10.15587/1729-4061.2016.58550.
34. Stoyan, Y.G., et al.: Method of balancing rotating discretely distributed masses. *Energo-mashinostroenie*, no. 2, pp. 4–5 (1982).

Trajectories Planning and Simulation of a Backhoe Manipulator Movement

Alexander Gurko^[0000-0001-9905-8584], Igor Kyrychenko^[0000-0002-2128-3500],
Aleksandr Yaryzhko^[0000-0001-6398-8472]

Kharkiv National Automobile and Highway University,
Ya. Mudrogo str., 25, Kharkiv, 61002, Ukraine
gurko@khadi.kharkov.ua, igk160450@gmail.com, yaryzhko@gmail.com

Abstract. An excavator is a highly widespread heavy-duty construction machine. The working equipment of the excavator can be thought of as a hydraulic manipulator mounted on a vehicle. To carry out the workflow effectively an operator is to move the bucket teeth along the given path with certain velocity and acceleration under the restrictions imposed by the kinematic parameters of the manipulator and the configuration of the working area. It requires very high skills of the operator. To accurately move the bucket teeth along the desired path the automatic control system can be used. This paper focuses on the excavator manipulator trajectories automatic planning and control. Firstly, the manipulator joints trajectories were obtained to perform digging and levelling operations. Then, a virtual model of the excavator equipment was built based on the MATLAB Multibody to simulate working operations. Finally, digital PID controllers were designed to improve the accuracy of the bucket teeth movement along the path required. As the example, the attached backhoe equipment of the excavator Boreks 2201 is considered.

Keywords: excavator manipulator, MATLAB Multibody model, kinematics, trajectory planning, movement control

1 Introduction

Despite the apparent simplicity of technological processes of road construction, it faces a number of difficulties due to the necessity of increasing the amount of performed works, improving their quality and reducing their cost, which can only be achieved by automation. One of the main reasons that hinder road construction automation is the limited data on the dynamic properties of objects and technological processes. The lack of this information leads to the fact that the hardware and software of road machine control systems are still developed without considering their interaction with each other and with the physical world. And then, after the control system has been developed, it is checked on the models and the impact of various uncertainties is eliminated by the special methods of adjustment. This process is expensive and labour-intensive, and it becomes practically impossible with the complication of the machines [1].

The situation began to change with development of cyber-physical systems (CPS), which are integrations of computation with physical processes [2, 3]. CPS has become an outstanding foundation for creating advanced industrial systems and applications by the integration of innovative features through the Internet of Things (IoT) and Web of Things (WoB) to enable the connection of real physical objects to computing and communication aids [4]. Hence, it is not a coincidence that CPS is one of the main technologies of the fourth industrial revolution, known as Industry 4.0 [4, 5].

As a part of CPS, a model-oriented approach to the design of automatic control systems for complex objects and processes was developed. According to this approach, a simulation model of the control object is created instead of a physical prototype, and it interacts with the physical world using real sensors and actuators [6-8]. It allows passing from simulation models to hybrid ones, which combine both the models of complex objects and the real physical devices, which provides the possibility to verify the concepts and technical solutions without creating physical prototypes and to reduce the number of full-scale tests. However, designing CPS requires more reliable models of physical processes occurring in control systems. The performance of CPS depends on how the model relates to reality.

The most effective software that allows building quite realistic models of complex technical systems, including road machines, are visual modelling tools that combine a graphical form of describing the model and a representation of the results as a 2D or 3D animation. One of the most powerful software providing such possibilities is MATLAB Simscape Multibody™. A valuable advantage of the Simscape Multibody is a CAD translator, which allows creating dynamic models of machines based on their solid models in CAD software like Autodesk Inventor, SolidWorks or Pro / Engineer. It allows relatively simple creating workable models since in CAD it is much easier to establish the correct connections between parts and nodes of a machine. Thus, Simscape Multibody is a perfect tool to investigate such a complex technical system as an excavator working equipment, whereas MATLAB affords an opportunity for connecting a Simscape model with the physical world.

2 Formal problem statement

The aim of the paper is to develop and investigate a control system, which plans the movements of an excavator manipulator in order to move the bucket teeth along a given path and to realize these movements. At this stage of the research, only the kinematics of the excavator manipulator is simulated without the connection of the model with the physical world. As an example, the attached working equipment of the backhoe Boreks 2201 is considered.

3 Literature review

Due to the mentioned advantages, Simscape Multibody is widely used for modelling construction and road machines, in particular, excavators, which are the most common among such machines [9-15]. For instance, the model of an excavator manipula-

tor was built in [10] using Simscape environment to analyse the spatial motion of the working equipment during the workflow.

In [11] the virtual model of the telescopic robotic excavator was built based on the SimMechanics software to simulate levelling and digging operation and to design the excavator manipulator motion controller. The validation of the model was verified experimentally. For this purpose, the hydraulic cylinder displacements on the model and on the real excavator were tracked, and then the x-axis and z-axis coordinate values of the bucket tip was calculated. The experimental results showed good consistency with simulation results. Hence, the SimMechanics model is feasible to study the real operation process.

The Simscape models of an excavator manipulator were also described in [12-14]. These models were used for the excavator boom, arm and bucket hydraulic actuators dynamics analysis and control to minimize vibrations of the excavator bucket during the excavation works. In [15] in order to investigate the skilled operator behaviour, an excavator model with the help of SimMechanics and SimHydraulics was obtained.

This paper continues to research the behaviour of an excavator working equipment with the help of Simscape Multibody models.

4 Model of the excavator manipulator

To study the kinematics and dynamics of the excavator, as well as to test the effectiveness of various control algorithms, a 3D model of Boreks 2201 working equipment was built. For this purpose, the model in Autodesk Inventor 2016 was originally built. Then, the sizes, mass and tensor of the moments of inertia of each element were imported as the mass and inertia of the solid body into Simscape Multibody. The general view of the model is given in Fig. 1.

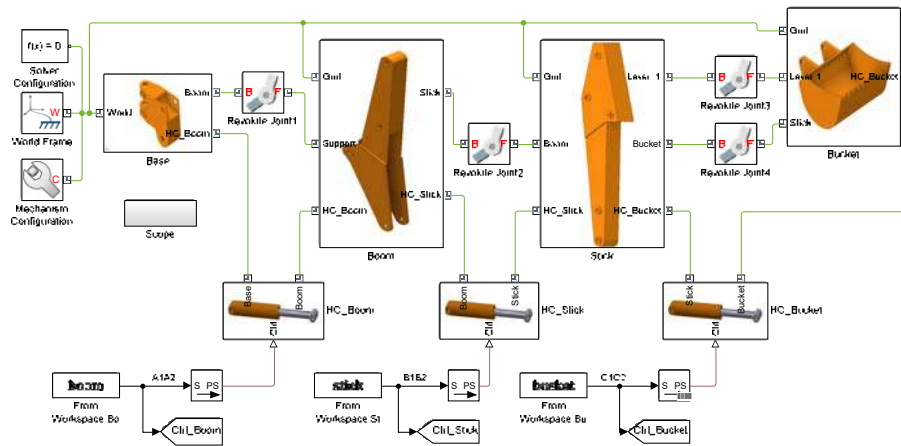


Fig. 1. General view of the excavator manipulator model

A visual representation of the movement of the mechanical part of the excavator manipulator model can be obtained using the built-in visualization function of SimScape (Fig. 2), which allows detecting errors in the model much faster.

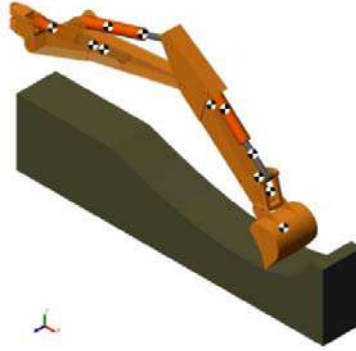


Fig. 2. A visual representation of the excavator manipulator model during digging

5 Relation of the joint angles and the displacement of the hydraulic cylinders rods

The ‘boom-stick-bucket’ system of the excavator can be considered as an open kinematic chain, which consists of three serial links connected by rotary joints and driven by hydraulic actuators (Fig. 3). Thus, either joint angles θ_j ($j = 2,3,4$) or displacements L_j of the hydraulic actuators rods can be taken as the generalised coordinates. Here we use the joint angles θ_j ($j = 2,3,4$) as the generalized coordinates since it is more convenient for planning the manipulator trajectories. It should be noted that we do not consider the swing angle θ_1 in this paper, since during digging operation θ_1 remains constant.

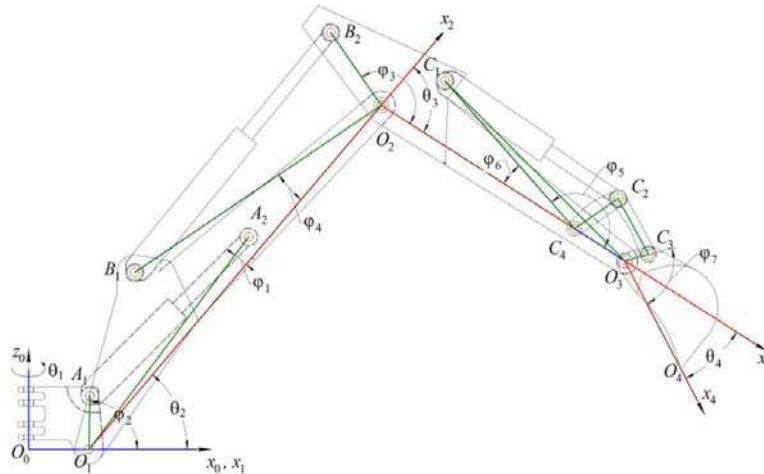


Fig. 3. Coordinate frames of the excavator Boreks 2201 manipulator

Since the change of the joint coordinates θ_j is carried out by displacing L_j the rods of the corresponding hydraulic cylinders, we will find the relations between these variables.

For the boom lengths O_1A_1 , O_1A_2 as well as angles φ_1 and φ_2 (Fig. 3) are constant, and their values depend on the features of a particular excavator model. At the same time length A_1A_2 and angle $\angle A_1O_1A_2$ have the variable values. From triangle $A_1O_1A_2$ (Fig. 3) length A_1A_2 is equal to:

$$A_1A_2 = \sqrt{O_1A_1^2 + O_1A_2^2 - 2O_1A_1 \cdot O_1A_2 \cos(\angle A_1O_1A_2)}, \quad (1)$$

where angle $\angle A_1O_1A_2$ is

$$\angle A_1O_1A_2 = \varphi_2 - \varphi_1 - \theta_2. \quad (2)$$

Similarly, for the stick, the values of angles φ_3 , φ_4 and length O_1A_1 are constant, and length B_1B_2 with angle $\angle B_1O_2B_2$ are variable (Fig. 3). Length B_1B_2 of the stick hydro cylinder can be found as:

$$B_1B_2 = \sqrt{O_2B_1^2 + O_2B_2^2 - 2O_2B_1 \cdot O_2B_2 \cos(\angle B_1O_2B_2)}, \quad (3)$$

where the value of angle $\angle B_1O_2B_2$ is determined from by following expression:

$$\angle B_1O_2B_2 = \theta_3 - \varphi_3 - \varphi_4 + \pi. \quad (4)$$

The relation between length C_1C_2 of the bucket hydro cylinder and the joint angle θ_4 is much more complicated. On the basis of the cosine theorem from triangle $C_3O_3C_4$ (Fig. 3), we find C_3C_4 :

$$C_3C_4 = \sqrt{O_3C_3^2 + O_3C_4^2 - 2O_3C_3 \cdot O_3C_4 \cdot \cos(\angle C_3O_3C_4)}. \quad (5)$$

Knowing this side angle $\angle O_3C_3C_4$ is:

$$\angle O_3C_3C_4 = \arccos\left(\frac{O_3C_3^2 + C_3C_4^2 - O_3C_4^2}{2O_3C_3 \cdot C_3C_4}\right). \quad (6)$$

From triangle $C_2C_3C_4$ angle $\angle C_2C_3C_4$ can be calculated as:

$$\angle C_2C_3C_4 = \arccos\left(\frac{C_3C_4^2 + C_2C_3^2 - C_2C_4^2}{2C_3C_4 \cdot C_2C_3}\right). \quad (7)$$

Then we can find angle $\angle C_2C_3O_3$. However, there is some difficulty here. At certain value θ_b of the joint angle θ_4 , point C_4 lies on the straight line C_1O_3 (Fig. 3). Therefore for the angles $\theta_4 < -\theta_b$ and $\theta_4 \geq \theta_b$ some formulas are different, e.g. if $\theta_4 < \theta_b$:

$$\angle C_3 O_3 C_4 = \pi + \theta_4 + \varphi_6 + \varphi_7, \quad (8)$$

and

$$\angle C_2 C_3 O_3 = \angle C_2 C_3 C_4 - \angle O_3 C_3 C_4. \quad (9)$$

Otherwise, i.e. if $\theta_4 \geq \theta_b$:

$$\angle C_3 O_3 C_4 = \pi - \theta_4 - \varphi_6 - \varphi_7, \quad (10)$$

and

$$\angle C_2 C_3 O_3 = \angle C_2 C_3 C_4 + \angle O_3 C_3 C_4. \quad (11)$$

This fact needs to be taken into account when determining the required length $C_1 C_2$ of the bucket hydro cylinder. From triangle $C_2 C_3 O_3$ length $C_2 O_3$ is:

$$C_2 O_3 = \sqrt{C_3 O_3^2 + C_2 C_3^2 - 2 C_3 O_3 \cdot C_2 C_3 \cdot \cos(\angle C_2 C_3 O_3)}. \quad (12)$$

From triangle $C_2 O_3 C_4$ angle $\angle C_2 C_4 O_3$ is:

$$\angle C_2 C_4 O_3 = \arccos \left(\frac{C_2 C_4^2 + O_3 C_4^2 - C_2 O_3^2}{2 C_2 C_4 \cdot O_3 C_4} \right), \quad (13)$$

Angle $\angle O_2 C_4 O_3$ can be calculated from triangle $O_2 C_4 O_3$:

$$\angle O_2 C_4 O_3 = \arccos \left(\frac{O_2 C_4^2 + O_3 C_4^2 - O_2 O_3^2}{2 O_2 C_4 \cdot O_3 C_4} \right), \quad (14)$$

From triangle $C_1 C_4 O_2$ we find angle $\angle C_1 C_4 O_2$:

$$\angle C_1 C_4 O_2 = \arccos \left(\frac{C_1 C_4^2 + O_2 C_4^2 - C_1 O_2^2}{2 C_1 C_4 \cdot O_2 C_4} \right), \quad (15)$$

Then

$$\angle C_1 C_4 C_2 = 2\pi - \angle C_1 C_4 O_2 - \angle C_2 C_4 O_3 - \angle O_2 C_4 O_3. \quad (16)$$

Knowing this angle, we find $C_1 C_2$:

$$C_1 C_2 = \sqrt{C_1 C_4^2 + C_2 C_4^2 - 2 C_1 C_4 \cdot C_2 C_4 \cdot \cos(\angle C_1 C_4 C_2)}. \quad (17)$$

Thus, the resulting equations (1) – (17) establish relations between the values of the joint angles θ_j of the excavator manipulator and lengths $A_1 A_2$, $B_1 B_2$ and $C_1 C_2$ of the corresponding actuators. Knowing the extreme positions of the rods of these cylinders, it is easy to obtain the expressions for their relative displacements L_2 , L_3 and L_4 .

6 Kinematic control of the excavator manipulator

In general terms, a robotic excavator works as follows. Input information is a desirable path of the edge of the bucket teeth, which can be determined either by an operator or by an on-board computer. Then, by one of the methods given in this section the manipulator trajectories $\theta_j(t)$ planning is performed, which are further converted into the desired displacements of hydro cylinder rods by formulas (1) – (17). Later, the task of realizing these movements under dynamic loads is solved.

In this section, we consider the solution of the problem of determining the change of the joint angles $\theta_j(t)$ at a certain time interval $t \in [t_0, t_f]$, that combines the initial and final configuration and satisfies the given velocities and accelerations constraints at the trajectories endpoints.

In robotics, for trajectories planning, high order interpolation polynomials are widely used, or the trajectory of the link is divided into several segments, each of which interpolates with a polynomial of the lower order [16]. The same methods are also used for robotic excavators [17, 18], although various numerical methods of kinematic control of excavator manipulators are also developed [19-21]. The minimal order polynomial, which satisfies the condition of smoothness and takes into account the constraints on position, velocity and acceleration of the link, is a fifth order polynomial:

$$\theta(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5. \quad (18)$$

The first and second order derivatives of (18) are also smooth polynomials:

$$\dot{\theta}(t) = a_1 + 2a_2 t + 3a_3 t^2 + 4a_4 t^3 + 5a_5 t^4, \quad (19)$$

$$\ddot{\theta}(t) = 2a_2 + 6a_3 t + 12a_4 t^2 + 20a_5 t^3. \quad (20)$$

To find the values of coefficients a_k , it is necessary to solve the following system:

$$\begin{pmatrix} \theta_0 \\ \dot{\theta}_0 \\ \ddot{\theta}_0 \\ \theta_f \\ \dot{\theta}_f \\ \ddot{\theta}_f \end{pmatrix} = \begin{pmatrix} 1 & t_0 & t_0^2 & t_0^3 & t_0^4 & t_0^5 \\ 0 & 1 & 2t_0 & 3t_0^2 & 4t_0^3 & 5t_0^4 \\ 0 & 0 & 2 & 6t_0 & 12t_0^2 & 20t_0^3 \\ 1 & t_f & t_f^2 & t_f^3 & t_f^4 & t_f^5 \\ 0 & 1 & 2t_f & 3t_f^2 & 4t_f^3 & 5t_f^4 \\ 0 & 0 & 2 & 6t_f & 12t_f^2 & 20t_f^3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix}. \quad (21)$$

When splitting trajectories into segments, the so-called Linear Segments with Parabolic Blends (LSPB) is mostly used. It provides a trapezoidal profile of the velocity which imposes a constant acceleration in the start phase, a cruise velocity, and a constant deceleration in the arrival phase. The essence of LSPB is the following. The desired trajectory of every link of the manipulator is divided into three parts. The first

part starts from time t_0 to time t_b and is described with a quadratic function. It leads to a gradual increase in velocity. At time t_b , called the blend time, the trajectory changes to a linear function. In the end, at the moment $t_f - t_b$, the trajectory changes again to the quadratic function when the velocity gradually decreases.

In terms of smoothness of accelerations, it is expedient to use the fifth order polynomials for the excavator manipulator trajectories planning, while the LSPB provides a greater speed of work operations execution. In this case, however, the laws of velocities and acceleration change do not satisfy the constraints in smoothness, which leads to jerks of working equipment.

In some cases, for example, at levelling, it is necessary to ensure the movement of the bucket teeth along the straight line in Cartesian space. In this case it is better to plan the trajectories directly in Cartesian space. In such a case, the initial and final points of the path are described by a homogeneous transformation matrix, which establishes a relationship between the bucket coordinate frame and the world coordinate frame. Then the values of the joint angles θ_j corresponding to these points are calculated using the manipulator inverse kinematics. Further, the trajectories between these points are interpolated in the joint space [16].

The described methods have been used for the trajectories planning of the Boreks 2201 manipulator. Two types of the bucket teeth path have been considered (Fig. 7):

- a path in the form of a parabolic line that simulates the movement of the bucket during the digging operation;
- a straight line path, that simulates the levelling operation.

As the boundary conditions, it has been assumed, that velocities v_0, v_f and accelerations a_0, a_f , of the manipulator links at the initial and final points of the path should be zero.

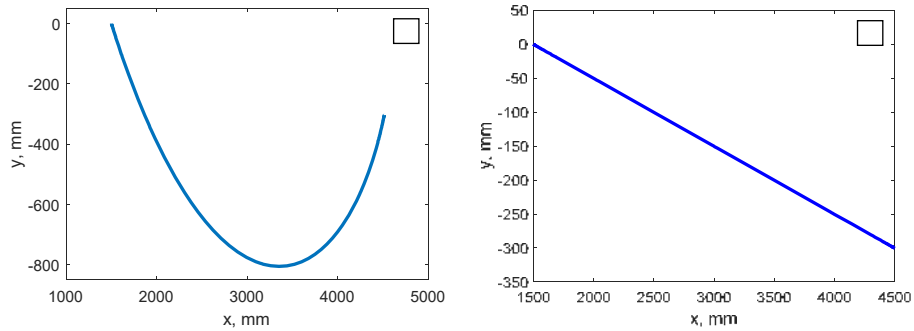


Fig. 4. Bucket teeth desired trajectories for digging (a) and levelling (b) operations

7 Manipulator movement simulation

7.1 Digging simulation

Determination of joint angles $\theta_j(t)$ ($j = 2, 3, 4$) changing and, consequently, the rods relative displacements $L_j(t)$, which provide the displacement of the bucket teeth along

the parabolic line (Fig. 4a), have been carried out according to the equation (21). The actual displacements $L_j(t)$ of the actuators rods obtained at the model (Fig. 1) are shown in Fig. 5, and their velocities $v_j(t)$ and acceleration $a_j(t)$ are shown in Figs. 6 and 7 accordingly. These figures show that in general, the obtained trajectories meet the requirements of smoothness, which minimizes the overloads in the kinematic chain of the excavator manipulator. The maximum velocities of the rods are 0.32 m/s.

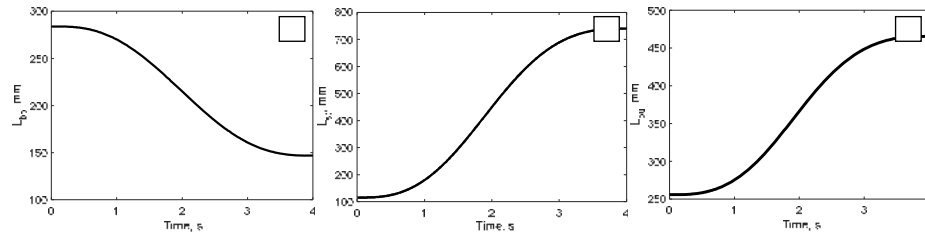


Fig. 5. Estimated displacement of the actuators rods of the boom (a), the stick (b) and the bucket (c) for the parabolic bucket path

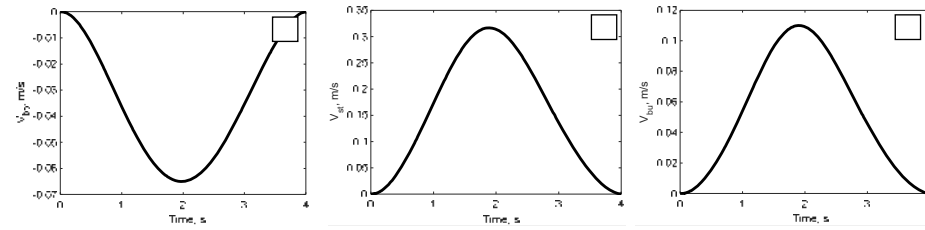


Fig. 6. Velocities of the actuators rods of the boom (a), the stick (b) and the bucket (c)

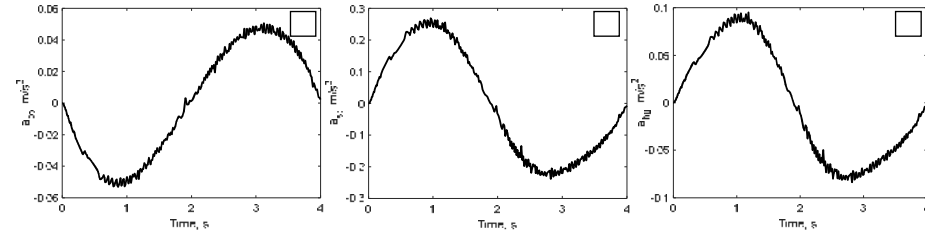


Fig. 7. Accelerations of the actuators rods of the boom (a), the stick (b) and the bucket (c)

Tracking errors of the boom and the bucket actuators rods are insignificant (Fig. 8): 1.3 mm (Fig. 8a) and 2.2 mm (Fig. 8c), respectively. However, the maximum tracking error of the hydraulic cylinder rod of the boom is quite large and modulo greater than 6 mm (Fig. 8b). These errors can be explained by the dynamic properties of the excavator manipulator.

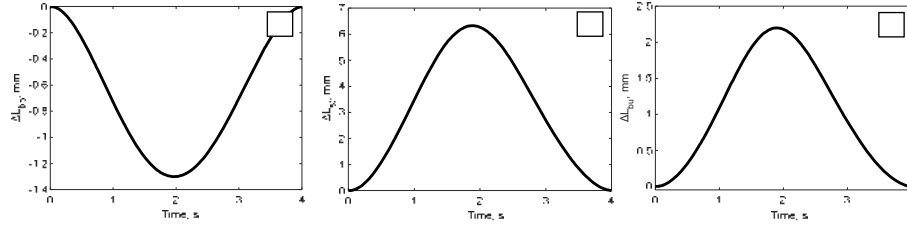


Fig. 8. Tracking errors of the actuators rods of the boom (a), the stick (b) and the bucket (c) for the parabolic bucket path

The presence of these tracking errors leads to some differences in desired and actual paths: the maximum error in the x -direction is 33 mm, and in the y -direction is 14 mm (Fig. 9). These are the acceptable digging errors for real conditions of excavation.

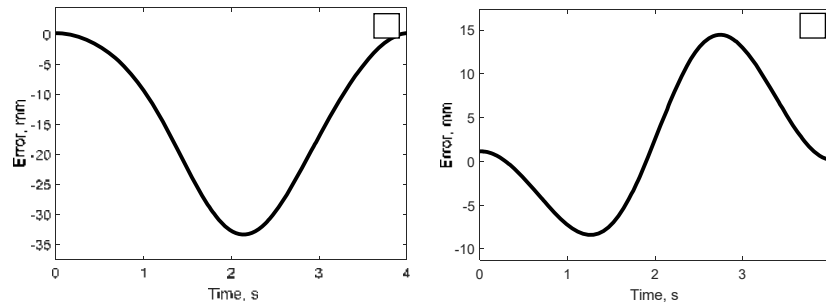


Fig. 9. Digging errors in the x -direction x (a) and in the y -direction (b)

It should be noted that in this paper only the kinematics of the excavator is considered, so dynamic loads are not taken into account. Obviously, in the presence of digging resistance forces, the digging errors will increase.

7.2 Levelling simulation

Planning of the excavator manipulator motion to move the bucket teeth along a straight line is made in Cartesian space. The obtained laws of the hydraulic cylinder rods displacements are shown in Fig. 10.

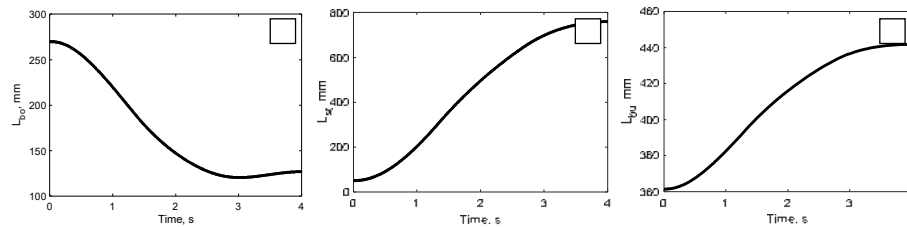


Fig. 10. Displacements of the rods of the boom (a), the stick (b) and the bucket (c) hydraulic cylinders

The velocities and accelerations of the rods are given in Fig. 11 and Fig. 12. The maximum velocity of the rods is 0.33 m/s, which does not exceed the allowed maximum velocity of 0.5 m/s. However, the laws of changing velocities and accelerations are not smooth, so jerks of acceleration can be seen in the graphs.

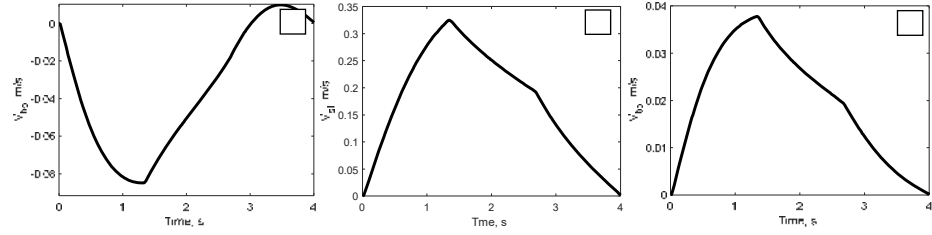


Fig. 11. Velocities of the movements of actuators rods of the boom (a), the stick (b) and the bucket (c) at levelling

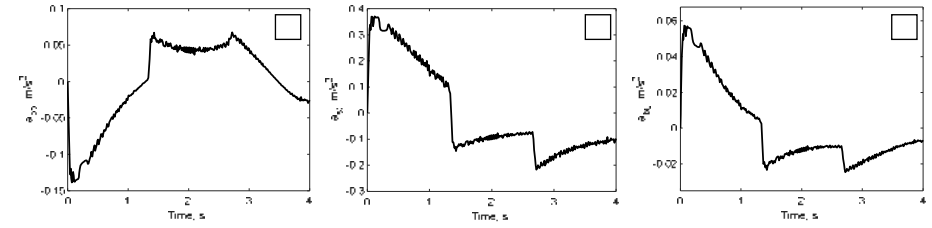


Fig. 12. Accelerations of the movements of actuators rods of the boom (a), the stick (b) and the bucket (c) at levelling

Tracking errors for the rods are shown in Fig.13. From these figures it is evident that, as in the case of the parabolic path, at levelling, the actuator of the stick has the maximum tracking error – about 7 mm.

The errors of the bucket teeth motion along the straight line are illustrated in Fig. 14; it can be seen that the maximum absolute error is 22 mm for the x -axis, and 3.2 mm for the y -axis. It is also worth to note that the straight line path of the bucket teeth is one of the most difficult paths to perform by an operator [22].

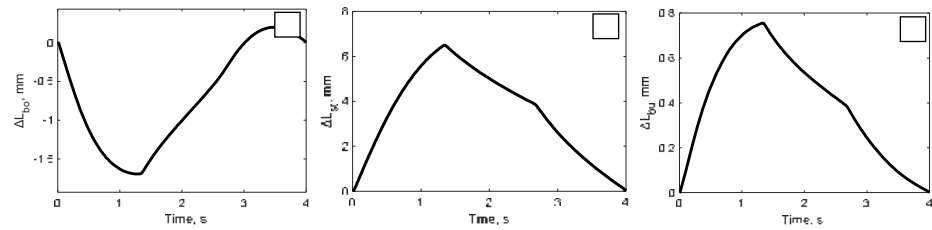


Fig. 13. Tracking errors of the actuators rods of the boom (a), the stick (b) and the bucket (c) for the parabolic bucket path at levelling

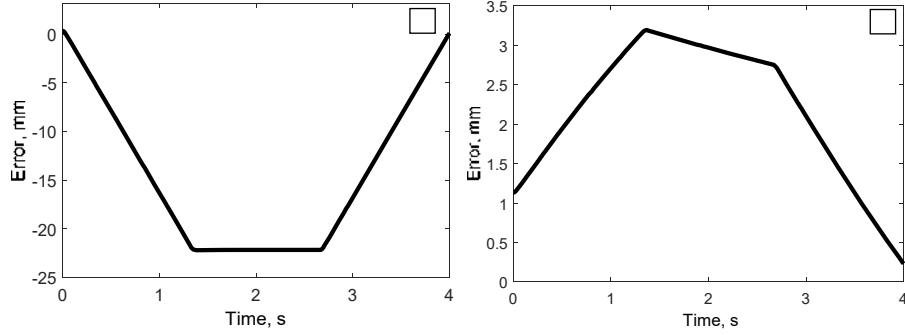


Fig. 14. Levelling errors in the x -direction (a) and in the y -direction (b)

8 Controller design

To improve the quality of the desired trajectories tracking by the excavator manipulator, the digital PID controllers are designed. The transfer function of the controller is:

$$C(z) = k_p + \frac{k_i T_s}{z-1} + \frac{k_d}{1 + \frac{NT_s}{z-1}}, \quad (22)$$

where k_p , k_i and k_d are the controller proportional, integral and derivative gains respectively; N is the filter coefficient and T_s is the sampling time.

A separate controller is used to control each joint. For this, the ‘Controller’ subsystem has been added (Fig. 15) to the model shown in Fig. 1. Tuning of the controllers parameters has been performed by means of Simulink Control Design. The sampling time is chosen $T_s = 5\text{ms}$. The results of the controllers tuning are given in Table 1.

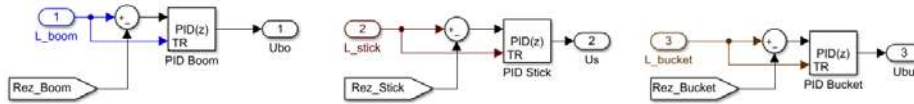


Fig. 15. The subsystem ‘Controller’ of the Boreks 2201 model

The use of PID controllers allowed reducing displacements errors of the hydraulic cylinders rods significantly. For example, the tracking error of the stick hydraulic cylinder rod when moving along the parabolic path decreased from 6 mm to 3 mm, and along the straight line – from 7 mm to 3 mm. It improved the accuracy of the joint angles execution accordingly, and, consequently, the accuracy of passing the bucket teeth along the given path (Table 1).

Table 1. The controllers parameters and path tracking errors

Controller parameters					Path tracking maximum error, mm			
	k_p	k_i	k_d	N	Parabolic Trajectory		Straight Line	
					x	y	x	y
Boom	4.4	159.7	0.015	337	2	4	2	5
Stick	2.7	119.5	0.009	225	7	2	6	2
Bucket	2.2	106.8	0.007	291	15	6	10	2

We note again that the given results are obtained in ideal conditions, that is, in the absence of external forces that appear, for example, when the bucket interacts with the soil. It should be expected that with the presence of these forces, as well as with the changing weight of the bucket during its loading, there will be more significant deviations between the desired and the actual paths. In order to avoid this, it is necessary to implement the appropriate controllers.

9 Conclusion and future work

Improving efficiency of excavators is inseparably linked with implementation of the working equipment automatic control systems. One of the main tasks of such control systems is to plan and execute such movements of the excavator manipulator links, which ensure the movement of the bucket along the given path.

In order to develop such a control system, in this paper the relationship between the position of hydraulic cylinder rods of the excavator manipulator and its joint angles has been determined. The task of the kinematic control of the excavator manipulator has been solved, that is, determination and provision of such laws for changing the angles of the links joint, their velocities and accelerations, which ensure the movement of the bucket teeth along the desired path in Cartesian space in the presence of constraints. When digging is accomplished, trajectories planning is desirable to perform in the joint space (in this paper we have used a fifth-order polynomial approximant), whereas, the bucket movement along a straight line is better to plan in Cartesian space.

In order to investigate the excavator manipulator movement, the 3D model of the Boreks 2201 manipulator was built in the MATLAB Simscape Multibody software. Experiments with the 3D model have shown that hydraulic actuators perform the desired trajectories with some errors due to the influence of mass-inertial parameters. As a result, the quality of the earthworks is decreasing.

The use of digital PID controllers increased the accuracy of the trajectories tracking by the manipulator links and, therefore, improved the precision of the bucket teeth movement along the desired path up to 67% in the case of the parabolic path and up to 33% in the case of the straight line path.

However, the given results are obtained under the ideal conditions, that is, in the absence of external forces that arise, for example, due to the bucket and the soil interaction. More significant deviations between the desired and actual paths should be

expected when these forces are considered, as well as the bucket weight change during its filling. In addition, uncertainties about the manipulator parameters values and the digging resistance forces values could well significantly influence the control system performance. Our future work is related with the study of the excavator manipulator dynamics, taking into account the indicated factors, as well as with a robust controller design to ensure the effective performance of the excavator workflow under the presence of variable and uncertain loads. Furthermore, our future research provides for connecting the model of the excavator manipulator with the physical world by using real-life actuators instead of their models. It should give more reliable data about the control system performance in real conditions.

References

1. Gurko, A. G., Plakhteev, A. P., Plakhteev, P. A.: Accuracy Increase of Dynamic Objects State Estimation by a Complex Matlab-Arduino when Cyberphysical Systems Designing. *Radio Electronics, Computer Science, Control* **1**, 84–91 (2016) (in Russian). doi: 10.15588/1607-3274-2016-1-10
2. Lee, E. A.: Cyber Physical Systems: Design Challenges. In: 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing, pp. 363–369. IEEE Press (2008). doi: 10.1109/ISORC.2008.25
3. Lee, E. A.: CPS Foundations. In: 47th Design Automation Conference, pp. 737–742 (2010). doi: 10.1145/1837274.1837462
4. Lu, Y.: Cyber Physical System (CPS)-Based Industry 4.0: a Survey. *Journal of Industrial Integration and Management* **2** (3), 1750014 (2017). doi: 10.1142/S2424862217500142
5. Lizárraga, M. L., et al.: Time Restriction Aspects in the Modeling of Cyber-Physical Systems for Industry 4.0. *Bulletin of Kharkov National Automobile and Highway University* **83**, 107–115 (2018). doi: 10.30977/BUL.2219-5548.2018.83.0.107
6. Paterno F.: *Model-Based Design and Evaluation of Interactive Applications*. Springer Science & Business Media (1999)
7. Jensen J. C., Chang D. H., Lee E. A.: A Model-Based Design Methodology for Cyber-Physical Systems. In: 7th IEEE International Wireless Communications and Mobile Computing Conference, pp. 1666–1671 (2011). doi: 10.1109/IWCMC.2011.5982785
8. Plakhteyev, A., Perepelitsyn, A., Frolov, V.: Edge Computing for IoT: An Educational Case Study. In: 9th IEEE International Conference on Dependable Systems, Services and Technologies, pp. 130–133 (2018). doi: 10.1109/DESSERT.2018.8409113
9. Xu, J., Yoon, H. S.: A Review on Mechanical and Hydraulic System Modeling of Excavator Manipulator System. *Journal of Construction Engineering* **2016**, 9409370 (2016). doi: 10.1155/2016/9409370.
10. Xu, G., Ma, Z., Lu, F., Hou, P.: Kinematic Analysis of Hydraulic Excavator Working Device Based on DH Method. In: International Conference on Applied Mechanics, Mechanical and Materials Engineering, pp. 8 (2016). doi: 10.12783/dtmse/ammme2016/6857
11. Hongxin, C., Ke, F., Huanliang, L., Jinhua, H.: Virtual Prototype and Experimental Research on Spatial Kinematics of Telescopic Robotic Excavator. *International Journal of Advanced Robotic Systems* **14** (3), 9 (2017). doi: 10.1177/1729881417705305
12. Curduman, L., Nastac, S., Debeleac, C.: On Active Control Of Transitory Regimes Within The Driving System Of A Single Bucket Excavating Equipment For the Bank-Sloping and

Differential Excavation Processes. In: 23rd International Congress on Sound&Vibration, pp. 1–8 (2016)

13. Curduman, L., Nastac, S., Debeleac, C., Modiga, M.: Computational Dynamics of the Rotational Heavy Loads Mastered by Hydrostatical Driving Systems. *Procedia Engineering* **181**, 509–517 (2017). doi: 10.1016/j.proeng.2017.02.427
14. Curduman, L., Debeleac, C., Nastac, S.: On Path Oscillations Analysis of Mechanical Multi-body and Hydrostatical Driving Units Coupled System, *Procedia Engineering* **181**, 518–525 (2017). doi: 10.1016/j.proeng.2017.02.428
15. Du, Y., Dorneich, M. C., Steward, B.: Virtual Operator Modeling Method for Excavator Trenching. *Automation in construction* **70**, 14–25 (2016). doi: 10.1016/j.autcon.2016.06.013
16. Siciliano B., Khatib O. (Eds.): *Springer Handbook of Robotics*. Springer (2016)
17. Gurko, A., Kolobova, I.: Simulation of Excavator Kinematics in Matlab Robotics Toolbox. *Bulletin of Kharkov National Automobile and Highway University* **60**, 59–65 (2013)
18. Gu, J., Ma, X. D., Ni, J. F., Sun, L. N.: Linear and nonlinear control of a robotic excavator. *J. Cent. South Univ.* **19**, 1823–1831 (2012). doi: 10.1007/s11771-012-1215-y
19. Sergiyenko, O. Yu., Hernandez-Balbuena, D., Gurko, A. G., et al.: Optimal Kinematic Control of a Robotic Excavator with Laser TVS Feedback. In: 39th Annual Conference of the IEEE Industrial Electronics Society, pp. 4239–4244. IEEE Press (2013). doi: 10.1109/IECON.2013.6699816
20. Gurko, A. G., Sergiyenko, O. Yu., Hipólito, J. I. N., et al.: Guaranteed Control of a Robotic Excavator During Digging Process. In: 12th International Conference on Informatics in Control, Automation and Robotics, **2**, pp. 52–59. SciTePress (2015).
21. Lee, B., Kim, H. J.: Trajectory generation for an automated excavator. In: 14th International Conference on Control, Automation and Systems, pp. 716–719. IEEE Press (2014). doi: 10.1109/ICCAS.2014.6987872
22. Chang, P. H., Lee, S. J.: A straight-line motion tracking control of hydraulic excavator system. *Mechatronics* **12(1)**, 119–138 (2002). doi: 10.1016/S0957-4158(01)00014-9

Method of Constructing Lexicographic Equivalence for Solving Linear Combinatorial Optimization Problems on Arrangements: Results of Computational Experiment

Tetiana Barbolina ^[0000-0002-4596-7907]

Poltava V.G. Korolenko National Pedagogical University, Ostrogradski str., 2, Poltava, 36003,
Ukraine
tm-b@ukr.net

Abstract. The method of constructing lexicographic equivalence for solving linear mixed combinatorial optimization problems on arrangements is considered. The software that implements algorithms of this method is described. The efficiency of algorithms is analyzed by means of computational experiment.

Keywords: Euclidean combinatorial optimization, problems on arrangements, computational experiments, method of constructing lexicographic equivalence.

1 Introduction

Actual trend of the modern theory of optimization is to study the problems of combinatorial nature. In particular, these problems are examined by [1–7]. Important results have been obtained as a result of immersion of combinatorial sets in Euclidean space and study the properties of such problems. Y.G. Stoyan started the theory of Euclidean combinatorial optimization, his disciples – O.O. Iemets, S.V. Yakovlev, I.V. Grebennik and numerous other representatives of this school carry out research in this sphere. They have studied in their works both properties of Euclidean combinatorial sets immersed in the arithmetical space, and extreme properties of the objective functions, and methods of solving Euclidean combinatorial optimization problems.

This article considers such an important class of Euclidean combinatorial optimization problems as problems on arrangement. The properties of the convex hull of the set of arrangements are explored in [8, 9], in [10, 11] and others — properties of the solution of certain classes of problems on arrangements, a number of results on the solving methods and algorithms of the optimization problems on arrangements in a rather general statement are summarized in [12]. In particular, method of constructing lexicographic equivalence for solving completely linear combinatorial problems on arrangements is substantiated. In [13] this method was extended to mixed combinatorial problems.

The aim of the paper is to describe the software implementation of the algorithms of the method of constructing lexicographic equivalence for solving linear mixed

combinatorial optimization problems on arrangements and to present results of computational experiments.

New scientific result obtained in the paper is investigation of an effectiveness of algorithms of method of constructing lexicographic equivalence. Since theoretical estimates are not obtained, then the effectiveness is analyzed by mean of computational experiments.

2 Formal problem statement

Let's consider necessary definitions and facts. As multiset G we understand set of elements, which can include and similar ones. Any multiset $G = \{g_1, g_2, \dots, g_\eta\}$ can be assigned by its base $S(G)$, i.e. the tuple of all its different elements, and by multiplicity – number of repetition of each element of the base. The tuple of multiplicities in the order that corresponds to the elements of the base is called primary specification and is defined by $[G]$. The set is called the Euclidian combinatorial set, the different elements of which are different ordered k -samples from the multiset $G = \{g_1, g_2, \dots, g_\eta\}$ of the representation

$$(g_{i_1}, g_{i_2}, \dots, g_{i_k}), \quad (1)$$

$g_{i_j} \in G$, $i_j \neq i_t \ \forall i_j, i_t \in J_n$, $\forall j, t \in J_k$ (here and after J_k defines set of k first natural numbers). Examples of Euclidian combinatorial sets are [9] general set of arrangements $E_\eta^k(G)$ — set of all k -samples of the representation from the multiset G , general multiset of permutations $E_\eta(G) = E_\eta^\eta(G)$.

The introduction of the concept of Euclidean combinatorial sets allows highlighting from problems of combinatorial nature the class of problems where the feasible set is Euclidean combinatorial set. In particular, the linear Euclidean problem of combinatorial optimization on arrangements is to find a pair $\langle L(x^*), x^* \rangle$ (consisting of maximum and maximal) such that

$$L(x^*) = \max_{x \in R^n} \sum_{j=1}^n c_j x_j, \quad x^* = \arg \max_{x \in R^n} \sum_{j=1}^n c_j x_j, \quad (2)$$

under the combinatorial condition

$$(x_1, x_2, \dots, x_k) \in E_\eta^k(G) \quad (3)$$

and additional linear constraints

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i \in J_m, \quad (4)$$

where $n \geq k$, $x = (x_1, x_2, \dots, x_k) \in R^k$, $c_j \in R^1 \quad \forall j \in J_k$, $E_\eta^k(G)$ is a general set of arrangements from the elements of multiset $G = \{g_1, g_2, \dots, g_\eta\}$. Variables $x_1, x_2, \dots, x_k$ are combinatorial, variables $x_{k+1}, x_{k+2}, \dots, x_n$ are continuous. The linear Euclidean problem of lexicographic combinatorial optimization on arrangements is to find pair $\langle L(x^*), x^* \rangle$ such that

$$L(x^*) = \operatorname{lexmax}_{x \in R^n} \sum_{j=1}^n c_j x_j, \quad x^* = \arg \operatorname{lexmax}_{x \in R^n} \sum_{j=1}^n c_j x_j \quad (5)$$

under conditions (3)–(4), i.e. $L(x^*)$ is the maximum value of $L(x)$ as the variables range over the set (3), (4) and x^* is lexicographically larger than any other maximal.

3 The method of constructing lexicographic equivalence

One of the approaches to solving problem (3)–(5) is in the partition of a polyhedron into equivalence classes followed by directed search of the obtained classes. For linear completely combinatorial optimization problems on arrangements, this approach (method of constructing lexicographic equivalence) is justified in [12, 14]. The study [13] generalizes the equivalence relation, which is used to partition the polyhedral set into equivalence classes, and proposes algorithms for constructing lexicographic equivalence to solve mixed combinatorial linear problems.

Lexicographic equivalence of points of the space with respect to k -arrangements (the relation λ_k) is used as the equivalence relation in the method of constructing lexicographic equivalence. Elements of the quotient set with respect to the equivalence λ_k are called λ_k -classes. Each element of set $E_\eta^k(G)$ defines separate λ_k -class. Such classes are called combinatorial. Algorithms for constructing lexicographic equivalence involve directed search of λ_k -classes. Let us describe these algorithms (validation of these algorithms for solving mixed combinatorial problems is presented in [13]).

The first algorithm is used to solve a problem of finding a pair (5) under conditions (3), (4) and $L(x) \in A$ where A is a discrete set. This algorithm involves search of λ_k -classes whose representatives satisfy

$$\sum_{j=1}^n c_j x_j = \sigma^h \quad (6)$$

where σ^h ($h = 0, 1, 2, \dots$) are sequential elements of set A .

According to the second algorithm, the direct search of combinatorial λ_k -classes in lexicographically increasing order and lexicographically decreasing order is car-

ried out. Classes whose representatives give the objective function a value smaller than result obtained at previous iterations, are excluded from the search.

The third algorithm is approximate and allows getting the objective function value that differs from the optimum by no more than a predetermined value. It involves search λ_k -classes whose representatives satisfy

$$\tau^h \leq \sum_{j=1}^n c_j x_j \leq \sigma^h \quad (7)$$

where difference $\sigma^h - \tau^h$ ($h = 0, 1, 2, \dots$) decreases.

4 Description of the software product

The proposed algorithms of the method of constructing lexicographic equivalence for solving linear conditional problems of combinatorial optimization on arrangement (3) - (5) were implemented as software and experimentally investigated in solving the problem.

Software was developed in the Turbo Delphi environment. The program allows you to generate a series of problems with the given parameters, to solve problems by selected algorithms, to save the results of problem solving, etc.

Input data (multiset, coefficients of the objective function and additional constraints) are presented as text files. This allows you to access the same set of data more than once. Data is generated using standard Delphi language procedure.

Multiset was assigned by its base and primary specification. Elements of the base are generated in increasing order, maximum difference between adjacent elements is given. Multiplicities of elements of the base were specified less or equal to the number k of combinatorial variables because the choice of larger multiplicities does not affect the solving problem.

When creating data files, the following parameters were set:

- the number of elements of the multiset base;
- the maximum difference between adjacent elements;
- the dimension of space n ;
- the number k of combinatorial variables;
- the number m of constraints.

The main characteristics of the algorithms studied during computational experiments are the running time and the number of iterations in the process of approaching the solution. Time is defined as the difference between the system time before the start of the procedure that implements the corresponding algorithm and after its completion.

Before starting the test, the user must select the directory in which the problem files are placed. You can choose to test all the files contained in the selected directory,

as well as individual files. You can also set the generation of files automatically before testing.

In the main window of the program (see Fig. 1), the user can select algorithms which will be used for solving problems. We assume that values of the objective function in the first algorithm of the method of constructing lexicographic equivalence should be integers (i. e. $A = Z$). If you choose the third algorithm you need to specify the exactness.

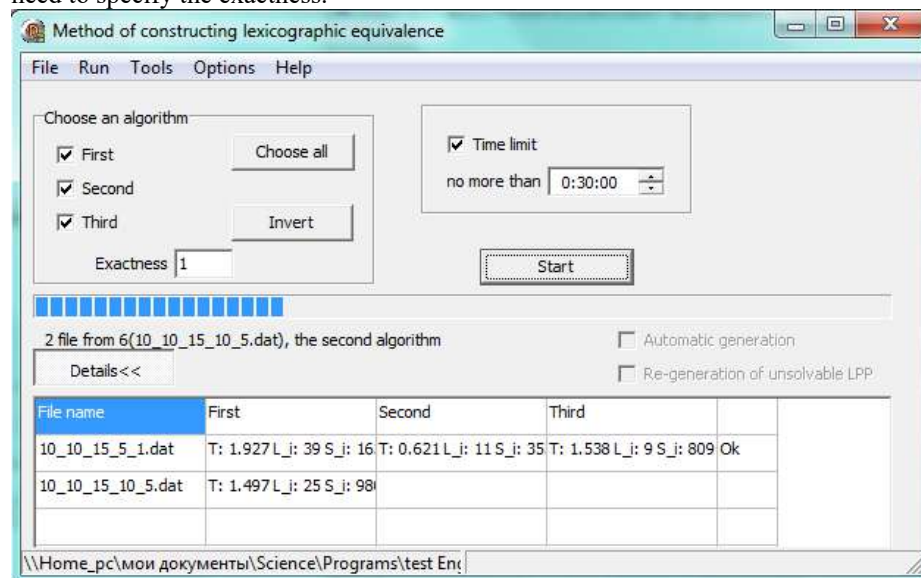


Fig. 1. The main window of software tool for solving linear combinatorial optimization problems on arrangements.

Since even with rather small dimensions there are problems whose solving requires a considerable amount of time, the program allows you to set the time limit for the execution of algorithms. In case the algorithm runs longer than the specified time, the execution is interrupted and the corresponding message is output to the result file. If necessary, you can try to solve such problems later.

During testing, the main window shows the percentage of problems that are completed, the name of the current problem file and the name of the algorithm. When the “Details” button is clicked, a table of results is displayed. This table contains the names of the tested files, the time of solving the corresponding problem by each of the algorithms and the number of iterations.

In addition to solving existing problems by selected algorithms, the software product allows automatically generate series of problems with given parameters (the dimension of space, the numbers of constraints and combinatorial variables, the number of elements of the multiset base, etc.). The maximum and minimum values of each parameter of the series are specified, as well as the number of problems with each set of parameters.

All algorithms of the method of constructing lexicographic equivalence involve solving linear programming problem which is obtained by replacing the combinatorial condition (3) with the condition $(x_1, x_2, \dots, x_k) \in \text{conv } E_\eta^k(G)$ where $\text{conv } E_\eta^k(G)$ is a convex hull of the set $E_\eta^k(G)$. If such a problem has no solution then the initial combinatorial optimization problem (3)–(5) also has no solution. In this case, the study of the effectiveness of the algorithms for constructing lexicographic equivalence does not make sense. Therefore the possibility of repeated generation is provided in the program. If a series received such problem that corresponding linear programming problem has no solution then file of this problem is generated again.

For the convenience of further analysis of the results of computational experiment, it is possible to create a summary file. This file contains for each problem information about the main parameters (the number of variables, the number of constraints, etc.) and the characteristics of the algorithms (the running time, the number of iterations). In the future, the data of such files can be systematized and grouped, for example, using a table processor.

5 Results of computational experiments

Several series of computational experiment were conducted using the software product described above. Testing of programs was carried out on a computer with a dual-core AMD A4-3400 processor, clock speed of each core 2700 MHz.

The first series of tests consists of 480 tasks, the generation of which parameters are given as follows: the dimension of space $n = 30$; number of combinatorial variables k accepts consecutive values from 15 to 30; number of constraints m accepts consecutive values from 5 to 10.

For each set of parameters, 5 problems were solved. Separate characteristics of the results of solving problems for which the running time does not exceed 30 min is given in the Table 1, Table 2, where such notation is used: qp is the percentage of problems solved in less than 30 minutes; st is the average running time; mt is the maximum running time.

Time is given in minutes:sec; time less than 1 second displayed as 0:01.

Table 1. The dependence of the running time of algorithms for constructing lexicographic equivalence on the number of constraints m (results of the first series of tests)

m	The first algorithm			The second algorithm			The third algorithm		
	qp	st	mt	qp	st	mt	qp	st	mt
5	100%	0:08	2:48	100%	0:40	12:58	100%	0:07	3:12
6	100%	0:11	4:41	100%	0:55	13:16	100%	0:10	4:16
7	100%	0:15	14:38	100%	0:51	15:33	100%	0:12	10:49
8	98,8%	0:33	10:56	97,5%	1:25	25:55	98,8%	0:33	14:27
9	97,5%	0:16	7:10	96,3%	1:04	16:09	97,5%	0:14	3:53
10	98,8%	0:43	9:28	93,8%	1:13	17:58	96,3%	0:25	7:07

Table 2. The dependence of the running time of algorithms for constructing lexicographic equivalence on the number of combinatorial variables k (results of the first series of tests)

k	The first algorithm			The second algorithm			The third algorithm		
	qp	st	mt	qp	st	mt	qp	st	mt
15	100%	0:01	0:06	100%	0:01	0:17	100%	0:01	0:04
16	100%	0:01	0:02	100%	0:02	0:13	100%	0:01	0:02
17	100%	0:01	0:12	100%	0:05	0:38	100%	0:01	0:10
18	100%	0:01	0:04	100%	0:03	0:22	100%	0:01	0:03
19	100%	0:01	0:06	100%	0:05	0:55	100%	0:01	0:06
20	100%	0:01	0:12	100%	0:05	1:38	100%	0:01	0:08
21	100%	0:01	0:10	100%	0:07	0:43	100%	0:01	0:10
22	100%	0:02	0:21	100%	0:19	4:09	100%	0:02	0:08
23	100%	0:02	0:27	100%	0:14	1:36	100%	0:02	0:12
24	100%	0:10	2:28	100%	1:03	12:25	100%	0:09	1:37
25	100%	0:11	2:44	93,3%	0:26	3:52	96,7%	0:05	0:44
26	100%	0:28	8:18	96,7%	1:44	9:36	93,3%	0:15	2:07
27	96,7%	0:32	9:28	96,7%	2:13	15:01	100%	0:29	7:07
28	96,7%	0:44	8:06	90%	2:04	12:58	96,7%	0:59	14:27
29	96,7%	1:24	7:36	100%	3:36	25:55	96,7%	0:49	3:53
30	96,7%	2:03	14:38	90%	4:43	17:58	96,7%	1:36	12:48

The results of the experiments presented in the Table 1, 2 show that the dependence of the running time of algorithms both on the number of combinatorial variables, and on the number of constraints, is characterized by a certain irregularity. In particular, all 30 problems with $k=27$ combinatorial variables were solved by the third algorithm in less than 10 minutes, whereas for two problems with $k=26$ the solution was not obtained in 30 minutes. The average running time of the second algorithm for problems with 22 combinatorial variables is greater than for problems with 23 combinatorial variables. Solutions 78 out of 80 problems with 9 constraints were obtained by the first algorithm in less than 30 minutes, while 79 out 80 problems with 10 constraints were solved in less than 30 minutes.

“Irregularity” confirmed Fig. 2, where each point corresponds to the results of the solving using the first algorithm for constructing lexicographic equivalence of one problem (the abscissa of the point is the number of combinatorial variables, and the ordinate is the corresponding time of solving the problem; if problem was not solved then time was equal to 30 min). For the second and third algorithms, similar images are obtained.

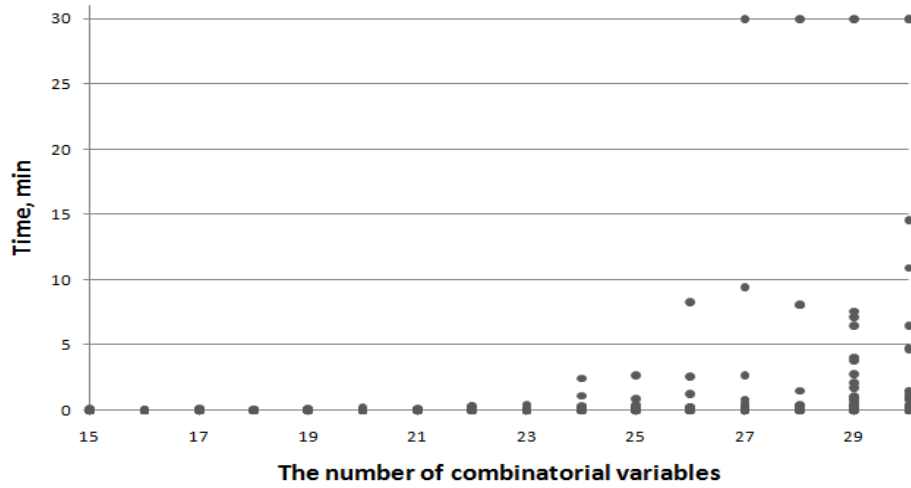


Fig. 2. The dependence of the running time of the first algorithm for constructing lexicographic equivalence on the number of combinatorial variables (results of the first series of tests)

The second series of tests consists of 420 problems (see Table 3).

Table 3. The dependence of the running time of algorithms for constructing lexicographic equivalence on the dimension n of space (results of the second series of tests)

n	The first algorithm		The second algorithm		The third algorithm	
	st	mt	st	mt	st	mt
30	0:11	2:16	0:43	8:57	0:11	2:56
31	0:16	4:49	0:53	10:46	0:11	3:26
32	0:01	0:08	0:13	1:54	0:01	0:09
33	0:02	0:20	0:19	4:04	0:02	0:17
34	0:10	4:37	0:34	4:43	0:04	1:00
35	0:02	0:14	0:12	2:03	0:02	0:15
36	0:02	0:21	0:18	1:53	0:02	0:12
37	0:01	0:04	0:17	4:44	0:01	0:06
38	0:01	0:13	0:20	3:19	0:02	0:17
39	0:01	0:05	0:09	1:28	0:01	0:14
40	0:01	0:09	0:18	3:26	0:01	0:13
41	0:01	0:06	0:15	1:59	0:01	0:09
42	0:01	0:23	0:08	2:00	0:01	0:16
43	0:01	0:02	0:05	1:03	0:01	0:03

The number of combinatorial variables is equal to 25, dimension accepts consecutive values from 30 to 43 and the number of constraints accepts consecutive values from 5 to 10, 5 problems were solved for each set of parameters. The running time was less than 11 minutes. The results are presented in the Table 3, where, as in the previous tables, st is the average running time; mt is the maximum running time. In

contrast to the first series of tests if the dimension of space increases (so the number of continuous variables increases) then the running time decreases.

Parameters of problems in the third series of tests were defined as follows: the number of continuous variables is equal to 7, the number of constraints is equal to 10, dimension accepts consecutive values from 25 to 48 (so the number of combinatorial variables changes from 18 to 40), the number of problems with identical parameters is equal to 10.

The analysis of Table 3 and Table 4 gives reason to assert that the dimension of space has less effect on the running time than the number of continuous variables, but the corresponding dependencies with an acceptable correlation coefficient can not be established.

Table 4. The dependence of the running time of algorithms for constructing lexicographic equivalence on the dimension n of space (results of the third series of tests)

n	The first algorithm			The second algorithm			The third algorithm		
	qp	st	mt	qp	st	mt	qp	st	mt
25	100%	0:02	0:17	100%	0:04	0:12	100%	0:01	0:04
26	100%	0:08	0:32	100%	0:13	0:37	100%	0:04	0:13
27	100%	0:02	0:08	100%	0:12	1:18	100%	0:01	0:04
28	100%	0:02	0:05	100%	0:08	0:31	100%	0:01	0:03
29	100%	0:13	1:51	100%	0:14	0:44	100%	0:04	0:21
30	100%	0:13	1:09	100%	0:55	3:27	100%	0:12	0:48
31	100%	0:07	0:22	100%	0:51	3:22	100%	0:06	0:17
32	100%	0:21	1:46	100%	1:06	2:19	100%	0:11	0:31
33	100%	0:08	0:26	90%	0:22	1:25	100%	0:08	0:41
34	100%	0:13	1:04	100%	1:52	6:58	100%	0:24	2:32
35	100%	0:59	7:02	100%	1:36	6:35	100%	0:31	2:60
36	100%	0:06	0:31	100%	2:49	12:04	100%	0:11	0:33
37	100%	0:03	0:11	100%	0:49	2:49	100%	0:04	0:16
38	100%	0:18	2:12	90%	1:36	7:20	90%	0:09	0:29
39	100%	0:10	0:32	80%	1:54	7:37	90%	0:27	2:27
40	100%	0:04	0:19	90%	0:29	2:08	100%	0:11	1:03
41	90%	0:14	0:59	90%	1:57	12:42	100%	2:57	25:37
42	100%	0:27	1:44	90%	7:19	26:38	100%	0:40	2:34
43	100%	1:26	4:48	70%	3:53	18:48	100%	1:36	6:52
44	80%	0:01	0:04	80%	2:27	8:17	80%	0:09	0:25
45	100%	0:58	4:41	90%	4:33	22:52	100%	1:14	5:25
46	100%	0:11	1:01	90%	2:37	12:53	100%	0:22	1:17
47	100%	0:11	0:28	90%	4:23	14:42	100%	0:28	1:53
48	90%	0:21	0:42	60%	12:01	28:54	90%	1:25	3:43

6 Conclusion

The paper considers the software implementation of the algorithms of the method of constructing lexicographic equivalence. This method is used for solving linear mixed

combinatorial optimization problems on arrangements and involves partition the space into equivalence classes followed by their direct search. The offered software allows you to generate a series of problems with the given parameters, to solve problems by selected algorithms, to save the results of problem solving. The computational experiment showed that the developed algorithms are effective for most problems with dimensions up to 50.

References

1. Papadimitriou, C.H., Steiglitz, K.: Combinatorial optimization: Algorithms and Complexity. Dover Publications, Mineola, New York (1982)
2. Grötschel, M., Lovász, L., Schrijver, A.: Geometric algorithms and combinatorial optimization. Springer-Verlag, Berlin Heidelberg (1993)
3. Korte, B., Vygen, J.: Combinatorial Optimization: Theory and Algorithms. Algorithms and Combinatorics, vol.21. Springer, Berlin, Heidelberg (2018) doi: 10.1007/978-3-662-56039-6
4. Nemhauser, G. L., Wolsey, L. A.: Integer and combinatorial optimization. John Wiley & Sons, New York (1988)
5. Pardalos, P.M., Du, D.-Z., Graham R.L. (eds.): Handbook of Combinatorial Optimization, Springer-Verlag, New York (2013)
6. Hulianyskyi, L., Riasna, I.: Formalization and Classification of Combinatorial Optimization Problems. In: Butenko S., Pardalos P., Shylo V. (eds): Optimization Methods and Applications. Springer Optimization and Its Applications, vol 130, pp. 239-250. Springer, Cham (2017) doi: 10.1007/978-3-319-68640-0_11
7. Zgurovsky, M.Z., Pavlov. A.A.: Combinatorial Optimization Problems in Planning and Decision Making. Studies in Systems, Decision and Control, vol. 173. Springer, Cham (2019) doi: 10.1007/978-3-319-98977-8
8. Emets', O. O., Roskladka, O. V., Nedobachii S. I.: Irreducible System of Constraints for a General Polyhedron of Arrangements. Ukr. Mat. Zh.. 55(1): 1-12 (2003) doi: 10.1023/A:1025060316418
9. Stoyan, Yu. G., Iemets, O. O.: Theory and methods of Euclidean combinatorial optimization [in Ukrainian]. Instytut systemnykh doslidzhen osvity, Kyiv (1993). <http://dspace.puet.edu.ua/handle/123456789/487>
10. Semenova, N. V., Kolechkina, L. N., Nagornaya, A. N.: One Approach to Solving Vector Problems with Fractionally Linear Functions of the Criteria on the Combinatorial Set of Arrangements. J. Automat. Inform. Sci. 42(2): 67-80 (2010) doi: 10.1615/JAutomatInfScien.v42.i2.50
11. Yakovlev, S. V., Grebennik, I. V.: Some classes of optimization problems on a set of arrangements and their properties. Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika 11: 74-86 (1991)
12. Iemets, O. O., Barbolina, T. M.: Combinatorial optimization on arrangements [in Russian]. Naukova dumka, Kyiv (2008) <http://dspace.puet.edu.ua/handle/123456789/473>
13. Barbolina, T. N. Solution of mixed combinatorial optimization problems on arrangements by the method of construction of lexicographic equivalence. Cybern. Syst. Analysis 49 (6): 137-149 (2013) doi: 10.1007/s10559-013-9582-4
14. Yemets, O. A., Barbolina, T. N.: Solution of euclidean combinatorial optimization problems by the method of construction of a lexicographic equivalence. Cybern. Syst. Analysis 40 (5): 76-734 (2004) doi: 10.1007/s10559-005-0010-2

Simulation of bio-like systems and processes using movable cellular automata

Volodymyr Zhikharevich¹, Kateryna Hazdiuk² and Serhiy Ostapov³

Yuriy Fedkovych Chernivtsi National University, Kotsyubynsky str., 2,
Chernivtsi, 58012 Ukraine

¹vzhikhar81@gmail.com

²kateryna.gazdyik@gmail.com

³sergey.ostapov@gmail.com

Abstract. In this work, the method of movable cellular automata is applied to the simulation of different biological systems and processes. A significant advantage of this method is the possibility of transition from a static grid to the concept of neighbours. Four examples are considered, namely the worm-like movement, the amoeba-like locomotion, the self-replication process and the biological engine. An object of simulation in all cases is divided into small parts, which are called cellular automata, that interact with each other by the rules of interactions. The basic principles of these interactions are the same, although the observable examples are diverse. According to the concept of neighbours, rules of cellular automata interactions were found for the constructed models. As a result, a computer models for each example were obtained. The models are qualitative, not quantitative, and allow you to demonstrate the fundamental possibility of movable cellular automata for such simulation.

Keywords: Movable cellular automata, Bio-like systems, Computer simulation, Worm-like movement, Amoeba-like locomotion, Self-replication, Biological engine.

1 Introduction

Nowadays computer modelling is one of the most powerful tools of scientific research and engineering applications. It provides an opportunity not only to reproduce the results of real experiments with a fairly high degree of confidence but also to reduce financial costs and eliminate of dangerous precedents. Moreover, in some cases, an alternative to computer models does not exist at all.

Simulation biological processes and systems is one of the most relevant and most promising areas of research. In contradiction to examples of physical phenomena, the mechanisms of many bioprocesses are unknown at this time. Thus, one of the tasks that are solved through simulation is the search for answers to questions that are still open. These issues, in particular, include: the path of self-organization and the evolution of living matter (Origin of Life); mechanisms of synergistic cooperation in the

colonies of unicellular organisms and their organization in the form of multicellular organisms, accompanied by cell differentiation (acquisition of specific functions by individual cell types within a single organism); self-organization of various signal flow in the neural subsystem of organisms and the formation of the phenomenon of consciousness. The last problem is particularly relevant in the field of artificial intelligence. It is obvious that all the achievements in this direction have nothing to do with the real intelligence of the human brain today, since their implementations are not based on the processes that take place in it. At the same time, the identification of key mechanisms leading to the self-organization and the evolution of consciousness will give impetus to the next scientific, technological and humanitarian revolution. In this case, it will be possible to talk not about artificial intelligence, but about the computer model of real intelligence.

The tools that are used to build models of biological objects and their dynamics are quite diverse – from describing investigating processes in terms of the theory of differential equations and then finding their numerical solution, to building imitational complex models that reflect the individual components of the corresponding processes or systems. Simulation by the method of movable cellular automata (MCA) can be attributed precisely to the simulation approach that has long been used to implement a wide range of artificial biosimilar structures. In general, this direction of simulation is named “Artificial Life” (Artificial Life), its purpose is to develop general approaches for constructing artificial models of living organisms. At the same time, the use of MCA as a simulating tool attracts with the possibility of creating simple models of rather complex systems.

The first researcher on “Artificial Life” was John von Neumann, who studied the possibilities of implementing self-replicating structures. The ability of self-reproduction (or self-replication) is a unique phenomenon that characterizes all living organisms. Neumann showed and described [1] the possibility of constructing a discrete deterministic self-replicating automaton, the certain analogue of the Turing machine. It was capable of copying the contents of a control program and fragments of the automaton structure.

However, within the framework of the Neumann and similar models, no processes of self-organization and self-complication (evolution) are observed. This is due to the fact that in such structures there is no implementation of the mechanisms that occur in real living organisms. It is well known, that biological objects are a combination of a large number of regulatory processes interconnected among them and organized in the form of extensive networks of various transformations of molecular complexes. The overwhelming majority of stimuli, that is, deviations from organic natural dynamics (if they are not destructive for the organism), are perceived by the corresponding receptor complexes that initiate certain chains of transformations and lead to the formation of an appropriate reaction. Thus, all regulatory processes have the nature of feedback. According to such principles, for example, the regeneration of damaged parts of organisms, the normalization of the state of the internal environment (homeostasis), the management of the organism's behaviour, its development, etc., occurs. The network of regulatory processes is determined by the organism's genotype. Moreover, the genotype is also subject to self-regulation (for example, the process of DNA re-

pair). But, at the same time, the biosystems are characterized by variability that associated with genotype mutations. These mutations are not renewed by regulatory mechanisms, because they harmoniously interact with them. During mutations processes, the genotype may become more complex, that leads to a complication of the phenotype of the organism and expands the network of regulatory processes. And, against the backdrop of Darwinian natural selection, this leads to the evolution of biological objects.

The basis of the above-mentioned dynamics of biological structures is also a process of self-organization, which means the absence of a single control centre. In contrast to the centralized management approach, each element of self-organizing systems acts by itself, interacting with only a small number of neighbouring elements. So, it turns out is enough to streamline chaotic structures. Based on the above, the method of mobile cellular automata is chosen as a simulation method. Or rather, the asynchronous stochastic MCA method with the symbol alphabet, which corresponds to the set of states of the MSA. This method is a further development of deterministic synchronous automata [2], which are used to simulate physico-mechanical processes in solid deformable bodies.

The aim of the work is not only to build a model in terms of cellular automata, but also to search for cellular automaton interactions that lead to the dynamics of a system imitating bio-like processes and structures. That, in turn, is an important both theoretical and practical task.

2 Simulation of worm-like movement

It is quite promising to use a multi-agent approach implemented on the basis of the MCA for simulation the processes of locomotion. To do this, the biosystem is decomposed into separate constituent elements and each element is represented by an MCA with a corresponding behaviour depending on the type of element and the state of the environment, that is, on the state of the neighbouring MCA. Similar artificial organism is also called an “animat” (from Artificial Animal). The advantages of using MCA in creating animats are: the possibility of a fairly flexible formation of an arbitrary morphology of an artificial organism and its organs; an imitation of the processes of differentiation, division and death of cells; a wide range of functional properties of cells, what form the corresponding tissues (muscular, nervous, skin, etc.).

These approaches are used by many researchers. For example, in [3], the authors describe the process of morphology formation of elementary animats by using the idea of encoding the forms of artificial organisms as an analogue of the genotype. The change in the content of the genotype is displayed as a corresponding change in the shape of the organisms. It is possible to use the operators of genetic algorithms (mutations, crossing, selection) to looking for various biosimilar forms. In addition, in models of animats, imitation of the spreading of signals along the diffusion nervous system is carried out, and muscle contraction take place. The whole body of the animat is a collection of cells and a kind of excitable medium in which nerve impulses can propagate in the form of waves. Under the influence of these waves, the distance

between cells is reduced or increased in order to imitate the contraction of the muscles of an artificial organism. In fig. 1.a) shows an example of a worm-like organism constructed in such a way.

The examples of other forms of animats (jellyfish, fish, turtles, hydras, etc.) are also given in [3]. Nevertheless, many researchers pay attention precisely on the worms. Nematode *Caenorhabditis elegans* was chosen as an object for computer modelling in the international project OpenWorm [4]. The aim of the project is to build a dynamic computer model of *C.Elegans* at the cellular level. It was expected that, knowing the complete information about the structure of the organism, including the structure of the connectome, it would be possible to create an adequate artificial analogue of a real organism with similar behaviour. The Open-Worm project has been actively developed since 2011, but the goals have not been already achieved.

On our opinion, it is not necessary to strive to realize the absolute equivalent of the real analogue, as the participants of the OpenWorm project try to do. Building a complex detailed model requires a lot of time and computational resources and, that most importantly, it can interfere with the search for regulatory mechanisms that maintain the wholeness and durability of an artificial organism, as well as ensure the natural harmony of its components. In addition, the development of a universal method for modelling bio-like systems, which could be applied to different objects, is quite relevant.

Regarding to this theory, the MCA method, what, in fact, was used by the authors of [3], is very promising, in our opinion. In addition, this method provides the possibility of a variety flexible modifications, partially described earlier. In fig. 1.c). an example of a model of the worm-like organism proposed by us is showed. This model is also based on the MCA method.

To implement the model, was chosen a scheme of neighbourhood. According the scheme each cell has four neighbours for a two-dimensional model and six in the case of a three-dimensional one. In this work we consider two-dimensional model, and in this case the neighbours are divided into two groups what simulate longitudinal and transverse muscles.

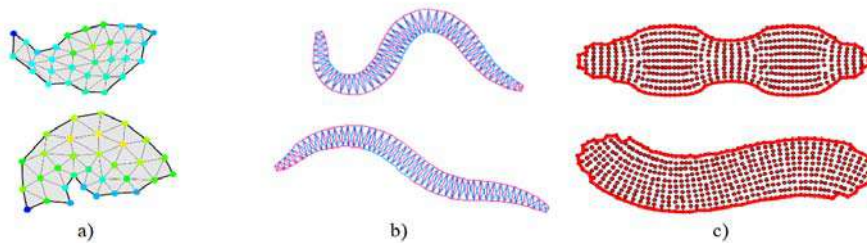


Fig. 1. Examples of models of worm-like organisms: a) the model given in [3]; b) the model from [4]; c) the model proposed by us.

When the transverse muscles are reduced, the corresponding body fragments should be increased in length and compressed, and with the reduction of longitudinal

vice versa - decrease in length and expand. The signal for muscle contraction is the state of the corresponding "nerve ending" of the neural subsystem, which is associated with the corresponding MCA. In fig. 2 shows examples of the mutual arrangement of MCA in interaction. Fig. 2.a) reflects the state of the MCA in cases of whether absence of muscle contraction signals for the selected cell, or attempts to reduce both muscles (longitudinal and transverse). Fig. 2.b) reflects the state of the MCA in the case of a signal of reduction of longitudinal muscle. And Fig. 2.c) reflects the state of the MCA in the case of a signal of reduction a transverse one.

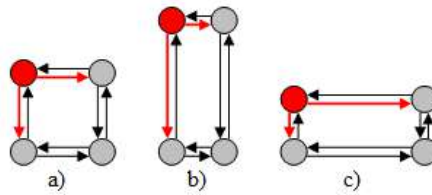


Fig. 2. Examples of the mutual arrangement of MCA

The essence of MCA interactions is to establish the distance between neighbours depending on the state of the corresponding efferent neuron, what also locate in the structure of the automaton. Moreover, according to the properties of incompressibility, the volume of the elements of the worm body is unchanged. The model provides for the possibility of the presence in the MCA structure, in addition to the efferent neuron, also of an afferent (receptor) one what responds to a specific state (form) of the surroundings. Connections between the neural parts of the MCA can be not only within the nearest neighbours, but also between distant ones. (fig. 3). Moreover, these connections can dynamically change, drifting from one cell to another, imitating the growth of dendrites and axons of real nerve cells. The signal flows in the nervous subsystem of an artificial organism lead to a corresponding contraction of the muscles and form a variety dynamic, examples of which are shown in Fig. 1.c).

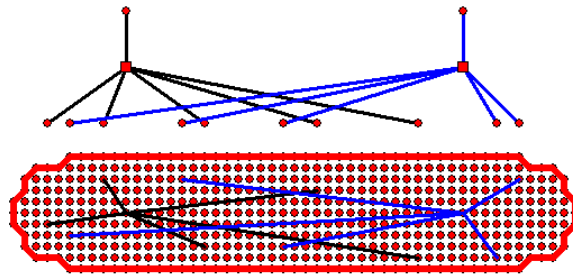


Fig. 3. Simple schematic representation of the neural subsystem

To describe the mathematical model of the neural network, we define the transformation function $f: X \rightarrow Y$, which forms one of three output values (0 is the absence of

a signal, 1 - reduce the longitudinal muscle, 2 - reduce the transverse muscle). It depends from the value of the amount of input signals multiplied by weight coefficients. Each input has its own synaptic weight what gives input the effect needing for the adder function of the processing element:

$$y_j = \begin{cases} 0, 0 \leq \sum_{i=1}^K x_i w_i < 2K/3, \\ 0, 2K/3 \leq \sum_{i=1}^K x_i w_i \leq 4K/3, \\ 2, 4K/3 < \sum_{i=1}^K x_i w_i \leq 2K, \end{cases} \quad (1)$$

where n – number of MCA in the modelled object; K – the maximum number of connections for neurons; x_i ($x_i = 0, 1, 2$) – input signals; $w_i \in [0 \dots 1]$ – weighted coefficients, what are the measure of the force of the input bonds and simulate the various synaptic forces of the biological neurons.

The weights of the substantial input are increased and, conversely, the weights of the non-essential input are reduced, which determines the intensity of the input signal. It should be noted that the values of the weight coefficients and the order of the conditional operators in formula (1) may change and they define a characteristic of the individual MCA.

3 Simulation of amoeba-like locomotion

Another interesting microorganism for research is the unicellular amoeba protea (Amoeba Proteus). It is moving in space by changing the shape of its body. At the same time certain protrusions, so-called pseudopodia ("false-feet") and posterior uropods, are formed. Amoeba recognizes various microscopic organisms that serve it as food. It goes away from bright light, mechanical irritation, and elevated concentrations of substances dissolved in water. In addition to these unconditioned reflexes, some researchers have documented the formation of conditioned reflexes in amoebas. Conditioned reflexes of a unicellular organism with a missing nervous subsystem! Many scientists, which watch about amoebas, ask themselves: how it is possible? It is also interesting to note that the size of the amoeba genome is almost 200 times the size of the human genome. This paradox indicates that there is not always a direct connection between the size of the genome and the complexity of the organism.

Let us focus on the problem of modelling the movement of the amoeba, which occurs, as already noted, due to the growth of the membrane in the form of so-called pseudopodia. At the same time some other parts of the membrane are reduced. Changes in the membrane are due to reversible transformations "sol" $\leftrightarrow$ "gel". These processes are guided by the cytoskeleton, whose structure is constantly changing (Fig. 4. a)). The result of complex membrane-cytoskeletal interactions is the harmonious dynamics of the movement of the microorganism.

All existing models of amoeba do not have taken into account the cytoskeletal subsystem of the body, but nevertheless quite realistically reflect the processes of movement, including in the direction of nutrients, or a leaving from lighting (phototaxis) or

from chemical stimuli (chemotaxis). One of the implementation options described in [5]. In this work, the amoeba body is a set of related spatial elements, where a gradient of a certain continuous parameter is formed. The internal environment of the unicellular organism (cytoplasm) moves along the gradient. This movement, leads to the growth of pseudopodia and contraction of the amoeba body in the opposite direction (Fig. 4.b)).

To build a model of the described microorganism, it is also possible to present it in the form of some multicellular analogue, and therefore use the MCA approach. This, in particular, will provide an opportunity to simulate membrane-cytoskeletal interactions. The MCA are able to change their state by defining interaction functions with neighbours as weakly or rigidly interconnected, when modelling the internal and membrane environment of a microorganism. That means forming the corresponding properties of the cellular environment, which form as "sol" or "gel". Also, the model includes self-assembly and self-destruction of cytoskeleton fragments. Self-assembly usually occurs in centrioles. Different types of filaments (filaments or microtubules) of the cytoskeleton can be distinguished, which have different effects on the transformation in the membrane (actin and myosin filaments). In cases of interacting with the appropriate types of the end of filament thread, a transition from the gel state to the sol state occurs and the membrane reduced, in the other positions the reverse process occurs and the membrane grows in the form of a pseudopodia (Fig. 4.c)).

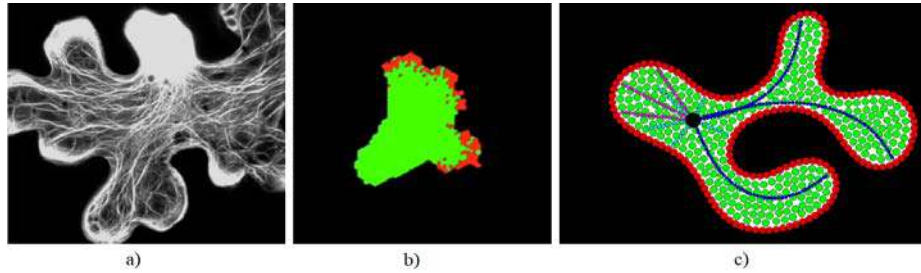


Fig. 4. Examples of models: a) an image of a real amoeba with a cytoskeleton; b) the model of amoeba-like organism described in [5]; c) the model proposed by us.

We used a dynamic index array, allowing not only to identify any MCA in the system, but also containing the indices of its nearest neighbours. As parameters, we choose: the type of MCA, which determines its properties; coordinates (for simplicity and clarity let us dwell on a two-dimensional model); indices of the nearest neighbours. The simulation used a hexagonal neighbourhood pattern ($M = 6$), providing for weak and rigid connectivity of pairs. In addition, to reduce the number of neighbours, we add zero values to the index of neighbours. The method of finding for the nearest neighbours under the conditions of their fixed value is described in [5]. In this paper, the criterion for finding neighbouring cells is the minimum distance between them, and the search is carried out by introducing an additional index array. Neglect of the inertial properties of the amoeba body allows one to abandon the content of the MCA velocity vector components in the index array (Fig. 5).

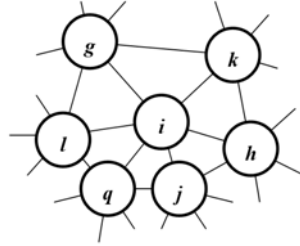


Fig. 5. The fragment of the MCA set.

The main essence of the algorithm is iterative modification of the contents of the index array. For this, an asynchronous approach is implemented, in which one of the N MCAs and one of the M of its nearest neighbours are selected in an equiprobable way. The choice of the asynchronous approach when developing the MCA algorithm is due to the fact that it allows to avoid collisions, that is, to satisfy the criterion of correctness (there will not be a single attempt to change the state of the same cell more than once at the same time t).

The cellular automaton model admits six possible states of movable cellular automata, which have a different purpose and functionality and are described using the alphabet $A = \{a_1; a_2; \dots; a_6\}$. For the difference are used circles of radii r_i , $i = 1, 2, 3$ ($r_1 < r_2 < r_3$). After selecting two adjacent MCAs, the interaction function is implemented. This function depends on the types of interacting MCA, and is convenient to present in a matrix form (Fig. 6). Naturally, the matrix should be symmetric. Wherein, in some cases described below, interactions are carried out with the participation of more than two MCA.

	a_1	a_2	a_3	a_4	a_5	a_6
a_1	F_2	F_2	F_0	F_0	F_0	F_0
a_2	F_2	F_3	F_1	F_4	F_5	F_1
a_3	F_0	F_1	F_1	F_7	F_7	F_7
a_4	F_0	F_4	F_7	F_6	F_0	F_0
a_5	F_0	F_5	F_7	F_0	F_6	F_0
a_6	F_0	F_1	F_7	F_0	F_0	F_6

Fig. 6. Matrix representation of the MCA interaction function

Each of the functions F is a composition of separate elementary operations carried out during interactions of the corresponding types. The functions' appointments are as follows:

- F_0 imitation of thermal fluctuations;
- F_1 repulsion of two MCA at their interpenetration;

- F_2 repulsion at move in and attraction at move away (imitation of the condensed state of the environment);
- F_3 MCA alignment (for modelling elastic properties of cytoskeleton threads and membrane surface);
- F_4 catalysis of the sol-gel transformation, in which the MCA of the internal environment a_1 is converted into the MCA of the membrane a_2 ;
- F_5 catalysis of the gel-sol inverse transformation.
- F_6 is an analogue of F_3 , but if the MCA is the end of the cytoskeleton thread (terminal), then it behaves as F_1 when interacting with the same type of MCA;
- F_7 imitation of the growth of thread from centrosome.

In this case, the following restrictions apply, which are taken into account during interactions:

1. MCA membranes a_2 cannot have more than two neighbours of the same type a_2 . This ensures the integrity of the membrane and prevents its adhesion. In the future, when modelling the processes of nutrition as well as the division of amoebas, this restriction will be modified;
2. The MCA of the a_4 , a_5 and a_6 cytoskeleton threads, if they are not terminal, cannot have more than two neighbours, which is ensured by entering four zero values in the neighbour indices field of the MCA array;
3. MCA adjacent to a_3 (centrosome cell) are considered related to it. Thus, the MCA of the centrosome can have an arbitrary number of cytoskeleton threads associated with it. In the future, when modelling the amoeba division process, it will be possible to divide the MCA of the centrosome, which, in turn, will initiate the process of amoeba body division into two separate unicellular organisms.

Research and modelling of biosimilar structures also obtain applied goals, besides the fundamental ones. In particular, the study of the features of locomotion of worm-like and amoeba-like organisms, discussed above, allows the use of appropriate approaches in the implementation of so-called soft-bodied robots. These works are quite promising when used for movement in hard-to-reach places, for example: pipelines of various configurations, sewage systems, wells, mines, caves, etc.

4 Simulation of self-replication processes

A separate area of research on the self-generation and evolution of bio-like structures is computer simulation of chemical processes that can lead to similar dynamics (Artificial Chemistry) [7, 8]. In particular, the authors of the work [7] managed to build a model of self-replication of cell-like structures (the so-called "Los Alamos Bugs"), which possess not only elementary metabolism, but also hereditary genetic information. The enzymes are kept inside the cell by a loop of membrane, thus ensuring that only the cell that produced them gets their benefit. A set of reaction rules, each simple and local, allows the cells to copy their genetic information and physically divide. The evolutionary possibilities of the cells are explored, and it is suggested that the system

provides a useful framework for testing hypotheses about self-driven evolution. In this case, researchers argue that the model lays the foundation for the first integrated spatially distributed computer modelling of the entire protocellular life cycle. The method for dynamics' modelling of three-dimensional dissipative particles was used, as a tool for constructing a fairly simple models of the jointed processes of diffusion, self-assembly and chemical reactions.

Analysing the mentioned above approaches to modelling self-replicating structures, we can propose an MCA method. The central problem of constructing a cellular automaton model of self-replication processes, as well as any arbitrary phenomenon or system, is the determination of the rules of automaton interactions. In the case of modelling of self-replicating structures, we propose to form a very wide range of different types of RCAs and establish equilibrium neighbourhood rules among them. These rules presuppose the existence of specific desirable sets of MCA types surrounding specific types of MCA. In such cases, the potential of MCA is minimal, and the structure, which they constitute, is robust. If the MCA is not in positions with its' equilibrium environment, then the structure is unstable, and therefore it can be arranged to move to equilibrium. At the same time, there are two possible options: either the MCA drift toward the potential decrease, and thus self-assembly process is implemented, or the MCA initiate the formation (synthesis) of the equilibrium environment and thus the self-regeneration process is realized. It can be possible to consider the combination of these two options.

As an example, we can consider the equilibrium interaction of two types of MCAs, one of which (S1) must be surrounded by six other RKAs (S2) in an equilibrium state. And for another the S2 type of MCA there must be three equilibrium neighbours – one S1 and two S2. These equilibrium neighbourhood rules can be expressed in tabular form (Tab. 1):

Table 1. Equilibrium Neighbourhood Rules for MCA

Number of neighbours	Type of sells	S1	S2
6	S1	-	6
3	S2	1	2

Demonstration of the structure's stability formed on the principles of equilibrium neighbourhood by the MCA method is shown in Fig.7. The deviations lead to a non-equilibrium state. At the next moments of time there is an arbitrarily comeback to the state of equilibrium and the initial structure is restored. That is, there is a process of self-regeneration.

Rules of the MCA interactions can be supplemented by the possibility of activation of different subsets that described transitions from one type to another, depending on some continuous parameters (analogue of nutrients supplying). In this way, an arbitrary periodic dynamic will be realized, the partial case of which the self-replicating is. In addition, the algorithm of the method can be supplemented by the mechanism of random addition of new types of MCA to a modelled cell-like structure (analogue of mutations), which can lead to increased stability of its dynamics, or vice versa - coun-

teract replication. Finally, by creating the conditions for natural selection, it is possible to organize the evolution of the studied self-replicating cell-like objects.

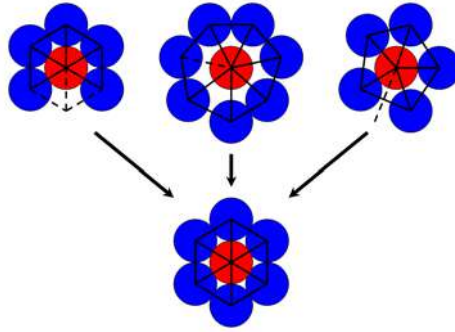


Fig. 7. A schematic demonstration of the stability of structures formed on the principles of equilibrium neighbourhood.

Consider an example of the elementary self-replicative dynamics of some cellular automaton structure. The rules of equilibrium neighbourhood of elements have the form (Tab. 2):

Table 2. The rules of equilibrium neighbourhood of self-replication structure

Number of neighbours	Type of sells	S1	S2	S3	S4	S5	S6
6	S1	-	6				
3	S2	1	2	1	1	1	1
4	S3		2		2		
6	S4		3	2	1		
5	S5		2			1	2
6	S6		4			2	

Here for a type S2 it is possible to see the peculiar ambiguity in sets of equilibrium neighbourhood. This suggests that any combination of three neighbours, which corresponds to the possible variations from Table 2, is equilibrium. For example, there may be different neighbours for S2: or one S1 and two S2, either one S3 and two S2, or one S1 and one S3 and one S4, etc.

In fig. 8 are depicted the stages of the process of self-replication of the cellular automaton structure. This process is accompanied by the transitions of the MCA from one type to another, in particular: $S1 \rightarrow S4 \rightarrow S6 \rightarrow S1$. Such transitions simulate the growth phases and the division of cell-like objects. Immediately after changing the type of the corresponding MCA structure, it is in a nonequilibrium state, what initiates the process of going to equilibrium, during which the spatial organization of the neighbourhood is rebuilt and new MCAs appear (self-generated).

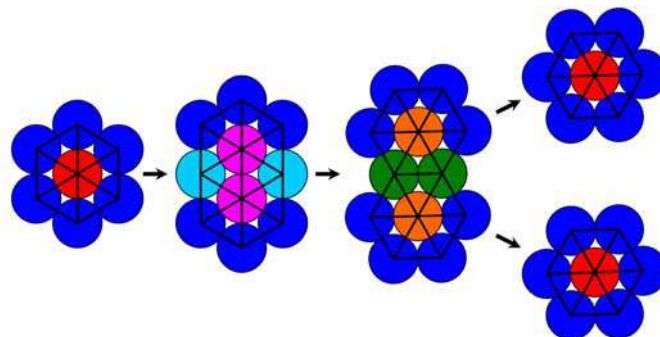


Fig. 8. Demonstration of the process of self-replication.

By changing the rules of equilibrium neighbourhood, it is possible to form various stable structures with arbitrary morphology, examples of which are shown in Fig. 9. Moreover, it should be noted that the rules of the neighbourhood may be asymmetric, that is, some MCA can establish a connection of the neighbourhood with the k -th, whereas k -th may not perceive it as its neighbour. Nevertheless, the structures will be stable, because the balance is not disturbed.

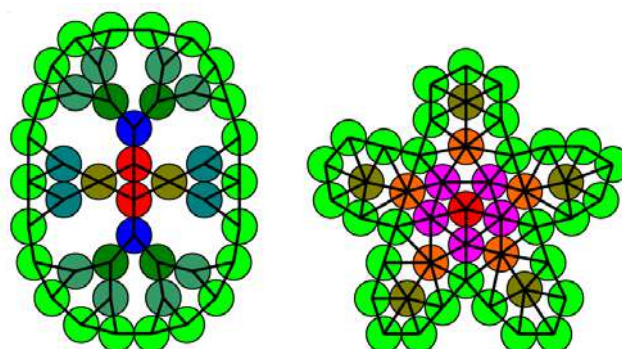


Fig. 9. Examples of structures with arbitrary morphology, formed on the principles of equilibrium neighbourhood of MCA.

5 Biological engine simulation

In addition to chemical reactions, where the process of imitation of molecular transformations takes place, MCA are also used to simulate the formation of various molecular complexes, as well as the dynamics of molecular and nanoscale structures [9]. It is also appropriate to use the MCA toolkit as an analogue of the molecular dynamics' method. As an example, we present a simple model of a molecular engine (Fig. 10 a)), where rotation is carried out due to the thermal energy of atomic vibrations. Therefore, the structure can be called a perpetual motion machine of the second kind. By the condition that the number of fibres of the external structure of the engine

and the teeth of the inner one is not the same, there is a probability of transition of one of the fibres tangents to the top of the tooth from one segment to another clockwise. This is shown in Fig. 10.b) and 10.c) The probability of the reverse transition is less.

The rules for MCA interactions in this case are as follows. All automata as elementary models of atoms carry out thermal oscillations. In this case, the angle and offset distance are randomly generated. MCA of the external structure, in parallel with thermal fluctuations, tend to align along one line and it pushing away the automata of the internal structure during this process. MCA external structure with three neighbours, tend to align the position of one of the neighbours relative to the other two. MCA internal structure, in parallel with the heat fluctuations, tend to push away the automata of the external structure.

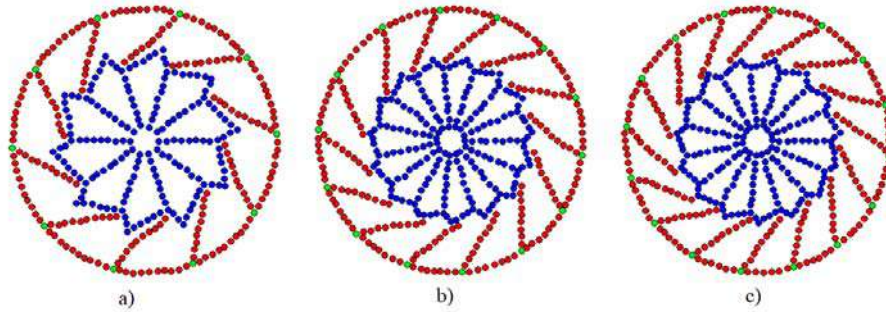


Fig. 10. Examples of molecular engine configurations: a) the number of external fibres and internal teeth coincides - there is no rotation; b) the fibres are one less than the number of teeth – there is rotation; c) one more teeth in the fibres – there is a rotation.

6 Conclusions

As we can see all of examples, mentioned above, leads to the same tasks, which can be solved by the MCA method. This paper describes some of the applications of the movable cellular automata method to the modelling of bio-like processes and systems. A significant advantage of this method is the possibility of transition from a static grid to the concept of neighbours.

For each model, the rules of cellular automaton interactions are found according to the concept of neighbourhood. As a result, computer models mimicking certain natural phenomena were obtained. It should be noted that the models are qualitative, not quantitative, and allow you to demonstrate the fundamental possibility of MCA method for such modelling. The novelty of the research lies in the use of a substantially new approach to modelling.

Bioinspired technologies are also actively used for optimization problems, in particular, for optimization of traffic flows. At the same time, in addition to the well-known ant algorithms, recently imitating models of the *Physarum Polycephalum*,

including using the MCA, have gained popularity. In the search for nutrients, the moulds first grow up with an almost continuous canopy, and then narrows, forming transport channels between the places where the nutrient medium is located. The structure of these transport channels is distinguished by its rationality and efficiency. It is a very interesting example for our future research.

In future, it is planned to model both simple and more complex multicellular organisms, collect data and study their behaviour, embryogenesis, self-organization, self-replication, their nervous system, muscular system, etc. Of course, this will be possible only with more detailed biological analysis of the subject area and with the development of software engineering.

References

1. J. von Neumann. The Theory of Self-Reproducing Automata. University of Illinois Press, Urbana (1966).
2. Shilko A. V., Psakhie S. G., Schmauder S., et al.: Overcoming the limitations of distinct element method for multiscale modeling of materials with multimodal internal structure. *Computational Materials Science*, vol. 102, 267-285 (2015).
3. Wrobel B.: Brainless Bodies: Controlling the Development and Behavior of Multicellular Animals by Gene Regulation and Diffusive Signals. *Artificial Life* 13, 349-356 (2012).
4. B. Sziget, P. Gleeson, M. Vella, S. Khayrulin, A. Palyanov, J. Hokanson, M. Currie, M. Cantarelli, G. Idili, S. Larson: OpenWorm: an open-science approach to modeling *Caenorhabditis elegans*. *Frontiers in Computational Neuroscience* 8, 1-7 (2014)
5. Nishimura S. I., Ueda M., Sasai M.: Non-Brownian dynamics and strategy of amoeboid cell locomotion. *Physical Review E* 85, 19-17 (2012)
6. Zhykharevych V. V., Hazdiuk K. P. Alhorytm vyznachennia susidnikh elementiv mnozhyzny rukhomykh klitynykh avtomativ za umov fiksovanoi kilkosti susidiv [Algorithm for determining the neighbouring elements of a set of movable cellular automata under the condition of a fixed number of neighbours]. *Bull. National Technical University Kharkov Polytechnic Institute. Ser. Computer Science and Modelling*, 33, 75–82 (2015) (in Ukrainian).
7. Hutton T.J.: Evolvable self-reproducing cells in a two-dimensional artificial chemistry. *Artificial Life*, 13 (1), 11-30, (2007).
8. Fellermann H., Rasmussen S., Ziöck H.-J., Solé R. V.: Life cycle of a minimal protocell – a dissipative particle dynamics study, *Artificial Life*, 13 (4), 319-345 (2007).
9. Peplow, M. The Tiniest Lego: A Tale of Nanoscale Motors, Switches and Pumps. *Nature*, 525, 18–21, (2015).

Method of statistical spline functions for solving problems of data approximation and prediction of objects state

Serhiy Babak¹[0000-0001-8805-1184], Vitaliy Babak²[0000-0002-9066-4307],
Artur Zaporozhets²[0000-0002-0704-4116], Anastasia Sverdlova²[0000-0001-8222-1357]

¹ University of Emerging Technologies, Kyiv, Ukraine

² Institute of Engineering Thermophysics of NAS of Ukraine, Kyiv, Ukraine

a.o.zaporozhets@nas.gov.ua

Abstract. The method of statistical spline functions is considered for problems of predicting the state of complex technical objects using the example of power transmission lines. The choice of parameters of spline fragments for building an adequate mathematical model is analyzed. Based on the experimental data, a short-term spline forecast of heating of overhead power lines has been created.

Keywords: spline functions, mathematical model, overhead power lines, prediction, technical state, diagnostics

1 Introduction

The basic idea of using the mathematical apparatus of statistical spline functions for processing the spatial characteristics of a field is that some numerical characteristics of physical processes change during the observation process (meteorological fields, airspace in the vicinity of energy facilities, navigation space, etc.) for any circumstances [1, 2, 3].

Regular observations of the history of gradual changes in these parameters over time provide an opportunity to obtain information on the trends of further changes in the studied parameters and to predict the behavior of the field at certain points [4]. For example, using the spline function algorithm, it can monitor and predict the temperature at a particular point in the airspace [5]. Such a problem arises during temperature control equipment of overhead power lines (OPL) using thermal imaging equipment installed on the unmanned aerial vehicles (UAV) [6, 7].

Using the spline functions, the problem of predicting electrical equipment failures is also solved. In this case, the main idea of predicting failures is that some numerical characteristics of physical processes occurring in certain nodes of electrical equipment change during the occurrence or development of faults and defects, which allows to identify them [8, 9]. Observations of their gradual change in time provide information on the development trends of the defect and provide a prediction of the possible moment of failure [10, 11].

In work the forecasting of values of temperature in certain points of an arrangement of the equipment of the OPL will be considered. Due to a significant tempera-

ture increase, failures of individual blocks of OPL are possible. So, the sharp heating of metal wires of OPL can lead to sagging and short circuit on the earth. The temperature factor may contribute to the occurrence of a breakdown or overlap of insulators of OPL. It is also possible breakdown of insulators due to contamination of their surface, or aging of the materials from which they are made. In addition, due to various reasons, accumulation of microdefects can occur in the insulator material, which contributes to their breakdown [12, 13, 14].

The listed defects of equipment of OPL in the course of their operation can lead to the emergence of so-called gradual failures. Accidental failures caused by unpredictable factors (for example, the overlap of insulation of OPL by birds or animals) will not be considered.

2 Features of spline functions

2.1 Mathematical model

To predict the possible failure of a selected node, you additionally need to have statistics on the values of the monitored parameter for a certain period of time, which is called the observation interval.

Let some functional dependence be given on the segment

$$y = f(x, A), \quad x \in [a, b], \quad (1)$$

where A is the deterministic vector of unknown real numeric parameters entering linearly in y and do not depend on x .

At given points $x_i \in [a, b]$, $i = \overline{1, n}$, random uncorrelated values of the function y are observed, which we denote as $\{y_i, i = \overline{1, n}\}$. For definiteness, we assume that these observations are distributed according to the normal law, and

$$\begin{cases} M y_i = x_{i0} a_0 + x_{i1} a_1 + \dots + x_{ir} a_r \\ D y_i = \sigma^2, \quad i = \overline{1, n} \end{cases}, \quad (2)$$

where $A = (a_1, a_2, \dots, a_r)$ and σ are unknown parameters,

$X = \{x_{ij}, i = \overline{1, n}; j = \overline{0, r}\}$ is a rectangular matrix of deterministic coefficients, functionally dependent from x_i and, as a rule, is called the planning matrix. It should also be noted that the dependence of these coefficients from x_i is not necessarily linear.

If we assume that Y and A are column vectors, which will be provided below, then (2) can be written in vector form: $MY = XA$, $DY = \sigma^2 I$. Here, M and D are the expectation and variance operators, and I is the n -th unit matrix order.

The paper considers the problem of constructing a statistical estimate of the unknown parameters $\{a_j, j = \overline{1, r}\}$ from the results of observations $\{y_i, i = \overline{1, n}\}$. In this case, an arbitrary choice of elements of the planning matrix X is allowed. If nothing is assumed about the distribution of observation errors or observations

$\{y_i, i = \overline{1, n}\}$, then as a result of solving this problem the largest, what can we expect, is the construction of point estimates for elements A . Under the assumption that the distribution law of observations is normal, we can obtain a confidence interval for the estimated parameters.

Considering that in the task set the question of introducing the planning matrix is solved in an arbitrary way, we specify this task by refining the choice of the planning matrix X . At the same time, we construct statistical estimates of the A parameters using the least squares method, which in the case of normally distributed y_i leads to the same result as maximum likelihood estimation.

Concretization in the choice of the planning matrix, first of all, is connected with the choice of the upcoming functions. The task was traditionally solved in the class of polynomials. But polynomial approximations have several disadvantages, the most significant of which is that the sequence of interpolation polynomials does not always converge to the interpolated function. Therefore, in many problems, the more natural and convenient approximation apparatus turned out to be splines, with the help of which we will solve the problem posed.

Splines are functions that are “glued together” from pieces of various functions in a specific pattern. Polynomial splines “stick together” from pieces of various polynomials in such a way as to ensure the necessary smoothness of the resulting spline. The simplest example of a polynomial spline is a broken line.

Let the grid be set on the segment $[a, b]$, $a, b \in R^+$, $a < b$ (partition):

$$\Delta_n : a = x_0 < x_1 < \dots < x_n < x_{n+1} = b, \quad (3)$$

where $n \in N$.

Let also P_m is the set of polynomials of degree not higher than m , $m \geq 0$, and $C^{(k)} = C^{(k)}[a, b]$ is the set of functions continuous on $[a, b]$ that have a continuous k -th derivative, $k \in Z^+$; R^+ is the set of positive numbers; N is the set of natural numbers; Z^+ is the set of positive integers.

The function $S_m(x) = S_{m,k}(x, \Delta_n)$ is called a polynomial spline of degree m of defect k ($1 \leq k \leq m$) with nodes (3), if

- $S_m(x) \in P_m$, $x \in [x_i, x_{i+1}]$, $i = \overline{0, r-1}$,
- $S_m(x) \in C^{m-k}[a, b]$.

The points $\{x_i\}$ are called spline nodes, the $(m - k + 1)$ -th derivative of $S_m(x)$ can be discontinuous on the segment $[a, b]$. Basically, it takes $k = 1$. There is a representation (with a fixed grid Δ_n):

$$S_m(x) = S_{m,k}(x, \Delta_n) = \sum_{s=0}^m a_s x^s + \sum_{s=1}^{n-1} \sum_{r=m-k+1}^m a_{r,s} (x - x_s)_+^r, \quad (4)$$

where $(x - x_s)_+^r = [\max(0, x - x_s)]^r$ is a Peano's core.

The coefficients a_s and $a_{r,s}$ can take arbitrary values from R ; the set $S_{m,k}(x, \Delta_n)$ with a fixed Δ_n is linear with dimension $m+1+nk$. Therefore, for an unambiguous definition of a spline, it is necessary to specify $m+1+nk$ independent conditions. For a linear spline this is $n+2$, that is, for the statistics obtained, the number of parameters that need to be estimated is found from condition $r < n+2$.

Thus, under the conditions of the formulated problem, it is necessary on the segment $[a, b]$ for given r to find a grid $\{\tilde{x}_j\}_+^r$ such that the spline $S_{k+1}(x) \in C^k[a, b], k=0,1,\dots$, constructed on the grid Δ_r provides the minimum (in terms of standard quadratic deviations) statistical estimate of the vector A , that is, we assume that

$$\sum_{j=1}^r x_{i1} a_1 + \dots + x_{ir} a_r = S_{k+1}(x_i), \quad i = \overline{1, n}. \quad (5)$$

The equations written in the matrix form

$$X^* X A = X^* Y \quad (6)$$

are called normal. They are obtained by minimizing the sum of squared differences between observations and their mathematical expectations:

$$\begin{aligned} (Y - X A)^* (Y - X A) &= \sum_{i=1}^n [y_i - x_{i1} a_1 - \dots - x_{ir} a_r]^2 = \\ &= \sum_{i=1}^n [y_i - S_{k+1}(x_i)]^2 \end{aligned} \quad (7)$$

As the estimated parameters of the spline, the a_j ordinates of the nodes of the grid Δ_r are chosen, whose estimates are now determined by solving the normal equations (6) in the matrix form

$$A = (X^* X)^{-1} X^* Y. \quad (8)$$

The planning matrix generally has the following form

$$X = \begin{pmatrix} \bar{x}_{10} & \bar{x}_{11} & 0 & 0 & \cdot & 0 & 0 \\ \bar{x}_{20} & \bar{x}_{21} & 0 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{x}_{k0} & \bar{x}_{k1} & 0 & 0 & \cdot & 0 & 0 \\ 0 & \bar{x}_{k_1 11} & \bar{x}_{k_1 12} & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \bar{x}_{k_1 k_2 1} & \bar{x}_{k_1 k_2 2} & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \bar{x}_{n-k_r r-1} & \bar{x}_{n-k_r r} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \bar{x}_{nr-1} & \bar{x}_{nr} \end{pmatrix} \quad (9)$$

This matrix is rectangular, it has n rows and $r+1$ columns. The components of the vectors $\{k_j, j = \overline{1, r}\}$ characterize the number of observations that fall in the j -th interval, $\sum_{j=1}^r k_j = n$. It is assumed that $n \geq r+2$.

Based on (9), we obtain the matrix in the form

$$C = X^{(*)} X = \begin{pmatrix} c_{00} & c_{01} & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ c_{10} & c_{11} & c_{12} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & c_{21} & c_{22} & c_{23} & 0 & \cdot & 0 & 0 & 0 \\ 0 & 0 & c_{32} & c_{33} & c_{34} & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot & c_{r-2r-2} & c_{r-2r-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdot & c_{r-1r-2} & c_{r-1r-1} & c_{r-1r} \\ 0 & 0 & 0 & 0 & 0 & \cdot & 0 & c_{rr-1} & c_{rr} \end{pmatrix} \quad (10)$$

where

$$\begin{aligned} c_{00} &= \sum_{i=1}^{k_1} \bar{x}_{i0}^2 = \sum_{i=1}^{k_1} \frac{(\tilde{x}_1 - x_i)^2}{(\tilde{x}_1 - \tilde{x}_0)^2} = \frac{\sum_{i=1}^{k_1} (\tilde{x}_1 - x_i)^2}{(\tilde{x}_1 - \tilde{x}_0)^2}, \\ c_{jj} &= \sum_{i=1+s_{j-1}}^{s_j} \bar{x}_{ij}^2 + \sum_{i=1+s_j}^{s_{j+1}} \bar{x}_{ij}^2 = \sum_{i=1+s_{j-1}}^{s_j} \frac{(\tilde{x}_{j-1} - x_i)^2}{(\tilde{x}_j - \tilde{x}_{j-1})^2} + \sum_{i=1+s_j}^{s_{j+1}} \frac{(\tilde{x}_{j+1} - x_i)^2}{(\tilde{x}_{j+1} - \tilde{x}_j)^2}, \\ c_{j+1j} &= c_{jj+1} = \sum_{i=1+s_j}^{s_{j+1}} \bar{x}_{ij} \bar{x}_{i,j+1} = \sum_{i=1+s_j}^{s_{j+1}} \frac{(\tilde{x}_{j+1} - x_i)(x_i - \tilde{x}_j)}{(\tilde{x}_{j+1} - \tilde{x}_j)^2}, \quad j = \overline{0, r-1}, \\ c_{j+mj} &= c_{jj+m} = 0 \text{ at } m \geq 2, \quad j = \overline{0, r-2}, \end{aligned} \quad (11)$$

$$c_{rr} = \sum_{i=1+n-k_r}^n \bar{x}_{ir}^2 = \sum_{i=1+n-k_r}^n \frac{(\tilde{x}_{r-1} - x_i)^2}{(\tilde{x}_r - \tilde{x}_{r-1})^2},$$

and x_i are the elements of the source data (experimental points), $S_j = \sum_{m=1}^j k_m$,

$j = \overline{1, r}$, $s_0 = 0$, $s_r = n$. The matrix C is rectangular $(r+1) \times (r+1)$.

The above relations are obtained by directly multiplying the matrices $X^{(*)}$ and X with (9) taken into account.

Let us briefly discuss some properties of the matrix C .

1. The sum of the diagonal elements of the matrix C is equal to the sum of the squares of all the elements of the planning matrix X , that is,

$$\begin{aligned} \sum_{j=0}^r c_{jj} &= \sum_{i=1}^n \sum_{j=1}^r \bar{x}_{ij}^2 = \sum_{i=1}^n \left\{ \left(\frac{\tilde{x}_1 - x_{i1}}{\tilde{x}_1 - \tilde{x}_0} \right)^2 I_{x_{i1}} (\Delta x_1) + \right. \\ &\quad \left. + \left(\frac{x_{r-1} - x_{ir}}{x_r - x_{r-1}} \right)^2 I_{x_{ir}} (\Delta x_r) + \right. \\ &\quad \left. + \sum_{j=1}^{r-1} \left[\frac{x_{j+1} - x_{ij+1}}{x_{j+1} - x_j} I_{x_{ij+1}} (\Delta x_{j+1}) \frac{x_{j+1} - x_{ij}}{x_j - x_{j-1}} I_{x_{ij}} (\Delta x_j) \right]^2 \right\}. \end{aligned} \quad (12)$$

2. The rectangular matrix C has all sums of elements in the rows equal to 1.
3. The matrix C (and C^{-1}) is only a function of the sums $\sum x_i$ and $\sum x_i^2$, which are taken over all intervals and do not change if the individual x_i change their values in such a way that these sums remain unchanged.

The determinant of the matrix C is denoted by $|C| = |C|_{r+1}$, where the index $r+1$ is equal to the number of rows or columns of the matrix C .

If the determinants $|C|_{r-1}$ and $|C|_r$ are calculated, then the determinant $|C|_{r+1}$ is found by the formula

$$|C|_{r+1} = c_{rr} |C|_r - c_{rr}^2 |C|_{r-1} \text{ at } r = 2, 3, \dots \quad (13)$$

The recurrence formula (13) can be obtained by direct calculation, by decomposing the C row (or column) in $r+1$ minors.

Thus, if for generalization we denote $|C|_1 = c_{00}$, then to calculate the determinant of the matrix C , we obtain the following recurrence relations

$$\begin{cases} |C|_0 = 1; & |C|_1 = c_{00}; \\ |C|_{j+1} = c_{jj} |C|_j - c_{jj}^2 |C|_{j-1}; \end{cases} \quad j = \overline{1, r}. \quad (14)$$

It can construct recurrent formulas for finding the elements of the inverse matrix C . Let us briefly discuss the construction of these formulas. The determinant $|C|_{r+1}$ is calculated by the recurrent formula (14).

The algebraic complement of the element $c_{\min}$ which located on the main diagonal is equal to the multiplication of two determinants similar in structure to the determinant of the matrix and calculated by recurrent formulas (14), which in this case take the form

$$\begin{cases} |C|_0 = 1; & |C|_1 = c_{00}; \\ |C|_{j+1} = c_{jj} |C|_j - c_{jj-1}^2 |C|_{j-1}; & j = \overline{1, m-1}; \end{cases}$$

$$\begin{cases} |\overline{C}|_0 = 1; & |\overline{C}|_1 = c_{rr}; \\ |\overline{C}|_{j+1} = c_{r-jr-j} |\overline{C}|_j - c_{r-jr-j+1}^2 |\overline{C}|_{j-1}; & j = \overline{1, r-m-1}, \end{cases} \quad (15)$$

where $|\overline{C}|_j$ is the determinant of the matrix obtained with C by crossing out the first $r-j+1$ rows and first $r-j+1$ columns in it.

The recurrence formulas (15) allow to find the corresponding determinants of the matrices $|\overline{C}|_{r-m}$ moving in the direction of the main diagonal from the periphery inward towards the center of the matrix. These formulas can also be rewritten in another form, using the movement from the middle to the periphery, namely, taking into account (14) we have

$$\begin{cases} |C|_0 = 1; & |C|_1 = c_{m-1m-1}; \\ |C|_{j+1} = c_{m-1-jm-1-j} |C|_j - c_{m-1-jm-j}^2 |C|_{j-1}; & j = \overline{1, m-1}; \end{cases}$$

$$\begin{cases} |\overline{C}|_0 = 1; & |\overline{C}|_1 = c_{m+1m+1}; \\ |\overline{C}|_{j+1} = c_{m+1+jm+1+j} |\overline{C}|_j - c_{m+1+jm+j}^2 |\overline{C}|_{j-1}; & j = \overline{1, r-m-1}, \end{cases} \quad (16)$$

and now $|C|_j$ is the determinant of the matrix obtained by crossing out the first rows $j-1$ and columns in the C matrix and $r-m+1$ last rows and columns, and $|\overline{C}|_j$ is the determinant of the matrix obtained from the matrix C by crossing out $m+1$ first rows and columns and $r-j+1$ last rows and columns, $|C|_j$ and $|\overline{C}|_j$ are the main minors of the matrix.

It should be noted that the final results of calculations by the recurrence relation (15) and (16) for a fixed matrix are the same, and the intermediate ones may differ.

Thus, the algebraic complement of the element of the matrix C , standing on the main diagonal, is determined by the ratio

$$C_{jj} = |C|_j |\overline{C}|_{r-1}. \quad (17)$$

The ratio for calculating the algebraic complements of the elements of the matrix C that are not on the main diagonal is calculated by the ratio

$$\begin{cases} c_{ji} = (-1)^{(i-j)} \left\{ 1 + \text{sign}(i-j) \left[\prod_{p=j}^{i-2} C_{pp+1} - 1 \right] \right\} |C|_j |C|_{r-1}, & j \leq i = \overline{0, r}; \\ c_{ji} = c_{ij}, & r \geq 1. \end{cases} \quad (18)$$

Denote by H the matrix, which is determined by the expression

$$H = X^{(*)} Y. \quad (19)$$

where $X^{(*)}$ is the matrix transposed with respect to the planning matrix, and Y (matrix-column) is the output of the observations.

The elements of the matrix H are defined as follows.

$$\begin{aligned} h_0 &= \sum_{i=1}^{k_1} \bar{x}_{i0} y_i = \sum_{i=1}^{k_1} \frac{(\tilde{x}_1 - x_i) y_i}{\tilde{x}_1 - \tilde{x}_0}; \\ h_j &= \sum_{i=1+s_{j-1}}^{s_{j+1}} \bar{x}_{ij} y_i = \sum_{i=1+s_{j-1}}^{s_j} \frac{(\tilde{x}_{j-1} - x_i) y_i}{\tilde{x}_j - \tilde{x}_{j-1}} + \sum_{i=1+s_{j-1}}^{s_{j+1}} \frac{(\tilde{x}_{j+1} - x_i) y_i}{\tilde{x}_{j+1} - \tilde{x}_j}; \\ h_j &= \sum_{i=1+n-k_r}^n \bar{x}_{ir} y_i = \sum_{i=1+n-k_r}^n \frac{(\tilde{x}_{r-1} - x_i) y_i}{\tilde{x}_r - \tilde{x}_{r-1}}, \quad j = \overline{1, r-1}. \end{aligned} \quad (20)$$

Then the estimates of the ordinates of the points of connection of the linear parts of the spline (the elements of the A matrix-column) are determined by the formula

$$a_i = \sum \{C^{-1}\}_{ji} h_j, \quad i = \overline{0, r}. \quad (21)$$

The accuracy of approximation of the desired dependence using the selected spline is estimated by the sum of the squares of the deviations of the ordinates of the observation points from the found dependence

$$d = \sum_{j=1}^r \sum_{u=1+s_{j-1}}^{s_j} \left[\frac{a_j (x_u - \tilde{x}_{j-1}) + a_{j-1} (\tilde{x}_j - x_u)}{\tilde{x}_j - \tilde{x}_{j-1}} - y_u \right]^2. \quad (22)$$

The confidence interval for the found estimates of a_j is calculated by the formula

$$I_{\beta}^{(j)} = \left[a_j \mp \gamma_{\beta} \sqrt{\{C^{-1}\}_{jj} \frac{d}{n-r-1}} \right]. \quad (23)$$

where γ_{β} is a value that satisfies the relation $\mathbf{P}\{|t_{n-r-1}| \leq \gamma_{\beta}\} = 1 - \beta$, if the random variable t_{n-r-1} is distributed according to Student's law with $n-r-1$ degrees of freedom; d is the sum of the squares of the deviations of the observations y_i s from the values of the resulting spline at the corresponding points.

The prediction of the confidence interval for the values of the function $x(t)$ at the point $\tilde{x}_{r+1}$ is carried out by enumerating all the splines and choosing one that minimizes this interval at the prediction point. In this case, the spline itself, in the mean-square sense, is closer to the points, and is observed experimentally.

The relation (23) allows to build confidence intervals in each node of the spline, and on the whole interval - a confidence corridor.

To obtain a forecast using a statistical spline, an additional node is introduced to the set of nodes of the spline, the abscissa of which corresponds to the forecast interval. Using a computer search method, a grid is selected that satisfies equation (8) and at the same time minimizes it on the set of possible non-uniform grids and the width of the confidence interval in the forecast node.

As a result, the expected value and the confidence interval for the selected controlled parameter at the end of the forecast interval will be obtained. The boundaries of the confidence corridor are formed by linear interpolation of the upper and lower boundaries of the confidence intervals in all nodes of the resulting spline, including the forecast node.

2.2 Experimental results

Based on the considered spline forecast method, algorithms are built and a computer program is developed that implements this method. It is a computer program that is the main part of the proposed method, which makes it possible to carry out a short-term forecast for determining the time interval in which the temperature of the OPL can reach critical values. All the basic information about this program, practical issues of its application are set out in the work, as well as in special documentation.

The application of the proposed method for constructing a forecast will be considered on a specific example.

For the implementation of such a forecast, it was necessary, first of all, to obtain experimental statistical data on the temperature state of the wires of OPL. In addition, the initial data for the program are:

- the number of temperature measurements $n = 11$ (in our case);
- the number of hours of the forecast $n_{for} = 3$;
- the number of intervals, interpolates spline $r = 5$;
- the number of intervals of the time interval, on which the temperature is measured (the breakdown is performed in order to find the optimal interpolation spline), $n_{break} = 10$;
- confidence probability of estimating the prediction of the length of the time interval to reach the critical temperature of the power lines, $P = 0.95$.

As a result of the calculation of the program we obtain the data that are presented in Table 1 and in Fig. 1.

Table 1. Experimental data

Hours	9.0	9.5	10.0	10.5	11.0	11.5	12.0	12.5	13.0	13.5	14.0
t^o, C	40.5	39.5	40.5	45	47.5	47.5	50.0	49.0	51.0	55.0	57.0

In the top line of the Table 1 shows the hours in which the temperature measurement of power lines was carried out. The bottom line contains the corresponding temperature values.

The graph presented in Fig. 1 is a spline approximation in the form of a channel consisting of six sections. The spline docking points at the upper boundary of the channel are designated 1u - 7u (these points are denoted by the symbol Δ), and at the lower boundary 1d - 7d (the symbol is used to designate these points). In the main line of approximation, the notation $\square$ is selected for the designation of docking points.

The width of the channel is determined by the adopted confidence probability $P=0.95$. Sections 1 - 6 of the graph characterize the observation interval ΔT_{observ} , and sections 6 - 7 characterize the forecast interval $\Delta T_{\text{forecast}}$. The forecast was carried out at 3:00 ahead regarding to 14:00 hours, the last point from the observation interval. As can be seen from the graph, with an increase in the time interval of the forecast, its upper and lower limits expand, that is, the probability of the forecast decreases with increasing time.

The statistical spline (Fig. 1) is based on the results of observation of the values of the temperature of wires in OPL, which is in operation. Temperature measurement was carried out on the time interval $\Delta T_{\text{observ}} \in 9 \dots 14$ every 0.5 hours using a thermal imager mounted on the UAV. The quantitative values of the measured temperature, given in the bottom line of the table, are shown on the graph as black dots located on the observation interval ΔT_{observ} .

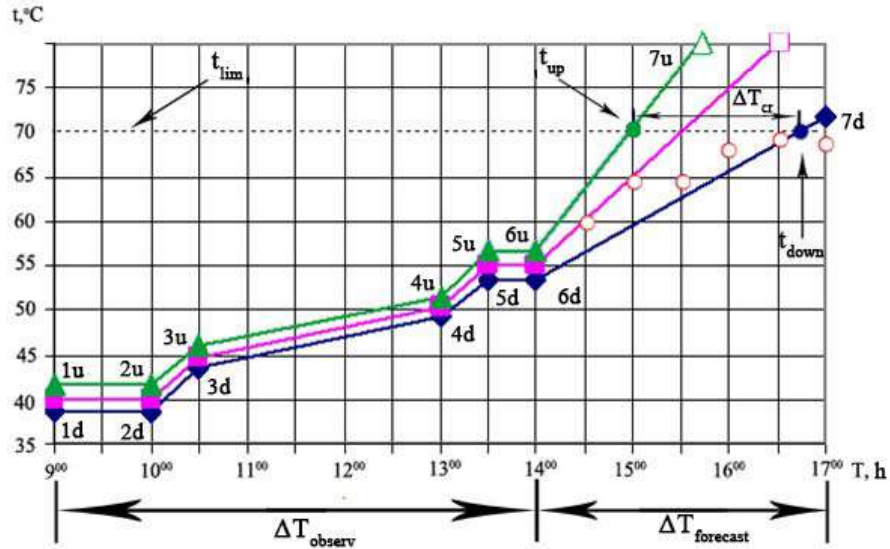


Fig. 1. Prediction using a statistical spline based on observations of temperature values of OPL

A spline was built (Fig. 1), divided into two sections, namely, the section where interpolation of the data on the measured temperature of the power transmission lines over the time interval $\Delta T_{\text{observ}} \in 9 \dots 14$ and section $\Delta T_{\text{forecast}} \in 14 \dots 17$ was carried out with the given probability ($P = 0.95$). The predicted value of the time interval ΔT_{cr} where the temperature of OPL can reach a critical level is determined by the

length of the time interval between the points of intersection of the temperature limit line with the upper t_{up} and lower t_d boundaries of the constructed spline. The threshold temperature value line $t_{lim}=70^{\circ}\text{C}$ is constructed parallel to the x-axis in accordance with the existing regulatory documents on the operation of OPL. In fig. 1 the predicted length ΔT_{cr} is indicated by a black line with arrows.

In order to verify the performance of the proposed method for predicting the possible values of the temperature of power lines, an experimental measurement of the temperature of these wires in the time period was carried out using the UAV, which corresponds to the forecast interval $\Delta T_{forecast}$. The data of the results of these measurements (6 points denoted by $\circ$) is plotted on a specified interval $\Delta T_{forecast}$ of the graph. Almost all (except for one point, the temperature value, which was measured at $T=17^{00}$), experimental data were obtained that did not exceed the upper and lower limits of the constructed spline forecast. This confirms the efficiency of the proposed method for predicting the values of the time interval in which the wires of OPL can reach unacceptable temperature values.

Based on the constructed spline forecast, the time interval when the power transmission lines can reach the maximum permissible temperature value of 70°C is between 3:00 pm and 4:45 pm with a confidence level of $P = 0.95$. It should also be emphasized once again that the achievement of such a temperature can occur only under constant conditions of operation of OPL.

During constructing a short-term spline forecast, it was assumed that the heating of the wires was caused by the random nature of changes in the load of OPL and the invariance of other (primarily meteorological) conditions during the entire observation interval.

3 Conclusions

It has been experimentally confirmed that using the method of statistical spline functions in combination with new information technologies implemented using UAV and maintaining unchanged meteorological conditions, allows a short-term forecasting of the time interval during which the OPL can reach critical temperatures with a given probability.

References

1. Velimirovic, L., Peric, Z., Stankovic, M.: Approximation of the optimal compressor function combining different spline functions in segments. 2013 11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services (TELSIKS), 16-19 October, 2013, Nis, Serbia. doi: 10.1109/TELSIKS.2013.6704436
2. Kabore, P., Wang, H.: A B-spline neural network based actuator fault diagnosis in nonlinear systems. Proceedings of the 2001 American Control Conference. (Cat No.01CH37148) 25-27 June, 2001, Arlington, VA, USA. doi: 10.1109/ACC.2001.945873
3. Zaporozhets, A.A., Eremenko, V.S., Serhiienko, R.V., Ivanov, S.A.: Development of an Intelligent System for Diagnosing the Technical Condition of the Heat Power Equipment.

- 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), 11-14 September, 2018, Lviv, Ukraine. doi: 10.1109/STC-CSIT.2018.8526742
4. Zaporozhets, A., Redko, O., Babak, V., Eremenko, V., Mokiychuk, V.: Method of indirect measurement of oxygen concentration in the air. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 2018, №5, pp. 105-114. doi: 10.29202/nvngu/2018-5/14
 5. Yao, L., Guan, Y.: Minimum entropy fault diagnosis and fault tolerant control for the non-Gaussian stochastic system. 2016 American Control Conference (ACC). 6-8 July, 2016, Boston, MA, USA. doi: 10.1109/ACC.2016.7526753
 6. Peng, X., Qian, G., Li, X., Gao, F.: Intelligent trip-out fault diagnosis of overhead transmission line. *High Voltage Engineering*, 2012, Vol. 38, Issue 8, pp. 1965-1972.
 7. Li, P., Li, Q., Cao, M., Gao, S., Huang, H.: Time series prediction for icing process of overhead power transmissions line based on BP neural networks. *Proceedings of the 30th Chinese Control Conference*. 22-24 July, 2011, Yantai, China.
 8. Kuts, Y., Eremenko, V., Monchenko, E., Protasov, A.: Ultrasound method of multi-layer material thickness measurement. *AIP Conference Proceedings* 1096, 2009, Vol. 1115. doi: 10.1063/1.3114079
 9. Zaporozhets, A., Eremenko, V., Serhiienko, R., Ivanov, S.: Methods and Hardware for Diagnosing Thermal Power Equipment Based on Smart Grid Technology, *Advances in Intelligent Systems and Computing III*, 2019, Vol. 871, pp. 476-492. doi: 10.1007/978-3-030-01069-0_34
 10. Babak, V., Mokiychuk, V., Zaporozhets, A., Redko, O.: Improving the efficiency of fuel combustion with regard to the uncertainty of measuring oxygen concentration. *Eastern-European Journal of Enterprise Technologies*, 2016, Vol. 6, №8, pp. 54-59. doi: 10.15587/1729-4061.2016.85408
 11. Klevtsov, S.: Identification of the state of technical objects based on analyzing a limited set of parameters. 2016 International Siberian Conference on Control and Communications (SIBCON), 12-14 May, 2016, Moscow, Russia. doi: 10.1109/SIBCON.2016.7491752
 12. Manusov, V.Z., Orlov, D.V., Frolova, V.V.: Diagnostic of Technical State of Modern Transformer Equipment Using the Analytical Hierarchy Process. 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), 12-15 June, 2018, Palermo, Italy. doi: 10.1109/EEEIC.2018.8493904
 13. Ming, J., Zhang, L., Sun, J., Zhang, Y.: Analysis models of technical and economic data of mining enterprises based on big data analysis. 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 20-22 April, 2018, Chengdu, China. doi: 10.1109/ICCCBDA.2018.8386516
 14. Dmitriev, S., Manusov, V., Ahyoev, J.: Diagnosing of the current technical condition of electric equipment on the basis of expert models with fuzzy logic. 2016 57th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 13-14 October, 2016, Riga, Latvia. doi: 10.1109/RTUCON.2016.7763126

The Methodology of Numerical Simulations of Separation Process in SPR-Separator

Oleksandr Liaposhchenko¹[0000-0002-6657-7051], Ivan Pavlenko¹[0000-0002-6136-1040],
Maryna Demianenko¹[0000-0002-4258-0379], Oleksandr Starynskyi¹[0000-0002-8437-2578]
and Jan Pitel²[0000-0003-1942-0438]

¹ Sumy State University, 2 Rymskogo-Korsakova St., 40007 Sumy, Ukraine

² Technical University of Kosice, 1 Bayerova St., 080 01 Presov, Slovakia
m.demianenko@omdm.sumdu.edu.ua

Abstract. The article describes the methodology of hydrocyclone SPR separator numerical simulation, which is used for oil lubricant and liquid impurities separation on technological lines of circulation lubrication of metallurgical machines. The methodology of hydrocyclone SPR separator numerical simulation was developed with taking into account up-to-date articles about the study of hydrocyclone equipment. It was done for determining the relation between the main structural dimensions of the device and the flow parameters, as well as the separation efficiency of the heterogeneous inlet mixture. Numerical simulations were performed by using ANSYS namely its module Fluent Flow. The methodology of the finite volume computational grid building was described. This methodology takes into account the geometry features of the model, such as the presence of surfaces curvature and small gaps compared to other linear dimensions. Settings of the Fluent module was described with design features of the SPR-separator and its working parameters. Eulerian multiphase flow model, k-ε RNG turbulence model with Swirl Dominated Flow settings was chosen. According to the results of the numerical simulation was determined the main hydrodynamic characteristics of the flow which presented as isolines in the middle section of the separator and value of the separation efficiency of the heterogeneous mixture. New design of a spiral nozzle was proposed for expansion of the range of effective work.

Keywords: Hydrocyclone, Heterogeneous Mixture, Research Methodology, Numerical Experiment, CFD.

1 Introduction

Nowadays, circulating oil lubrication systems are widespread in the metallurgical equipment. According to this method, the liquid lubricant moves through oil pipes, which form a closed ring circuit, lubricates the units and cools of the friction pairs. It should be noted that the lubricant insignificant amount loose during its movement. Therefore, the main advantage of the circulating oil lubrication systems is their economic efficiency. One of the required element of the circulating oil lubrication sys-

tems is separation devices since oil lubricant become contaminated with liquid and solid particles after passing the circuit oil pipes. For this purpose, hydrocyclones are widely used.

By one of the highly efficient hydrocyclones are considered to be SPR-separators with the spiral nozzle (turbo-spiral), which is designed for the incoming flow directing in a radial way and as a result creating of centrifugal forces. The main advantage of the SPR-separator is the high fluid speed in the enclosure of the centrifugal element, which provides a high-intensity separation process. On the other hand, we can observe that high fluid speed may cause a negative influence on the separation process, such as re-mixing of the dispersed and the continuous phase. Therefore, technological regimes causing the above phenomena must be eliminated at the design stage to ensure reliable and efficient operation of this unit.

The main aim of this work is relations determining between the main structural dimensions of the SPR-separator (Fig. 1) and the flow parameters, as well as the separation efficiency of the heterogeneous inlet mixture.



Fig. 1. The scheme of the SPR-separator with the main flows [1].

2 Literature Review

It should be noted that hydrocyclones have commonly used equipment in various technological lines of chemical and food industries [1–3], at oil production from wells [4–6], for gas purification in titanium dioxide productions [7, 8], etc. Thus, first and foremost, the equipment characteristics are determined depending on its purpose. At

the same time, the development of new hydrocyclone designs continue by creating favorable hydrodynamic regimes for separating [1–5], adding new propelling power of the process in addition to centrifugal [7] and methods of them manufacturing [5, 7–10], as well as their new appointments opening [11].

In each of the above works, the flow parameters were determined by experimental [3, 5, 8, 10–14] and numerical methods [1, 2, 4, 7–10, 15, 16]. Herewith, separation efficiency is determined by flow parameters. A significant advantage of the numerical methods is the ability to studying of a large number of hydrocyclone designs as a consequence determining of the basic design parameters, which effect on the separation efficiency of the input heterogeneous mixture, without significant capital expenditures associated with the creation of experimental samples. In works [1, 2, 4, 7–10, 15, 16] various software packages were used for carrying out numerical calculations among which ANSYS, OPENFoam, etc. They are based on the finite volume method.

For example, in works [1, 4, 7, 10, 15, 16] the ANSYS software package namely its module Fluent Flow was chosen for simulation. Wherein, various models of multiphase mixtures were defined depending on the medium under study. The Mixture model was selected as the model of a multiphase flow in [6] for carrying out numerical studies of magnetic hydrocyclon. In other cases [1, 15–17], the VOF model was chosen. Its feature is the solution of the momentum equation and temperatures determining for each phase separately.

Most scientists for describing of turbulent flows in hydrocyclone enclosure prefer the Reynolds Stress Model (RSM) since this model gives the best correlation with experiment for swirling flows. Thus, in works [6, 9, 15, 17] this model was chosen. Herewith, it is resource-intensive, so in [4] the Swirl model RNG k- ϵ turbulence was used as an alternative.

In summary, the ANSYS Workbench software package, namely its module Fluent Flow was chosen for carrying out numerical calculations on flow parameters in the SPR separator since it has different models for describing both the multiphase flow and the turbulence.

3 Research Methodology

As mentioned above, the ANSYS Workbench software package, namely its Fluent Flow module was chosen to determine the flow characteristics in the SPR-separator under study. The first stage of calculations in this module is the building of a three-dimensional design geometry, which is used for determining the boundary conditions as shown in Fig. 2.

As can be seen from Fig. 2, the boundary condition of the “wall” type was not specified since the Fluent Flow module automatically assigns it to surfaces, which aren’t defined as other boundary conditions.

The second stage of the calculation was the building of the finite volume grid, which takes into account the features of the model, such as the presence of surfaces curvature and small gaps comparatively to other linear dimensions. To that end, the global setting Proximity and Curvature was chosen for considering the above model

features. The number of elements in the gap 5 was additionally specified by the Proximity Size Function – Face and Edges. According to the above function, 5 elements are built always, even as the main specified element sizes deny building a required number of elements on the surfaces/faces. The boundary layer was built on the surfaces, which aren't specified as inlet and outlet, for ensuring a qualitative description of the flow near the wall. The Inflation Option – Last Aspect Ratio was used for reducing the height of subsequent layers as appropriate (for example, in the gap), with First Layer Heights equivalent to 50 μm . The Advanced Options namely Collision Avoidance – Layer Compression were also used. It allows reducing the height of the first layer in the gaps and prevents Stair Stepping. A volumetric computational grid with 1.5 million elements was obtained as a result of the above settings using. This grid has following quality indicators: the maximum Skewness value is 0.93, which is not exceed the permissible value 0.95; the minimum Orthogonal Quality value is 0.13, which does not less than the permissible value 0.1; Aspect Ratio – 809, which is satisfactory when using Double Precision. Last indicated setting is necessary for multi-phase flow calculations. The computational grid that was obtained is shown in Fig. 3.

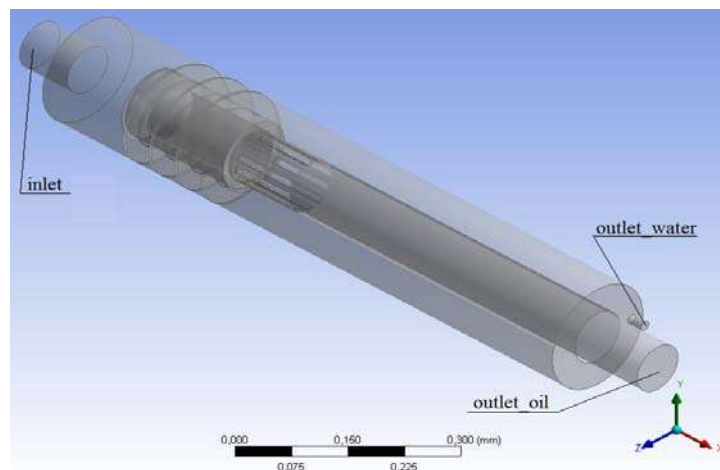


Fig. 2. Three-dimensional design model, with surfaces for specifying boundary conditions.

The next (third) stage of the numerical calculations is the setting of the Fluent Flow module, which begins with the introduction of General Setting, namely Double Precision and Gravity. First activated setting is needed for multiphase flows, second - for taking into account the gravity force, which is necessary for consideration draining of the separated liquid film directed vertically downwards (from the inlet to the outlet) along the X-axis. It makes sense to use the Eulerian approach for description of multiphase flow since we consider water particles with an average size of 100 μm , their volume fraction of dispersed phase is 0.1. Thus, the Eulerian model with Formulation – Implicit was chosen. The k- ϵ RNG turbulence model was chosen with the following settings: Swirl Dominated Flow, Curvature Correction and Near Wall Treatment – Enhanced Wall Treatment. It was done accordingly to the value of the Rey-

nolds number $2.9 \cdot 10^4$ and from the viewpoint of reducing the computation resource-intensiveness.

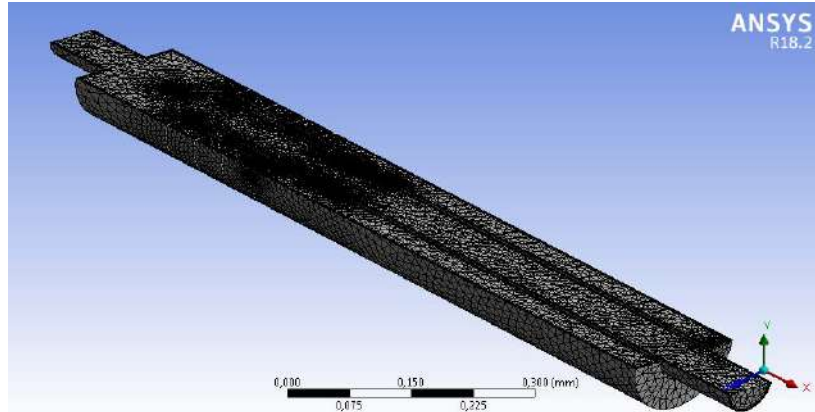


Fig. 3. The obtained finite-volume computational grid

The necessary stage was materials definition for the phases from the library:

- primary phase – kerosene-liquid, with properties: density $\rho = 780 \text{ kg/m}^3$, dynamic viscosity $\mu = 0.0024 \text{ Pa}\cdot\text{s}$;
- secondary phase – water liquid, with properties: density $\rho = 998 \text{ kg/m}^3$, dynamic viscosity $\mu = 0.001003 \text{ Pa}\cdot\text{s}$.

The following boundary conditions are specified at the inlet:

- for the mixture – a gauge pressure of 0.4 MPa, turbulence parameters such as hydraulic diameter 76 mm and the average value of turbulence intensity – 5 %;
- for the main phase (kerosene-liquid) - inlet speed 1.18 m/s;
- for the dispersed phase (water) - volume fraction 0.1.

The following boundary conditions are specified at the lubricant outlet:

- for the mixture – a gauge pressure of 0.4 MPa, turbulence parameters such as hydraulic diameter 76 mm and the average value of turbulence intensity – 5 %;
- for the dispersed phase (water) - volume fraction in the possible reverse flow 0.1.

As the boundary conditions at the water outlet are given:

- for the mixture – a gauge pressure of 0.4 MPa, turbulence parameters such as hydraulic diameter 15 mm and the average value of turbulence intensity – 5 %;
- for the dispersed phase (water) - volume fraction in the possible reverse flow 1.

Standard initialization was used for calculations. Firstly, the problem was solved by the Method-Velocity Coupling - Simple method with the following Spatial Discretization parameters: Gradient – Least Square Cell Based; Momentum – First Order

Upwind; Volume Fraction – First Order Upwind; Turbulent Kinetic Energy – First Order Upwind; Turbulent Dissipation Rate – First Order Upwind. The Pressure-Velocity Coupling-Coupled method was used after problem convergence for the improvement of the results accuracy, wherein following Spatial Discretization parameters were used: Gradient-Least Square Cell Based; Momentum – Second Order Upwind; Volume Fraction – Second Order Upwind; Turbulent Kinetic Energy – First Order Upwind; Turbulent Dissipation Rate – First Order Upwind.

4 Results

The isolines of a dispersed phase in the cross-section of the unit (Fig. 4), the isolines of gauge pressure (Fig. 5) and the velocity vectors of water (Fig. 6) were obtained as a result of the calculations.

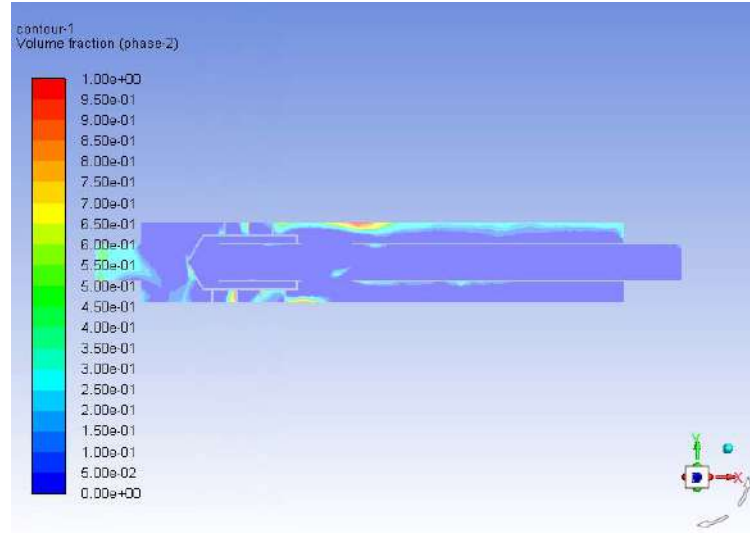


Fig. 4. The isolines of the water volume fraction in the middle section of the SPR-separator.

Based on the results of the SPR-separator hydrodynamics numerical simulation, namely isolines and velocity vectors of the dispersed phase, it can be argued, that the proposed modeling methodology corresponds to the theoretical concepts of the separation of two-component streams in the centrifugal field. Fig. 4 shows that the component with a large value of density (water) is located near to the walls and is absent in the central part of the apparatus. Fig. 6 shows the velocity vectors of the dispersed phase, which are directed from the axis to the periphery and along to the walls of the SPR-separator. This indicates that the separation process is efficient. The value of the separation efficiency ($\eta = 95 \%$) was calculated according to the equation:

$$\eta = \left(1 - \frac{c}{c_0}\right) \cdot 100\%, \quad (1)$$

c_0 – volume fraction of the secondary phase at the inlet; c – volume fraction of the secondary phase at the outlet of the SPR-separator.

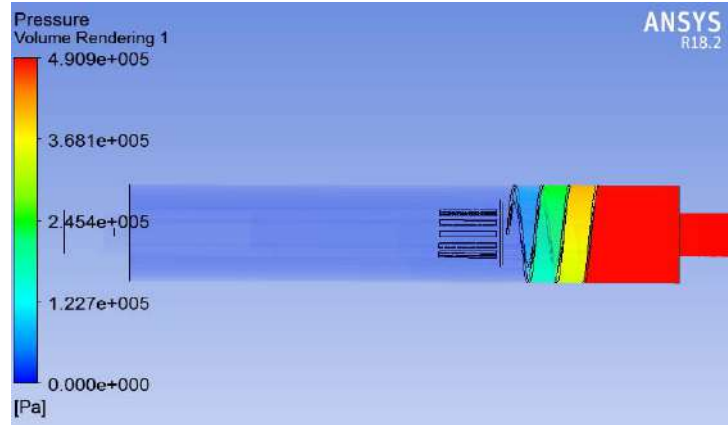


Fig. 5. Isolines of the gauge pressure.

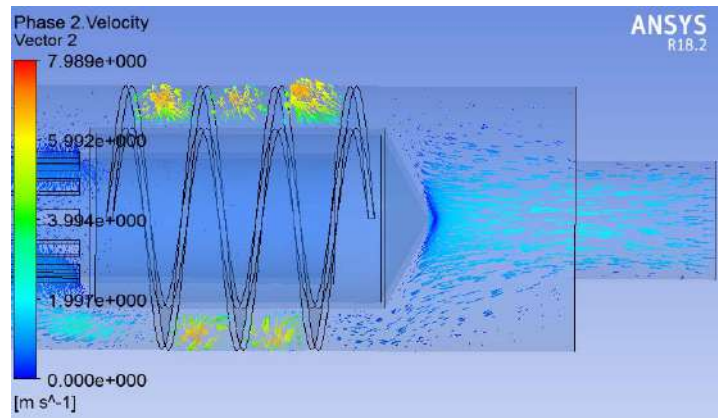


Fig. 6. Water velocity vectors.

As mentioned above, the separation efficiency value is 0.95. However, this value was determined for the following working parameters: velocity at the inlet – 1.18 m/s, gauge pressure at the inlet and outlet – 0.4 MPa, discrete phase volume fraction at the inlet – 0.1, but high fluid speed may cause a negative influence on the separation process, such as re-mixing of the dispersed and the continuous phase, and on the other hand low fluid speed may cause insufficient centrifugal force for separation. Similar problem takes place at process of a gas-liquid separation, which is based on the usage of inertia of particles (droplets). The main drawback of this method is the secondary

splashing, resulting from a possible increasing of the flow velocity to a critical value and disruption of the trapped liquid, with following occurrence of the highly dispersed particles [19]. The avoidance of this problem is realized in dynamic separation elements. The related certificates of the authorship are “The method of capturing highly dispersed dropped liquid from the gas-liquid flow” Sumy State University, Ukraine, bulletin No. 20 with No. u201505124 and No. u201605061. The above elements work as an automatic control system, in which the regulating action is elastic forces, and the object of regulation is the hydraulic resistance.

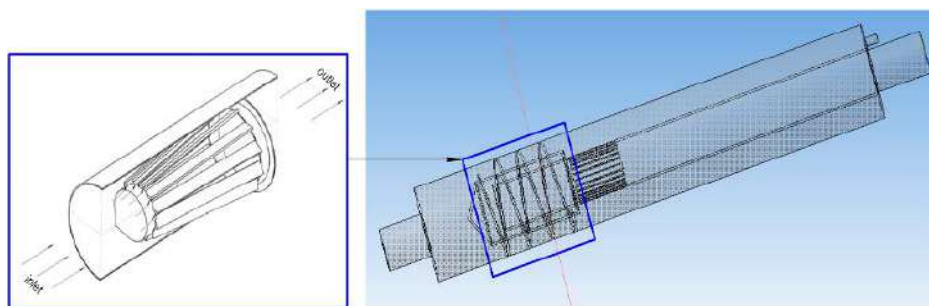


Fig. 6. New design of the spiral nozzle.

New design of the spiral nozzle was proposed and it shown in Fig. 7. This spiral nozzle design was developed accordingly to the results of numerical simulations and based on similar principles of the dynamic separation elements. The related certificate of the authorship is “The device for separation of disperse liquid from gas flow” Sumy State University, Ukraine, bulletin No. 23 with No u 2018 06182.

For future research main aim will be developing of engineering method for calculation of new SPR-separator design. Herewith, main problem is necessarily of solving a complex problem of the hydroaeroelasticity, which analytical solution has been already obtained for similar case in [19]. The articles [20–22] will be used for father improving of SPR-separator, due to creating regularly-structured elements with well-organized liquid drainage [20, 21] and mounting filtering or their thin-layer modules [22].

5 Conclusions

The method of numerical research of the SPR-separator is proposed. This technique allows to determining the main hydrodynamic parameters of the separation processes, the efficiency of separation, hydraulic resistance, value and direction of the velocities of the disperse phase, etc.

The efficiency of separation was calculated, which in turn confirms the adequacy of the proposed methodology and the possibility of using it for future research. Main disadvantages of separation process in SPR-separator are re-mixing of the dispersed and the continuous phase, insufficient centrifugal force for separation. New design of

the spiral nozzle was developed accordingly to the results of numerical simulations and based on similar principles of the dynamic separation elements. For future research main aim will be developing of engineering method for calculation of new SPR-separator design.

6 Acknowledgments

Results of the research were achieved within the project “Development and implementation of energy efficient modular separation devices for oil and gas purification equipment” (Ministry of Education and Science of Ukraine, State Reg. No. 0117U003931) due to the close cooperation between the Department of Process and Equipment of Chemical and Petroleum-Refineries and the Department of General Mechanics and Machine Dynamics of the Faculty of Technical Systems and Energy-Efficient Technologies of Sumy State University.

Numerical simulations using ANSYS software were provided by the Faculty of Manufacturing Technologies with a seat in Presov of Technical University of Kosice within the research project “Identification of Parameters for Technological Equipment using Artificial Neural Networks” supported by the National Scholarship Programme of the Slovak Republic.

References

1. Xu, Y., Tang, B., Song, X., Sun, Z., Yu, J.: Simulation analysis on the separation characteristics and motion behavior of particles in a hydrocyclone. *Korean Journal of Chemical Engineering* 35(12), 2355–2364 (2018). <https://doi.org/10.1007/s11814-018-0171-0>
2. Razmi, H., Goharrizi, A. S., Mohebbi, A.: CFD simulation of an industrial hydrocyclone based on multiphase particle in cell (MPPIC) method. *Separation and Purification Technology* 209, 851–862 (2019). <https://doi.org/10.1016/j.seppur.2018.06.073>
3. Vehmaanpera, P., Safonov, D., Kinnarinen, T., Hakkinen, A.: Improvement of the filtration characteristics of calcite slurry by hydrocyclone classification. *Materials Engineering* 128, 133–140 (2018). <https://doi.org/10.1016/j.mineng.2018.08.042>
4. Im, I.-T., Gwak, G. D., Kim, S. M., Park, Y. K. A numerical study of the flow characteristics and separation efficiency of a hydrocyclone. *KSCE Journal of Civil Engineering* 22(11), 4272–4281 (2018). <https://doi.org/10.1007/s12205-018-1780-1>
5. Shi, S.-Y., Xu, J.-Y. Flow field of continuous phase in a vane-type pipe oil-water separator. *Experimental Thermal and Fluid Science* 60, 208–212 (2015). <https://doi.org/10.1016/j.expthermflusci.2014.09.011>
6. Lin, L., Zhao, L., Yang, X., Wang, Y., Xu, B., Liang, B.: Innovative design and study of an oil-water coupling separation magnetic hydrocyclone. *Separation and Purification Technology* 213, 389–400 (2018). <https://doi.org/10.1016/j.seppur.2018.12.051>
7. Vega-Garcia, D., Brito-Parada, P. R., Cilliers, J. J.: Optimising small hydrocyclone design using 3D printing and CFD simulations. *Chemical Engineering Journal* 350, 653–659 (2018). <https://doi.org/10.1016/j.cej.2018.06.016>

8. Huang, L., Deng, S., Guan, J., Hua, W., Chen, M.: Separation performance of a novel liquid-liquid dynamic hydrocyclone. *Industrial and Engineering Chemistry Research* 57, 7613–7623 (2018). <https://doi.org/10.1021/acs.iecr.8b00137>
9. Kyriakidis, Y. N., Silva, D. O., Barrozo, M. A. S., Vieira, L. G. M.: Effect of variables related to the separation performance of a hydrocyclone with unprecedented geometric relationships. *Powder Technol* 338, 645–653 (2018). <https://doi.org/10.1016/j.powtec.2018.07.064>
10. Li, Y., Liu, C., Zhang, T., Li, D., Zheng, L.: Experimental and numerical study of a hydrocyclone with the modification of geometrical structure. *The Canadian Journal of Chemical Engineering* 96, 2638–2649 (2018). <https://doi.org/10.1002/cjce.23206>
11. Pecarevic, M., Mikus J., Prusina, I., Juretic, H., Cetinic, A. B., Brailo, M.: New role of hydrocyclone in ballast water treatment. *Journal of Cleaner Production* 188, 339–346 (2018). <https://doi.org/10.1016/j.jclepro.2018.03.299>
12. Crowley, S. V., Molitor, M. S., Kalscheuer, R., Lu, Y., Kelly, A. L., O'Mahony, J., Lucey, J. A.: Size classification of precipitated calcium phosphate using hydrocyclone technology for the recovery of minerals from deproteinised acid whey. *International Journal of Dairy Technology* 72(1), 142–151 (2018). <https://doi.org/10.1111/1471-0307.12570>
13. Chang, Y.-L., Wang, H.-L., Jin, J.-H., Liu, Z.-M., Lv, W.-J.: Flow distribution and pressure drop in UZ-type mini-hydrocyclone group arranged in compact parallel manifolds. *Experimental Thermal and Fluid Science* 100, 114–123 (2019). <https://doi.org/10.1016/j.expthermflusci.2018.07.014>
14. Lv, W.-J., Chen, J.-Q., Chang, Y.-L., Liu, H.-L., Wang, H.-L.: UU-type parallel mini-hydrocyclone group separation of fine particles from methanol-to-olefin industrial wastewater. *Chemical Engineering Processing – Process Intensification* 131, 34–42 (2018). <https://doi.org/10.1016/j.cep.2018.03.015>
15. Xu, Y., Tang, B., Song, X., Yu, J.: A high-efficiency hydrocyclone designed by response surface methodology for acid hydrolysis residue recycling. *Royal Society Open Science* 6, 172339 (2019). <https://doi.org/10.1098/rsos.172339>
16. Matvienko, O. V., Andropova, A. O.: Separation of nonspherical particles in a hydrocyclone. *Journal of Engineering Physics and Thermophysics* 91(3), 712–730 (2018). <https://doi.org/10.1007/s10891-018-1794-z>
17. Yue, T., Chen, J., Song, J., Chen, X., Wang, Y., Jia, Z., Xu, R.: Experimental and numerical study of upper swirling liquid film (USLF) among gas-liquid cylindrical cyclones (GLCC). *Chemical Engineering Journal* 358, 806–820 (2019). <https://doi.org/10.1016/j.cej.2018.10.018>
18. Liaposhchenko, O. O., Sklabinskyi, V. I., Zavialov, V. L., Pavlenko, I. V., Nastenkenko, O. V., Demianenko, M. M.: Appliance of inertial gas-dynamic separation of gas-dispersion flows in the curvilinear convergent-divergent channels for compressor equipment reliability improvement. In: *IOP Conference Series: Materials Science and Engineering*, vol. 233, 012025 (2017). <https://doi.org/10.1088/1757-899X/233/1/012025>
19. Pylypaka, S., Klendiy, M., Zaharova, T.: Movement of the particle on the external surface of the cylinder, which makes the translational oscillations in horizontal planes. In: *Advances in Design, Simulation and Manufacturing. DSMIE 2018. Lecture Notes in Mechanical Engineering*. Springer, Cham, pp. 336–345 (2019). https://doi.org/10.1007/978-3-319-93587-4_35
20. Liaposhchenko, O., Nastenkenko, O., Pavlenko, I.: The model of crossed movement and gas-liquid flow interaction with captured liquid film in the inertial-filtering separation chan-

- nels. *Separation and Purification Technology* 173, 240–243 (2017). <https://doi.org/10.1016/j.seppur.2016.08.042>
21. Liaposhchenko, O., Khukhryanskiy, O., Moiseev, V., Ochowiak, M., Manoilo, E.: Intensification of foam layered apparatus by foam stabilization. *Journal of Engineering Sciences* 5(2), F13–F18 (2018). [https://doi.org/10.21272/jes.2018.5\(2\).f3](https://doi.org/10.21272/jes.2018.5(2).f3)
 22. Sklabinskyi, V., Liaposhchenko, O., Pavlenko, I., Lytvynenko, O., Demianenko, M.: Modelling of liquid's distribution and migration in the fibrous filter layer in the process of inertial-filtering separation. In: *Advances in Design, Simulation and Manufacturing. DSMIE 2018. Lecture Notes in Mechanical Engineering*. Springer, Cham, pp. 489–497 (2019). https://doi.org/10.1007/978-3-319-93587-4_51

Building a Spatial Model of Destructive Processes Based on Fuzzy Rough Soft Topology

Maryna Zharikova^[0000-0001-6144-480X], Volodymyr Sherstjuk^[0000-0002-9096-2582]

Kherson National Technical University, 24 Berislav Road, Kherson, 73008, Ukraine
marina.jarikova@gmail.com, vgsherstyuk@gmail.com

Abstract. This work presents a spatial model for the real-time GIS-based decision support systems based on dynamic fuzzy rough soft topology, which represents a spatial structure that contains a multitude of interacting processes, which evolve in space and time. The dynamics of destructive processes are modeled using the spread model. The area of interest is represented as an approximation by a grid of cubic cells. This allows taking into account the peculiarities of the initial information obtained using remote sensing techniques and having a significant uncertainty. As a result, boundaries of contours of spreading destructive processes are blurred using fuzzy rough soft topology. The proposed model reduces the computational complexity and provides the acceptable performance.

Keywords: destructive processes, spatial model, fuzzy rough soft set, fuzzy-rough soft topology, grid of cells, blurred boundaries

1 Introduction

Nowadays, the society faces the problem of increasing loss of lives and damages to properties caused by natural disasters that rise steadily due to population growth, urbanization, deforestation, environmental degradation, and global climate change. An effective way to overcome this problem is a proper risk management strategy that calls for disaster analysis consisting of spatiotemporal modeling of disaster in the area of interest (AOI).

The authors are concerned with the areas containing natural and artificial objects among which are valuable objects requiring disaster protection. AOI with a multitude of interacting disasters, which evolve in space and time giving rise to danger and risk to some valuable objects is considered a dynamic system. The paper deals with real-time disaster spatial modeling.

However, the most of the destructive processes are poorly observed and their spreading within the AOI is weakly modeled, so real-time disaster modeling is a complex and non-trivial task, which becomes more complicated due to uncertainty of information, a wide geographically distribution of events and, as usual, a lack of time [1]. The efficiency of decision-making strongly depends on the availability of online disaster monitoring tools aimed at the real-time computation of the most important parameters related to the spreading of the destructive processes.

Today, a suite of the most advanced methods and techniques, such as remote sensing, GIS, geospatial analysis, unmanned aerial vehicles (UAV), can be synergistically used for GIS-based disaster modeling. Remote sensing techniques play a crucial role, as they provide powerful tools for the rapid acquisition of relevant data for disaster monitoring [2] in a form of streams of great volumes that come from sensors on a continuous basis at a high rate and should be analyzed in a real-time [3]. These data can be used for representing spatial distribution and properties of disaster, and for supporting forecasting disaster models.

This paper presents a spatiotemporal disaster model in the context of the most common types of disasters such as wildfires. The authors consider fire monitoring as a continuous or discrete process of observing a status and changes in an active fire directly or indirectly and determining some fire parameters such as intensity, size, the rate of spread, and others relevant to respond operation and important for the decision maker. UAVs can effectively perform long-time missions to obtain remote sensing data [4]. However, due to the instrumental inaccuracy and distortions caused by vibrations, remote sensing information obtained from UAVs is incomplete, imprecise, vague, and often blurred [5]. The dynamics of wildfire spreading depends on the accuracy of determining the boundaries of its dynamic contour. However, the uncertainty of observations significantly reduces the accuracy of determining the boundaries of such contours [6]. Obtained remote sensing data should be correctly transferred to wildfire spread model, geolocated and mapped to the AOI.

Numerous methods have been developed to model disasters based on a huge array of remote sensing and other data gathering techniques. Using the well-established traditional approaches for spatial modeling such as statistical methods do not provide the required performance and acceptable efficiency of GIS-based real-time wildfire modeling [7]. A key aspect to achieve the desired performance is to build an approximate spatial model of wildfire spreading, taking into account partial observability and uncertainty of observations. Thus, we need to soften the requirements for the accuracy of remote sensing data representation, which will give us the opportunity to improve modeling performance. In this case, the boundaries of the dynamic contours of the spreading processes can be vague and blurred.

There are several well-known approaches to deal with the uncertainty and vagueness in the spatial models, such as fuzzy set theory [8], rough set theory [9] and soft set theory [10]. Each of these approaches has its inherent difficulties as pointed out in [10]. It should be noted that due to the absence of some important information a priori, such as membership functions for fuzzy sets, equivalence relations for rough sets, or parameterizations for soft sets, these approaches cannot ensure the adequacy of the spatial model of the destructive process independently. Therefore, many researchers combine some of these approaches. Some authors proposed to use for spatial modeling the combinations of rough and fuzzy sets [1], rough and soft sets [11], fuzzy and rough sets [12]. In [13], the authors proposed the concepts of rough fuzzy soft sets and fuzzy rough soft sets, which have a number of advantages to build a blurred spatial model. Based on this, we can use soft topological spaces to build a spatial model of the destructive process, as well as the fuzzy rough method for its blurring.

The aim of this work is to develop the approximate spatiotemporal disaster model

within AOI [14] in the context of forest fires. To overcome the computational complexity problem, we build a topological spatial model and soften the effects of discretization using the fuzzy rough sets. The developed model allows analyzing big data streams coming from remote sensors and representing them in a user-friendly style.

2 Modeling dynamics of destructive processes

Let us consider the AOI as an open connected subspace X of three-dimensional Euclidean space endowed with the topological properties [15, 16]. Firstly, the considered AOI is divided into a finite set of disjoint spatial objects represented as geometric shapes, which outline boundaries of certain areas. Such objects are named as geotaxons and represent geo-referenced natural parts of the terrain with the same characteristics. GIS can contain an unlimited number of geotaxons' layers. To build a topological space on X we use an equivalence relation $\mathfrak{R}_X \subseteq X \times X \times X$ (reflexive, symmetric, and transitive) [1]. Then the pair $apr_X = (X, \mathfrak{R}_X)$ is called the approximation space. The family of all composite sets is denoted by $Def(apr_X)$ and uniquely determines the topological space $T = (X, Def(apr_X))$.

Suppose that each point $x \in X$ has a non-empty finite set of attributes A , V_a is a domain of $a \in A$ and f is a function such that $f : X \times A \rightarrow V$. Let's impose a metrical grid of coordinate lines with $\delta = \Delta\alpha_1 = \Delta\alpha_2 = \Delta\alpha_3$ within X , which form a set C of cubic cells with the size being $\delta \times \delta \times \delta$. Thus, space X is discretized by a grid C of isometric cubic cells $c \in C$. Assume that a cell $c \in C$ is a spatial homogeneous object of minimal size. The grid C approximates AOI and constitutes a certain GIS layer.

Each cell $c \in C$ is associated with a set of attribute values, which is called the cell state, via the value function $f(c, A)$. The proposed discretization assigns equal values of the attributes to each point belonging to a certain cell c , therefore each cell $c \in C$ represents a homogeneous area of the AOI in terms of attribute values A , so it can be reduced to a point of X . It's suggested to model disaster dynamics by means of a change of states of the cells covered by the disaster.

Suppose the set of attributes A can be divided into subsets [16]: not changing over time (static) attributes A_s , time-varying (dynamic) attributes A_d , slowly changing (environmental) attributes A_e , $A = A_s \cup A_d \cup A_e$. Suppose $W = \{w_0, \dots, w_i, \dots, w_F\}$ is an ordered set of the cell states (phases), where w_0 is the initial phase, w_F is the final phase, and each w_i is the transitional phase. We consider each significant change of the cell attribute's value, which forces the cell to change its state, as an event. Assume, during the destructive process, the cell moves through a sequence of qualitatively different categories of states, which should be evaluated during continuous remote sensing. It is clear that the model of the destructive process can be represented as a model of dynamic change of states of a subset of cells covered by the process within the spatial model. Thus, the spatiotemporal structure of AOI can be represented as a topological space, which includes subspaces of cells of the same phase and makes it possible to assess the position and boundaries of the dynamic contour of the

process. Since the belonging of each cell to a certain phase is determined approximately due to the uncertainty of remote sensing, the topological space describing the structure of the spatial model as well as the boundaries of the contour of the spreading destructive process is blurred.

3 Soft sets and soft topologies

Dynamics of destructive processes can be described by a blurred structure of AOI containing the sets of cells, which belong to a certain phase at any given time t . It is proposed to represent such a structure as a soft set, which blurring in different ways makes it possible to obtain blurred structures. The most common way of blurring the boundaries between subsets of cells belonging to different phases is to represent them in the form of a fuzzy set, but in practice, this method is impossible to implement. The soft set allows us to represent the AOI as a blurred topological space, which can be created by blurring the boundaries between the sets of cells corresponding to different phases.

Suppose destructive processes at each time gives rise to a certain state of AOI represented by a blurred structure, which consists of plausible sets of cells that belong to a certain phase. Such a structure is proposed to be presented as a soft set.

Consider the concept of soft sets in general. Let $W = W_S \cup W_D$ is a union of the sets of environmental conditions and the set of phases of the cells, $W_i \subseteq W$ is any of its subsets, and 2^W is the set of all subsets of W , $W_i \subseteq 2^W$ [16].

A couple $\Upsilon_{W_i} = (\Upsilon, W_i)$ is called a soft set on the set of cells C if Υ is a mapping $\Upsilon: W_i \rightarrow 2^C$, where 2^C is the set of all subsets of C [16]. In other words, the soft set is a parameterized family of subsets of cells C . Each set $\Upsilon(w)$, $w \in W_i$ of this family can be considered as a set of w -elements of soft sets (Υ, W_i) .

A soft set can be defined by a plurality of pairs $\Upsilon_{W_i} = \{(w, \Upsilon_{W_i}(w)) : w \in 2^W, \Upsilon_{W_i}(w) \in 2^C\}$. A set $\Upsilon_{W_i}(w)$ is called a w -element of the soft set and is determined for each $w \in W_i$. The soft set is associated with a set of equivalence classes generated by the Pawlak's indiscernibility relation. As mentioned above, we can identify an A_i -indiscernibility relation in the set of cells.

If a certain set of parameters A_i determines the class $w \in W_i \subseteq W$, then A_i -indiscernibility relation can be substituted by w -indiscernibility relation. Thus, we assume that the soft set Υ_{W_i} splits the cell set C into equivalence classes generated by w -indiscernibility relation $\mathfrak{R}_C^w$, where $w \in W_i$. In other words, a parameterized family of subsets of cells C , which forms the soft set Υ_{W_i} , is a factor set $C / \mathfrak{R}_C^w$ consisting of all equivalence classes of the set C generated by the relation $\mathfrak{R}_C^w$.

Thus, the soft set Υ_{W_i} can be used to generate equivalence classes in the set C instead of the equivalence relation $\mathfrak{R}_C^w$, $w \in W_i$. $\mathfrak{R}_C^w$ can be generalized and represented

as a similarity or a tolerance relation such that the soft set Υ_{W_i} splits the plurality of cells C onto the vague sets (fuzzy or rough). Of particular interest for building the structure of AOI is the A_S - indiscernibility relation, which can be replaced by the w_S - indiscernibility, as well as A_D - indiscernibility relations, which can be replaced by the w_D - indiscernibility relation. The approximation spaces generated by these relations can be represented as soft sets Υ_{W_S} and Υ_{W_D} respectively.

Consider the w_S - indiscernibility relation in a plurality of cells and the corresponding soft set Υ_{W_S} . The set of cells defining the AOI can be represented as the soft set, which divides the set of cell onto classes with respect to terrain conditions as $\Upsilon_{W_S} = \left\{ (w, \Upsilon_{W_S}(w)) : w \in 2^W, \Upsilon_{W_S}(w) \in 2^{W_C} \right\}$, where $\Upsilon_{W_S}(w)$ is a set of cells, which corresponds to a class $w \in W_S$, maps it into a set of cells, and describes the static component of AOI as $\Xi = \Upsilon_{W_S} = \left\{ (w, \Upsilon_{W_S}(w)) : w \in 2^W, \Upsilon_{W_S}(w) \in 2^C \right\}$ [16]. While A_S - indiscernibility relation generates static equivalence classes in a set of points $x \in X$, which constitute a topological space $T_X^{A_S}$, and geotaxons are their connection components, W_S - indiscernibility relation defined on the set of cells also generates static equivalence classes. Their connection components are subsets of cells, which approximate geotaxons, and they constitute a topological space $T_C^{w_S} = (C, \text{Def}(\text{apr}_C^{w_S}))$.

The decomposition of the subspace C of approximation subspace X using geotaxons, approximated by cells, is a topological space that represents the static component of the spatial model Ξ as $T_C^{w_S} = (C, \text{Def}(G_C^{w_S}))$. Each i - class of equivalence $(\text{apr}_C^{w_S})_i$ of approximation space $\text{apr}_C^{w_S}$ can be represented as the value of the soft set $\Upsilon_{W_S}(w_i)$, $w_i \in W_S$, i.e. $(\text{apr}_C^{w_S})_i = \Upsilon_{W_S}(w_i)$. Let $\text{Def}(\Upsilon_{W_S})$ is a family of all composite sets of the soft set Υ_{W_S} . Obviously, $\text{Def}(\Upsilon_{W_S}) = \text{Def}(\text{apr}_C^{w_S})$. Thus, the topological space can be represented as the soft set $T_C^{w_S} = (C, \text{Def}(\text{apr}_C^{w_S})) = (C, \text{Def}(\Upsilon_{W_S}))$.

The special role is related to W_D - indiscernibility relation, which splits the set of cells into phases and generates dynamic equivalence classes. A dynamic topological space $T_C^{w_D}$ is built upon dynamic equivalence classes and determines the dynamic behavior of the destructive processes. As well, each i - class of equivalence $(\text{apr}_C^{w_D})_i$ can be represented as the value of the soft set $\Upsilon_{W_D}(w_i)$, $w_i \in W_D$. At any time t , the set of cells can be represented as a dynamic soft set that splits the set of cells into phases (Fig. 1) $\Upsilon_{W_D}(t) = \left\{ (w, \Upsilon_{W_D}(w, t)) : w \in 2^W, \Upsilon_{W_D}(w, t) \in 2^C \right\}$, where $\Upsilon_{W_D}(w, t)$ is a set of cells, which belong to the phase $w \in W_D$ at the time t . This set describes the state of the process $F : \text{State}_F' = \Upsilon_{W_D}(t) = \left\{ (w, \Upsilon_{W_D}(w, t)) : w \in 2^W, \Upsilon_{W_D}(w, t) \in 2^C \right\}$.

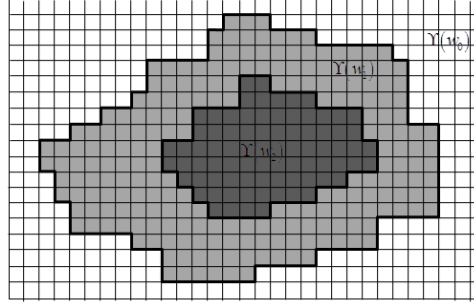


Fig. 1. State of the destructive process in the form of a soft set

Fig. 1 reflected the state of destructive processes in the form of soft sets, which divides the set of cells into three subsets: w_0 -elements (in white), w_1 -elements (in dark gray), and w_2 -elements (in light gray). If w_i and w_j elements related by achievability relation, then the areas approximated by cells, which are w_i - and w_j -elements of the soft set, must be adjacent to each other. The decomposition of the subspace C of approximation subspace X using W_D -indiscernibility relation is a topological space $T_C^{w_D} = (C, Def(apr_C^{w_D}))$ superimposed on topological space $T_C^{w_0}$. Each i -class of equivalence $(apr_C^{w_D})_i$ can be represented as the value of the soft set $F_{w_D}(w_i)$, $w_i \in W_D$.

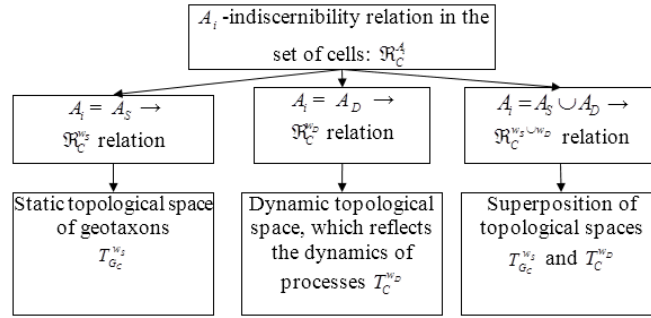


Fig. 2. The indiscernibility relations used to construct of the spatial model

Let $Def(\Upsilon_{w_D})$ be a family of composite sets of the soft set Υ_{w_D} . Obviously, that $Def(\Upsilon_{w_D}) = Def(apr_C^{w_D})$. Thus, the topological space can be represented as the soft set $T_C^{w_D} = (C, Def(apr_C^{w_D})) = (C, Def(\Upsilon_{w_D}))$ [16].

Consider q -indiscernibility relation on the set of cells that generates dynamic equivalence classes, which constitute a dynamic topological space T_C^q representing

homogeneous regions in respect of relative hazard, threat, or risk assessments at any given time t . A subset of cells belonging to one class of equivalence forms a zone, each of which does not necessarily have to be connected. At any time t the set of cells can be represented as a dynamic soft set, which splits the set of cells into zones from the set Q : $\Upsilon_Q(t) = \{(w, \Upsilon_Q(q, t)) : w \in 2^Q, \Upsilon_Q(q, t) \in 2^C\}$, where $\Upsilon_Q(q, t)$ is a set of cells, which are within a zone $q \in Q$ at a time t . The decomposition of the subspace C of approximation space X using q -indiscernibility relation is a topological space $T_C^q = (C, \text{Def}(\text{apr}_C^q))$.

Each i -class of equivalence $(\text{apr}_C^q)_i$ can be represented as the value of the soft set $\Upsilon_Q(w_i)$, $q \in Q$. Let $\text{Def}(\Upsilon_Q)$ be a family of composite sets of the soft set F_Q . Obviously, that $\text{Def}(\Upsilon_Q) = \text{Def}(\text{apr}_C^q)$. Thus, the topological space can be represented as a soft set $T_C^q = (C, \text{Def}(\text{apr}_C^q)) = (C, \text{Def}(\Upsilon_Q))$. The set Υ_{W_S} is static while the sets Υ_{W_D} and Υ_Q are dynamic. The spatial model Ξ can be represented as a multilayer topological space, which is a superposition of topological spaces in the form of soft sets:

$$T = (C, \{\text{Def}(\Upsilon)\} \mid \Upsilon \in \{\Upsilon_{W_S}, \Upsilon_{W_D}, \Upsilon_Q\}),$$

where $\text{Def}(\Upsilon)$ is a family of composition sets generated by the soft sets Υ_{W_S} , Υ_{W_D} , and Υ_Q , each of which constitutes a separate layer of the spatial model. Fig. 2 shows three types of A_i -indiscernibility relations in the set $(A_i \subset A)$ [16]. Table 1 reflects the properties of the considered topological spaces, which are shown in Fig. 3 [16].

4 Approximate topological space

Since the spatial model of AOI is blurred, in order to build an approximate topology we need to generalize (blur) a strict indiscernibility relation $\mathfrak{R}_C^{w_D}(t)$. Using the approximated soft sets we can represent topological spaces of terrain conditions (geotaxons), dynamic conditions (phase), estimations, etc. The blurring of topological spaces will be considered on the example of the topological space of the dynamic conditions, which is important for obtaining risk assessment.

Let us build a generalization of the relation $\mathfrak{R}_C^{w_D}(t)$ into the similarity relation $\mathfrak{R}_C^{w_D}(t)$, which can be replaced with the fuzzy soft set $\tilde{\Upsilon}_{W_D}$ [17]. As a result, at each time t we obtain a fuzzy approximation space $\text{apr}_C^{w_D}(t) = (C, \mathfrak{R}_C^{w_D}(t)) = (C, \mathfrak{R}_{W_D}^{w_D}(t))$ and a fuzzy topology, which represents a partition of all cells in C into fuzzy sets of cells $\mathcal{C}^{w_i}(t)$, $i = 0, \dots, n-1$ that enumerate all possible phases of the set W_D [18].

Let L denotes the interval $[0, 1]$, 2^C denotes a family of crisp subsets of C , and L^C denotes a family of all fuzzy subsets of C , where each fuzzy set is a mapping

$\mathcal{C}^w(t): C \rightarrow L$. Thus, we can represent the fuzzy soft set, which divides the set of cells into phases and define the state at time t as

Table 1. Properties of the topological spaces [16]

		Cartographic object or zone			
		Geotaxons	Cells	Processes	Assessments
1	topological space	$T_{G_C}^{A_S}$	T_C	$T_C^{w_D}$	T_C^q
2	The base of formation	A_S - indiscernibility relation $\mathfrak{R}_C^{A_S}$ and soft set Υ_{W_S}	Sampling space on equal objects	w_D - indiscernibility relation $T_C^{w_D}$ and soft set Υ_{W_D}	q - indiscernibility relation T_C^q and soft set Υ_Q
3	Connection components	Geotaxon	Cell	A set of cells, which belong to the same phase w_D	A set of cells, which corresponds to the assessment q
4	Attribute values of connection components	Static	Dynamic	Dynamic	Dynamic
5	The composition of connection components set (variability in space)	Static	Static	Dynamic	Dynamic
6	Elements of the equivalence class	Homogeneous with respect to a set of certain attributes of A_S	Homogeneous with respect to a set of certain attributes of A	Homogeneous with respect to a certain phase w_D	Homogeneous with respect to a certain assessment q

$\mathfrak{Y}_{w_D}^0(t) = \{(w, \mathfrak{Y}_{w_D}^0(w, t)) : w \in 2^{W_D}, \mathfrak{Y}_{w_D}^0(w, t) \in L^C\}$, where $\mathfrak{Y}_{w_D}^0(w, t) = \{(c, \mathfrak{Y}_{w_D}^0(w, t)(c)) : c \in C\} = \mathcal{C}^w(t) = \{(c, \mathcal{C}^w(c)(t)) : c \in C\}$ is the fuzzy set of cells, which belong to the phase $w \in W_D$ at time t , and $\mathfrak{Y}_{w_D}^0(w, t)(c) = \mathcal{C}^w(c)(t)$ is a degree of membership of the cell c to fuzzy set of cells $\mathcal{C}^w$, which belong to the phase w (w - element of the fuzzy soft set $\mathfrak{Y}_{w_D}^0$) at time t . In this relation, we use the fuzzy w - elements instead of crisp ones, so the soft set becomes the fuzzy soft set.

5 Fuzzy-rough soft topology

The above-considered model of the topological space is based on the fuzzy splitting of the set of cells into phases, the number of which in the general case can be unlimited. However, it is not always possible to find a way to determine the degrees of membership of cells to certain phases. If such degrees are not known, then instead

of fuzzy sets, it is convenient to use rough sets defined by a lower approximation (as a subset of cells that uniquely belong to an approximate set), an upper approximation

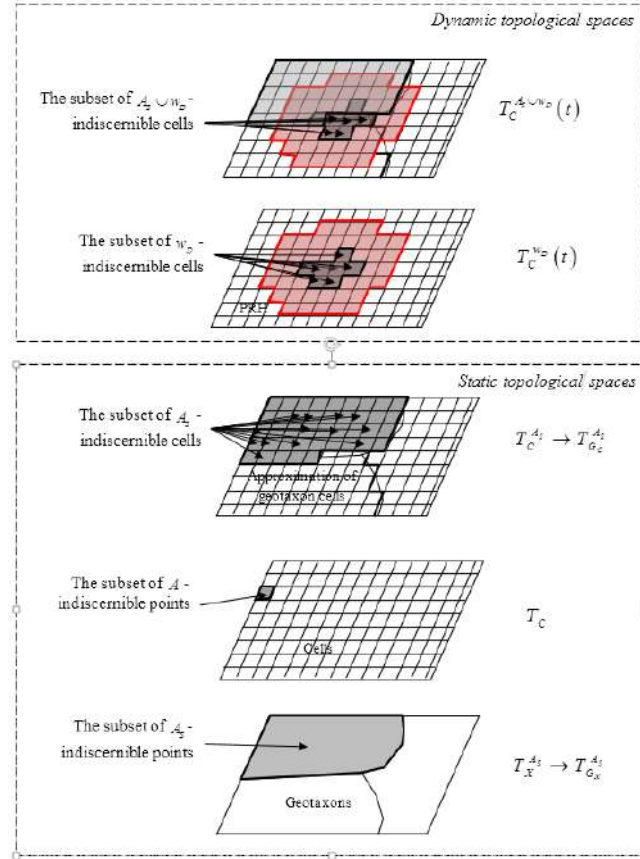


Fig. 3. The topological spaces of the spatial model

(as a subset of cells that may belong to an approximate set), and a boundary region (as a subset of cells, whose degree of membership is unknown with respect to the approximated set).

The approximate indiscernibility relation at a time t generates an approximation space $apr_C(t) = (C, \mathfrak{R}_C^{W_D}(t))$ and an approximate topology, that is, the partition of the set of cells C on the approximate subset of cells $\mathfrak{C}^{W_D}(t)$, $i = 0, \dots, n-1$, which belong to each of the possible phases of the set W_D . To build an approximated soft set of cells, we should blur the crisp soft set by introducing the Pawlak lower and upper rough approximations.

Let $apr_C^{W_D} = (C, \mathfrak{R}_C^{W_D})$ be Pawlak space approximation, and $\Upsilon_{W_D} = (\Upsilon, W_D)$ be a soft

set within C . Denote the lower and upper rough approximation of the soft set Υ_{w_D} in $(C, \mathfrak{R}_C^{w_D})$ by $\underline{\Upsilon}_{w_D} = (\underline{\Upsilon}, W_D)$ and $\overline{\Upsilon}_{w_D} = (\overline{\Upsilon}, W_D)$ respectively. Clearly, they are the soft sets $\underline{\Upsilon}_{w_D}(w) = \{c \in C \mid \mathfrak{R}_C^{w_D}(c) \subseteq \Upsilon_{w_D}(w)\}$ and $\overline{\Upsilon}_{w_D}(w) = \{c \in C \mid \mathfrak{R}_C^{w_D}(c) \cap \Upsilon_{w_D}(w) \neq \emptyset\}$ for all $w \in W_D$. In the case, when $\underline{\Upsilon}_{w_D}(w) = \overline{\Upsilon}_{w_D}(w)$, the w -element of soft sets Υ_{w_D} is the crisp set, otherwise, it is the rough set.

The above definition is a rough approximation of the soft set [9]. The approximate set of cells that belong to a certain phase w at time t is determined by two approximations as $\hat{\Upsilon}_{w_D}(w, t) = \{\underline{\Upsilon}_{w_D}(w, t), \overline{\Upsilon}_{w_D}(w, t)\}$, where $\underline{\Upsilon}_{w_D}(w, t)$ is a lower approximation, which contains all cells that belong to the set $\hat{\Upsilon}_{w_D}(w, t)$ clearly and necessarily (i.e., they belong to the phase w); and $\overline{\Upsilon}_{w_D}(w, t)$ is an upper approximation, which contains all cells that may belong to the set $\hat{\Upsilon}_{w_D}(w, t)$.

The negative region of the rough set $\hat{\Upsilon}_{w_D}(w, t)$ is called a set of cells of the universe C , which do not reliably belong to $C^w(t)$: $NEG(\hat{\Upsilon}_{w_D}(w, t)) = C - \overline{\Upsilon}_{w_D}(w, t)$. A boundary region of the rough set $\hat{\Upsilon}_{w_D}(w, t)$ is called a set of cells of the universe C , which belong to the upper approximation $\overline{\Upsilon}_{w_D}(w, t)$ but does not belong to the lower approximation $\underline{\Upsilon}_{w_D}(w, t)$: $BND(\hat{\Upsilon}_{w_D}(w, t)) = \overline{\Upsilon}_{w_D}(w, t) - \underline{\Upsilon}_{w_D}(w, t)$.

During the monitoring of destructive processes, it is often possible to obtain information for the cells about the graduation of their degree of membership with respect to the boundary region of the certain rough set. For this purpose, it is convenient to represent the state of the destructive process as a fuzzy rough soft set of cells, which divides the set of cells into phases at each time t and can be represented as a triple consisting of the upper and lower approximations of the rough set, and the boundary region of the rough set represented as the fuzzy set: $\hat{\Upsilon}_{w_D}^{\mathcal{F}}(t) = \langle \underline{\Upsilon}_{w_D}(t), \overline{\Upsilon}_{w_D}(t), BND(\hat{\Upsilon}_{w_D}(t)) \rangle$ [19].

Fuzzy-rough soft set splits the set of cells into w -elements, each of which is a fuzzy rough set of cells, which belong to a certain phase w , $w \in W_D$: $\hat{\Upsilon}_{w_D}^{\mathcal{F}}(w, t) = \langle \underline{\Upsilon}_{w_D}(w, t), \overline{\Upsilon}_{w_D}(w, t), BND(\hat{\Upsilon}_{w_D}(w, t)) \rangle$, where $BND(\hat{\Upsilon}_{w_D}(w, t))$ is a fuzzy set of cells, which belong to the boundary region of the w -element of the rough set $\hat{\Upsilon}_{w_D}(t)$: $BND(\hat{\Upsilon}_{w_D}(w, t)) = \{c, BND(\hat{\Upsilon}_{w_D}(w, c, t)) : c \in BND(\hat{\Upsilon}_{w_D}(w, t))\}$. $BND(\hat{\Upsilon}_{w_D}(w, c, t))$ is the degree of membership of the cell c , which belongs to the boundary region of the rough set $\hat{\Upsilon}_{w_D}(t)$, to the fuzzy set $BND(\hat{\Upsilon}_{w_D}(w, t))$ at a time t .

Fig. 4 shows the blurring of the boundaries of w -elements of the soft set Υ_{w_D} using fuzzy rough soft set $\hat{\Upsilon}_{w_D}^{\mathcal{F}}$. The top of the figure shows the state of the destructive process in the form of the soft sets Υ_{w_D} , which splits the set of cells into three subsets

(w_0 - elements, w_1 - elements, and w_2 - elements), each of which is a crisp set.

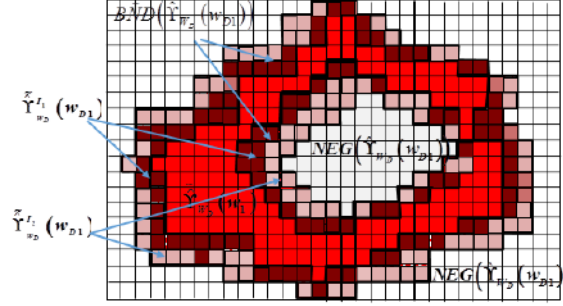


Fig. 4. Blurring the boundaries between the elements of the soft set using the fuzzy rough sets

We obtain a fuzzy rough soft set by generalizing the soft set $\mathcal{Y}_{w_D}^{\phi}$ and blurring the boundaries between its elements. Two lower figures show fuzzy rough sets, which are elements of the fuzzy rough soft set: w_2 - item ($\mathcal{Y}_{w_D}^{\phi}(w_2)$) and w_1 - item ($\mathcal{Y}_{w_D}^{\phi}(w_1)$). The boundary regions of the approximate sets are represented by fuzzy sets. Cells having different degrees of membership to the rough set are represented in different colors. Thus, the state of the process at a time t can be defined as a fuzzy rough soft set of cells $\mathcal{Y}_{w_D}^{\phi}(t): \mathcal{Y}_{w_D}^{\phi}(t) = \left\{ (w, \mathcal{Y}_{w_D}^{\phi}(w, t)) : w \in W_D \right\}$, where each fuzzy rough set of cells $\mathcal{Y}_{w_D}^{\phi}(w, t)$ is the w -element of the fuzzy rough soft set $\mathcal{Y}_{w_D}^{\phi}(t)$. Let $C / \mathbb{R}_C^{\phi_{w_D}}(t)$ be a factor-set, consisting of fuzzy sets of cells $\mathcal{Y}_{w_D}^{\phi}(w, t) = \mathbb{C}_{w_D}^{\phi}(t), w \in W_D$, generated by fuzzy rough relation $\mathcal{H}_C^{\phi_{w_D}}(t)$. In this case, $\mathcal{A}^{\phi_{w_D}}_C = (C, \mathbb{R}_C^{\phi_{w_D}}(t)) = (C, \mathcal{Y}_{w_D}^{\phi}(t))$ is a fuzzy rough approximation space and $Def(\mathcal{A}^{\phi_{w_D}}_C) = Def(\mathcal{Y}_{w_D}^{\phi})$ is a family of fuzzy rough sets representing the cells, which belong to a certain phase; $\forall t \in T$ the set $\mathcal{Z}^{\phi_{w_D}}(t) = Def(\mathcal{Y}_{w_D}^{\phi}(t))$ is a fuzzy rough topology on C . At any $t \in T$ a couple $\mathcal{T}_C^{\phi_{w_D}}(t) = (C, \mathcal{Z}^{\phi_{w_D}}(t))$ is fuzzy rough topological space. Each element of $\mathcal{Z}^{\phi_{w_D}}(t) = Def(\mathcal{Y}_{w_D}^{\phi}(t))$ is the fuzzy rough open set in C [20].

6 Developing a spatial model of the destructive process

In order to diagnose the situation during destructive processes in real-time systems, in terms of time limit there is no need to allocate a large number of phases of a cell. It is quite enough for a certain time to allocate a subset of cells not yet covered by the destructive process, a subset of the cells covered by the destructive process, and a

subset of cells destroyed as a result of the destructive process (which were covered at the previous moments of time). To do this, we use three possible values of the cell phase c in the spreading area of the destructive process F (Fig. 4) [16]:

- "Not covered by F " ($w_D(c, t) = w_{D0}$)
- "Covered by F " ($w_D(c, t) = w_{D1}$)
- "Phase not defined" ($w_D(c, t) = w_{D2}$).

As a rule, information on being cells in these phases can be obtained during monitoring with different sensors. The cells covered with the destructive process belong to the phase w_{D1} and form a zone limited by the internal and external contour of the process. These contours are blurred and can be represented as boundary regions of an approximate set of cells in this phase. During the monitoring of the dynamics of the process with various sensors, it is often possible to obtain information about the possibility of covering the cells that belong to the blurred contour. Often, we can also determine the gradation of the possibility of covering cells that belong to a blurred contour of the destructive process. The state of the process can be represented as the fuzzy rough set of cells, covered by the destructive process, given by a triple, consisting of the upper and lower approximations of the rough set, as well as the boundary region of the rough set, presented as the fuzzy set:

$$\tilde{Y}_{w_D}(w_{D1}, t) = \left(\hat{Y}_{w_D}(w_{D1}, t), \overline{\hat{Y}_{w_D}}(w_{D1}, t), B\tilde{N}D(\hat{Y}_{w_D}(w_{D1}, t)) \right), \quad \text{where}$$

$$B\tilde{N}D(\hat{Y}_{w_D}(w_{D1}, t)) = \left\{ \left(c, \tilde{Y}_{w_D}(w_{D1}, c, t) \right) \right\}.$$

For some destructive processes, the concept of the inner contour does not make sense (e.g., for floods). Therefore, it is often advisable to consider only the outer contour of the process, which we will call simply a contour. The location and dynamics of the outer contour are decisive for assessing the time-level threat, representing the time for which the outer contour of the process can reach a certain object (Fig. 5).

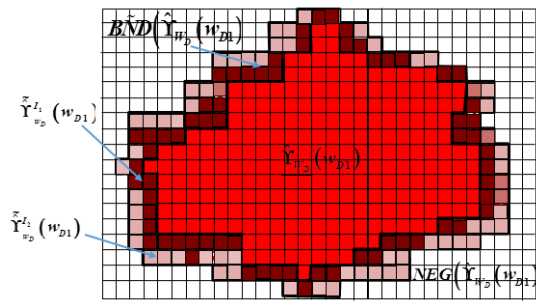


Fig. 5. State of the destructive process defined by the fuzzy rough soft set

In this case, we consider two phases of cell dynamics:

- "Not covered by F " ($w_D(c,t) = w_{D0}$)
- "Covered by F " ($w_D(c,t) = w_{D1}$).

During monitoring it is possible to determine the areas covered and destroyed as a result of the destructive process (lower approximation $\hat{Y}_{w_D}(w_{D1}, t)$, that is, the set of cells belonging to the phase w_{D1}), and areas not yet covered by the destructive process (negative region $NEG(\hat{Y}_{w_D}(w_{D1}, t))$, that is, the set of cells belonging to the phase w_{D0}). There is a blurred territory between these areas, which constitute a blurry contour of the destructive process (fuzzy set $B\tilde{N}D(\hat{Y}_{w_D}(w_{D1}, t))$). Based on the monitoring data it is not difficult to construct a fuzzy rough soft approximation space $\tilde{apr}_c(t) = (C, \tilde{Y}(t))$ and a fuzzy rough soft topology $Def(\tilde{apr}_c(t)) = \tilde{\tau}(t)$, that is, partitioning the set of cells C at each time t on the fuzzy rough subset of cells $\tilde{Y}_{w_D}(w_{Di}, t)$, which belong to each of the possible phases w_{Di} of the set W_D .

7 Experiment Results

The proposed spatial model has been implemented using Visual C and tested on computer based on the Pentium i5-7400 3 GHz processor and 16 GB RAM. The developed spatial model of the spreading destructive processes based on the fuzzy rough soft topology was used in the GIS-based real-time DSS providing the geospatial analysis of emergencies in real time disaster situations [21]. The developed DSS allows evaluating a number of indicators, e.g. danger degrees, threats, and risks, for target objects, as well as providing the geospatial analysis of emergencies in real time disaster situations. To obtain such estimates, it is necessary to build a spreading model of the destructive process and track the movement of its contour in real time by monitoring using UAVs.

To examine the developed model, we use real-time DSS in the forest fire response operations. Fig. 6 shows the representation of the forest fire front based on the fuzzy rough soft topology, which has been obtained during the monitoring. Fig. 7 depicts a fuzzy-rough cut of the forest fire front evaluated by the possibility of burning obtaining during the forest fire monitoring.

The results of the experiment show that the proposed spatial model provides acceptable performance in terms of accuracy and speed for all kind of topology. The fuzzy rough soft topology shows sufficient results on the speed with enough accuracy.

8 Conclusions

The approximate spatial model for the real-time GIS-based DSS based on the fuzzy-rough soft topology is proposed. The model of the destructive process is repre-

sented as the model of dynamic change of states of the subset of cells covered by the process within the spatial model. As a result, the spatiotemporal structure of AOI is represented as a topology space, which includes subspaces of cells that belong to the same phase.

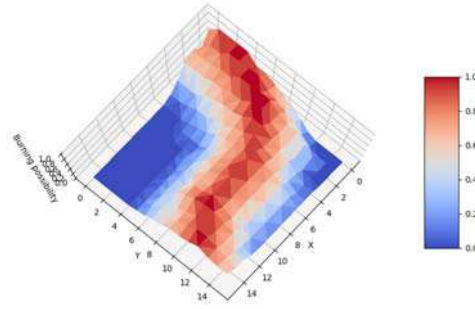


Fig. 6. Representation of the forest fire front based on the fuzzy rough soft topology

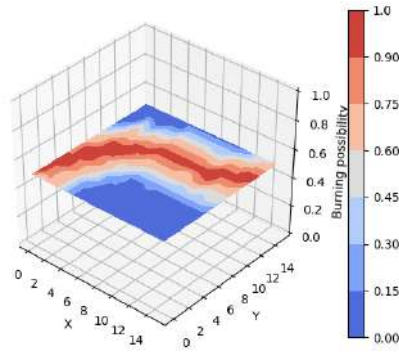


Fig. 7. A fuzzy-rough cut of the forest fire front assessed by the possibility of burning

The soft topological spaces are used to build a spatial model, as well as the fuzzy-rough method is used for its blurring. Since the belonging of each cell to the certain phase is approximately determined due to the uncertainty of remote sensing, the topological space is blurred and the boundaries of the dynamic contour of the destructive process are also blurred. The proposed spatial model representing uncertain information about the disaster reduces the computational complexity and provides flexible and timely decision-making in real time.

9 References

1. Sherstjuk, V., Zharikova, M., Sokol, I.: Approximate spatial model based on fuzzy-rough topology for real-time decision support systems. In Proc. on IEEE First Ukraine

- Conference on Electrical and Computer Engineering (UKRCON), pp 1037–1042, Kyiv (2017).
2. Chen, G., Zhao, J., Yuan, L., Ke, Z., Gu, M., Wang, T.: Implementation of a geological disaster monitoring and early warning system based on multi-source spatial data: a case study of Deqin Country, Yunnan Province. *Hazards Earth Syst. Sci.*, pp. 15 (2017).
 3. Chi, M., Plaza, A., Benediktsson, J.A., Sun, Z., Shen, J., Zhu, Y.: Big data for remote sensing: challenges and opportunities. In *Proc. of the IEEE*, 104(11), 2207–2219 (2016).
 4. Yuan, C., Zhang, Y., Liu, Z.: A Survey on Technologies for Automatic Forest Fire Monitoring, Detection and Fighting Using UAVs and Remote Sensing Techniques. *Canadian Journal of Forest Research* 45(7), 783–792 (2015)
 5. Martínez de Dios, J., Arrue, B., Merino, L., Ollero, A., Gómez-Rodríguez, F.: Computer vision techniques for forest fire perception. *Image and Vision Comp.* 26(4), 550–562 (2007).
 6. Merino, L., Caballero, F., Martínez-de-Dios, J., Maza, I., Ollero, A.: An Unmanned Aircraft System for Automatic Forest Fire Monitoring and Measurement. *Journal of Intelligent & Robotic Systems* 65, 533–548 (2012).
 7. Zharikova, M., Sherstjuk, V.: Development of the Model of Natural Emergencies in Decision Support System. *Eastern European Journal of Enterprise Technologies* 4(1), 62–69 (2015).
 8. Zadeh, L.A.: Fuzzy sets. *Information and Control* 8, 338–353 (1965).
 9. Pawlak, Z., Jerzy, W., Slowinski, R., Ziarko, W.: Rough Sets. *Comm. of ACM* 38(11), 88–95 (1995).
 10. Molodtsov, D.A.: Soft set theory – first results. *Computer & Mathematics with Applications*, pp. 19–31 (1999).
 11. Maji, P. K., Roy, A. R., Iswas, R. B.: An application of soft sets in a decision-making problem. *Computers and Mathematics with Applications* 44, 1077–1083 (2002).
 12. Maji, P. K., Iswas, R. B., Roy, A. R.: Fuzzy soft sets. *Journal of Fuzzy Mathematics* 9, 589–602 (2001).
 13. Meng, D., Zhang, X., Qin, K.: Soft rough fuzzy sets and soft fuzzy rough sets. *Computers and Mathematics with Applications* 62, 4635–4645 (2011).
 14. Zharikova, M., Sherstjuk, V.: Threat assessment method for intelligent disaster decision support system. *Advances in Int. Systems and Computing* 512, 81–99 (2016).
 15. Allam, A.A., Bakeir, M.Y., Abo-Tabl, E.A.: Some Methods for Generating Topologies by Relations. *Bull. Malays. Math. Sci. Soc.* 2(31), 35–45 (2008).
 16. Zharikova, M.: Methodological basis of geoinformation technology of decision support in combined natural and man-made systems in destructive processes conditions. Dr Thesis, KNTU, Kherson, p. 503 (2018)
 17. Varol, B.P., Aygun, H.: Fuzzy soft topology. *Hacettepe Journal of Mathematics and Statistics* 41(3), 407–419 (2012).
 18. El-Diafy, S.N.: Comparative study of fuzzy topology. Thesis submitted in partial fulfillment of the requirement for the degree of master of mathematics, Gaza, The Islamic University of Gaza, 76 p. (2014).
 19. Meng, D., Zhang, X., Qin, K.: Soft rough fuzzy sets and soft fuzzy rough sets. *Computers and Mathematics with Applications* 62, 4635–4645 (2011).
 20. Tang, W., Wu, J., Zheng, D.: On fuzzy rough sets and their topological structures. *Mathematical problems in engineering* 4, 17 (2014).
 21. Sherstjuk, V., Zharikova, M., Sokol, I.: Forest Fire-Fighting Monitoring System Based on UAV team and Remote Sensing. In *Proc. of 2018 IEEE 38th International Conference on Electronics and Nanotechnology*, pp. 663–668 (2018).

Inductive technology of the target clusterization of enterprise's economic indicators of Ukraine

Irina Lurie¹[0000-0001-8915-728X], Andrii Podlevskyi²[0000-0002-3166-7487],
Natalia Savina²[0000-0001-8339-1219], Maria Voronenko¹[0000-0002-5392-5125],
Anna Pashnina²[0000-0002-1425-1615], Volodymyr Lytvynenko¹[0000-0002-1536-5542]

¹Kherson National Technical University, Department of Informatics & Computing
Technology, Kherson, Ukraine
E-mail: lurieira@gmail.com immun56@gmail.com,
mary_voronenko@i.ua

²National University of Water and Environmental Engineering, 11, Soborna Street,
Rivne, Ukraine, 33000
E-mail: a.a.podlevskyi@nuwm.edu.ua, n.b.savina@nuwm.edu.ua

Abstract. The article presents an inductive model of objects clustering economic indicators based on the method of arguments group accounting. The basic principles of creating objective clustering inductive model are formed, the ways and prospects for the possible model implementation are shown, the advantages of an objective clustering model compared to traditional data clustering methods are defined.

Keywords. inductive modeling, economic objects clustering, method of arguments group accounting, k-means algorithm, external balance criterion.

1 Introduction

Relevance of the work. The economic objects clustering is currently receiving much attention. This is primarily due to the increased requirements for the accuracy of the recognition and identification systems under various conditions for obtaining information. Currently, there are a large number of diverse clustering algorithms for economic entities, each of which has its own advantages and disadvantages and is focused on a specific data type. One of the existing clustering algorithm drawbacks is their subjectivity, i.e. getting good results of clustering objects on one set does not guarantee to get similar results on another similar set. One of the ways to improve the clustering objectivity is the development of hybrid models based on the method of complex systems inductive modeling, which is a logical continuation of the method of group accounting of arguments (MGUA) [1-3]. In this regard, the development of hybrid models and clustering objects methods based on the methods of complex systems inductive modeling is an important problem from both theoretical and practical points of view.

2 Literature review

The use of a cluster approach is of particular importance for the Ukrainian enterprise's economic indicators clustering. Cluster policy is aimed at combining the capabilities, knowledge, and capacities of structures number with the aim of solving joint and private tasks. The immediate results from solving such problems will form the basis for the economic, social and technological development of the region.

Cluster policy is aimed at combining the capabilities of a structures number in order to solve joint and private tasks. The immediate results from solving such problems will form the basis for the economic, social and technological development of the region.

The theory and methodology of creating economic clusters in the regions are reflected in the writings of many authors. In scientific studies, it is noted that the successful development of the national economy depends on the development of the local concentration of specialized industries (industrial districts), which are the basis of the cluster approach [4]. This provision was first described in [5], where the synergistic effect of a merger of enterprises was first identified and analyzed.

In [6], such areas of knowledge as the new economic geography, business research of firms, regional studies and innovations that influence the development of cluster theory in the economy were identified. In [7], the authors proposed three clustering models: classical agglomeration models, industrial complex models, and network interaction models. In work [4], such five key concepts as externalities, innovative environment, interfirm struggle, competition of cooperation, dependency path, which constitute cluster theory, were considered. The authors of this work also focus on the geographical concentration and specialization of enterprises, the diversity of cluster participants, the critical mass and life cycle of the cluster, innovation and competition.

In [8], the authors succeeded in systematizing the extensive theoretical and empirical accumulated earlier, where the advantages of using national competitive relations in the economy were shown.

The authors of [9] believe that the cluster approach in the economy represents the synthesis of several areas, including local industrial specialization, spatial economic agglomeration, and regional development, as well as the provisions of strategic and venture management.

The diversity of the cluster theory indicates the absence of the only correct approach to its practical operationalization and makes it relevant to use the cluster approach for clustering the economic indicators of Ukrainian enterprises.

This study proposes inductive models of clustering objects of the economy to justify their creation at the regional level. The basic concepts of creating an inductive model of clustering objects based on the method of group accounting of arguments are described in [1-3, 10] and further developed in [11-16]. The authors formulated the basic principles of creating an inductive model of objective clustering, showed ways and prospects for a possible implementation of the model, determined the advantages of the model of objective clustering compared with traditional methods of data clustering. However, it should be noted that, despite the achievements achieved in this subject area, the inductive model of objective clustering currently has no practical implementation.

In the paper [16, 17] an inductive model of objective clustering of objects based on the k-medium clustering algorithm was developed, an estimation of the stability of the model to the noise component was made, ways of further improvement of the proposed model with purpose of increasing the objectivity of the clustering of the studied data. Approbation of the work of the proposed model was performed using the data "Compound" and "Aggregation" of the database of the computer school of the East Finnish University. It is presented studies to assess the stability of the model to the noise component using data "Seeds".

In this paper, a more improved version of the k-means inductive clustering algorithm is used. In contrast to [16], the centers of masses of the clusters are not calculated, and the Silhouette, Entropy, Dunn's index and Calinski-Harabasz index are used as an internal criterion for the quality of clustering (in [16] uses only the Calinski-Harabasz index).

3 Formal problem statement

The unresolved parts of a common problem include:

- efficient algorithms lack for extracting equal-power subsets from the initial data set;
- lack of research on the influence of external and internal criteria on the clustering quality;
- insufficient implementation of the objective clustering inductive model in various areas of society, especially economic.

The aim of the article is to develop and study the influence of internal and external criteria on the quality of the objective clustering inductive model of objects based on the k-means clustering algorithm in the study of the enterprise's different types economic data in Ukraine.

4 Method Description

Let $A = \{x_{ij}\}, i = 1, \dots, n; j = 1, \dots, m$ – the matrix of the objects features under study, where n is the number of rows or observed objects, m is the number of features characterizing the object. The clustering task is reduced to splitting the set of objects into non-empty subsets of disjoint clusters, and the plane separating the clusters can take any form:

$$\begin{aligned} K &= \{K_s\}, s = 1, \dots, k; K_1 \cup K_2 \cup \dots \cup K_k = A; \\ K_i \cap K_j &= \emptyset, i \neq j; i, j = 1, \dots, k \end{aligned} \quad (1)$$

where k – the number of clusters.

The methodology of complex systems inductive modeling is based on three fundamental principles borrowed from various scientific fields:

1. The principle of heuristic self-organization, the i.e. search of applicants various models with a choice of the best from the point of view of the relativity external criterion, the value of which is determined on two equally powerful data sets;

2. the external addition principle, the idea of which is the need for objective verification of the model using additional “fresh” information;
3. The inconclusive decisions principle, i.e. generation of certain solutions set with the optimal variant subsequent choice.

The implementation of these principles in the framework of an inductive objective clustering model implies the following steps:

- normalization of the studied objects signs, i.e. bringing them to the same range with the same median of the attribute space attributes;
- splitting the original objects set into two equally powerful submultiples;
- determination of an external criterion or relevance criteria group for choosing the optimal clustering on two equally powerful subsets;
- selection or development of a basic clustering algorithm used as a component in an objective clustering inductive model of objects.

Data normalization was performed according to the signs in accordance with the formula:

$$x'_{ij} = \frac{x_{ij} - med_j}{\max(|x_{ij} - med_j|)} \quad (2)$$

where x_{ij} – the value of attribute i in the column j , x'_{ij} – normalized value of this feature, med_j – column median j . The choice of this normalization method was determined by the fact that as a result, the data features set in all columns had the same median with a maximum variation attributes range from -1 to 1, while the data volume for each column falling into the inter-quantile distance (50%) is the largest compared to other normalization methods.

Algorithm for the separation of the original objects set Ω to 2 equipotent disjoint subsets Ω^A and Ω^B consists of the following steps [13]:

1. calculation $n \cdot (n-1)$ pairwise distances between objects in the original data sample;
2. selection of objects pair X_s, X_p , the distance between which is minimal:

$$d(X_s, X_p) = \min_{i,j} d(X_i, X_j) \quad (3)$$

3. object distribution X_s into a subset Ω^A , and object X_p into a subset Ω^B ;
4. repetition steps 2–3 for the remaining objects. If the objects number is odd, the last object is distributed into both subsets.

As internal criteria (IC) for the quality of clustering used:

1. Silhouette [18]

$$SWC = \frac{1}{K} \sum_{j=1}^K S_{x_j} \quad (4)$$

where K - number of clusters, S_{x_j} - the "best" element belonging x_j to cluster p .

The best partition is characterized by the maximum SWC, which is achieved when the distance inside the cluster is small, and the distance between the elements of the neighboring clusters is large.

2. Dunn's index [19]

Compares intercluster distance with cluster diameter. The higher the index value, the better the clustering.

$$DI(k) = \min_{i \in k} \quad (5)$$

3. Calinski – Harabasz index [20]

$$QC_{CH} = \frac{QCB \cdot (N - K)}{QCW \cdot (K - 1)} \rightarrow \max \quad (6)$$

where N - number of objects, K - number of clusters. The maximum index value corresponds to the optimal cluster structure.

4. Entropy [21]

$$PE = \frac{\sum_{q=1}^Q \sum_{k=1}^K \ln(u_{qk})}{Q}, PE \in [0, \ln Kk] \quad (7)$$

Entropy is known as the numerical expression of the system orderliness. The entropy of the partition reaches a minimum with the highest orderliness in the system (in the case of a clear partition, the entropy is zero). That is, the greater the belonging degree of an element to one cluster (and the smaller the belonging degree to all other clusters), the smaller the entropy value and the more qualitatively clustering is performed. The main disadvantage of these methods is that their computation becomes more and more complex, both with an increase in the clusters number k and with an increase in the objects number included in the data. To calculate the external criterion of balance, the approach taken from [16] was taken as a basis. In this paper, the external criterion of balance (ECB) of controlled clustering is defined as the normalized optimal value of the sum of squared deviations between the values of the internal criteria of clustering quality (4) – (7):

$$ECB = \sqrt{\frac{(IC_A - IC_B)^2}{(IC_A + IC_B)^2}} \rightarrow opt \quad (8)$$

To create equal clustering conditions on subsets and using the k-average clustering algorithm, an initial number of clusters is determined at the initialization stage, and the initial value of criterion (8) is zero. The experiment showed that at subsequent iterations the criterion value at the first step increases, then monotonously changes until it reaches saturation, which corresponds to stable clustering on two equally powerful subsets. The block diagram of the inductive clustering model based on the k-means algorithm is shown in Fig. 1.

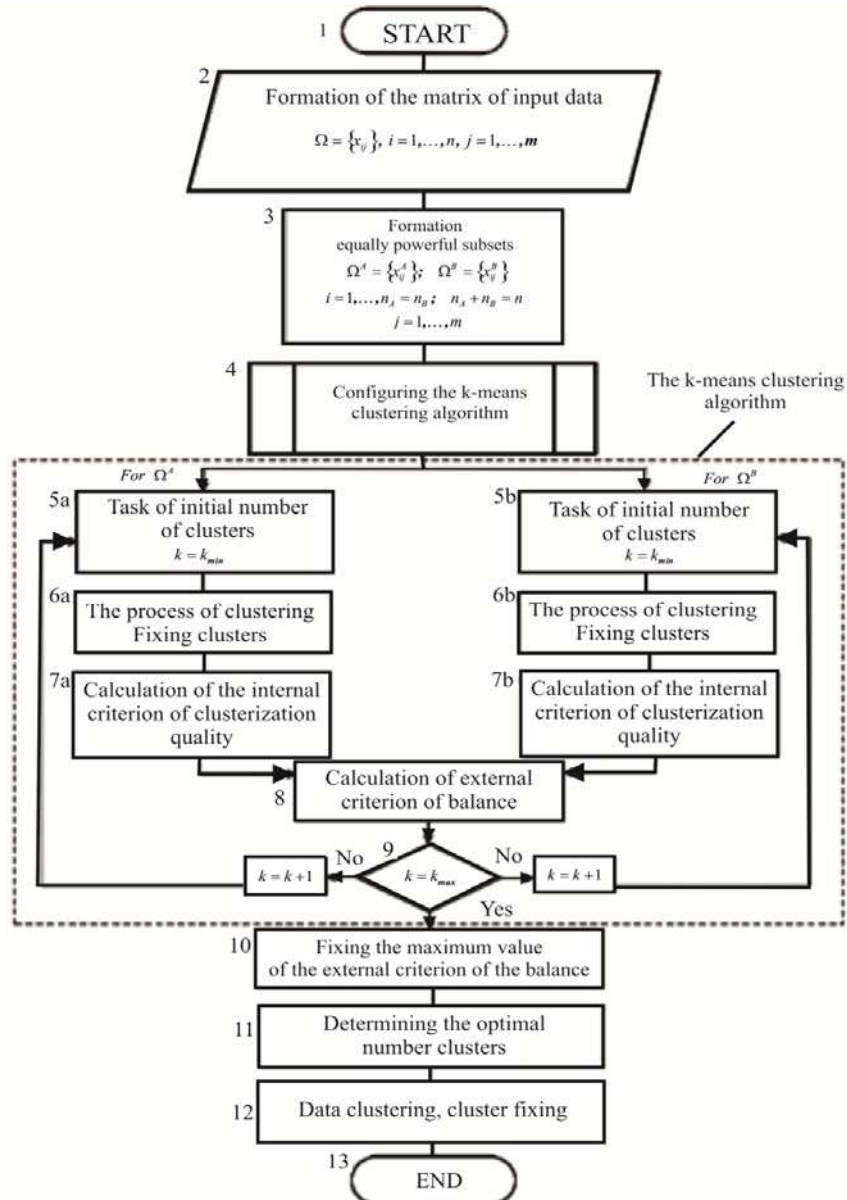


Fig. 1. Block diagram of an inductive model of objective clustering based on the k-means algorithm

The implementation of the algorithm requires the following steps:

Step 1. Start

Step 2. Formation of the initial set Ω of objects under study. Data preprocessing (filtering, normalization, analysis for missing values). Representation of data in the

form of a matrix $n \times m$, where n - the number of rows or the number of objects studied, m - the number of columns or the number of signs characterizing the objects.

Step 3. The division $\mathcal{Q}$ into two equally powerful subsets in accordance with the above algorithm. The resulting subsets $\mathcal{Q}^A$ and $\mathcal{Q}^B$ formally can be represented as follows:

$$\mathcal{Q}^A = \{x_{ij}^A\}, \mathcal{Q}^B = \{x_{ij}^B\}, j = 1, \dots, m$$

$$i = 1, \dots, n_A = n_B, n_A + n_B = n \quad (9)$$

Step 4. Configure the k-means clustering algorithm.

For each equally powerful subset:

Step 5. Select the number of clusters.

Step 6. Sequential clustering and cluster fixing

Step 7. Calculation of the internal criterion of clustering quality.

Step 8. Calculation of the external balance criterion in accordance with formula (8).

Step 9. If the value of the balance criterion reaches the optimum, then:

Step 10 Fixation of the received clustering is performed.

otherwise, the number of clusters is increased by 1 and repeated Step 5-9

Step 11. Determining the optimal number of clusters.

Step 12. Clustering data (sets $\mathcal{Q}$ of objects under study), fixing clusters.

Step 13. End

5 Characteristics of the data used

5.1 Analysis of the ratio of small, medium and large enterprises, and their importance for economic development in the regions of Ukraine

For Ukraine, unbalanced, resource-intensive, with significant territorial-branch disproportions is a model of a national economic complex that requires significant financial and organizational efforts to remedy the situation. Small, medium and large enterprises play an important role in the Ukrainian economy and have well-known advantages and disadvantages. To strengthen the benefits and reduce the impact of shortcomings for each group of companies in the context of accelerated socio-economic national development can be considered a system of production cooperation state regulation, which has its own characteristics in the sectoral and regional dimension.

Therefore, it is important to study the main indicators of large, medium and small businesses of Ukraine with the help of cluster analysis tools, which will reveal certain differences in the functioning of entrepreneurship in a regional dimension and offer more effective and flexible mechanisms of regions socio-economic development.

The matrix of the objects under study contained 26 rows (objects) and 18 columns (signs characterizing the objects). In accordance with the goals of the problem to be solved, clustering was carried out according to signs, i.e. after transformation, the data matrix under investigation had a dimension of 18×26 . When dividing this set into

two closed subsets in accordance with the algorithm described above, we get two subsets with dimensions 9×13 (9).

6 Clustering Results

Using the NbClust package from the programming language and the environment for developing, analyzing data and statistical calculations R, we construct a diagram that shows the dependence of the clusters number on the NbClust criteria (Fig. 2).

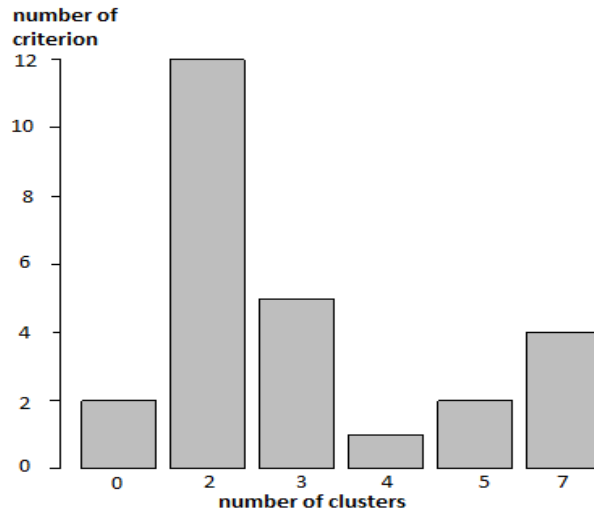


Fig. 2. Cluster number diagram for the k-means clustering algorithm

Analysis of the chart allows us to conclude that, from the point of view of the criteria used, it is optimal to divide the test signs into two, three or seven groups. However, since, in accordance with the goals set, the division of the features set into a large clusters number is inexpedient, we consider the most optimal clustering when dividing objects into three groups. The results of the work of objective clustering inductive models based on the above algorithm are presented in Table 1. Analyzing the values of the criteria, it is clear that splitting into three clusters shows the best results for the values of Silhouette, Dunn index and Calinski-Harabasz index. The minimum Entropy value is achieved when split into two clusters.

Table 1. The results of the work of objective clustering inductive models

Internal criterion of clustering quality	Number of clusters		
	2	3	7
	Exterior balance creature	Exterior balance creature	Exterior balance creature
Silhouette	0,0243161	0,1375	0,0766
Dunn;s index	0,15297984	0,3497	0,0369
Calinski-Harabasz index	0,01035773	0,0204	0,0175
Entropy	0	0,2115	0,0353

Table 2 shows the results of clustering into two, three and seven clusters. As can be seen from the table, the divisions into three and seven clusters distinguish the economic indicators of Kiev city into a separate cluster. Given that Kyiv is the capital of Ukraine not only in the administrative-territorial field, but also in the international and economic, in most cases it is clustered with clustering as it is in a cluster, because in most economic indicators of the enterprises it is significantly ahead of the nearest cities-"competitors" (Kharkiv, Dnipro, Odessa, Lviv, etc.) and the whole area. The remaining groups are formed, taking into account the economic development and territorial regions location.

7 Economic interpretation of the results

2-cluster model. The results of such clustering can clearly identify 2 clusters with clear socio-economic characteristics. The first cluster contains Ukraine regions, which traditionally are centers of economic Ukraine macro-regions (Dnipropetrovsk, Donetsk, Lviv, Odessa, Kharkiv, Kyiv and Kyiv). This cluster can also be positioned as industrial since in these regions, the main industries are concentrated and the historically concentrated largest share of large enterprises, which also interact with a significant number of small and medium enterprises. Accordingly, this cluster is more productive in socio-economic terms.

The second cluster focuses on the remaining Ukraine regions with greater economic diversification of enterprises various types (mostly medium and small) and relatively lower economic indicators of regional development.

3-cluster model. The main results are duplicated by the above-mentioned model, the main difference is the isolation of the Ukraine capital -- the city of Kiev - as a separate cluster, given its significant socio-economic potential.

7-cluster model. In our opinion, this model is more relevant in interpreting the impact of enterprises different types on the Ukrainian regions economic development. To a large extent, such a division is correlated with studies of Ukraine regions by the integrated indicator of specialization [19,20]. In the first cluster, only Kyiv is represented, because according to the main indicators of social and economic development, and as the capital of Ukraine, it is positioned as a separate cluster.

Table 2. Clustering results for the k-means algorithm

Number of clusters		2				3				7			
Cluster number		1	2	1	2	3	1	2	3	4	5	6	7
Region	Vinnys'tka	+		+				+					
	Volyns'tka	+		+									+
	Dnipropetrovs'tka		+		+						+		
	Donetska		+		+						+		
	Zhytomyrs'tka	+		+					+				
	Zakarpats'tka	+		+								+	
	Zaporiz'tka	+		+				+					
	Ivano-frankivs'tka	+		+								+	
	Kievs'tka		+		+						+		
	Kirovohrads'tka	+		+								+	
	Luhans'tka	+		+								+	
	L'vivs'tka		+		+					+			
	Mykolayivs'tka	+		+								+	
	Odes'tka		+		+					+			
	Poltavs'tka	+		+				+					
	Rivnens'tka	+		+					+				
	Sums'tka	+		+					+				
	Ternopils'tka	+		+									+
	Kharkivs'tka		+		+					+			
	Khersons'tka	+		+								+	
	Khmelnys'tka	+		+					+				
	Cherkas'tka	+		+				+					
	Chernivets'tka	+		+					+				
	Chernivhivs'tka	+		+								+	
	city of Kiev		+			+	+						

The second cluster (Vinnitsa, Zaporozhia, Poltava, Cherkassk) contains fairly balanced enterprises (small, medium and large) and their indicators of the region, and is characterized mainly by industrial and agricultural profile.

The third cluster (Zhytomyr, Rivne, Sumy, Khmelnytsky, Chernivtsi) has a more pronounced agrarian profile with significant influence of the economic potential of small and medium enterprises.

The fourth cluster (Lviv, Odessa, Kharkiv) is represented primarily by regions in which all types of enterprises are harmoniously represented. In addition, significant developments in these areas of information industries and high-tech industries can be

noted, indicating their informational and post-industrial profile.

The fifth cluster (Dnipropetrovsk, Donetsk, Kiev) has the largest number of large enterprises (from 45 to 61) and is characterized by significant industrial potential.

The sixth cluster (Zakarpattia, Ivano-Frankivsk, Kirovograd, Luhansk, Mykolayiv, Kherson, Chernihiv) contains areas with a well-developed small and medium business and a widespread service area. They can be attributed equally to the industrial and agricultural profile, and to the agrarian-industrial.

The seventh cluster (Volyn, Ternopil) is represented by regions with lower socio-economic development and agronomic indicators than the average in Ukraine and the dominance of small and medium-sized enterprises of a non-industrial type.

8 Conclusions

The results of the model work showed the high efficiency of the developed clustering models on the novel inductive method of simulation of complex systems.

In this paper, the k-medium algorithm was used as the basic algorithm, while the effect of four internal criteria (Silhouette, Dunn's index, Calinski-Harabasz index, Entropy) on the quality of clustering was studied. This choice was determined by the simplicity of its implementation. The advantage of the proposed model lies in its stability, which is determined by using an external balance criterion on two coherent samples.

It should be noted that the use of inductive simulation methods does not eliminate the main disadvantage of the k-mean algorithm: the result of clustering depends on the selection of the source centers of the clusters, but with other things being equal, the proposed model gave better clustering results compared to the traditional k-medium algorithm implemented in the software environment R.

Interpreting the results of clustering using an inductive model showed a significant effect of using such a methodology to identify the degree of influence of small, medium and large enterprises on the socio-economic development of regions, which is largely confirmed by the results of other studies. However, some of the detected clusters (for instance 6) contain ambiguous characteristics that do not allow them to be clearly interpreted. This indicates the need to use a larger array of output data or to combine this technique with other approaches to provide the most rational and reliable result that could be the subject of further scientific research.

References

1. Ivakhnenko, A.G.: Group method of data handling – a competitor of stochastic approximation method // *Automatics*, 1968. - №3. - P. 58-72. [In Ukraine] (1968)
2. Ivakhnenko, A.G.: Inductive method for self-organization of complex systems models.– Kiev: Scientific Thought, 1982.– 296 p. [In Russian] (1982)
3. Ivakhnenko, A.G.: Objective self-organization based on the theory of self-organization models // *Automatics*, 1987. - №5. - P. 6-15. [In Russian] (1987)
4. Bergman, E.M., Feser, E.J.: *Industrial and Regional Clusters: Concepts and Comparative Applications*, Regional Research Institute, WVU (1999)
5. Marshal, A. : *Principles of Economics*. London: Macmillan (1999)
6. Humphrey J., Schmitz H.: *Governance and upgrading: linking industrial clusters and global value chain research*. IDS Working Paper 120. Institute of Development Studies (2000)

7. Gordon, I. R., McCann, P. : Industrial Clusters: Complexes, Agglomeration and/or Social Networks? *Urban Studies*, 37(3), pp. 513–532 (2000)
<https://doi.org/10.1080/0042098002096>.
8. Andersson T., Schwaag-Serger, S., Sorvik, J.: Emily Wise Hansson. The Cluster Policies Whitebook, IKED (2004)
9. Porter, M. E.: The Competitive Advantage of Nations. New York: Free Press, 1990. (Republished with a new introduction, 1998.) (1990)
10. Madala, H.R., Ivakhnenko, A.G.: Inductive Learning Algorithms for Complex Systems Modeling. – CRC Press, 1994. – 365 p (1994)
11. Stepashko, V.S.: Theoretical aspects of GMDH as a method of inductive modeling // Managing Systems and Machines, 2003. - №2. - P. 31-38. [In Russian] (2003)
12. Stepashko, V.S.: Elements of the inductive modeling theory / State and prospects of informatics development in Ukraine: Monograph / Team of authors. - Kiev: Scientific Thought, 2010. - 1008 p. - P. 471-486. [In Ukraine] (2010)
13. Osypenko, V.V.: Two approaches to solving the problem of clustering in the broad sense from the standpoint of inductive modeling // Power and Automation, 2014. - №1. - P.83-97. [In Ukraine] (2014)
14. Osypenko, V.V., Reshetjuk, V.M.: The Methodology of Inductive System Analysis as a Tool of Engineering Researches Analytical Planning / V.V. Osypenko, V.M. Reshetjuk // Ann. Warsaw Univ. Life Sci. – SGGW. – № 58, 2011. – P. 67-71 (2011)
15. Sarycheva, L.V.: Objective cluster analysis of the data on the basis of the Group Method of Data Handling // Problem of Management and Informatics, 2008. - №2. - P. 86-104. [In Russian] (2008)
16. Babichev, S., Lytvynenko, V., Taif M.: Estimation of the inductive model of objects clustering stability based on the *k*-means algorithm for different levels of data noise // Radio electronics, informatics, management. - Zaporizhzhya, ZNTU. – 2016. - №4, C. 55-60 (2016)
17. Reid, R., Schein, S., Wilson, H.: Industrial Clusters in Jackson and Josephine Counties. – School of Business Southern Oregon University, November 30, 2006. URL: http://www.ashland.or.us/files/SOU_IndustrialClustersReport.pdf (2006)
18. Kaufman, L., Rousseeuw, P. : Finding Groups in Data. An Introduction to Cluster Analysis. Wiley (2005)
19. Bezdek, J.C., Dunn, J.C.: Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions // IEEE Transactions on Computers, 835–838pp (1975)
20. Calinski, R.B., Harabasz, J.: A dendrite method for cluster analysis // Comm. in Statistics, 3:1.27 (1974)
21. Sripada, S. Ch., Rao, M. S.: Comparison of purity and entropy of *k*-means clustering and fuzzy *c* means clustering , Indian journal of computer science and engineering; Vol 2 no.3 June 2011; ISSN:0976-5166 (2011)

Decentralized system for run services

Galina Kirichuk^{1[0000-0002-0405-7122]}, Volodymyr Tymoshenko^{1[0000-0001-8161-5119]},
Oleksii Rudkovskyi^{1[0000-0002-5267-2525]}, Svitlana Hrushko^{1[0000-0002-0064-408X]},

¹Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063,
Ukraine
kirgal08@gmail.com, PigerDeus@gmail.com, alexteen12@gmail.com,
grushko_ss@i.ua

Abstract. In this work proposed to use decentralized network for executing desktop applications which developed using of own programming languages and developed to run on complete isolated operating system. It allows to process data faster and without having to have own computing power. Objectives of work are development of new protocols for transfer information, instead of IPv6 and transition to decentralized model the work of services.

Keywords: data processing, algorithm, encryption, hashing, isolated environment, decentralization, database

1 Introduction

In modern world it's impossible to imagine work of computers, laptops, tablets and smartphones without a network connection. At the same time, amount of traffic and data that needs to be stored or processed increases.

Issue of further development of networks and their security is becoming more acute with development of new technologies. Therefore, the issue of information availability, increasing speed of its receipt and use, improving efficiency and quality of computing is very relevant at the present time.

Solution to all of above problems is the decentralization of networks and applications. Therefore, many studies in this area are being carried out and practical applications are already in place [1-3].

2 Formulation of problem

Currently, almost all applications work as client-server services. They allow you to exchange messages, send files, receive news and publish them. Services interact with both users and other services that are often used by intruders when stealing public information, bank card data, private data, files, etc. Therefore, data protection is one of priority concerns for service delivery companies.

Another problem is rapid development of the internet of things. According to latest forecasts in the coming years, the number of devices connected to the Internet will increase several times, so issue of their personal identification is important. According

to growth rate of the Internet of things, the range of IPv6 addresses may end much earlier [3]. Therefore for reliable operation of the Internet these tasks need to be solved: the development of new protocols for the transfer of information and the transition to a decentralized model of service [4-6].

Replacement of services, their decentralized analogues, will solve the following issues: improve the reliability of services itself; increase the speed of services and to make the transition to new protocol for information transferring.

Purpose of work is to implement a system for support of computing in decentralized networks which supports the work of application services. Object of study is the process of implementing a system for support of computing in decentralized networks. Subject is the models, methods, and tools for support of computing in decentralized networks.

The main tasks of work: construction of models and algorithms of work of individual modules of the system; development of decentralized network, which provides data exchange, decentralized storage of information and the launch of services in the network; development of the protocol for transmission and processing of information for decentralized networks of different purposes; implementation of the system for supporting the work of services and conducting its testing on performance [3].

3 Software tools and solutions

Hashing algorithms are common and have several implementations (MD5, SHA, RIPEMD, etc.). The main characteristics of the hashing algorithms are: size; complexity of calculation; cryptographic stability; the speed of the calculation and the number of collisions [7].

Currently, the MD5 algorithm is not crypto-resistant. SHA combines various algorithms such as SHA-1, SHA-2, SHA-3, which are very popular and widely used in the modern world. SHA algorithms are characterized by the fact that making even small changes in input string significantly change the hash. Having looked at the hash algorithms SHA-2 group was chosen because they are protected, sufficiently researched, built on the basis of the structure of the Merkle-Demgard and have a high speed of execution [8-10].

Currently there are many encryption algorithms such as AES, RSA, 3DES, GOST 28147-89 and others. Different decentralized networks use different encryption algorithms. Often you can find the use of two or more algorithms for transmitting and storing information. The most secure and popular encryption algorithms are symmetric - AES, 3DES, Twofish and asymmetric - RSA, Diffie-Hellman, DSA, ECC and ECDH.

There are two encryption algorithms - ECDH for generating a shared key and AES for further encryption of data, which used in the work. This combination gives the system a high degree of stability, since: encryption keys are not transmitted due to unprotected communication lines; AES algorithm is most-protected at the moment; input data for generating a shared key is stable and computed very quickly; there is a possibility to encrypt large volumes of data [11-15].

There are two protocols for information transfer - TCP and UDP. Both protocols work on the transport layer of OSI model. Each protocol has its own peculiarities and

advantages. Sometimes applications support work using both protocols. Advantages of TCP are: reliable delivery of segments; streamlining segments when receiving; work with sessions; transferring speed control.

The advantage of UDP is high speed of transmission relative to TCP. The UDP protocol is used when it is necessary to receive data quickly. When using UDP, a logical connection is not created, so anyone can send data to the application. After reviewing both protocols, it is resolved to use TCP because of its reliability, which is very important in decentralized networks [6].

4 System development

System of computing support is decentralized and overlay. It has high security requirements for data stored and transmitted in system. Since the system is overlay, there is a need for customer identification, so you can use any sequence of characters of any length as the address. This avoids ending the range of available addresses by adding additional characters without changing the entire system.

Reliable encryption algorithms are used to achieve the required security level. Since the AES encryption algorithm, which is a symmetric algorithm, is selected, there is a weak point in the transmission of encryption keys. To solve this, we use an algorithm that generates a shared key without passing it through the network [16].

Using the Elliptic Curve algorithm, customers generate their own private and public keys. Private key is stored by the client and never transmitted by network, and open key is the client's address in a decentralized network. Every client knows the public key of each client with whom the link is established. Using the ECDH algorithm, clients, having their public and private keys and public key of a second client, can generate a shared key that is the same when generating on other side, when using another pair of keys and another public key.

This approach eliminates the need to pass the key through the network and allows you to combine the advantages of symmetric and asymmetric encryption algorithms while eliminating their shortcomings. In this case, the use of a public key as a client's address ensures that the virtual address can't be mapped to the physical location of the client or Internet. This algorithm generates a public key of large size, which for a long time, if necessary, will ensure a significant expansion of the address range by adding only one character to address.

When communicating information, customers use different operating systems (interaction between smartphone and computer) and data formats to achieve one result. Two methods are used simultaneously to solve these issues: serialization and encapsulation. When serialized, data structures are converted into one common format - text.

Having received text, the client converts it into the necessary structure. For serialization, we use the JSON data format.

```
import sys, os
import json
from serialization import Encoder
sys.path.insert(0, os.path.dirname(os.getcwd()))
class Serializer(object):
```

```

__instance = None
@staticmethod
def get_instance():
    if Serializer.__instance is None:
        Serializer.__instance = Serializer()
    return Serializer.__instance
def __init__(self):
    self.types = {
        # messages types
    }
    self.encoder = Encoder.Encoder()
def serialize(self, obj):
    return json.dumps(obj, default=self.encoder)
def deserialize(self, message):
    return json.loads(message, default=self.encoder)
def get_type(self, obj):
    return obj.__class__.__bases__[0].__name__
def add_type(self, obj, tag):
    obj_type = self.get_type(obj)
    self.types[obj_type] = tag

```

It is quite easy to use and has support in modern programming languages. For serialization of complex data structures, we use encapsulation: data themselves are enclosed in an object that further has a description of these data. With this description, the client will uniquely identify type of data and correctly handle it [17-21].

Since various applications are implemented by different people, and recipient indicates the sender itself, then to solve problems using a single description for different services, the programmer himself should take care of description uniqueness.

For example, the customer sends data that is stored in a separate class.

```

from dc.package import package
class weather (package):
    Def __init __ (se lf):
        self.country = ''
        self.city = ''
    def fields (self):
        return ['country', 'city']

```

Client pre-creates the association of this class with its code, for example, "weather.request". In the example, the country field contains the value "UA", and the city is "zaporizhzhia". In the serialized form, this class is described below.

```

{"Type": weather.request, "date": {country: UA, city,
zaporizhzhia}}

```

After receiving the data in this form, the client easily reproduces the object of the desired class. The data is encapsulated in a separate object, which contains additional

information that is necessary for the client (the specific data with which the client uses in a separate implementation). The final package view in abbreviated form is given.

```
{"sender": "...", "nodes": ["...", "..."], "length": "...",
"date": {"date": { ... }, "id": "...", " number ": 1, "total": 1,
"crc": ... }}
```

Package has a lot of additional information, so the transmission of small volumes of data occurs at a lower speed. Packets data is encrypted and available only to the recipient and sender, and route information is publicly available [22, 23].

Therefore, anyone can receive the package and send it to the recipient without access to the data themselves. The system model is shown in Fig. 1.

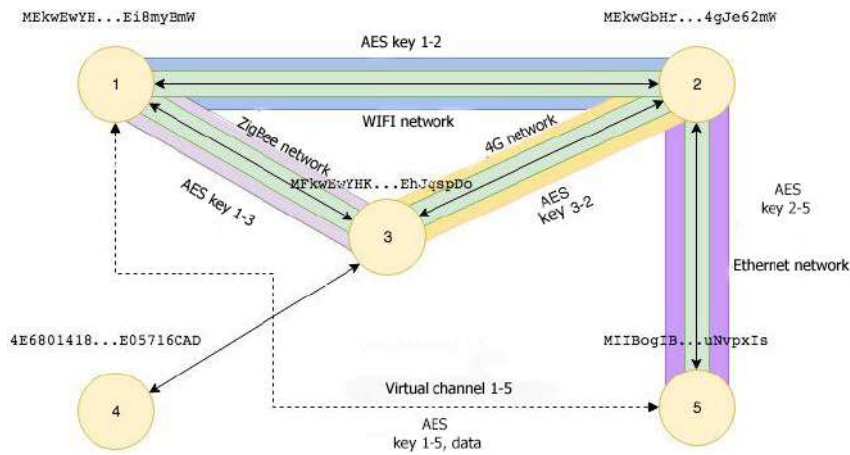


Fig. 1. System model

Important part of system are commands. The command is a regular class that is passed within framework of system and additionally has necessary methods that implement correct behavior.

System work is implemented on commands: routing, launching and publishing applications, receiving data from memory, etc. This makes testing easier and expand functionality and greatly reduces errors in commands and system as a whole [3]. Main purpose of commands - perform certain actions for clients. And the important difference from applications is that they do not work decentralized. To start the command, the client must have an address of another client who will execute this command.

Command can be sent by someone who knows the necessary information on command (data format and its code). Only client who has a registered command with this code can execute the command. In case of receiving a command with the same code but assigned to another client, the behavior is unknown, because the client can't determine if command was sent correctly. With the help of commands, a chain of routing is implemented, applications are published and information from distributed memory is obtained.

For correct delivery of data, client must manually build the necessary chain before transfer. To do this, he performs following: defines client from which to start a search; sends request on presence of the recipient in the list of connected clients; handles the response from the client. If recipient is found, the process is interrupted, otherwise the sender receives a list of all connected customers and begins to search again using another client as a final one. Using this algorithm allows you to control the chain of customers. Client can remove intermediate nodes to increase reliability of system. Implemented chain operates for some time after which it needs to be rebuilt. Therefore, chain changes over time and it makes it more reliable because of physical impossibility of intercepting data in certain areas.

Important part of system is distributed memory that stores all the information that system is operating with, user data and applications itself. Data is stored in memory "as is", so the command or application cares about encryption and decryption of these data. Anyone can get data, but each memory entry has a unique identifier. Therefore, you need to know this identifier to receive data.

Another feature of data in memory is their immutability. After entering data in memory, it becomes impossible to edit them. Distributed memory is a distributed hash table. Physical data is a separate file on storage device, so maximum size of distributed memory depends on number of clients and on how much each client allows to occupy space on data storage device for memory usage. The minimum and maximum data size that can be stored in memory does not exist.

The application in this system is program written using any programming language that is unique in creating decentralized applications. Programs functioning in a decentralized system and on the local computer is no different as they are identical. Applications interact with system as follows: send replies and new data to customers; cause commands; launch new applications; read and write data from distributed memory. Additionally, they can use graphic adapters for complex calculations, make requests to the Internet or work with the file system.

To ensure security of the data, applications are launched in a special, isolated environment. We are using Docker to support running of applications developed in any programming language and different operation systems. Using an isolated environment allows you to separate an application from a physical computer during execution by protecting data and user files.

This approach allows you to run several applications in parallel, provide them with isolated work from each other and from the client operating system. Docker also allows maximum application portability and uses open source technology to easily test applications in a seamless environment before publishing to the system.

Implementation of the system has been performed in form a set of different modules that work together to fulfill the tasks. Using the modular structure allows you to change parts of the system without downtime of system as a whole in cases: transferring the application to another operating system, or not supporting the already existing code in whole or in part [3].

System has the following modules: networks; hashing; encryption; serialization; messages; commands; applications; memory; extensions and routing module. Some modules contain little code so they can be replaced by direct calls to required functions in system body, but process of transferring code to other systems will become more complex and require much more time.

These modules include encryption, hashing, serialization and network modules. Extension module can also be attributed to this set of modules, but it does not integrate into other system modules and should work separately. Other modules are much more complex and perform the basic system functionality. These are message modules, commands, applications, memory, extensions, and routing. Any system module is used in another module or use other modules for calculations or data (Fig. 2).

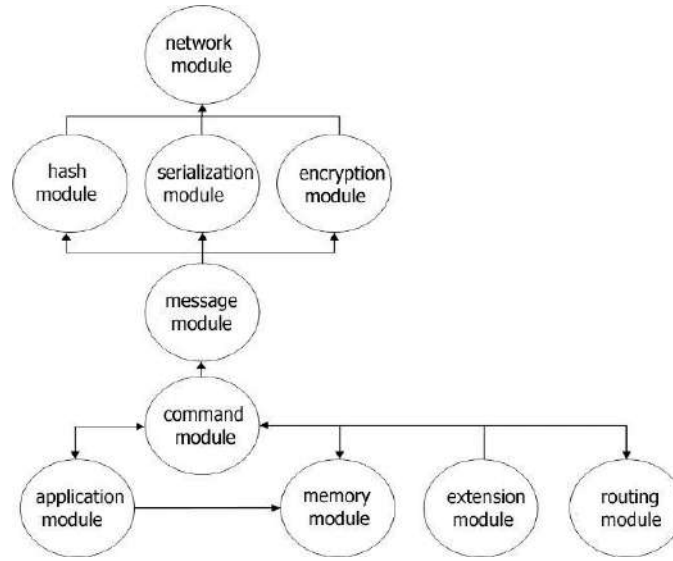


Fig. 2. Interconnection of system modules

Splitting system into separate modules allows you to easily test behavior of modules. Since each separate module is small compared to entire system, its development time and complexity are minimal and this greatly increases the system reliability [3].

System can work in two modes: transit and client. In transit mode, application only transmits the data from sender to recipient, system does not change or receive the data and works only with header. In client mode, application receives data from other users and initializes transfer. It also works simultaneously in transit mode for transmission of messages that are not assigned to current client.

The routing module is one of the most important on network. Algorithm for on-ion routing is chosen routing base. With one difference - not all encryption package, but only the chain address. This approach allows the transfer of information safely and unexpectedly, since the transfer route that is defined by the client and route is securely encrypted.

The chain segments are encrypted the previous segment from end, in the way that a separate key is used for each segment. The closer the segment is to the end, the more times it is encrypted, so the number of encryption can be calculated by the formula (1):

$$N_e = i - 1, \quad (1)$$

where N_e – number of encryption iterations; i – number of current chain segment.
 Algorithm of building the chain is given in Fig.3.

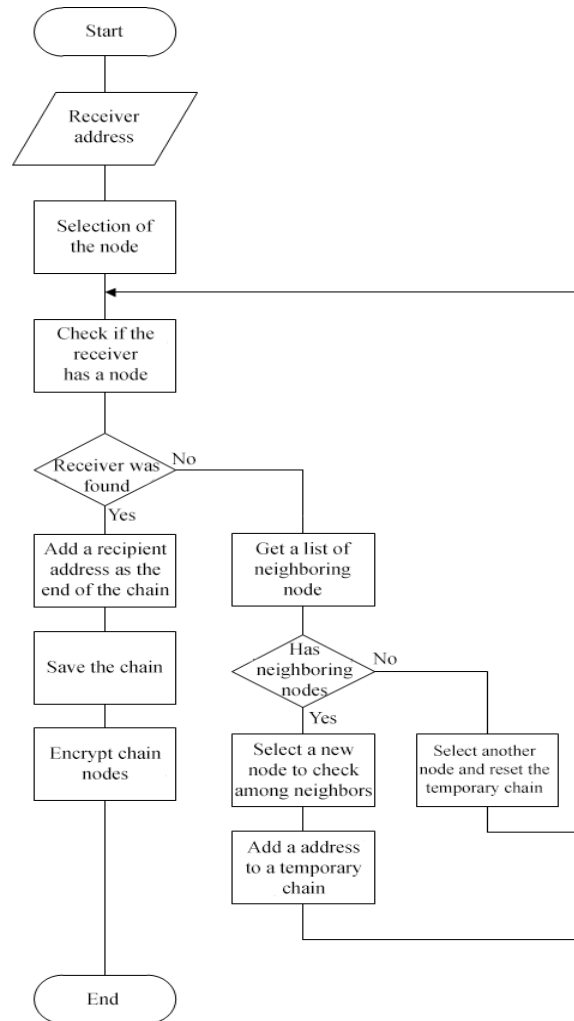


Fig. 3. Algorithm of building a nodes chain for information transmission

For system to work properly, individual modules should only work with module interface. Each module consists of the minimum set of classes. For correct modules operation programmer must implement them correctly. Modules functioning is closely linked to implementation of dependencies. For example, we have hash module interface.

```

from abc import abstractmethod
class hashing ():

```

```
@abstractmethod
def encode (self, object):
    passport
```

And also an example of interface implementation.

```
from dc.module.hashing import hashing
from Crypto.Hash import SHA256
class HashingImplementation (Hashing):
def encode (self, object):
    try:
        h = SHA256.new (data = object)
    except ValueError as e:
        h = SHA256.new (data = bytes (object, 'utf-8'))
    return h.hexdigest ()
```

Dependencies registration and usage example is given below.

```
from dc.injection import inject
from dc.module.hashing import Hashing
from module.hashing import HashingImplementation
inject.get_instance().bind(Hashing, HashingImplementation)
# using:
inject.get_instance().get(Hashing).encode(...)
```

Implementation is completely separate from interface and if there is an error in module it can be corrected without touching other modules and without affecting algorithm of system as a whole.

5 System testing

When connecting to the system, the client must have an IP address already connected to client system. If a connection with client who does not have a connection to other clients occurs, a new isolated network with its own memory, applications, and users is created. After opening a socket and establishing a connection, each client sends own public key to another client. After receiving the key, which is the client's address on the network, a shared key is generated [24].

When sending data, client collects it in a certain class, which represents this data in a virtual form, and receives address of recipient. After filling class, sender initiates the preparation for dispatch. At this time, data is converted into a package – it is serialized and segmented if necessary. After the segmentation, the data is transmitting to the routing module, which defines chain of nodes through which route of their delivery is laid. If the chain is already built and it is relevant, each segment is individually encapsulated into a transit packet and transmitted to network module, which transmits the data.

If the chain is not found, the connected client from the list of clients chooses another client, taking into account the necessary criteria and asks for connection with

the recipient. In case of communication with the recipient, sender adds this client to the list of nodes and decides whether the received chain corresponds to certain criteria. Data is encrypted as required. Segment is encrypted with shared key with the recipient. When receiving a package by a transit client, it reads the chain of nodes, removes itself from chain and sends package further according to route. Once packet is received, the recipient can decrypt it by generating a shared key. After receiving all segments, those segments are merged into one package and deserialized. The following is sent to a command or application, or executed if data is a command.

If it is necessary to start the service, client uses the command, sends the request for an execution to the system closest to client with whom he has a direct connection. This system is a "router" and acts as an intermediate element between client who initiated the launch and the working program. Through this node that exchange of information between the client and the application takes place. Intermediate client copies application from memory and sends a broadcast packet containing an application information request. Upon request, each client decides whether to execute the application or not. If it can, then the client sends answer to "router".

After receiving response, the application is sent to this client, while others will be denied. Client who runs the application runs an isolated environment and copies the application to it. After that, application starts up, provides the necessary service, completes execution and informs the router.

Upon receiving the message about the completion of the application, "router" informs the client and is deleted. While running the application, he can send responses to client requests about the implementation of the application to the "router" and from the "router" to the client. The chain in this case is not being built, and transfer process is the same as during the transfer between different clients.

For testing purpose, an application that executes multiplication of matrices is created, runs in same thread and is written in Java programming language. According to the testing scenario, user received a matrix, which is the result of multiplying two input matrices of 200x200 size. Testing was carried out on local and remote nodes. On the local node - the application was launched in usual environment of operating system, and on the remote node - in an isolated environment. During the testing, 50 matrix multiplications were performed. Matrix elements are eight bytes long.

```
import java.util.Random;
public class Main {
    public static void main(String[] args) {
        (new Main()).run();
    }
    private void run() {
        int round = 50;
        int rounds = 100;
        int matrixSize = 200;
        do {
            long[][] matrixA = this.createMatrix(matrixSize, matrixSize);
            long[][] matrixB = this.createMatrix(matrixSize, matrixSize);
            this.fill(matrixA, matrixB, matrixSize);
            long startedAt = System.currentTimeMillis();
            long[][] matrixC = this.multiply(matrixSize, matrixA, matrixB);
```

```

        long endedAt = System.currentTimeMillis();
        System.out.println(endedAt - startedAt);
        round++;
    } while (round < rounds);    }
    private long[][] createMatrix(int n, int m) {
        return new long[n][m];    }
    private void fill(long[][] matrixA, long[][] matrixB, int
matrixSize) {
        Random rand = new Random();
        for (int i = 0; i < matrixSize; i++) {
            for (int j = 0; j < matrixSize; j++) {
                matrixA[i][j] = rand.nextLong();
                matrixB[i][j] = rand.nextLong();
            }    }
        private long[][] multiply(int size, long[][] matrixA, long[][]
matrixB) {
            long[][] matrix = this.createMatrix(size, size);
            for (int i = 0; i < size; i++) {
                for (int j = 0; j < size; j++) {
                    for (int k = 0; k < size; k++) {
                        matrix[i][j] += matrixA[i][k] + matrixB[k][j];
                    }
                }
            }
        }
        return matrix;
    }
}

```

To construct the graph, the portion of measurements was used. A non-optimal algorithm was used for testing, since it can clearly be seen the difference in performance of the applications. Results of testing in isolated and normal environments are shown in Fig.4, where: dotted line is an isolated, continuous - usual environment.

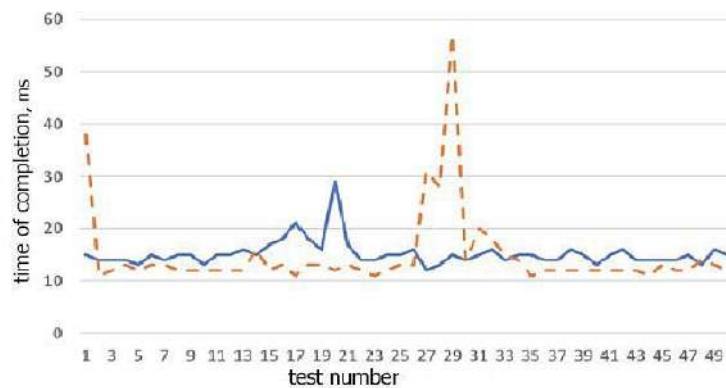


Fig. 4. Test results

Speed of information transfer in the system falls significantly due to some reasons: maximum transmission speed can't exceed the minimum speed between any segment of the transmission chain; there is an additional time for the building of chain; data transmitted by the system is encrypted; additional data is transmitted together with the payload.

During testing, a chain of four segments was used, two of which were intermediate. The following data was obtained: average size of additional data is 222 bytes; size of encrypted data is 256 bytes. It should be noted that through of use of encryption, minimum amount of data transmitted was about 159 bytes, even when transmitting several bytes of data. Thus, when using default MTU on Fast Ethernet (1500) network and TCP protocol, instead of transmitting 1476 bytes of information, it is possible to transfer no more than 1000 bytes (system limitation).

On pair with encryption and additional information, volume of information is 1745 bytes. Thus, the drop of transmission speed is additionally about 30 percent. It should be noted that while transmitting small volumes of information, transmission speed is much smaller, since the significant part of data transmitted is the useless data.

Relatively low transmission speed is compensated by speed of application execution. Comparing overall speed with existing systems is impossible because of their principle of operation, namely because of their strict adherence to data processing schedule and delays in information sending and receiving.

As it can be seen from the results, the execution time of one and the same application is different [24]. Most applications in an isolated environment run faster.

6 Conclusion

In this paper the stages of implementation computing support system in decentralized networks of different purposes is presented. It is proposed to use a decentralized system to run services over an isolated operating system, which makes it possible to process data faster and without the need to have its own computing power.

Proposed system can be used as a typical example in many spheres of human activity: in the legal and financial spheres when automating the signature of contracts as a third independent party; in medicine for formation of single patients database, with further development of the "virtual doctor" module; in the Internet of things, as a platform for operation of devices and their applications, using the transparent interaction of different parts of the system among themselves; in military sphere, as a platform for the automation of machines and safe coordination of military units, using open communication channels.

In the future, it is planned to expand the system by integrating it into mobile devices and other existing Internet services.

References

1. IETF Tools. Mobile Ad hoc Networking (MANET). <http://tools.ietf.org/html/rfc2501> (1999). Accessed 03 Feb 2019

2. Murarka, N.: Decentralized and Trustless Networks. <https://hackernoon.com/decentralized-and-trustless-networks-f881671fae4e> (2018). Accessed 10 Feb 2019
3. Kirichek, G., Rudkovsky, O., Timoshenko, V.: Decentralized computing support system. Scientific notes of TNU named V.I. Vernadsky Series: Engineering. Kyiv, vol. 29 (68), no.6, pp. 161-166 (2018)
4. A Next-Generation Smart Contract and Decentralized Application Platform. <https://github.com/ethereum/wiki/wiki/White-Paper> (2019). Accessed 10 Feb 2019
5. Akhmetov, B., Korchenko, A., Sidenko, V. et al: Applied Cryptology: Encryption Methods. Almaty: KazNRTU (2015)
6. Kirichek, G., Kurai, V.: Implementation quadtree method for comparison of images. In: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018, Slavske, pp.129-132 (2018) doi: 10.1109/TCSET.2018.8336171
7. Klima, V.: Tunnels in Hash Functions: MD5 Collisions Within a Minute. <https://eprint.iacr.org/2006/105.pdf> (2006). Accessed 10 Feb 2019
8. Rivest, R.: The MD5 Message-Digest Algorithm. <https://www.ietf.org/rfc/rfc1321.txt> (1992). Accessed 10 Feb 2019
9. Wang, X., Feng, D., Lai, X.: Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEMD. <https://eprint.iacr.org/2004/199.pdf> (2004). Accessed 10 Feb 2019
10. Eastlake, D., Jones, P.: US Secure Hash Algorithm 1 (SHA1). <https://tools.ietf.org/html/rfc3174> (2001). Accessed 10 Feb 2019
11. Eastlake, D., Hansen, T.: US Secure Hash Algorithms. <https://tools.ietf.org/html/rfc4634> (2006). Accessed 10 Feb 2019
12. Announcing the advanced encryption standard (AES). <http://www.impic.org/papers/Aes-192-256.pdf> (2001). Accessed 10 Feb 2019
13. Karn, P., Metzger, P., Simpson, W.: The ESP Triple DES Transform. <https://tools.ietf.org/html/rfc1851> (1995). Accessed 10 Feb 2019
14. Rivest, R., Shamir, A., Ademan, L.: Cryptographic communications system and method. <https://people.csail.mit.edu/rivest/pubs/RSA83b.pdf> (1983). Accessed 10 Feb 2019
15. Rescorla, E.: Diffie-Hellman Key Agreement Method. <https://tools.ietf.org/html/rfc2631> (1999). Accessed 10 Feb 2019
16. Ship, M., Yiqun, L.: Cryptanalysis of Twofish (II). <https://www.schneier.com/twofish-analysis-shiho.pdf> (1999). Accessed 10 Feb 2019
17. Lafore, R.: Data Structures and Java Algorithms, St. Petersburg: Peter (2013)
18. Deitel, P., Deitel, H.: Visual C# How to Program. Pearson (2016)
19. Burger, W., Burge, M.J.: Digital Image Processing: An Algorithmic Introduction Using Java. Springer (2016)
20. Heineman, D., Pollis, G., Selkov, S.: Algorithms. Reference with examples in C, C ++, Java and Python. Moscow: Alpha-book (2017)
21. Malik D.S.: C++ Programming: From Problem Analysis to Program Design, ISBN-13: 978-1337102087 (2017)
22. Corner, A.: Elliptic Curve Cryptography: a gentle introduction. <http://andrea.corbellini.name/2015/05/17/elliptic-curve-cryptography-a-gentle-introduction/> (2015). Accessed 10 Feb 2019
23. Recommendation for Pair-Wise Key Establishment Schemes Using Discrete Logarithm Cryptography. <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-56a.pdf> (2007). Accessed 10 Feb 2019
24. Kirichek, G., Ovchar, A.: The system of data protection at transmission in the network. Scientific notes of Chernovtsy National University. Series: Computer Systems and Components, t.7, vol.2, pp. 32-39 (2016)

Soft Decoding Based on Ordered Subsets of Verification Equations of Turbo-Productive Codes

Alexandr Kuznetsov ¹[0000-0003-2331-6326], Anastasiia Kiian ¹[0000-0003-2110-010X],

Kateryna Kuznetsova ¹[0000-0002-5605-9293], Tetiana Ivko ¹[0000-0003-1772-0074],

Oleksii Smirnov ²[0000-0001-9543-874X], Dmytro Prokopovych-Tkachenko ³[0000-0002-6590-3898]

¹V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, 61022, Ukraine
kuznetsov@karazin.ua, nastyak931@gmail.com,
kate.kuznetsova.2000@gmail.com, t.ivko@outlook.com

²Central Ukrainian National Technical University, avenue University, 8, Kropivnitskiy, 25006,
Ukraine, dr.smirnovoa@gmail.com

³University of Customs and Finance, st. Volodymyr Vernadsky, 2/4, Dnipro, 49000, Ukraine,
omega2@email.dp.ua

Abstract. Methods of soft decoding of cascade code constructions based on the schemes-products of linear block codes (Turbo Product Codes) are considered. An approach is being developed based on the iterative exchange of soft solutions between block codes constituting a cascade design. It is shown that a sequential execution of procedures for the formation of ordered subsets of test equations and the logarithms estimation of a likelihood ratio allows decoding of turbo-productive codes according to the criterion of minimizing the erroneous reception of code symbols.

Keywords. Cascade Structures, Turbo Product Codes, Soft Decoding, Verification Equations, Noise Immunity.

1 Introduction

A promising area in the development of noise-resistant coding theory is cascade code structures [1-7, 32-33], methods and algorithms for their decoding with an iterative exchange of soft solutions that allow to provide a required noise immunity of discrete message transmission [8-14].

It should be noted that the implementation complexity of decoding methods based on the use of decision functions increases with length of the code and the correcting capacity [14-17]. Decoding complexity can be reduced by using decision functions defined on a preformed subset of check equations [18-20]. At the same time, this decrease also leads to a decrease in the energy gain [19, 20].

Thus, an actual direction of research is a development (improvement) of decoding methods with soft solutions based on decisive functions, which, without significantly reducing the energy gain from coding, would significantly reduce the complexity of

practical implementation. A promising direction in this sense is the formation of ordered subsets of test equations and decoding methods based on them.

2 Theoretical substantiation of the proposed decoding method

The theoretical basis for soft decoding methods is a criterion for testing hypotheses, the mathematical justification for which is based on the total probability formula and the Bayes theorem [18-20].

Suppose that one can make mutually M exclusive assumptions (hypotheses) $H_1, H_2, \dots, H_M$ about the setting of the experience, and an event A can appear only with one of these hypotheses.

Then the probability of an event is calculated by the formula of total probability:

$$\begin{aligned} P(A) &= P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots + P(H_M)P(A|H_M) = \\ &= \sum_{i=1}^M P(H_i)P(A|H_i), \end{aligned}$$

where $P(H_i)$ is the probability of the hypothesis H_i ; $P(A|H_i)$ - conditional probability of an event A with this hypothesis.

If prior to the experiment, probabilities of the hypotheses were $P(H_i), i = 1, 2, \dots, M$ and as a result of the experiment an event A occurred, then the a posteriori (experimental, subject to the occurrence of the event A) hypotheses probabilities are calculated using the Bayes formula:

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{i=1}^M P(H_i)P(A|H_i)}, \quad i = 1, 2, \dots, M.$$

The Bayes formula makes it possible to calculate the conditional probabilities of occurrences of the following events, taking into account the posterior probabilities of hypotheses, $P(H_i|A), i = 1, 2, \dots, M$. So, if after the first experiment in which an event A occurred, the next experiment B is performed, in which an event may occur, the conditional probability $P(B|A)$ is calculated using the formula of total probability, into which not a priori probabilities $P(H_i)$ are substituted, but a posteriori, calculated after the occurrence of the event A , probabilities $P(H_i|A)$, i.e. we will receive:

$$P(B|A) = \sum_{i=1}^M P(H_i|A)P(B|H_iA), \quad i = 1, 2, \dots, M,$$

where $H_i A$ is an event A under the hypothesis H_i , $P(B|H_i A)$ is the conditional probability of co-being B under the hypothesis H_i and event A .

Suppose now that the demodulator, based on the observation of the received signal and noise interference, estimates which of the possible signals $S_i \in \{S_1, S_2, \dots, S_M\}$ (from an ensemble of signals with power M) was transmitted. You will make mutually exclusive assumptions M (hypotheses) that the corresponding signal S_i , $i = 1, 2, \dots, M$ has been transmitted. We calculate the posterior probability of the hypothesis, subject to admission: S^*

$$P(S_i|S^*) = \frac{P(S_i)P(S^*|S_i)}{\sum_{i=1}^M P(S_i)P(S^*|S_i)}, \quad i = 1, 2, \dots, M, \quad (1)$$

where $P(S_i)$ - a priori probability of formation S^* of a signal S_i by the transmitter; $P(S^*|S_i)$ - conditional probability of reception under the condition that the signal S_i is formed by the transmitter.

It is usually S^* represented as a continuous random variable underlying the hypothesis testing criteria. Consider the probability distribution function $P(S^*)$:

$$P(S^*) = \sum_{i=1}^M P(S_i)P(S^*|S_i).$$

$P(S^*)$ - is a probability distribution function of the mixture of signal and interference S^* , which gives test statistics in the full signal space $\{S_1, S_2, \dots, S_M\}$.

In equation (1), the value of the function $p(S^*)$ is the scaling factor, since the value $P(S^*)$ is obtained by averaging over the entire space of the signals.

Consider a case for two signals. Let binary logic elements 1 and 0 be represented by signals $S_1 = 1$ and $S_2 = -1$. A rigid decision rule, called as a *maximum likelihood rule*, determines a choice of one of the hypotheses (corresponding to the transmission of signals S_1 and S_2 , accordingly) based on the comparison of probabilities values $P(S^* = x|S_1)$ and $P(S^* = x|S_2)$ the choice of the larger one. For each data bit transmitted, it is decided that the signal S_1 was transmitted if $S^* = x$ falls on the right side of the decision line (indicated γ), or that the signal S_2 was otherwise transmitted.

A similar decision rule, known as the *maximum a posteriori probability* (MAP), can be represented as a minimum error probability rule, taking into account the prior probability of data. In general, the MAP rule is expressed as follows:

$$S = \begin{cases} S_1, & \text{if } P(S^* = x|S_1) > P(S^* = x|S_2) \\ S_2, & \text{if } P(S^* = x|S_1) < P(S^* = x|S_2) \end{cases}, \quad (2)$$

where S - value of the signal corresponding to the decision.

Thus, expression (2) establishes the rule for choosing one of the hypotheses corresponding to the signals S_1 and S_2 . Using expression (1), we obtain the equivalent expression:

$$S = \begin{cases} S_1, & \text{if } P(S_1)P(S^*|S_1) > P(S_2)P(S^*|S_2) \\ S_2, & \text{if } P(S_1)P(S^*|S_1) < P(S_2)P(S^*|S_2) \end{cases},$$

where probability

$$P(S^*) = \sum_{i=1}^M P(S_i)P(S^*|S_i)$$

in both parts of inequality reduced.

Using (2) we introduce a function as a ratio of likelihood functions $P(S^* = x|S_1)$ and $P(S^* = x|S_2)$:

$$F = \frac{P(S_1)P(S^*|S_1)}{P(S_2)P(S^*|S_2)}, \quad (3)$$

then the rule for choosing one of the hypotheses is written as

$$S = \begin{cases} S_1, & \text{если } F > 1 \\ S_2, & \text{если } F < 1 \end{cases}. \quad (4)$$

Let us translate the expression (3), we get:

$$\ln F = \ln \left(\frac{P(S_1)}{P(S_2)} \right) + \ln \left(\frac{P(S^*|S_1)}{P(S^*|S_2)} \right).$$

Thus, a logarithm of the ratio of likelihood functions $\ln F$ is a real representation of the soft solution at the decoder input, with first term on right side of the equality being the logarithm of the relations of a priori probabilities $P(S_1)$ and $P(S_2)$

$$L_S(S_1, S_2) = \ln \left(\frac{P(S_1)}{P(S_2)} \right),$$

and the second term is the essence of the logarithm of the posterior probability ratio $P(S^*|S_1)$ and $P(S^*|S_2)$:

$$L_{DS}(S_1, S_2) = \ln \left(\frac{P(S^*|S_1)}{P(S^*|S_2)} \right)$$

as a result of channel measurements in the receiver.

So, the logarithm of the likelihood function $L_{FS} = \ln F$ is rewritten as

$$L_{FS}(S_1, S_2) = L_S(S_1, S_2) + L_{DS}(S_1, S_2). \quad (5)$$

It should be noted that for AWGN channels, the logarithm of the likelihood function as the result of channel measurements of the received mixture of signal and noise in the receiver will be as follows:

$$\begin{aligned} L_{DS}(S_1, S_2) &= \ln \left[\frac{P(S^*|S_1)}{P(S^*|S_2)} \right] = \ln \left[\frac{\frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{S^*-1}{\sigma} \right)^2 \right]}{\frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{S^*+1}{\sigma} \right)^2 \right]} \right] = \\ &= -\frac{1}{2} \left(\frac{S^*-1}{\sigma} \right)^2 + \frac{1}{2} \left(\frac{S^*+1}{\sigma} \right)^2 = \frac{2}{\sigma^2} S^*. \end{aligned}$$

Considering the ratio

$$\frac{1}{\sigma^2} = \frac{2E_b}{N_0},$$

where $\frac{E_b}{N_0}$ - is the ratio of energy of a binary signal E_b to the spectral power density of the noise N_0 , we obtain:

$$L_{DS}(S_1, S_2) = 4 \cdot \frac{E_b}{N_0} S^*,$$

those value of a logarithm of posterior probability ratio $P(S^*|S_1)$ and $P(S^*|S_2)$, as a result of channel measurements at the receiver, depends exclusively on the signal-to-noise ratio and the value of the received signal and noise mixture S^* .

In [20], it was shown that for systematic codes, the soft decision at the decoder output (on a logarithmic scale) about received symbol is written in the form of expression

$$L_{FDK}(S_1, S_2, C_1, C_2) = L_{FS}(S_1, S_2) + L_{DK}(c_1, c_2), \quad (6)$$

where $L_{DK}(C_1, C_2)$ is the logarithm of the likelihood function relation on the received symbol, obtained as a result of decoding.

Substituting (5) into (6) we get:

$$L_{FDK}(S_1, S_2, C_1, C_2) = L_S(S_1, S_2) + L_{DS}(S_1, S_2) + L_{DK}(c_1, c_2), \quad (7)$$

those the soft decision at the decoder output depends on three values: $L_S(S_1, S_2)$ - a logarithm of the ratio of the prior probabilities of the signals S_1 and S_2 ; -a logarithm of the ratio of the posterior probabilities of the signals S_1 and S_2 (the result of channel measurements) and $L_{DK}(C_1, C_2)$ - a logarithm of ratio of the likelihood functions of binary code symbols C_1 and C_2 as the result of decoding.

To get $L_{FDK}(S_1, S_2, C_1, C_2)$, you need to sum up the individual contributions, since all three components are statistically independent [20]. Soft decoder output $L_{FDK}(S_1, S_2, C_1, C_2)$ is a real number, providing both the hard decision itself and its reliability. The sign $L_{FDK}(S_1, S_2, C_1, C_2)$ sets a hard decision, i.e.:

$$c_i = \begin{cases} C_1 = 1, & \text{if } L_{FDK}(S_1, S_2, c_1, c_2) > 0 \\ C_2 = 0, & \text{if } L_{FDK}(S_1, S_2, c_1, c_2) < 0 \end{cases}, \quad (8)$$

where c_i is the value of the i -th bit corresponding to the taken decision.

An eigenvalue $L_{FDK}(S_1, S_2, C_1, C_2)$ determines the reliability of the decision.

As a rule, a value $L_{DK}(C_1, C_2)$ has the same sign as $L_{FDK}(S_1, S_2, C_1, C_2)$, thus increasing the reliability of the decision.

For statistically independent values x and y , the sum of two logarithmic likelihood ratios $L(x)$ and $L(y)$ is determined by the following expression:

$$\begin{aligned} L(x)[+]L(y) &= L(x \oplus y) = \ln \left[\frac{e^{L(x)} + e^{L(y)}}{1 + e^{L(x)}e^{L(y)}} \right] \approx \\ &\approx (-1) \times \text{sgn}[L(x)] \times \text{sgn}[L(y)] \times \min(|L(x)|, |L(y)|), \end{aligned} \quad (9)$$

where function $\text{sgn}[z]$ returns a sign of its argument z , and the sign " $\oplus$ " is used to denote the sum of data modulo 2 represented by binary digits. The sign $[+]$ is used to denote the sum of the logarithms of the likelihood functions, which is defined as the logarithm of the likelihood function of the sum modulo 2 of the corresponding arguments.

An implementation of the turbo decoding procedure involves the use of decoding methods with a soft solution at the input and a soft solution at the output. During the first iteration on such a decoder, the data is considered equally probable, which gives the initial a priori value $L_S(S_1, S_2) = 0$ in equation (7). Channel measurement gives the value $L_{DS}(S_1, S_2)$ that is obtained by taking the logarithm of the ratio of the values $P(S^* = x|S_1)$ and $P(S^* = x|S_2)$ for certain values and is the second member of equation (7). The decoder output $L_{DK}(C_1, C_2)$ is information derived from the decoding process. For iterative decoding, the external likelihood is fed back to the input (of another composite decoder) to update the prior probability of the next iteration information, i.e. updates a priori probability:

$$L_S(S_1, S_2) = L_{DK}(C_1, C_2).$$

Thus, the decision in the final decoding of each character of the code sequence and information about its reliability depends on the value $L_{FDK}(S_1, S_2, C_1, C_2)$. Based on equation (7), we write the algorithm that gives an estimate of the soft output of the decoder $L_{DK}(C_1, C_2)$ and the resulting estimate $L_{FDK}(S_1, S_2, C_1, C_2)$.

1. Install $L_S(S_1, S_2) = 0$.
2. We decode with the soft solution the first composite code, i.e. find a soft solution $L_{FDK}(S_1, S_2, C_1, C_2)$.
3. Based on equation (7) we calculate

$$L_{DK}(C_1, C_2) = L_{FDK}(S_1, S_2, C_1, C_2) - L_S(S_1, S_2) - L_{DS}(S_1, S_2)$$
4. For the following composite code install $L_S(S_1, S_2) = L_{DK}(C_1, C_2)$.
5. With a soft solution, we decode the following composite code, i.e. find a soft solution $L_{FDK}(S_1, S_2, C_1, C_2)$.
6. For all composite codes, repeat steps 3-5.
7. The result of turbo decoding is a hard decision about a code symbol c by expression (8) based on the soft decision obtained in the last step $L_{FDK}(S_1, S_2, C_1, C_2)$.

Thus, as the analysis of above algorithm shows, the main task in implementation of turbo decoding is a development of efficient soft decoding procedures for composite codes, i.e. development of soft decision $L_{DK}(C_1, C_2)$ calculation procedures for an iterative exchange procedure in the process of turbo decoding.

We study the procedures for finding the soft solution $L_{DK}(C_1, C_2)$ at the decoder output, analyze the possible ways to calculate the last term on the right side of equality (7) - the logarithm of the ratio of the likelihood functions of binary code symbols C_1 and C_2 as a result of decoding.

Consider a linear (n, k, d) block code over a finite field $GF(2)$. A linear code as a subspace $GF^k(2) \subseteq GF^n(2)$ is defined by the generator matrix G , the lines of which

form the basis of the linear space $GF^k(2)$. By definition, for each linear code there is an orthogonal completion - a subspace $GF^{n-k}(2) \subseteq GF^n(2)$, all elements of which are orthogonal to the elements of $GF^k(2)$. The basis of the linear space $GF^{n-k}(2)$ is given by the check matrix H , and the mutual orthogonality condition implies equality $GH^T = 0$, where by "0" is meant the $k \times r$ matrix of zero elements $GF(2)$.

We write the last equality in the form $cH^T = 0$, where $c = (c_0, c_1, \dots, c_{n-1})$ is the arbitrary code word of the linear block (n, k, d) code under consideration, i.e. $c \in GF^k(2)$ $c_i \in [0, 1]$.

Taking into account the fact that all elements $GF^{n-k}(2)$ can be expressed in terms of a linear combination of rows of a check matrix H , we have $ch_i^T = 0$:, where $h_i = (h_{i_0}, h_{i_1}, \dots, h_{i_{n-1}})$ is an arbitrary vector obtained by a linear combination of rows of a matrix H , $i = 0, 1, \dots, 2^{n-k} - 1$.

In other words, the last equality holds for all 2^{n-k} vectors from $GF^{n-k}(q)$ and we have a system of test equations:

$$\begin{cases} c_0 h_{0_0} + c_1 h_{0_1} + \dots + c_{n-1} h_{0_{n-1}} = 0; \\ c_0 h_{1_0} + c_1 h_{1_1} + \dots + c_{n-1} h_{1_{n-1}} = 0; \\ \dots \\ c_0 h_{(2^{n-k}-1)_0} + c_1 h_{(2^{n-k}-1)_1} + \dots + c_{n-1} h_{(2^{n-k}-1)_{n-1}} = 0. \end{cases} \quad (10)$$

Suppose now that the code word $c = (c_0, c_1, \dots, c_{n-1})$ is taken by the criterion of the maximum a posteriori probability, i.e. the values of the log-rhymes of the posterior probabilities $P(S^*|S_1)$ and $P(S^*|S_2)$:

$$L_{DS}(c_j) = L_{DS}(S_1, S_2) = \ln \left(\frac{P(S^*|S_1)}{P(S^*|S_2)} \right)$$

about each code symbol c_j , $j = 0, 1, \dots, n-1$ as a result of channel measurements of the corresponding signals in the receiver.

The logarithms of the relations of a priori probabilities $P(S_1)$ and $P(S_2)$, corresponding to each of the code symbols c_j , $j = 0, 1, \dots, n-1$ we denote

$$L_S(c_j) = L_S(S_1, S_2) = \ln \left(\frac{P(S_1)}{P(S_2)} \right).$$

Then, taking into account (7) and rule (9) for the i -th checking equation, we have

$$L_{DK_i}(c_j) = \begin{cases} \left((L_S(c_0) + L_{DS}(c_0))h_{i_0} [+] (L_S(c_1) + L_{DS}(c_1))h_{i_1} [+] \dots [+] \right. \\ \left. [+] (L_S(c_{j-1}) + L_{DS}(c_{j-1}))h_{i_{j-1}} [+] (L_S(c_{j+1}) + L_{DS}(c_{j+1}))h_{i_{j+1}} [+] \dots [+] \right. \\ \left. [+] (L_S(c_{n-1}) + L_{DS}(c_{n-1}))h_{i_{n-1}} \right) = \sum_{\substack{l=0, \\ l \neq j}}^{n-1} (L_S(c_l) + L_{DS}(c_l))h_{i_l} & \text{if } h_{i_j} = 1; \\ \left((L_S(c_0) + L_{DS}(c_0))h_{i_0} [+] (L_S(c_1) + L_{DS}(c_1))h_{i_1} [+] \dots [+] \right. \\ \left. [+] (L_S(c_{n-1}) + L_{DS}(c_{n-1}))h_{i_{n-1}} \right) = \sum_{l=0}^{n-1} (L_S(c_l) + L_{DS}(c_l))h_{i_l} & \text{if } h_{i_j} = 0, \end{cases} \quad (11)$$

where the summation of "[+]" and " $\sum$ " is carried out according to the rule of adding likelihood logarithms, i.e. by expression (9).

If we assume that all the estimates $L_{DK_i}(c_j)$ $j = 0, 1, \dots, n-1$ are statistically independent (for example, if the test equations are mutually orthogonal), then the resulting estimate $L_{DK}(c_j)$ will be written as:

$$L_{DK}(c_j) = \sum_{i=0}^{2^{n-k}-1} L_{DK_i}(c_j), \quad (12)$$

where the summation is performed according to the usual arithmetic rule of addition of real numbers.

The soft output of the decoder $L_{FDK}(c_j) = L_{FDK}(S_1, S_2, C_1, C_2)$ is a real number, and is determined by the expression (7):

$$\begin{aligned} L_{FDK}(c_j) &= L_S(c_j) + L_{DS}(c_j) + L_{DK}(c_j) = \\ &= L_S(c_j) + L_{DS}(c_j) + \sum_{i=0}^{2^{n-k}-1} L_{DK_i}(c_j). \end{aligned} \quad (13)$$

The sign $L_{FDK}(c_j)$ sets a tough decision according to rule (8):

$$c_j = \begin{cases} C_1 = 1, & \text{if } L_{FDK}(c_j) > 0; \\ C_2 = 0, & \text{if } L_{FDK}(c_j) < 0. \end{cases}$$

Expressions (11), (12) and (13) define the decisive function based on using logarithms of the ratio of likelihood functions of received signals (calculated using a priori and a posteriori probabilities), as well as the logarithm of the ratio of likelihood functions of binary code characters as a result of decoding. The corresponding sum (12) defines the decision function based only on the use of the decoding result.

Let us analyze the expression (12). Expanding the summation sign according to rule (11), we obtain that expression (12) contains 2^{n-k} terms, each of which is the result of summation of the n logarithms of the likelihood of code symbols. In turn, the likelihood logarithms of code symbols are the sum of the likelihood logarithms of the received signals (calculated using a priori and a posteriori probabilities). It is obvious that with an increase in the code parameters (n, k, d) , the number of terms increases rapidly and already with the application $n - k > 32$ of the considered approach it becomes computationally inexpedient. A promising direction in this sense is the development of a rule for the formation of ordered subsets of check equations and a theoretical substantiation on their basis of decisive functions for decoding methods with soft solutions.

3 Conclusions

As a result of the conducted research, the method of soft decoding of cascade code constructions with iterative exchange of soft solutions was improved which differs from the known methods by the accelerated procedure of selecting test equations with the most reliable symbols, which allows realizing decoding of code words by the criterion of minimizing the erroneous reception of code symbols and speeding up the process of turbo decoding of concatenated codes.

The obtained results may be useful in constructing information security code schemes [21-26], for example, as a real alternative to traditional cryptography for post-quantum applications [27]. In addition, research results may be useful for optimizing computing in modern telecommunications networks [28-31, 34-35].

References

1. Gomtsyan, H.A.: Computer simulation of cascade codes for CDMA systems. In: Proceedings of the Second International Symposium of Trans Black Sea Region on Applied Electromagnetism (Cat. No.00TH8519), Xanthi, Greece, 2000, p. 105. (2000) doi:10.1109/AEM.2000.943262
2. Perez, L.C., Costello, D.J.: Cascaded convolutional codes. In: Proceedings of 1995 IEEE International Symposium on Information Theory, Whistler, BC, Canada, 1995, p. 160. (1995) doi:10.1109/ISIT.1995.531509
3. Permuter, H.H., Weissman, T.: Cascade source coding with side information at first two nodes. In: IEEE Information Theory Workshop on Information Theory (ITW 2010, Cairo), Cairo, 2010, pp. 1–5. (2010) doi:10.1109/ITWSPS.2010.5503190
4. Zhang, S., Song, R., Yang, F.: Joint design of QC-LDPC codes for cascade-based multi-source coded cooperation. In: International Conference on Wireless Communications &

- Signal Processing (WCSP), Nanjing, 2015, pp. 1–4. (2015) doi:10.1109/WCSP.2015.7340967
5. Kuznetsov, A., Serhiienko, R., Prokopovych-Tkachenko, D.: Construction of cascade codes in the frequency domain. 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T 2017), Kharkov, 2017, pp. 131–136. (2017) doi:10.1109/INFOCOMMST.2017.8246366
 6. Chen, H., Ling, S., Xing, C.: Quantum codes from concatenated algebraic-geometric codes. *IEEE Transactions on Information Theory*, 2005, vol. 51(8), pp. 2915–2920, (2005) doi:10.1109/TIT.2005.851760
 7. Changuel, S., Le Bidan, R., Pyndiah, R.: Iterative Decoding of Block Turbo Codes over the Binary Erasure Channel. In: *IEEE International Conference on Signal Processing and Communications*, Dubai, 2007, pp. 1539–1542. (2007) doi:10.1109/ICSPC.2007.4728625
 8. Landolsi, M.A.: A Comparative Performance and Complexity Study of Short-Length LDPC and Turbo Product Codes. In: *2nd International Conference on Information & Communication Technologies*, Damascus, 2006, pp. 2359–2364. (2006) doi:10.1109/ICTTA.2006.1684775
 9. Nakajima, S., Sato, E.: Trellis-coded 8-PSK scheme combined with turbo and single-parity-check product codes. In: *Proceedings IEEE 56th Vehicular Technology Conference*, Vancouver, BC, Canada, 2002, vol. 3, pp. 1782–1786. (2002) doi:10.1109/VETECF.2002.1040523
 10. Stasev, Yu.V., Kuznetsov, A.A., Nosik, A.M.: Formation of pseudorandom sequences with improved autocorrelation properties. *Cybernetics and Systems Analysis*, 2007, vol. 43(1), pp. 1–11. (2007) doi:10.1007/s10559-007-0021-2
 11. Berrou, C., Glavieux, A., Thitimajshima, P.: Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes. In: *Proceedings of ICC '93 - IEEE International Conference on Communications*, Geneva, Switzerland, 1993, vol. 2, pp. 1064–1070. (1993) doi:10.1109/ICC.1993.397441
 12. Stasev, Yu.V., Kuznetsov, A.A.: Asymmetric code-theoretical schemes constructed with the use of algebraic geometric codes. *Kibernetika i Sistemnyi Analiz*, 2005, vol. 3, pp. 47–57. (2005)
 13. Wang, F.G., Tang, Y., Yang, F.: The iterative decoding algorithm research of Turbo Product Codes. In: *The 2010 International Conference on Apperceiving Computing and Intelligence Analysis Proceeding*, Chengdu, 2010, pp. 97–100. (2010) doi:10.1109/ICACIA.2010.5709859
 14. MacKay, D.J.C., Neal, R.M.: Near Shannon limit performance of low density parity check codes. *IEEE Electronics Letters*, 1996, vol. 33(6), pp. 457–458. (1996) doi:10.1049/el:19970362
 15. Turbo Product Code Encoder / Decoder. www.aha.com
 16. IEEE 802.16 Broadband Wireless Access Working Group. Turbo Code Comparison (TCC v TPC). <http://ieee802.org/16>
 17. Turbo Product Code FEC. Comtech EF Data Corp. www.comtechefdata.com
 18. MacWilliams, F., Sloane, N.: *The Theory of Error-Correcting Codes*. Elsevier (1977)
 19. Proakis, J.: *Digital communications*. McGraw Hill (2001)
 20. Sklar, B.: *Digital Communications: Fundamentals and Applications*. Prentice Hall Communications Engineering and Emerging Techno. Pearson Education (2016)
 21. QC-MDPC KEM: A Key Encapsulation Mechanism Based on the QC-MDPC McEliece Encryption Scheme”, NIST Submission, 2017. <https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-1-Submissions>

22. Wang, Y.: RLCEKey Encapsulation Mechanism (RLCE-KEM) Specification. NIST Submission. (2017) <http://quantumca.org>
23. Melchor, C.A., Aragon, N., Bettaieb, S., Bidoux, L., Blazy, O., Deneuville, J.-C., Gaborit, P., Zemor, G.: Rank Quasi-Cyclic (RQC). NIST Submission. (2017) <http://pqc-rqc.org>
24. Post-Quantum Cryptography, Round 1 Submissions, 2017. <https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-1-Submissions>
25. Kuznetsov, A., Kiian, A., Lutsenko, M., Chepurko, I., Kavun, S.: Code-based cryptosystems from NIST PQC. In: 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT), Kyiv, Ukraine, 2018, pp. 282–287. (2018) doi:10.1109/DESSERT.2018.8409145
26. Kuznetsov, A., Pushkar'ov, A., Kiyan, A., Kuznetsova, T.: Code-based electronic digital signature. In: 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT), Kyiv, Ukraine, 2018, pp. 331–336. (2018) doi:10.1109/DESSERT.2018.8409154
27. Bernstein, D., Buchmann, J., Dahmen, E.: Post-Quantum Cryptography. Springer-Verlag, Berlin-Heidelberg (2009)
28. Rassomakhin, S.G.: Mathematical and physical nature of the channel capacity. Telecommunications and Radio Engineering, 2017, vol. 76(16), pp. 1423–1451. (2017) doi:10.1615/TelecomRadEng.v76.i16.40
29. Krasnobayev, V.A., Koshman, S.A.: A Method for Operational Diagnosis of Data Represented in a Residue Number System. Cybernetics and Systems Analysis, 2018, vol. 54(2), pp. 336–344. (2018) doi:10.1007/s10559-018-0035-y
30. Gorbenko, I.D., Zamula, A.A., Semenko, A.E., Morozov, V.L.: Method for synthesis of performed signals systems based on cryptographic discrete sequences of symbols. Telecommunications and Radio Engineering, 2017, vol. 76(17), pp. 1523–1533. (2017) doi:10.1615/TelecomRadEng.v76.i17.40
31. Tenth UK Teletraffic Symposium. Performance Engineering in Telecommunications Network. In: Tenth UK Teletraffic Symposium, 10th. Performance Engineering in Telecommunications Network, Martlesham Heath, UK
32. Kavun, S.: Conceptual fundamentals of a theory of mathematical interpretation. Int. J. Computing Science and Mathematics, 2015, vol. 6(2), pp. 107–121. (2015) doi:10.1504/IJCSM.2015.069459
33. Kuznetsov, A., Kavun, S., Panchenko, V., Prokopovych-Tkachenko, D., Kurinniy, F., Shoiko, V.: Periodic Properties of Cryptographically Strong Pseudorandom Sequences. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 129–134. (2018) doi:10.1109/INFOCOMMST.2018.8632021
34. Zamula, A., Kavun, S.: Complex systems modeling with intelligent control elements. Int. J. Model. Simul. Sci. Comput., 2017, vol. 08(01). (2017) doi:10.1142/S179396231750009X
35. Kavun, S., Zamula, A., Mikheev, I.: Calculation of expense for local computer networks. In: Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 2017 4th International, Kharkiv, Ukraine, 2017, pp. 146–151. (2017) doi:10.1109/INFOCOMMST.2017.8246369

Internet-of-Things Device Set Configuration for Connection to Wireless Local Area Network

Yaroslav Krainyk^{1[0000-0002-7924-3878]}, Andrii Razzhyvin¹, Olena Bondarenko¹ and Irina Simakova¹

¹ Petro Mohyla Black Sea National University 10. 68 Desantnykiv str., Mykolaiv, Ukraine
yaroslav.krainyk@chmnu.edu.ua, greatwwgg@gmail.com,
elena.igorevna2205@gmail.com, darkling6667@gmail.com

Abstract. In this paper, new model of interaction in IoT-infrastructure for the set of devices is investigated. Integration into wireless network based on WiFi-technology, software identification, data exchange, and software upgrade problems have been considered for the end-user product that consists of group of devices. The established solution utilizes assignment of the server and client role to devices in the set. The server deploys temporary network so the user device can access it and enter necessary information about network to connect to. Communication mechanisms within customer's network for device set have been also considered. Using proposed models, customer can control device set, send and receive data (particular case for audio data is considered), update firmware and resources of devices directly from his/her own device via specially designed software. We have used set of ESP8266 boards and NodeMCU firmware to implement and test the developed model. The developed prototype system is able to receive configuration of the user network, receive audio data from client device and play audio. This proof-of-concept system can be further extended for the new functionality.

Keywords: device, device set, Internet-of-Things, configuration.

1 Introduction

In the era of Internet-of-Things (IoT) smart devices are emerging to provide different services to customers. They are built up to execute set of functions clients are interested in. Usually, the device is delivered to the customer as single ready-to-use gadget. Its configuration can be performed as any other device in the local network. However, it is not an obligatory and if the consumer gets set of devices rather than single device, individual configuration of every device in the set could be time-consuming. Moreover, if the set of devices are to be moved into another network, the manual configuration task should be performed repeatedly. Thus, apparatus for automatic configuration for the set of devices is necessary to satisfy best user experience practices.

ESP8266 WiFi-module [1] has been recognized as one of the most useful devices for IoT-applications. It has tremendous performance (runs on frequency up to 160

MHz) as for embedded device and combines it with low price. Due to the availability of digital interfaces (UART, SPI, I2C), it is a perfect candidate to perform sensor readings and provide network connectivity. Low-power consumption modes are available and, as a result, it can work even with battery power supply. Numerous novel investigations in IoT are based on usage of ESP8266. Hence, in this paper, we selected ESP8266 device as a sample WiFi-module for further description of the process and its implementation. As ESP8266 relies on the WiFi-technology, the proposed approach can be applied to modules of different producers.

After the devices from the set are configured, they can perform necessary programmed functions. In spite of that, usually, IoT-system development supposes that device constraints should be taken into consideration. As a consequence, it limits list of functions that can be implemented. However, ESP8266 modules can handle even simple multimedia data processing such as audio networking or graphics functions. In this investigation ESP8266 modules act as receivers of audio data. As they have been configured, their main function is to play audio data sent from server device.

Therefore, the main goal of the current paper is development of configuration methods for integration of IoT-devices into customer's wireless network and design of the communication scheme for further interactions between devices within the network where each device from the set is considered as an equal part.

2 Review of contemporary scientific sources

Despite the fact that at the first sight device set configuration seems a mere problem, we found out that few researches paid thorough attention to it. The known technical solutions for wireless devices are described by the patent documents [2, 3]. The authors of [3] offer the idea of usage of master or server device that identifies existing peripheral devices. The master is able to replicate configuration to other devices. However, the main presumption for this method is that all devices are already connected to the network. In opposite, we consider the case when device set is connected to the network by default. Moreover, the apparatus from [3] assumes that dedicated server node is always present while in the imposed challenge devices in the set are treated as clients to avoid usage of additional device. [2] concentrates on the communication scheme for wireless clients software update procedure. Nevertheless, it does not go further and does not formalize procedure of data exchange on the application level.

The results of [4] explain communication scheme for specific protocol and specific hardware type from single producer. Therefore, it cannot be considered as a universal method of configuration and communication for wireless IoT-device set.

The idea described in [5] supposes division of area on separate zones each controlled by separate master controller. Each of master controllers connects to main controller in the network that supposes at least three levels of hierarchy in the device network. This architecture is specifically oriented on HVAC application and can be used for Wireless Sensors Network construction but it is hard to apply it multimedia data transmission. Also the system supposes that system is deployed only once and

the configuration is stable. Lots of additional manipulations are necessary to configure system to work in new environment.

Authors of [6] claimed the development of system for home appliances that occupies technique based on ID-mechanism. The appliances are identified by dedicated IDs sent to controller of the network. This approach supposes that every device have pre-assigned ID. However, it can not work with appliances that have no ID.

In [7] method and apparatus for client identification were proposed. They are oriented on network devices that has no non-volatile memory to store configuration. Configuration is also ID-based. However, the device can receive ID during configuration stage. Nevertheless, it demands numerous configuration steps to be performed that is not applicable for lightweight device configuration.

As has been stated above, ESP8266 modules are the state-of-the-art modules in IoT field. They have been used for home automation, sensor readings, etc.

The update mechanism for ESP8266 has been proposed in [8]. Over the Air (OTA) update procedure is preferable in the end-user device because it does not require additional devices. The problem of upgrading software is quite important due to possibility to change firmware in case bugs have been detected or new features have been implemented.

Security is one of the main concerns in IoT. Security of ESP8266 has been considered in [9]. It is a problem of the special importance in conjunction with software updating. Therefore, this paper also touches the point of reliable software update with security concerns.

The paper [10] considers the problem of video data transmission from the multiple sources in the dynamic networking infrastructure. It uses multi-agent based approach to coordinate actions of devices in the set. However, the main shortcomings in context of problem imposed in this paper is an assumption about pre-configured network structure with fixed connection points.

Authors of the [11], the problem of IoT integration into existing solutions is investigated. The proposed solution is based on the open standards and offers integration of various modules. However, the paper [11] considers the situation in general by developing a universal approach with scalable components responsible for all actions within a network. Configuration of IoT device set does not require presence of such complex system.

The contribution of the paper is the developed model of configuration for IoT-device set, proposed communications schemes for device identification and client software upgrade and data transmission. In this paper, we proposed model of device set connection into consumer wireless local area network and method for integrating device set into the network. The devices in the set are able to communicate with other user devices in the network according to the developed approach for their identification. In the considered examples, the devices can process audio data and can be upgraded OTA.

3 Configuration approach and communications organization for IoT-device set

3.1 Integration of device set into existing network

In order to start consuming device in the network, the customer have to connect it into the network. In the proposed model, we distinguish two types of devices in set and we denote them as following:

- Server device (deploys service network and sends data to client nodes).
- Client device (connects to the server).

According to the proposed approach, the connection of the device set is performed with the following sequence of steps:

1. One of the devices (selected as server) deploys its own network, starts inner web-server and waits for the incoming connection. All client devices from the set connect to the server.
2. The customer connects to the network and can access web-page to write credentials to the device.
3. As customer has sent credentials information to the server device, it stores it into internal memory and is able to connect to the customer's network.
4. The server device sends data to the other devices from the set.

The steps described above are performed only under circumstance that the devices cannot connect to the previously selected network. Information about last connection is stored in the devices' memory. Device set reconfiguration is only necessary in case when we move it to other network. It is also worth to mention, that when devices are connected to the client's network, they all behave as clients (no server device among them).

3.2 Identifying server device

When the set is connected to the customer's network, it can interact with application software installed on other customer's devices (smartphones, laptops, PCs, etc.). However, it knows nothing about where the software is deployed. To identify device with the software we suppose usage of the next procedure in our model:

- Clients start listening for incoming message on dedicated port.
- When the software is running it acts like a server and sends multicast request to the specified port (UDP-server). In addition, we need to accept connection from the client so TCP-server is also started on the customer's device.
- As clients receive message, they can instantly identify IP-address to connect.

Thus, the software starts two servers to organize connection with clients. Both servers run on different ports to avoid collision.

3.3 Exchanging data with application

Constrained amount of resources (RAM and storage memory) should be taken into consideration to develop communication principles for the system. The end devices can not store and process large amount of data. Abuse of memory resources causes the device restart that can break all the communication states for the software and other clients and, furthermore, can cause crash of the whole system. This situation is not the first-class user experience and there is a demand for careful tracking of resource usage to avoid such situation.

As a consequence, of the previous statements, in case of implementation of multimedia functionality in the system, the server software is responsible for the processing of the multimedia data. In addition, the main dataflow for such features is supposed to go from the server to the clients.

Another point of interest concerned with software is the software architecture. Several devices (clients) connect to the device with server software. Therefore, it has to utilize multithreading concepts to control them simultaneously.

3.4 Updating client's software

In the proposed model, we suppose that end device functionality is defined by the scripts stored in the internal memory. Scripts are executed according to the external events (WiFi-connection, UDP data reception, etc.) and data received from server side. Software update in this case supposes updating script files or creation of the new ones. This operation could be performed without any wired connection. We denote this technique as “ad-hoc over-the-air software update” - ad-hoc OTA. Special caution should be taken to implement such feature because incorrect changes into the script could break the whole system. In the simplest case, the server initiates a new connection by sending command with specific token. It signals clients to create another connection to read scripts from the server. As a result, scripts stored on the file will be overwritten. To ensure that vital parts that is responsible for this action are always available, they should be placed into the separate script file with no manipulation allowed for it in network mode. With an additional requirement for security, the proposed mechanism should also utilize verification of the server identity. Thus, server has to confirm that it is not malicious device by providing necessary information. The procedure might be used only for software update process, however, it is also eligible for any interactions between server and clients.

3.5 Processing audio data

When the client devices are connected to the server, they can interact with server as well as with each other. As demonstration of this process, we selected transmission of audio data to the client devices. Despite the fact that audio processing at the end-user device is really straightforward, several peculiarities should be taken into account. First of them is memory constraint of the device. It should be noted, that several dozens of kilobytes of RAM are available to store data. Thus, it is necessary to free up the

memory as soon as data has been processed. The second one is the processing flow organization. At this point we assume that NodeMCU software is used in the module and class to play audio data via Pulse-Code Modulation (PCM) interface is responsible for this action.

The class supposes usage of callback mechanism to organize data processing. The data is received using network interface and is immediately appended into the list of data chunks to play. When the number of chunks is higher than lower threshold, data are sent into audio interface. Therefore, two processes run simultaneously:

- Reception of the data via the network interface and appending it into list of chunks.
- Transfer of data to PCM-interface.

Under these conditions, there is possible situation when the number of packets is close to the limit and addition and reception of one more new chunk into the list is not feasible (causes crush and system restarts). Thus, it is obligatory to control the upper threshold of the list. When the list length matches selected value it stops communication over the network interface and waits for the time when the list length is equal to lower threshold.

3.6 Implementation of the system

The system actions performed during configuration stage can be described in UML-notation. Visual Paradigm 12.1 software [12] has been used to create the appropriate diagrams for the developed system. In Fig. 1, sequence diagram for the process of connection to the server is depicted.

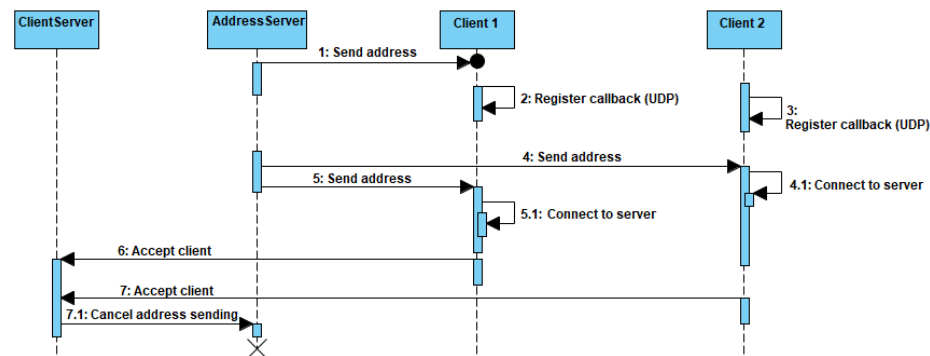


Fig. 1. Sequence diagram for the process of connection to the server

The software for the server side has been developed in JetBrains IDEA Community Edition 2017 development environment [13]. The main activity class contains communication-control nested classes that operate in ordinary client-server manner. Class diagram is shown in Fig. 2. Only the most important operations are illustrated on the diagram.

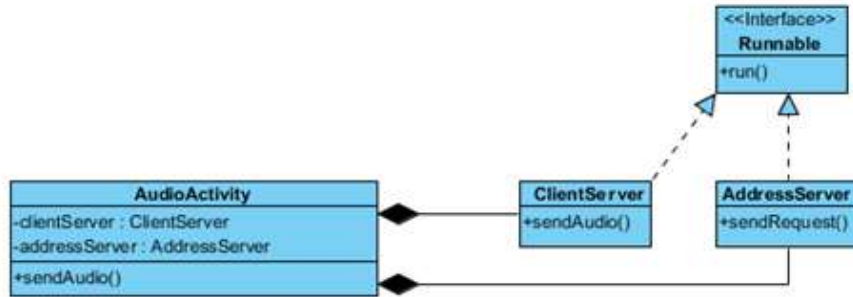


Fig. 2. Class diagram for the server software

To function properly, software components are to be deployed on server and client devices. While server device, e.g. smartphone, requires only application file to be installed, client devices store several script files. Deployment diagram for the system is shown in Fig. 3.

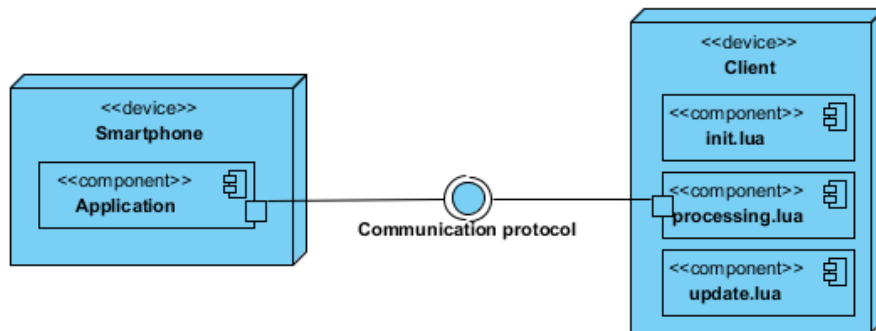


Fig. 3. Deployment diagram for the developed system

ESP8266 are Wi-Fi-modules that support different execution models. They could be pre-programmed using Arduino software. They also can run tiny versions of language interpreters such as Lua or MicroPython adopted for the requirements of the platform (small RAM available, constrained disk storage space, etc.). In the second case program code is distributed among the files (scripts) that are called depending on the user actions. The first way provides performance to the system while the second one is determined to be used for rapid prototyping and development of flexible solutions.

The system has been prototyped using NodeMCU v2 board that includes ESP8266 and simplifies power supply problem for the developer because it can be delivered directly via micro-USB interface. NodeMCU firmware for the client devices was obtained from cloud build service [14]. NodeMCU is a Lua-base firmware that includes interpreter of the Lua instructions. The instructions should be distributed among the .lua-files. NodeMCU programming model supposes intensive usage of

callbacks and this way it reduces resources consumption and simplifies program implementation.

In the implemented platform the security level matches security level of WiFi-protocol. It is enough to prototype and test the overall concept but it should be reviewed thoroughly to prevent possible attacks. For now, they can be performed almost on every stage of communication between device and server. Special attention should be paid to the upgrade stage due to its importance and possible negative consequences (up to fully non-functional device).

4 Testing of the system

The first stage is to perform setup for user network, so devices can connect to it. First, it is required to connect to the device set access point and provide information about user network. In the Fig 4, connection screen for the setup network is shown.

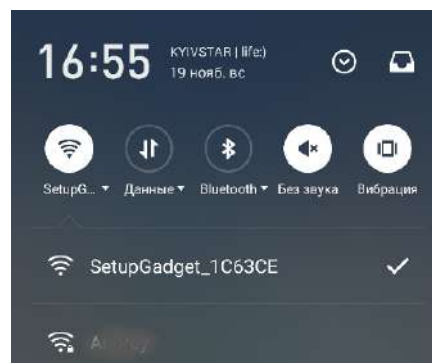


Fig. 4. Selection of the configuration network on smartphone device

The network to connect has name that follows the pattern “SetupGadget_xxxxxx” where the last part can vary according to the device. The setup network requires no credentials to enter.

Built-in functionality of NodeMCU library (enduser_setup) has been used to implement the setup stage. It allows to integrate setup functionality with single function call and takes responsibility for all the further internal processing. From under the hood, it provides a web-page deployed on the device and returns it in case of incoming request. User connects to the device using web-browser and can access configuration page shown in Fig. 5.

16:58

WiFi Login

Device: 1C63CE
Connect device to your Wi-Fi

Scan for Networks

Wi-Fi Name

Password

Save

Need a valid Network and Password

Fig. 5. Page for entering parameters of WiFi-network

As necessary information is filled into the fields, connection attempt can be performed. In case of successful configuration, the device updates page and shows the following message that is demonstrated on Fig. 6.

16:59

WiFi Login

Device: 192.168.0.107

Success!

Your device has successfully
connected to the Wi-Fi network.

You may now close this web page.

Wi-Fi successfully connected!

Fig. 6. Notification about successful connection to the user's network

As main device receives network parameters, it immediately sends them to the client devices in the set. When the devices are turned on next time, they will try to connect to the last configured network automatically. However, if they are supposed to work in different network (even if they have been connected previously), the user should repeat setup procedure.

The developed application has been tested on Meizu M2 smartphone in the network with TP-Link router and two ESP8266 devices where the first device performs the role of temporary server and the second one is a client-only device. The user interface of the application is shown on figure. It sends UDP-messages to help device set to identify network location of the running program. The program notifies user about connected devices showing up an information message. Then, it can transmit audio data to the clients through TCP-connection when user presses the button. The screenshot of the launched application with single connected device is shown in Fig. 7.

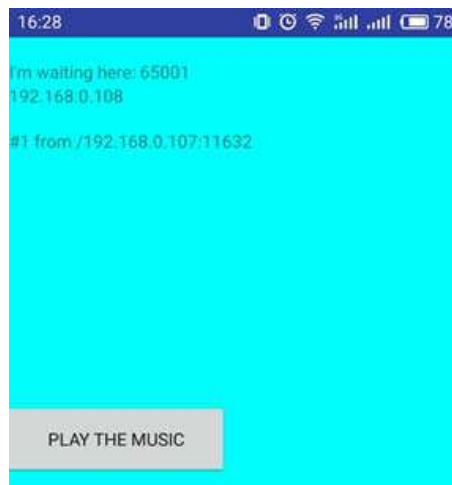


Fig. 7. User interface of the developed application for testing device connections

Fig. 7 demonstrates an application developed according to the previously mentioned class diagram (see Fig. 2). This is a multi-threading application that supports work with multiple clients. The click on the button initiates transmission of the prepared music data to all connected devices.

To summarize the results of the investigation we can state that proposed configuration procedure reduces time of configuration of IoT device set because they are configured simultaneously. Network mechanisms for identification of the server device in customer network have been proposed. The implemented software for both server and client side allowed us to prove correctness of the outlined developments.

5 Conclusions

In the presented paper, we have devised an approach for network configuration of the IoT device set. The approach is based on picking out server device within a set that serves as an access point at the initial stage. It receives information from the user about network to connect to and distributes it among other devices in the set. Thus, user does not need to configure each device individually that reduces time of configuration for full set of devices. Then, stage of communication with server device has been outlined. The server software is to await incoming connections and, in addition, has to send advertising broadcast message via UDP about server address in the network. With the reception of the message, devices can identify network node they need to connect to. The transmission of data for the software level relies on the TCP and stable connections between the server and clients. The case of audio data transmission has been considered in the work and implemented in the application for Android platform using Java programming language. The testing of proposed statements has been done using a pair of ESP8266 devices. All stages of connection procedure have been checked and shown the correctness of the apparatus presented in the paper.

References

1. ESP8266 Overview, <https://www.espressif.com/en/products/hardware/esp8266ex/overview>, last accessed 2019/03/27.
2. Bovell, M.C., Devlin, T.P., Goodner, A.S., Northway, T.N., Usery, P.H.: Wireless device configuration management. US Patent US8180860B2, 15 May 2012.
3. Lehotsky, D.A., Mannion, M., Nobutani, T.: System, apparatus and method for automated wireless device configuration. US Patent US7940744B2, 20 May 2011.
4. Lane, N.: Development and Integration of Honeywell's One-Wireless Network. Engineering Honours Thesis, Murdoch University (2018).
5. Estrin, D., Girod, L., Pottie, G., Srivastava, M.: Instrumenting the world with wireless sensor networks. In: International Conference on Acoustics, Speech, and Signal Processing, vol.4, IEEE (2001). doi: 10.1109/ICASSP.2001.940390
6. Cerpa, A., Elson, J., Estrin, D., Girod, L., Hamilton, M., Zhao, J.: Habitat monitoring: Application driver for wireless communications technology. In: ACM SIGCOMM Computer Communication, rev. 31(2), pp. 20-41, ACM (2001). doi: 10.1145/844193.844196
7. Ding, J., Cheung, S.Y., Tan, C.-W., Varaiya, P.: Signal processing of sensor node data for vehicle detection. In: The 7th International IEEE Conference on Intelligent Transportation Systems, Washington, USA, 3-6 October, 2004, IEEE (2004). doi: 10.1109/ITSC.2004.1398874
8. Gore, S., Kadam, S., Mallayanmath, S., Jadhav, S.: Review on Programming ESP8266 with Over the Air Programming Capability. International Journal of Engineering Science and Computing, 7(4), 6684-6686 (2017).
9. Naikoti, A., Kodali, R.K.: ECDH based security model for IoT using ESP8266. In: 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kumaracoil, India, 16-17 December 2016, IEEE (2016). doi: 10.1109/ICCICCT.2016.7988026

10. Burlachenko, I., Zhuravska, I., Musiyenko, M.: Devising a method for the active coordination of video cameras in optical navigation based on the multi-agent approach. *Eastern-European Journal of Enterprise Technologies*, 1, 9(85), 17-25 (2017). doi: 10.15587/1729-4061.2017.90863
11. Dave, B., Buda, A., Nurminen, A., Framling, K.: A framework for integrating BIM and IoT through open standards. *Automation in Construction*, 95, 35-45 (2018). doi: 10.1016/j.autcon.2018.07.022
12. Ideal Modelling & Diagramming Tool for Agile Team Collaboration, <https://www.visual-paradigm.com/>, last accessed 2019/03/25.
13. IntelliJ IDEA: The Java IDE for Professional Developers by JetBrains, <https://www.jetbrains.com/idea/>, last accessed 2019/03/27.
14. NodeMCU Custom Builds, <https://nodemcu-build.com/>, last accessed 2019/03/27.

Statistical analysis of data on the traffic intensity of Internet networks for the different periods of time

Olga Malyeyeva^{1[0000-0002-9336-41821]}, Yuriy Davydovsky^{2[0000-0003-2813-41692]},

Viktor Kosenko^{3[0000-0002-4905-85083]}.

^{1,2}National Aerospace University "Kharkiv Aviation Institute", Chkalov str., 17, Kharkiv, 61070, Ukraine

o.maleyeva@khai.edu, davidovskyi2350@gmail.com

³State Enterprise "Kharkiv Scientific-Research Institute of Mechanical Engineering Technology", Sumy str., 130a, Kharkiv, 61023, Ukraine
kosvv@ukr.net

Abstract. The article is devoted to the problems of the analysis of the regularity of changes in the intensity of telecommunications traffic for the different periods of time. A visual analysis of statistics on the intensity of traffic in different countries is conducted during the day, month, year, and five years. With the use of nonparametric criteria (chi-square, Mann-Whitney, Kruskal –Wallis), a pairwise and general comparison of data was conducted to identify their common trends. In addition, with the use of statistical methods of time series processing, data are analyzed for a month and five years. By constructing an autoregression model of the time series, according to observation data, during the month a periodicity of a series with a seasonal lag has been proved. Thus, the existence of a clear periodicity in daily data has been proved and the frequency of the year is shown. The obtained results can be used to model the network parameters taking into account the predicted traffic.

Keywords: traffic analysis, intensity, periodicity, statistical criteria, hypotheses

1 Introduction

The present time can be characterized by the rapid growth of traffic volumes of computer networks. Such a trend is due to the rapid development of such computer technologies as a cloud computing, a cloud storage, "streaming" services for movies, music or games. However, this growth is not a new one for the Internet providers, according to available statistics, the volume of traffic over the past 10 years has increased 8-10 times [1], and the number of people connected to the Internet as a percentage increased from 23% to 55% [2, 3]. All this leads the Internet providers to solve the obvious task of increasing and upgrading the existing computer networks

The primary task when modernizing a computer network is to analyze network traffic, namely its structure and volume [4]. That is why this article is devoted to the

analysis of computer network traffic, the detection and description of its laws for a certain time.

Based on the abovementioned, it is obvious that computer traffic has been repeatedly investigated and formalized using different methods and approaches [5, 6]. From the point of view of network modernization, the most worthwhile attention are the following:

- fractal (self-similar) analysis of network properties [7, 8];
- analysis of the source-destination streams of network traffic (Origin-Destination flows) [9]

It should be noted that these methods come to a similar conclusion that computer traffic has the clearly expressed fractal properties that are especially evident in the large time intervals such as weeks, months.

Therefore, the purpose of this work is to analyze the amount of traffic from different networks for different periods of time. It is necessary to find regularities in the time series and to prove the statistical significance of these hypotheses.

The persistent properties of web traffic allow you to build simulation models for forecasting network parameters [10, 11], to model server status and load on communication channels [12 - 15].

An important factor for maintaining the adequacy of the model is the input data for modeling. Special attention should be paid to the quantitative estimates of the volume of traffic, its marginal (maximum and minimum) and the average value for a certain period of time.

2 Materials of the research

An important component of today's computer networks is the collection of statistical data; however, if the specific indicators of Internet service providers in most cases are closed information, then the information collected on the Internet hubs is open for analysis and observation. The first objective of this study is to analyze international traffic on similarity and self-similarity. That is, you should show the similarity of network traffic for some time, as well as demonstrate the independence of the picture of the network dynamics from the specific location of the computer network, and its size.

To accomplish this, we turned to the hubs of such cities in Europe and America as: Frankfurt, Hamburg, Amsterdam, New York. Also, in order to show the relevance of the revealed properties in Ukraine, the relevant statistics collected in Ukraine will be provided.

We analyze the daily load on the network in different countries, that is, with a significantly different average traffic volume. It is necessary to analyze the hypothesis about a certain regularity of the change in the volume of traffic during the day and its independence from the mean value, which shows the independence of the properties of the operation of computer networks from their geographical location.

Below there is a statistics for network traffic in Germany, Ukraine, and America. All graphs are depicted in the form of a diagram with two axes, the X axis depicts a certain period of time (hours, days, weeks), along the Y axis - the amount of data transmitted at the specified time.

Figure 1 shows the web traffic of one of the largest European hubs in Frankfurt [16]. You can see the clearly expressed daily traffic regularity for two days, namely its growth in the so-called "peak hours" (around 8 P.M.), and the decline in traffic at night (around 4 A.M.).

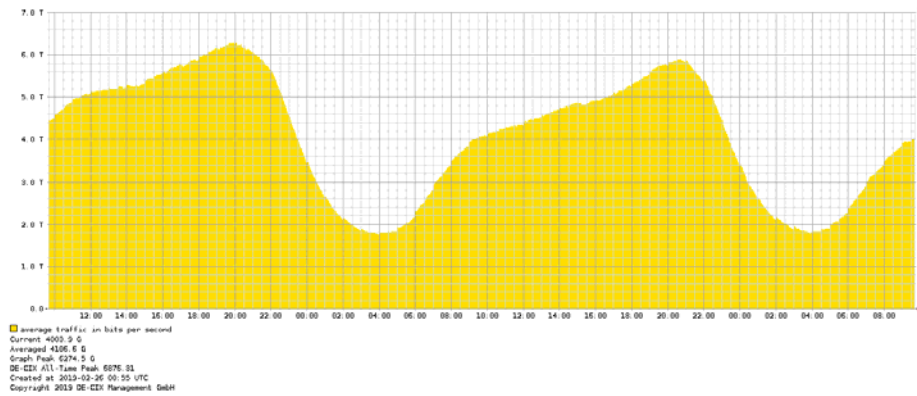


Fig. 1. Two-day web hub traffic in Frankfurt city

Figure 2 depicts web traffic for the same time period in New York [17]. Here you can observe the similar properties, as in the Frankfurt hub. However, additionally there is a peak load at about 12 o'clock.

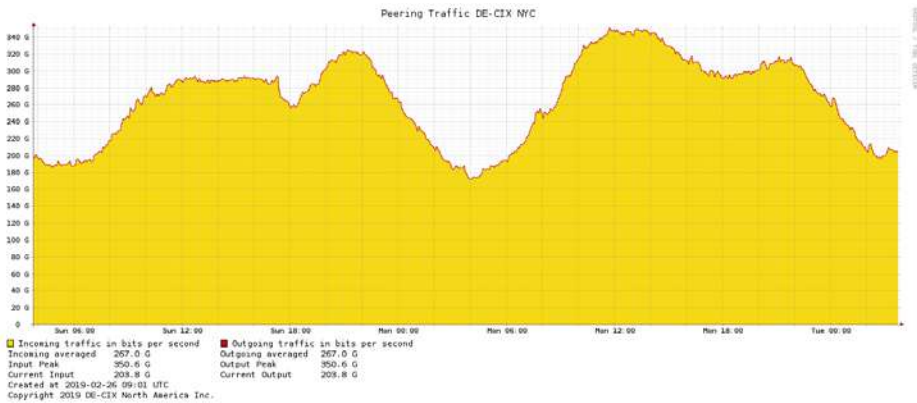


Fig. 2. Two-day web hub traffic in New York City

To illustrate a similar picture in Ukraine, Figure 3 shows the generalized web traffic collected from all hubs of Ukrainian cities [18]. Here the peak charge is centered between 8 P.M. and 9 P.M., and the minimum is between 4 A.M. and 5 A.M.

Although the resource UA-IX allows seeing the statistics only for the current time, the similarity of Web traffic in Ukraine with all the above-mentioned is evident.

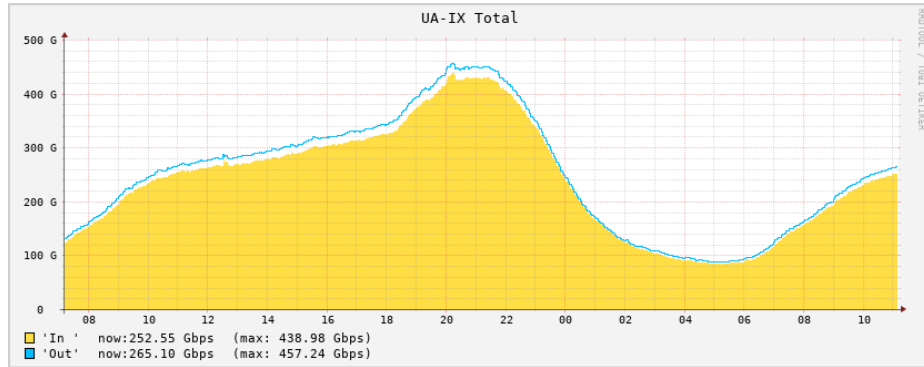


Fig. 3. Daily web traffic in Ukraine

We saw a visual similarity, but noticed some differences. It should be proved that the law is essential, and the differences are random. To revise this hypothesis, it is necessary to compare several graphs, that is, several types of distribution of the random variable of the intensity of traffic for a certain hour.

All Internet hubs that were selected vary not only in geographic distance, but also in traffic volumes. We can conditionally split these hubs into large, medium and small ones. Therefore, the hub in Frankfurt is the largest, with peak loads in several terabytes of information. The Hub of New York City is a mid-day sized and has a peak load of 250-300 gigabytes of data. Ukraine is depicted not by specific cities, but by the total volume of data in the country; therefore, hubs of Ukraine can be considered small, since the total volume of the country does not exceed the amount of data transmitted in New York City, that is, 400 gigabytes in peak hours.

Consider the task of analyzing the similarity of computer network traffic over a period of time. This study aims to demonstrate the similarity of computer traffic over a significant period of time, which in turn allows the use of simulation to predict the state of the computer network with a high probability.

To simulate the behavior of a computer network with the use of simulation models, it is important to use predictable traffic [19]. Therefore, it is necessary to prove that the trends indicated in the preliminary data (Fig. 1-3) are stored for a long time. Using the information gathered in these Internet hubs, it can be shown that relative volatility of network traffic remains unchanged despite the apparent and continuous increase in network traffic.

Figure 4 shows the network traffic of the city of Frankfurt during the month, it is possible to clearly indicate the storage of the above trends in this period of time.

However, this schedule cannot demonstrate the growth of traffic volumes due to the modernization of computer networks. To do this, you need to use annual reports, or even longer periods of time. For example, the statistics of the city of Frankfurt for the year (Fig. 5) and over 5 years are shown below (Fig. 6). An annual graph shows traffic growth of about 0.5 terabytes.

The best dynamics of growth can be seen in this figure. In just two years the amount of traffic has doubled from 2 terabytes to 4.

Despite the volatile dynamics of computer traffic and its continuous-out growth, one can distinguish user's party behavior.

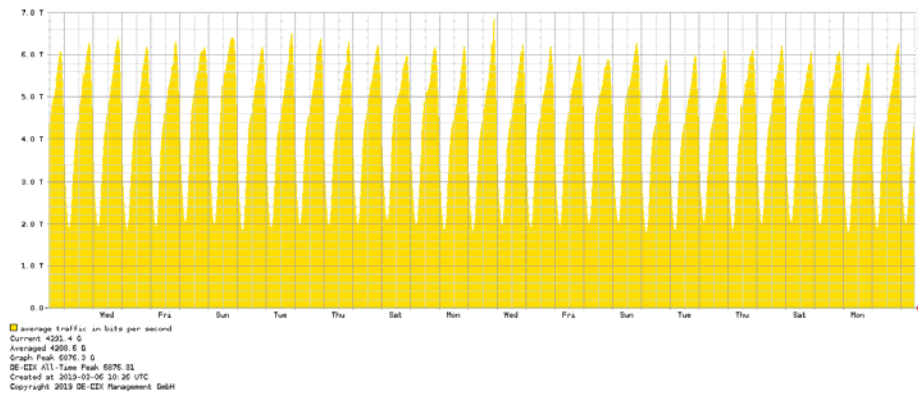


Fig. 4. Monthly web traffic in Frankfurt

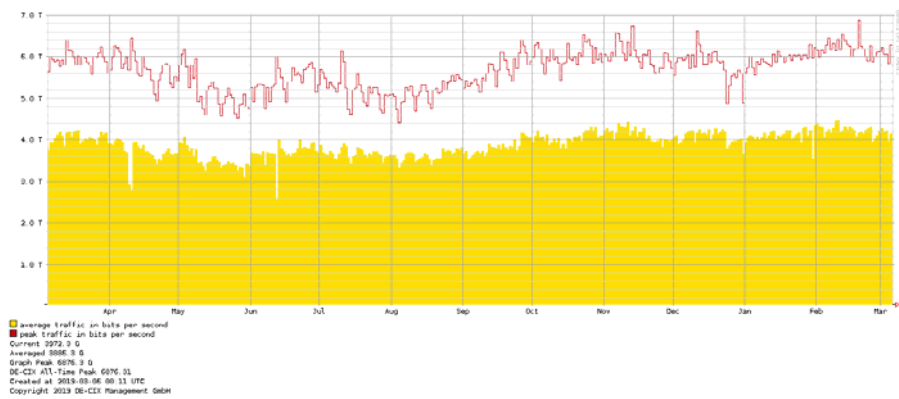


Fig. 5. Annual Web Traffic in Frankfurt

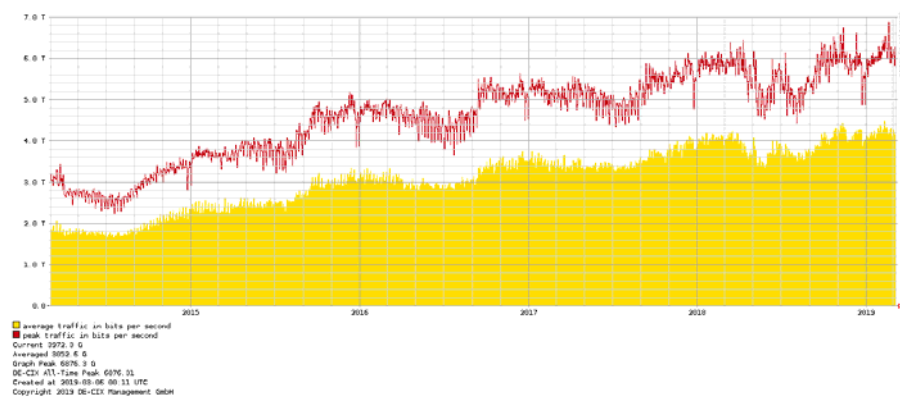


Fig. 6. Five-year web traffic in Frankfurt

3 Results and discussions

According to figures shown in Fig. 1-3, select a daily trend and reduce it to one linear graph (Figure 7). It can be seen that the graphs are similar to changes in volumes at certain parts of the time, but somewhat different.

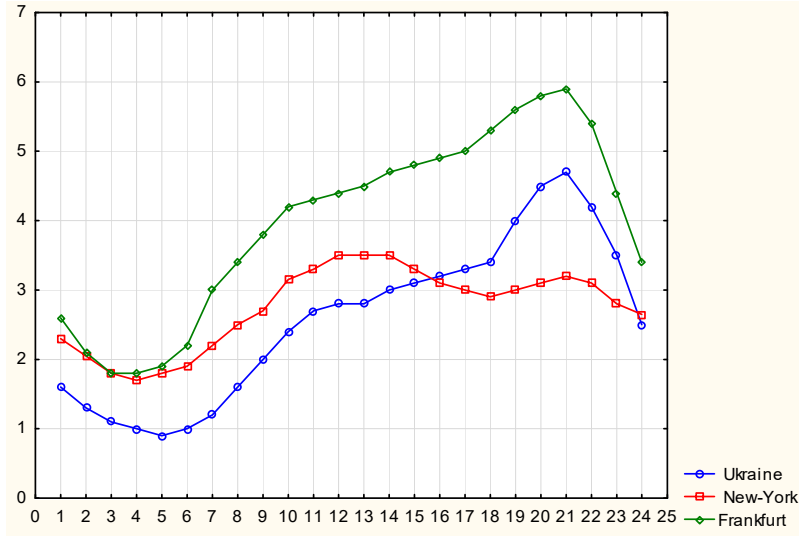


Fig. 7. Outgoing comparable graphs of intensity of traffic during the day

To make the analysis invariant relative to the absolute value of traffic, we normalize the output timelines, that is, let us give the average value of traffic to zero value (Fig. 8). It is seen that the second graph (data in New York) is slightly different. To compare the intensity distribution over the course of the day, use the non-parametric chi-square [21, 22] criterion, making three pairs of comparisons:

(Ukraine-Frankfurt): $\chi^2 = 2,019$, $df = 23$, $p = 1,000$

(Ukraine-New York): $\chi^2 = 17,300$, $df = 23$, $p = 0.794$

(New York-Frankfurt): $\chi^2 = 11.670$, $df = 23$, $p = 0.975$.

In the first comparison, the data coincide very well. The hypothesis of the randomness of convergence is confirmed (the level of significance approximates to one). In the second case, the discrepancy is significant ($p = 0,794$), the error is about 21%. But the third comparison also gave a meaningful result ($p = 0,975$).

Thus, under certain assumptions, we can conclude that the trend of changing the intensity of traffic for a long time is statistically significant. To predict the maximum load and simulate loading, you can use averaged traffic, but average values are calculated with the certain weights: the New York data graph should be considered with less weight, according to the value of the degree of confidence received. Averaged traffic (with output) is shown in Fig. 8.

Since the chi-square criterion does not provide a confident answer to the incidence of discrepancies in relation to all three charts, we will apply another verification procedure.

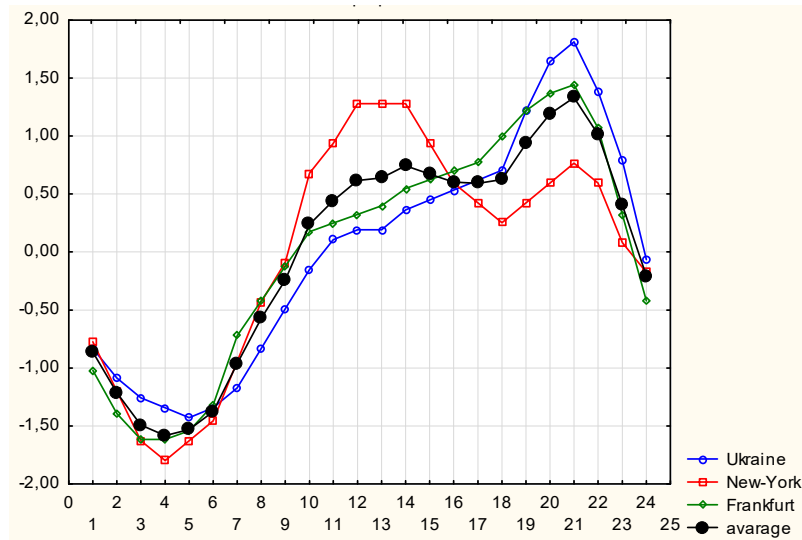


Fig. 8. Normalized linear graphs of traffic intensity and averaged overnight traffic

For a pairwise comparison of the three graphs, apply the Mann-on-Whitney criterion for independent samples [21]. The comparison results are given in Table 1, 2. The indicated criteria are significant.

Table 1. Results of the Mann-Whitney's Criteria for samples (Ukraine-New York)

	Sum rank gr 1	Sum rank gr 2	U	Z	p- value	Z - adjust.	p- value	N gr.1	N gr.2	2-sided exact p
Traf 1-2	584	592	284	-0,072	0,942	-0,072	0,942	24	24	0,943

Table 2. Results of the Mann-Whitney's Criteria for samples (Ukraine - Frankfurt)

	Sum rank gr 1	Sum rank gr 2	U	Z	p- value	Z - adjust.	p- value	N gr.1	N gr.2	2-sid. exact p
Traf 1-3	585	591	285	-0,05	0,958	-0,051	0,958	24	24	0,959

The results of the comparison of the samples (New York-Frankfurt) fully coincide with the results of Table 2, because the selected criterion is ranked, and the calculated ranges in these groups coincide, despite different values of the output data.

As can be seen from the results of the estimates of the significance of the criterion, all comparisons are based on the randomness of the difference in groups: confidence level $p = 0,942$, and 2-sided exact $p = 0,943$; in the second comparison $p = 0.958$, and 2-sided exact $p = 0.959$. These values confirm the hypothesis of the significance of the difference. So, in the first comparison, the error is about 5.6%, in the other two about 4%.

Now let's conduct simultaneous comparison of three samples with the help of the criterion of Kruskal-Wallis (Table 3).

Table 3. Results of the Kruskal-Wallis criterion for three samples

H (2, N= 72) = 0,008 p = 0,996			
Cod	N	Sum Ranks	Mean Rank
1	24	869,0000	36,20833
2	24	877,0000	36,54167
3	24	882,0000	36,75000

The results of this test fully confirm the hypothesis of the incident of discrepancies, level of trust is close to one ($p = 0,996$).

Based on the entire above, one can conclude that the structure of web traffic is similar throughout the world, regardless of the geographical location of the computer network and the volume of web traffic. Thus, it can be used in constructing simulation models of computer networks, as inputs to achieve model adequacy and forecast accuracy.

The next step in the study is to analyze the change in traffic over the course of a month. The graph of its change in accordance with Fig. 4 is shown in Fig. 9.

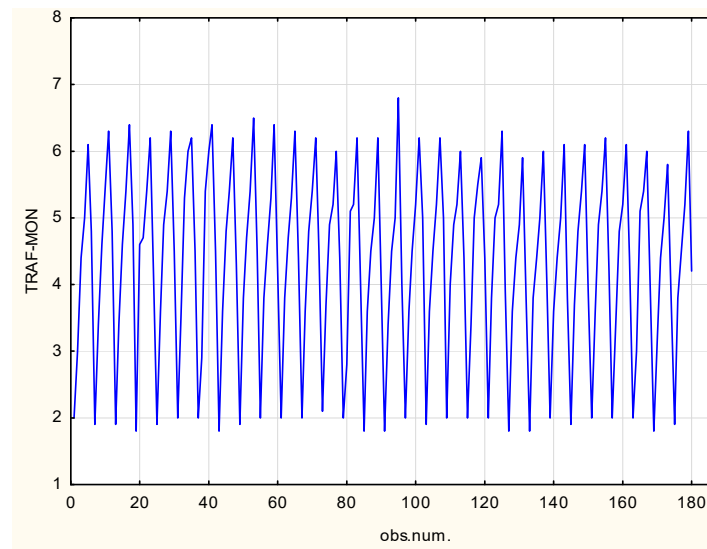


Fig. 9. Change the intensity of traffic for a month

Output data form a stationary time series (under the condition of converting the normalization of values to normalized values) [23]. Frequency of changes is not in doubt. Let us prove this by constructing the model of auto regression taking into account the seasonal component (Table 4) [24, 25]. A model with three parameters was selected: q - autoregressive parameter, $Ps(1)$ and $Ps(2)$ - seasonal parameters of auto regression. It is seen that the estimates of the parameters are very precise (the significance levels p are close to zero).

Table 4. Estimation of parameters of autoregressive data model for five years

Model (1,0,0)(2,0,0) Seasonal lag: 24 MS Residual = 0,3041				
Param.		Asimpt. - Std.error	Asimpt. - t(116)	p
p(1)	0,7932	0,0473	16,7417	0,0000
Ps(1)	0,7400	0,0753	9,8263	0,0000
Ps(2)	0,2361	0,0756	3,1202	0,0021

Thus, the resulting model accurately describes the time series of observations. The periodicity of a series with a seasonal lag of 24 hours is confirmed.

The randomness of the model errors is confirmed by the residue graph, which is close to the normal distribution, as illustrated by the normal probabilistic graph (Fig. 10). It is seen that the remnants are not substantially deviating from the straight line, only some deviations in the upper part of the graph are observed.

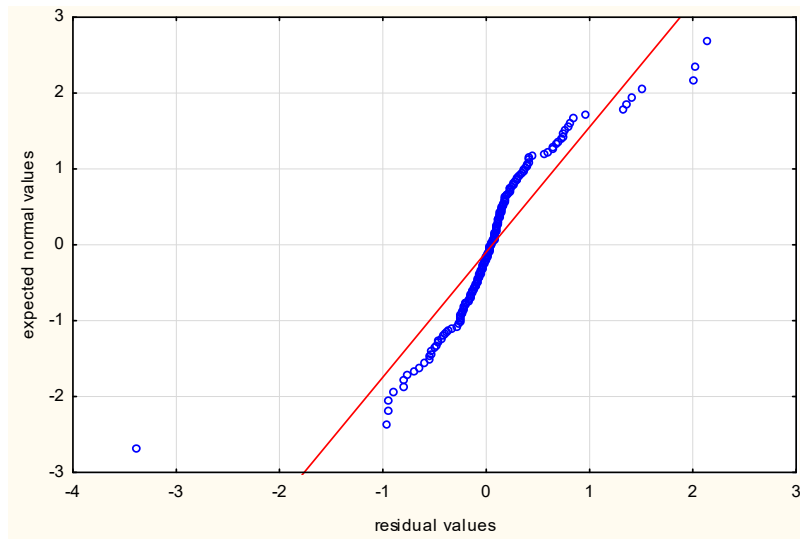


Fig. 10. Normal probabilistic distribution graph of residues

The frequency of the analyzed time series is also confirmed by the graph of the auto regression function.

The third stage of the study is the analysis of traffic changes over the year based on the data presented in Fig. 5 (data on the maximum load was applied). A linear graph of data, forming a time series, is presented in Fig. 11. It should be noted that the data are presented from March to February of the year following the year.

It is seen that the range is not stationary. For its analysis, a number of significant transformations should be carried out [26]. When applying the exponential smoothing of a range you can see the tendency of its change, which should pay attention to the months of maximum load. One can assume that there is an annual periodicity of traffic changes. It can be analyzed only on the basis of data for several years (Fig. 6). The time range of values of maximum loads in the network is shown in Fig. 12.

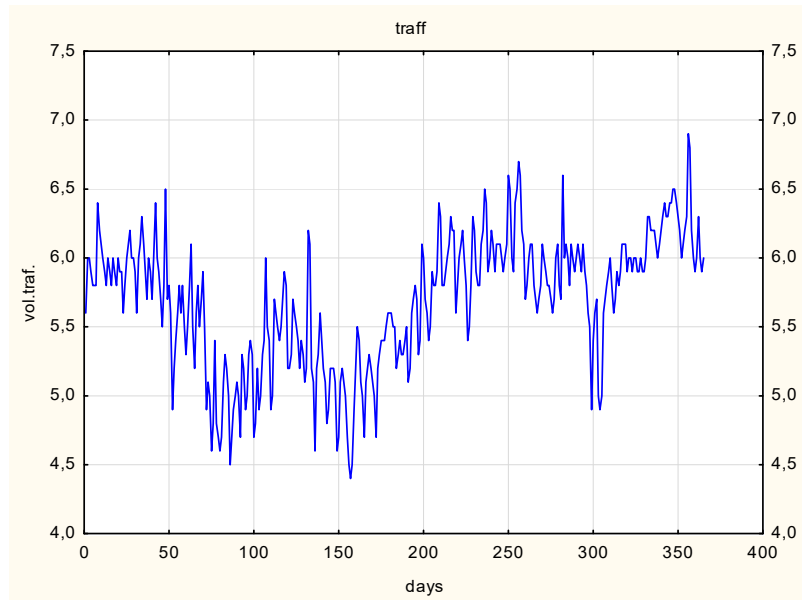


Fig. 11. Time range of traffic intensity throughout the year

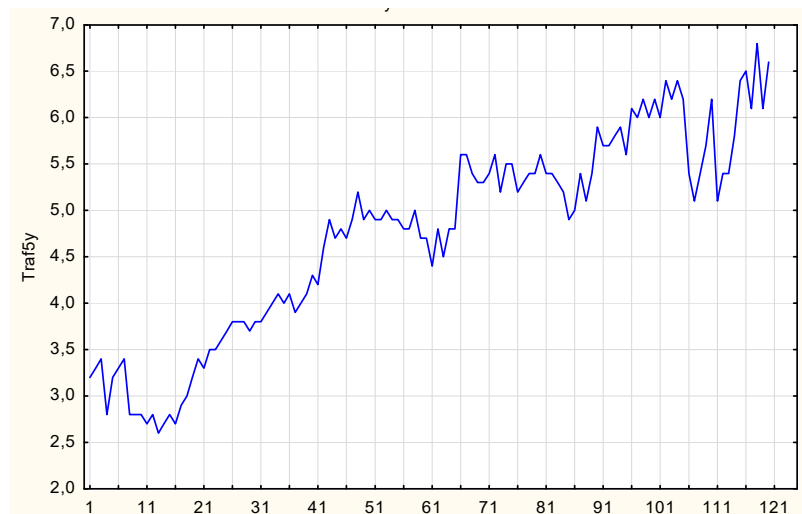


Fig. 12. Time range of maximum traffic loads for five years

The frequency of this range during the year visually expressed is not clear. In order to check its periodicity, using the autocorrelation model, it is necessary to increase the number of stationary ones. To do this, the shift was applied to one lag (Figure 13).

A model of auto regression with three parameters was selected (Table 5), here additionally Q_s is the seasonal parameter of the sliding average. It is seen that the parameters are estimated fairly accurately (the levels of significance are close to zero),

the function of autocorrelation (Figure 14) also confirms the correctness of the model. The figure shows that for 15 logs the value of autocorrelation is random: both positive and negative values do not exceed the line of the permissible limit. One can conclude that the seasonal annual frequency is significant. That is, it is a pattern that the lowest traffic is observed in June and August, and the highest in February (this is evident from Figure 12).

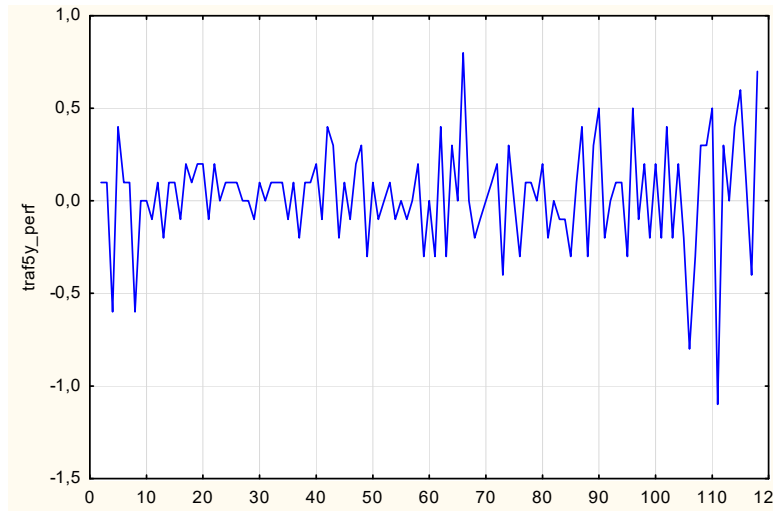


Fig. 13. Converted time range of maximum traffic loads for five years

Table 5. Estimation of parameters of autoregressive data model for five years

Transform.: D (1) Model (0,1,1) (1,0,1) Seasonal lag: 24 MS Residual = 0,0662				
Param.		Asimpt. - Std.error	Asimpt. - t(116)	p
q(1)	0,4865	0,0906	5,366837E+00	0,0000
Ps(1)	0,9998	0,0000	2,734817E+19	0,0000
Qs(1)	0,6742	0,1017	6,623994E+00	0,0000

According to these same data one can investigate the trend of increasing traffic intensity (Fig. 15). The trend was obtained by smoothing with the help of the weighted least squares method.

The resulting trend can be used to predict the maximum load on the network for future years [27, 28]. At the same time, it should be made possible not to take into account the sharp changes in the trend associated with qualitative changes (for example, due to the explosion of new technologies) in the field of telecommunications in the future.

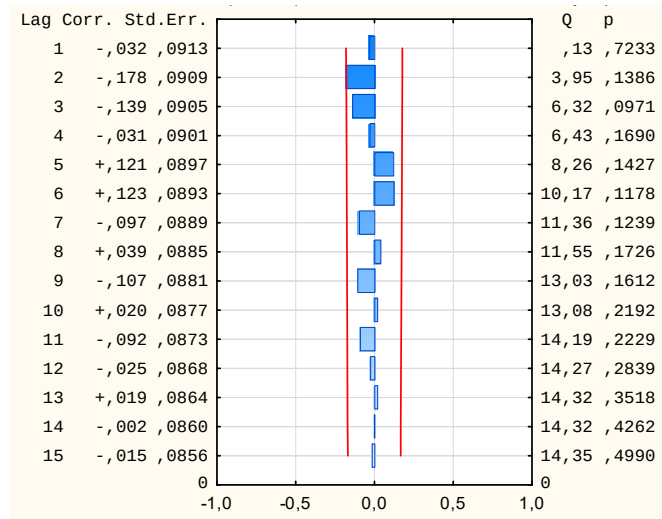


Fig. 14. Autocorrelation function for auto regression model

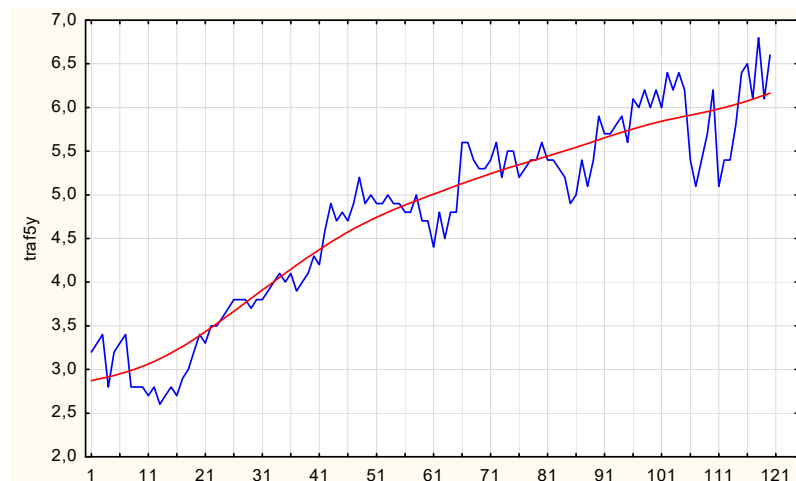


Fig. 15. Trend increase traffic for five years

4 Conclusion

The task was to find and analyze the patterns in the on-load of computer networks by studying international traffic at different time intervals (day, month, year and five years). The following results are obtained:

1) Analyzed the daily load of computer networks. Comparison of international traffic on similarity and self-similarity with the help of three statistical criteria is carried out. It is concluded that the trend of changing the intensity of traffic over the course of the day is statistically significant. The daily periodicity is proved for a cer-

tain time and the hypothesis of the independence of the network dynamics from the specific location of the computer network and its size is confirmed. Consequently, we can conclude that the structure of web traffic is similar throughout the world, regardless of the geographical location of the computer network and the volume of traffic. Thus, you can use averaged traffic as input data when simulating the loading of computer networks.

2) Traffic data is analyzed within a month. Taking into account the considerable number of observations, methods of analysis of time series were used. By constructing an auto-regression model of the time series, according to observations of observations during the month, the frequency of a series with a seasonal lag of 24 hours has been proved.

3) Traffic data for the year is analyzed. By smoothing random data, a graph of intensity changes was received, from which no periodicity was visually revealed.

4) The observation of intensity of traffic for five years is analyzed. The seasonal annual frequency of traffic intensity was proved on the basis of an autoregressive model with three parameters. In addition, on the basis of five-year observations, a traffic growth trend has been obtained, which can be used to predict the maximum load in the future for years to come.

Thus, the researches are characterized by a scientific novelty, which consists in the fact that the hypothesis of self-similarity, periodicity and the presence of a trend in the intensity of traffic of modern computer networks has been proved for the first time, which allows scientifically reasonably predict maximum network load at its modernization.

Further research is to be carried out in the direction of forecasting of the maximal load on the network and simulation of network parameters with consideration of predicted traffic.

References

1. Reva, A., Davydovskiy, Y.: Method of the network topology transformation to quasihomogeneous structure. *Radioelectronic and computer systems*, 2, 43-51 (2018), doi: 10.32620/reks.2018.2.05
2. Internet World Stats: Usage and Population Statistics. <https://www.internetworldstats.com/stats.htm>
3. Zhuang, Y., Cappos, J., Rappaport, T., McGeer, R.: Future Internet Bandwidth Trends: An Investigation on Current and Future Disruptive Technologies. Secure Systems Lab, Dept. Comput. Sci. Eng., Polytech. Inst. New York Univ., New York, NY, USA, Tech. Rep. TR-CSE-2013-0411/01/2013 (2013)
4. Lavrut, O. O.: The quality control of information flows in the telecommunication system of critical use. *Systems of Arms and Military Equipment*, 4 (40), 89-93 (2014)
5. Stepanov, S. N.: Basics of teletraffic multiservice networks. *Eco-Trendz* (2010)
6. Paramonov, A. I.: Models of traffic flows for M2M networks. *Telecommunications and radio engineering*, 4, 11-16 (2014)
7. Leland, W., Taqqu, M., Willinger, W., Wilson, D.: On the self-similar nature of Ethernet traffic. *IEEE/ACM Transactions on Networking*, 1, 1-15 (1994)
8. Peters, E. E.: Fractal analysis of financial markets. *Application of Chaos Theory in Investment and Economics*. Internet-trading, Moscow (2010)

9. Lakhina, A., Papagiannaki, K., Crovella, M., Diot, C.: Structural Analysis of Network Traffic Flows. *ACM SIGMETRICS Performance Evaluation Review*, 32(1), 61-72 (2004)
10. Cisco: Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper. Cisco VNI, (2018)
11. Bunimovich, L., Smith, D., Webb, B.: Specialization Models of Network Growth. *Journal of Complex Networks*, (2018), doi: 10.1093/comnet/cny024
12. Mestres, A., Alarcon, E., Cabellos-Aparicio, A.: Understanding the Modeling of Computer Network Delays using Neural Networks. *Big-DAMA@SIGCOMM* (2018), doi: 10.1145/3229607.3229613
13. Kosenko, V.: Principles and structure of the methodology of risk-adaptive management of parameters of information and telecommunication networks of critical application systems. *Innovative technologies and scientific solutions for industries*, 1 (1), 45–51 (2017) doi: <https://doi.org/10.30837/2522-9818.2017.1.046>
14. Kosenko, V., Persyanova, E., Belotskyy, O., Malyeyeva, O.: Methods of managing traffic distribution in information and communication networks of critical infrastructure systems. *Innovative technologies and scientific solutions for industries*, (2 (2), 48-55 (2017) doi: 10.30837/2522-9818.2017.2.048
15. Kuchuk, G. A., Kirillov, I. G., Pashnev, A. A.: Modeling of traffic of multiservice distributed telecommunication network. *Processing Systems Information*, 9 (58), 50–59 (2006)
16. DE-CIX Frankfurt statistics. <https://www.decix.net/en/locations/germany/frankfurt/statistics>
17. DE-CIX New York statistics. <https://www.de-cix.net/en/locations/united-states/new-york/statistics>
18. UI-EX. Ukrainian Internet Exchange. Statistics. <https://www.ix.net.ua/pro-kompaniyu/statystyka>
19. Poshtarenko V. M., Andreev A. Yu., Amal M.: Ensuring the quality of service at critical sites of the multiservice network. *Bulletin of the NTU "KhPI"*, 60, 94–100 (2013)
20. Sultanov, A. Kh., Sultanov, R. R.: Method of assessing the quality of service indicators of hierarchical multiservice networks. *Bulletin USATU*, Vol. 12, 1 (30), 175–181 (2009).
21. Borovikov, V. P.: *STATISTICA. The art of analyzing data on a computer*. Piter, SPg. (2001)
22. Blake, A. P., Kapetanios, G.: Pure Significance Tests of the Unit Root Hypothesis Against Nonlinear Alternatives. *Journal of Time Series Analysis*. Vol. 24, 3, 253–267 (2003)
23. Kendal, M.: *Time Series. Finance and Statistics*, Moscow (1981)
24. Box, J., Jenkins, G.: *Time series analysis, forecast and management*. Mir, Moscow (1974)
25. Borovikov, V. P.: *Prediction in the STATISTICA system in the Windows environment. Fundamentals of the theory and intensive practice on the computer*. Finance and statistics, Moscow (2000)
26. Orlov, Yu. N., Osminin, K. P.: *Non-stationary time series. Forecasting methods with examples of analysis of financial and commodity markets*. Librokom, Moscow (2011)
27. Lukashin, Yu. P.: *Adaptive methods for short-term forecasting of time series*. Finance and Statistics, Moscow (2003)
28. Hanke, D. E., Wichern, D. W., Reitsch, A. J.: *Business Forecasting*. Williams, Moscow (2003)

Project-oriented management of adaptive teams' formation resources in multi-project environment

Nataliia Dotsenko¹[0000-0003-3570-5900], Dmytro Chumachenko¹[0000-0003-2623-3294] and Igor Chumachenko²[0000-0003-2312-2011]

¹ National Aerospace University "Kharkiv Aviation Institute", Chkalov str., 17, Kharkiv, Ukraine
nvdotsenko@gmail.com
dichumachenko@gmail.com

² O. M. Beketov National University of Urban Economy in Kharkiv, Marshal Bazhanov str., 17, Kharkiv, Ukraine
ivchumachenko@gmail.com

Abstract. The subject of the research is the processes of project-oriented resource management of the formation of adaptive teams in a multi-project environment. The aim of the work is to create models and methods of project-oriented resource management of the formation of adaptive commands in a multi-project environment. A model of project-oriented resource management in a multi-project environment has been developed. A set of methods for determining human resource management strategies in a multi-project environment is proposed. It is proposed to use a stratified representation of resource management processes in a multi-project environment, which will allow an analysis of human resource management processes. A method for stakeholder-based human resource management of projects in a multi-project environment has been developed. A model of the process of constructing a functional-backup team has been developed.

Keywords. project management; multi-project environment; project portfolio; project team; project-oriented management; stakeholders; reservation.

1 Statement of the Problem (Introduction)

Resource management in a multi-project environment is a complex multi-criteria task that requires managers to apply modern management approaches. The growth in the number of projects implemented by the company, fierce competition for resources, the need to ensure the flexibility of project teams, the increasing interconnection of projects, additional constraints imposed by stakeholders, lead to the need to develop effective methods for managing resources in a multi-project environment.

2 Analysis of recent research and publications

By a multi-project environment (MPE), we mean a set of projects that are not necessarily functionally related, but use the same resources from the company's general pool of resources. Multi-project management is a short-term tactical management of many projects that share the same resources [1].

In papers [1-3] an analysis of approaches to human resource management in a multi-project environment is presented. The criteria for the classification of approaches were the degree of centralization, the degree of focus on human resources, optimization algorithms. The main approaches are:

- Heuristic OR;
- Buffer management approaches like critical chain;
- Agile by using scrum of scrums;
- Multi-agent auction based;
- Resource sharing policies with dedicated and core teams and shared resource pools;
- Systems management.

As a result of the analysis, it was determined that, in most cases, approaches to human resource management in a multiproject environment are based on automated centralized decision making. The use of Agile Scrum, Scrum-of-Scrums flexible methodologies will ensure decentralization and self-regulation, which is especially important when managing projects in IT companies.

The use of flexible project management methodologies has led to the need to solve a number of tasks caused by the self-organization of teams [4-5]:

- changing resource requirements;
- engaging senior management in project-level sponsorship;
- achieving cross-functionality and effective team-level evaluations;
- the statement of independence and independent assignment of resources at the individual level;
- lack of eligibility criteria and dependencies at the task level.

Resource management in a multi-project environment includes planning, distributing, balancing and coordinating resources in individual projects, as well as resolving resource conflicts between different projects in order to achieve an optimal allocation of resources between projects in a multi-project environment.

The use of Critical Chain Method (CCM) for multi-project resource management is based on the assumption that resource constraints between different projects mainly occur in the critical paths of projects that form a multi-project environment. The redistribution of resources in this approach occurs in accordance with the priority of each project. To prevent delays in performing work that lies on a critical path, it is proposed to use a resource buffer [6].

Within the framework of promoting sustainable use of resources in business organizations (Green human resource management) and ensuring environmental sus-

tainability, additional requirements are put forward for project teams: high level of technical and managerial skills among collaborators, environmental values, environmental interests of stakeholders. Accounting for these restrictions in the formation of teams (Green selection) will provide further induction and improve the efficiency of the organization [7-8].

When forming Agile teams, it is necessary to take into account factors [5]:

- competences (availability and level);
- team climate;
- team variety;
- team innovation;
- characteristics of team members: technical and personal competencies of team members;
- presence of corporate culture;
- values;
- group leader behavior;
- top management support: project management commitment;
- emotional intelligence;
- communication skills.

Additional requirements for the project team can be formed when identifying risks [9]. Risk identification will allow to take into account the need to ensure the reliability of the project team's operation by introducing redundancy into teams and forming functionally-reserved teams.

According to the Standard for Portfolio Management, PMI, human resource management is responsible for maintaining the right balance of staff and competencies based on strategic plan drivers and market pressures [10].

The use of competence-based approach in the formation of the project team [11-15] allows to provide the required level of competence in the teams formed.

The adaptability of the team leads to the ability to respond to changes in the list of currently available business requirements and technical requirements for the project and its product [11].

Since the stakeholder-oriented approach in determining the success of a project (maximally satisfying the expectations of stakeholders) is typical for managing a portfolio of projects, human resource management in a multi-project environment implies finding a balance of interests of the stakeholders in human resource management processes and a compromise that ensures the achievement of the company's strategic goals [16].

In order to automate the decision-making process when choosing a project team, a number of authors propose to use the decision support system for choosing a project team [17].

3 Isolation of previously unresolved parts of a common problem. Objective

The application of the considered methods allows to form project teams, but does not take into account the need to apply an integrated approach in the management of human resources in a multi-project environment. Actual scientific problem is the development of theoretical foundations and tools of human resource management software in the formation, development and management teams of the projects and programs.

The purpose of this article is to create models and methods of project-oriented resource management of the formation of adaptive teams in a multi-project environment.

The article solves the following tasks:

- creation of a model of project-oriented resource management in a multi-project environment;
- development of methods for determining human resource management strategies in a multi-project environment;
- development of a method for stakeholder-based human resource management of projects in a multi-project environment;
- development of a model for the construction of a functional-redundant adaptive team.

In the study, system analysis was used to formalize the processes of formation and functioning of project teams, the apparatus of optimization theory for formulating and solving the tasks of forming teams of projects in a multi-project environment under given constraints, Boolean algebra and set theory to solve problems of forming teams and redistributing resources.

4 The main research material

The task of providing resources to projects in a multi-project environment can be represented as follows:

- defining a strategy for managing human resources in a multi-project environment;
- formation of requirements for resources;
- formation of adaptive project teams with given resource requirements and constraints;
- monitoring of resource requirements;
- redistribution of resources;
- management of critical competencies;
- analysis of the involvement of resources in the project portfolio.

4.1 Model of project-oriented resource management in a multi-project environment

A model of project-oriented resource management in a multi-project environment (Fig. 1) has been developed, reflecting the relationship of the proposed methodology with project resource management processes (PMI PMBoK), program resource management (PMI The Standard for Program Management), project portfolio resource management (PMI The Standard For Portfolio management) [10, 18-19].

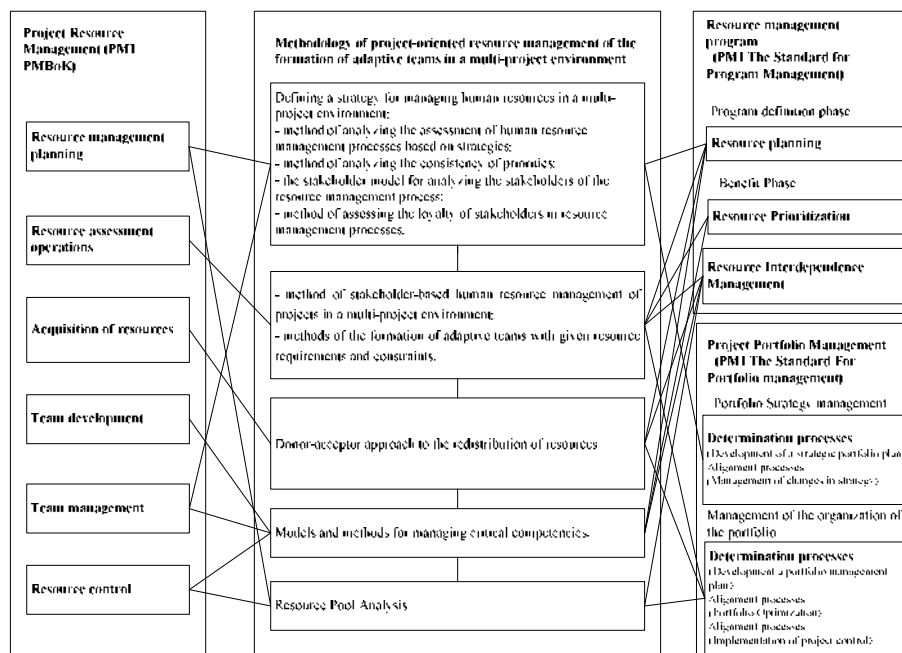


Fig. 1. Model of project-oriented resource management in a multi-project environment

4.2 Defining a strategy for managing human resources in a multi-project environment

Strategic human resource management in a project-oriented organization is aimed at achieving a strategic alignment of the organization's intentions and plans concerning the strategy, policy and practice of human resource management of projects (hiring, searching and selecting employees, building project teams, managing resources in a project portfolio, training and development, motivation) [20].

Depending on the company implementing the portfolio of projects, leadership style, human resource management strategies can be distinguished [21]:

- resource oriented management;
- high performance management;
- management focused on a high level of commitment;

- participatory management.

Since human resource management in a multi-project environment is associated with the influence of certain project stakeholders, the project portfolio, the project organization as a whole, the definition of a coherent human resource management strategy is an important task affecting the efficiency of project implementation and ensuring the viability of the company.

When determining the strategy of human resource management in a multi-project environment, it is proposed to apply a set of methods:

- a method for analyzing the assessment of human resource management processes in the basis of strategies, a method for adapting the strategy of a project-oriented organization to exogenous changes [22];
- a method for analyzing the consistency of priorities [23];
- a method for assessing the loyalty of stakeholders in resource management processes [24].

The use of the stakeholder model for analyzing the stakeholders of the human resource management process [25] allows identifying stakeholders, determining their interest in managing human resources, and evaluating their loyalty.

Each project process that is executed in a multi-project environment is characterized by:

- belonging to the project ($Pr = \{Pr_1, \dots, Pr_n\}$, n is number of projects implemented in a multi-project environment);
- indicator value of resource management processes in a project $LS_{i,j}$ (showing the level of the j -th process in the i -th project);
- criticality $K_{i,j}$ (the number of critical factors influenced by the j th process in the i -th project).

According to the PMI PMBoK standard, multiple resource management processes for the i -th project can be represented as follows:

$$P_i = \{P_{i,1}, P_{i,2}, P_{i,3}, P_{i,4}, P_{i,5}, P_{i,6}\}, \quad (1)$$

where $P_{i,1}$ – resource management planning;

$P_{i,2}$ – resource assessment operations;

$P_{i,3}$ – resource acquisition;

$P_{i,4}$ – team development;

$P_{i,5}$ – team management;

$P_{i,6}$ – resource control.

When constructing a stratified representation of resource management processes in a multi-project environment, it is proposed to use a three-dimensional coordinate system (system cube) that reflects the achievement of a specific resource management strategy in projects running in a multi-project environment (Fig. 2).

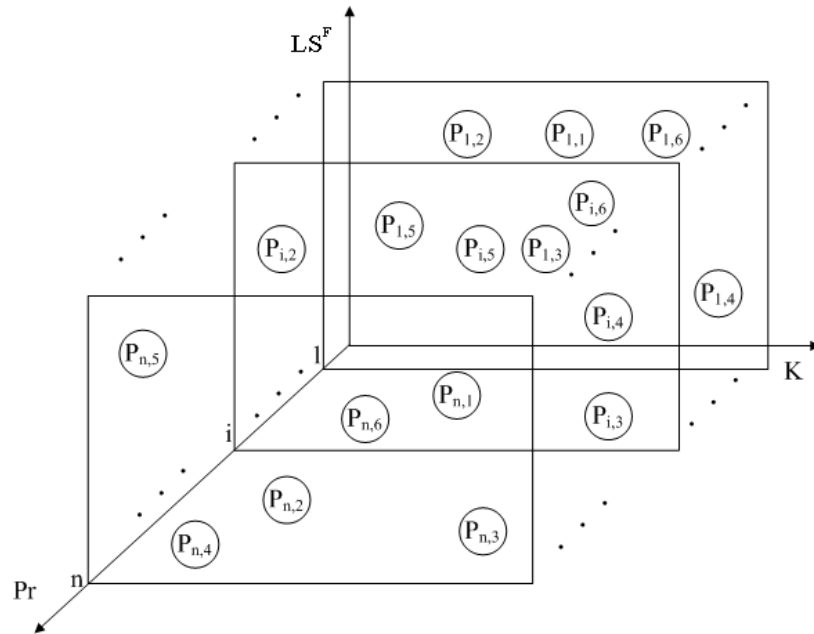


Fig. 2. A stratified representation of resource management processes in a multi-project environment

An analysis of the projections of the stratified representation of human resource management processes in a multi-project environment allows us to determine the direction of development of the processes, as well as identify the processes that need improvement (Table 1).

Table 1. The results of the analysis of the projections of human resource management processes in a multi-project environment

Projection	Description
Portfolio project process $Pr_{i,j}$ – level of resource management processes in a project $LS_{i,j}$	Allows you to analyze the level of indicators of the processes of human resources management projects in the portfolio
Level of resource management processes in a project $LS_{i,j}$ – criticality $K_{i,j}$	Allows you to assess the processes of human resource management in the bases of the level of indicators and criticality of processes
Portfolio project process $Pr_{i,j}$ – criticality $K_{i,j}$	Allows you to assess the process of human resource management in terms of their impact on critical success factors

Based on the result of the analysis of projections, a change vector is formed, which can be used in the reengineering of human resource management processes.

4.3 Method of stakeholder-based human resource management of projects in a multi-project environment

Accounting for the relationship of resource management processes with the processes of management of stakeholders in project management in a multi-project environment can improve management efficiency by reducing the risks associated with the human factor [24-25].

Stakeholders of resource management processes in a multi-project environment can be company executives, human resources director (ensures availability and availability of qualified human resources for managing a portfolio and its components), individual project managers, investors, customers, etc.

When managing a project portfolio, the stakeholders are [15]:

- supreme management body of the company, making strategic decisions on the development of the company's portfolio;
- portfolio management team;
- director of the portfolio;
- portfolio board;
- managers of programs;
- project managers;
- program / project office;
- program / project teams;
- customers;
- suppliers and partners.

The director of the portfolio, along with the analysis of financial indicators, performs an analysis of the availability of resources necessary for the implementation of the portfolio components.

The proposed method of stakeholder-based human resource management projects in a multi-project environment.

The main stages of the method:

Stage 1. Identification of stakeholders in the human resource management process:

- identification of stakeholders for each project;
- determining the projection of stakeholder interest on project management processes;
- identifying the stakeholders of the project's human resource management process and determining their interest;
- determining loyalty of interested parties;
- formation of stakeholders' expectations from the process of human resource management.

Stage 2. Formation of generalized requirements for project resources in a multi-project environment:

- formation of requirements for project resources;
- consistency check;
- formation of requirements for project portfolio resources;
- defining critical competencies;
- formation of a pool of critical competencies;
- checking resource availability.

Stage 3. Formation of adaptive commands:

- formation of adaptive teams with given resource requirements and constraints (reservation, prohibition on combination, prohibition on participation in projects, preferences of stakeholders, cost, etc.);
- consideration of adaptation;
- optimization by specified criteria;
- selection of command options.

Stage 4. Team activity analysis:

- changes analysis;
- requirements update;
- team adaptation.

Stage 5. Correction:

- resource pool correction;
- correction of the pool of critical competencies;
- correction of the register of stakeholders.

Stage 6. Stakeholder satisfaction analysis.

Depending on the chosen strategy of human resource management, there is a choice of criteria by which candidates are selected and a project team is formed.

To minimize the impact of resource constraints, it is proposed to use redundancy (as an alternative to the resource buffer). The formation of adaptive project teams with functional redundancy was considered in [11]. Under the functional reservation understand the ability of several performers to perform a specific function. The redundancy factors determine the minimum required number of performers capable of performing this function to ensure the implementation of the project.

The use of functional redundancy in the project will lead to the fact that a specific executor will be assigned to perform a specific function. The reserve for this function will be allocated to the execution of other functions, so that, if necessary, it can be reassigned to the performance of the function. Depending on the existing restrictions allocate "cold" and "hot" backup. In case of "cold reservation", the reserve is entered only if necessary and some time should be taken into account for the redistribution of the project team. The use of a "hot reserve" minimizes the time to put in a reserve, but

it leads to an increase in the staff of the project, an increase in the budget of the project.

The model of the process of building a functional-redundant team is shown in Fig. 3

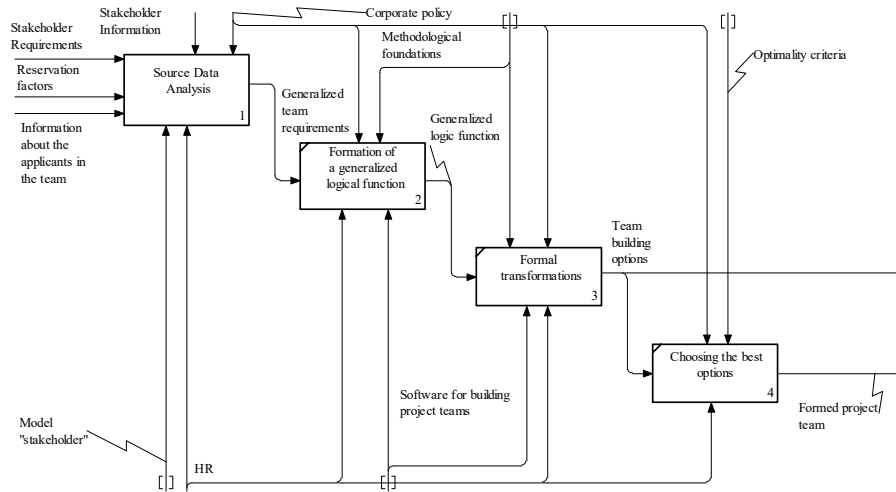


Fig. 3. Model of the process of building a functional redundant adaptive team

Resolving the task of redistributing resources is based on the use of the donor-acceptor approach. Donor-acceptor resource interaction in a multi-project environment is due to the presence of a single pool of organization resources, within which there is a redistribution of resources between the work of donor projects and acceptor projects [27].

5 Conclusions and prospects for further development

The article discusses and analyzes the features of human resource management in a multi-project environment.

A model of project-oriented resource management in a multi-project environment has been developed. A set of methods for determining human resource management strategies in a multi-project environment is proposed. It is proposed to use a stratified representation of resource management processes in a multi-project environment, which will allow an analysis of human resource management processes. Developed a method of stakeholder-based human resource management projects in a multi-project environment. A model of the process of constructing a functional-backup team has been developed.

The proposed models and methods of project-oriented resource management of the formation of adaptive teams in a multi-project environment will improve the efficiency of using human resources in the formation of multi-project teams and programs.

A promising direction is the development of a software package based on the proposed methods, which will automate the formation of adaptive commands and reduce the influence of the subjective factor on human resource management processes.

6 References

1. Ponsteen, A., Kusters, R.J.: Classification of Human- and Automated Resource Allocation Approaches in Multi-Project Management. *Procedia – Social and Behavioral Sciences*, Vol. 194, pp. 165-173 (2015).
2. Besikci, U., Bilge, U., Ulusoy, G.: Different Resource Management Policies in Multi-Mode Resource Constrained Multi-Project Scheduling. *Proceedings of 2011 World Congress on Engineering and Technology*, Vol. 2, pp. 64-67 (2011).
3. Adhau, S., Mittal, M.L., Mittal, A.: A multi-agent system for distributed multi-project scheduling: An auction-based negotiation approach. *Engineering Applications of Artificial Intelligence*, vol. 25(8), pp. 1738-1751 (2012).
4. Hoda, R., Murugesan L.K.: Multi-level agile project management challenges: A self-organizing team perspective. *The Journal of Systems and Software*, vol. 117, pp. 245–257 (2016).
5. Lalsing, V., Kishnah, S., Pudaruth, S.: People factors in agile software development and project management. *International Journal of Software Engineering & Applications (IJSEA)*, vol.3, no.1, pp. 117-137 (2011).
6. Li, X.B., Nie, M., Yang, G.H.: The Study of Multi-Project Resource Management Method Suitable for Research Institutes from Application Perspective. *Procedia Engineering*, vol. 174, pp. 155 – 160 (2017).
7. Arulrajah, A., Opatha, H., Nawaratne, N.: Green Human Resource Management Practices: A Review. *Sri Lankan Journal of Human Resource Management*, vol. 5, no.1, pp.1-17. (2015).
8. Renwick, D.W.S., Redman, T., Maguire, S.: Green Human Resource Management: A Review, and Research Agenda. *International Journal of Management Review*, vol. 15, pp. 1-14. (2013).
9. Kharitonov, Yu.N., Yelgina, E.V.: Identified risks when choosing a project team. *Eastern-European Journal of Enterprise Technologies*, vol. 3(2), pp. 39-41. (2013).
10. The standard for portfolio management. 4th ed. Newtown Square, PA : Project Management Institute, 127 p. (2017).
11. Dotsenko, N.V., Sabodosh, L.Yu., Chumachenko, I.V.: Methods of human resources management in the formation of multi-project and program teams, 201 p. (2015).
12. Yesina, V., Matveeva, N., Chumachenko, I., Manakova, N.: Method of data openness estimation based on user experience in infocommunication systems of

municipal enterprise. 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), pp. 171-176. (2018).

13. Araujo, C., Pedron, C., Quevedo-Silva, F.: IT Project Manager Competencies and Team Commitment: A new scale proposal. *Revista de Gestão e Projetos*, vol. 9, pp. 39-57. (2018).
14. Mathieu, J., Tannenbaum, S., Kukenberger, M., Donsbach, J., Alliger, G.: Team role experience and orientation: a measure and test of construct validity. *Group and organization management*, vol. 40 (1), pp. 6-34. (2015).
15. Obradovic, V., Montenegro, A., Bjelica, D.: Digital era and project manager's competencies. *European project management journal*, vol. 8, iss. 1, pp. 4-9. (2018).
16. Barros, A., Oliveira, S.: Benefits of using program and project portfolio management on the execution of deliberate and emerging strategies. *European Journal of Scientific Research*, vol. 147, no. 4, pp. 369-377 (2017).
17. Bazilevych, K. et al.: Stochastic modelling of cash flow for personal insurance fund using the cloud data storage. *International Journal of Computing*, vol. 17, iss. 3, pp. 153-162 (2018).
18. A Guide to the Project Management Body of Knowledge (PMBOK Guide). Newtown Square, Pa: Project Management Institute, (2004).
19. The Standart for Program Management – Fourth Edition. Newtown Square, Pa: Project Management Institute, (2017).
20. Armstrong, M., Taylor, S.: Armstrong's handbook of human resource management practice, 824 p. (2014).
21. Chumachenko, D.: On Intelligent Multiagent Approach to Viral Hepatitis B Epidemic Processes Simulation. *Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018*, pp. 415-419. (2018).
22. Babaiev, V.M., Kadykova, I.M., Husieva, Yu.Yu, Chumachenko, I.V.: The method of adaptation of a project-oriented organization's strategy to exogenous changes. *Scientific Bulletin of National Mining University*, vol. 2 (158), pp. 134-140. (2017).
23. Dotsenko, N.V., Sinitskaya, N.V.: Application of the method of analyzing the consistency of priorities of strategies to identify the starting point of business process reengineering. *Radioelectronic and computer systems*, vol. 1, pp. 150-156. (2014).
24. Dotsenko, N.V., Skrynnik, A.I., Lysenko, A.I.: Application of the stakeholder model for analyzing the effectiveness of project management by stakeholders. *Radioelectronic and computer systems*, vol. 1, pp. 115-119. (2015).
25. Dotsenko, N.V.: Analysis of the impact of stakeholders on human resource management in the project. *Information technology and innovation in economics, project management and programs*, pp. 307-315. (2016).
26. Dotsenko, N., Chumachenko, D., Chumachenko, I.: Modeling of the processes of stakeholder involvement in command management in a multi-project environment, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 – Proceedings, pp. 29-32 (2018).

27. Dotsenko N. Application of the donor-acceptor approach to resource providing in a multi-project environment. Transformational processes the development of economic systems in conditions of globalization: scientific bases, mechanisms, prospects: collective monograph, vol. 1., pp. 173–181. (2018).

A Method of Common Signal Monitoring in FPGA-Based Components of Safety-Related Systems

Oleksandr Drozd^[0000-0003-2191-6758], Viktor Antoniuk^[0000-0001-8436-5338],
Svetlana Antoshchuk^[0000-0002-9346-145X], Myroslav Drozd^[0000-0003-0770-6295]

Odessa National Polytechnic University, Ave. Shevchenko 1, 65044 Odessa, Ukraine
drozd@ukr.net, viktor.v.antoniuk@gmail.com, asgonpu@gmail.com,
myroslav.drozd@opu.ua

Abstract. Traditional solutions in ensuring the functional safety of safety-related systems and their digital components based on methods and means of testing and on-line testing, as well as fault-tolerant structures, including majority schemes using multi-version technologies to counter common cause failures are considered. The limitation of these approaches by the logical checkability of digital circuits in the structural, structurally functional, and dual-mode versions is shown. Multi-version solutions are aimed at countering common cause failures, including common control faults related to reset, synchronization signals and other common signals that can block digital components and their checking circuits in a state identified as working. However, faults in chains of common signals can also be addressed to hidden faults, which remain a problem in safety-related systems. The logical checkability of the circuits decreases from structural to dual-mode and increases with the reduction of matrix structures. The maximum reduction is achieved in bitwise pipelines. The successes of green and FPGA technologies created the conditions for the development of on-line testing methods based on an assessment of energy consumption. These methods can significantly complement the logical checking. A method for monitoring common signals by estimating consumption currents in circuits of bitwise pipelines using the example of a shifting register is proposed. The results of experimental confirmation of the effectiveness of the proposed method is achieved.

Keywords: safety-related system, component, FPGA design, logical and power-oriented checkability, hidden faults, common signal, matrix structure, consumption current, bitwise pipeline

1 Introduction

One of the most important areas in the development of computer systems is associated with their critical applications, where they are used as instrumentation and control safety-related systems for managing high-risk objects that are widely represented in transport, energetics and other manufacturing and service industries [1, 2]. International standards governing the design and use of safety-related systems, impose on them the task of ensuring functional safety of the object and the system itself to prevent accidents and reduce losses in the event of their occurrence [3, 4].

Safety is provided by a set of test maintenance procedures for system components during work breaks and their continuous automatic monitoring based on on-line testing methods and tools [5, 6]. Components are designed using fault-tolerant solutions based on correction codes, reconfiguration [7, 8], as well as majority structures and multi-version technologies to counter common cause failures. Such failures include design errors and faults in chains of common signals that provide reset, synchronization, and other general control functions. Multi-version technologies are based on various types of diversity and are aimed at eliminating a common cause or the same effects in different versions of performing calculations [9].

As a rule, testing and on-line testing is performed by analyzing the result for an error, which relates the appropriate methods and means to the logical checking of digital circuits [10, 11]. The source of error is the fault of the circuit. Therefore, the logical checking is performed if the circuit fault can be observed in the calculated result, causing an error in it. This property of the suitability of the circuit to a certain type of checking of its faults is called checkability [12, 13].

Thus, the logical checking of the circuit can be performed only within its logical checkability. This conclusion also applies to the effectiveness of fault-tolerant structures and the multi-version technologies used in them [14]. A fault that does not cause an error in the analyzed result is not dangerous for computer systems working only in the operating mode, but creates the problem of hidden faults in safety-related systems designed for operation in two modes: normal and emergency. These faults, including faults in chains of common signals, can accumulate over the course of an extended normal mode and manifest themselves in emergency mode, violating the fault tolerance of the circuits and the safety of systems and control objects [15, 16].

The need to support logical checking with other forms of checking and checkability finds a response in the modern level of development of green technologies [17, 18] and FPGA design [19, 20]. At the intersection of these areas, the evaluation of circuits for their energy consumption is improving.

On-line testing methods based on measuring the temperature of the circuits in the course of their work are known [14]. We propose a method that continues this direction in relation to faults in chains of common signals based on the capabilities of modern CAD systems in estimating consumption currents.

Faults in chains of common signals have a significant impact on the energy consumption of the circuit. Monitoring of common signals is carried out within the frame of power-oriented checkability, which marks the bottom level of current consumption with proper functioning of the circuit. The proposed method detects the impossible (with the correct operation) decrease in dynamic component of the current consumption in case of a fault in the chain of common signal.

Section 2 describes logical checking problems that limit the ability to detect faults in chains of common signals for safety-related systems and highlights bitwise pipelines that show the greatest logical checkability. Section 3 notes the possibilities of modern FPGA design in the evaluation of circuits in their power consumption and formulates the main points of the method of common signals monitoring. Section 4 presents the results of experiments on the use of the suggested method in a bitwise pipeline using the example of detecting synchronization errors of a shifting register.

2 Problems of Logical Checking in Common Signal Monitoring

Logical checking problems begin with its checkability. At the same time, the restrictions imposed on the checking of the circuit by appropriate checkability are also the motivation for its increase.

Logical checkability can be refined for testing and on-line testing as forms of logical checking.

Circuit testing is limited to testability, i.e. the suitability of the circuit for developing tests that detect faults. Testability refers to logical checkability, which is characterized as structural, since it is completely determined by the structure of the circuit. Structural checkability is enhanced by methods and means of testable design, which is aimed at the formation of the structure of the scheme with a high degree of controllability and observability of its internal points [21, 22].

On-line testing is limited to logical checkability, which is characterized as structurally functional, since it is determined not only by the structure of the circuit, but also depends on its input data. Therefore, an increase in the structurally functional checkability of the circuit can be achieved on the basis of two approaches: the improvement of its structure, and a change in the character of the input data.

The matrix structure orients the digital circuit to receive and process of input data in parallel codes. This introduces major limitations to the controllability and observability of the interior points of the circuit. Therefore, the first approach, as a rule, is based on the reduction of matrix structures. The processing of approximate data in floating-point formats [23, 24] demonstrates this approach in the use of truncated arithmetic operations with mantissas [25]. Residue checking of abbreviated operations performed in matrix arithmetic nodes is described in [26, 27].

Both approaches are implemented in bitwise pipelines, which reduce the matrix of operational elements in the pipeline section to one element. At the same time, the nature of the input data changes, which are represented by sequential codes. The structure of bitwise pipelines is based on scanning registers, which are elements of testable design [28].

The fault-tolerant circuits used in the digital components of safety-related systems are also limited by logical checkability, which reflects the characteristics of these systems to separate the operating mode into normal and emergency [29].

The result of this separation is the following chain of consequences:

- Various input circuit data in these modes;
- Various structurally functional checkability, characterized as dual-mode;
- A new type of faults hidden in normal mode due to the lack of input data showing them and violating fault tolerance of circuits on other input data in emergency mode.

The problem of hidden faults has not yet received a safe solution. Detection of hidden faults that can accumulate in the chains of common signals is still performed using simulation modes, that recreate emergency conditions and not once led to them as a result of unauthorized activation by malfunction or with human participation [30, 31].

It should be noted that testability, as the simplest form of logical checkability, is the greatest. Structurally functional checkability is limited from above by testability

because it has an additional limiting dependence on the input data. Dual-mode checkability is the least. It is bounded above by the structurally functional checkability of the normal mode, since does not consider faults hidden in emergency mode.

Therefore, it is important to increase dual-mode checkability both at the level of testability and at the level of structurally functional checkability. This solution is to perform calculations on bitwise pipelines:

- Their register structures combine scanning register functions that increase testability in testable design;
- Minimal matrix structures in the sections of the pipeline increase the structurally functional checkability of the circuit;
- Processing data in sequential codes evens out the variety of input data and structurally functional checkability of the circuit in normal and emergency modes and in this way improves dual-mode checkability.

Thus, logical checking is not sufficiently ensured by the checkability of circuits in the components of safety-related systems and needs to be supplemented with other forms of checking, in particular, to detect hidden faults in chains of common signals. It should be noted that from the point of view of logical checking, the most checkable solution for safety-related systems is the design of their digital components based on bitwise pipelines.

3 Monitoring of the Common Signals

The method of monitoring of the common signals is based on power-oriented checkability, which is provided by the circuit with the support of FPGA design methods and tools developed in green technologies. Modern FPGA-oriented CAD systems contain utilities that model a FPGA project to estimate its power consumption. The simulation results are estimates of currents consumption: the total current and its dynamic and static components. These results can be obtained for different activity of the input signals, which is given as a percentage of the activity of the clock signals. Reducing the activity of the input signals leads to a decrease in current consumption.

Fault detection in the chain of common signals is based on the manifestation of this fault in the form of a decrease in current consumption in its dynamic component due to a decrease in the number of switching signals in the circuit. Such a manifestation of a fault is characteristically for synchronization signals, reset signals and other general control signals. Fault of the common signal blocks the operation of the circuit or its part, as a result of which we can expect a significant decrease in switching activity, as well as the dynamic component and the total current consumption.

The method detects a fault if the current consumption drops below the minimum possible level when properly functioning. This threshold $I_{PT.MIN}$ of fault detection is marked by power-oriented checkability, which takes into account:

- The minimum total current consumption I_{PT} , which is determined by simulation at zero activity of input signals;
- The error δI_{PT} of simulation when estimating total current consumption;
- The error δI_M of measuring the current consumption by the sensor in the process of common signals monitoring.

The fault detection threshold is determined by the following formula:

$$I_{PT.MIN} = I_{PT} - \delta I_{PT} - \delta I_M.$$

As a rule, CAD systems simulate and sensors measure the current consumption with an error not exceeding $\delta I_{PT} = \delta I_M = 0.5\%$, which determines the fault detection threshold at the level: $I_{PT.MIN} = 0.99 I_{PT}$.

We consider the proposed method for a bitwise pipeline on the example of a shifting register to the right, which consists of n series-connected triggers, numbered from 1 to n from left to right. The dynamic component I_{PTD} of current consumption consists of two parts I_{DC} and I_{DI} , which are determined by switching the clock signals and information signals, respectively: $I_{PTD} = I_{DC} + I_{DI}$. The switching of information signals is taken into account when setting the activity of the input signals as a certain number of z as a percentage of the activity of the clock signals: $I_{DI} = z I_{DC}$.

Then the dynamic component is represented as $I_{PTD} = I_{DC} (1 + z)$.

Let the fault turn off the clock signals for d triggers. Then we can expect the decrease of the current I_{DC} by value $\Delta I_{DC} = I_{DC} r$, where $r = d / n$. I_{DI} current decreases by value $\Delta I_{DI} = I_{DI} (n - x + 1) / n$, where x is the smallest number within the disabled triggers numbers, $x \leq n - d + 1$, since all triggers starting from the disabled x trigger will have zero information activity signals.

The decrease ΔI_{PT} of the total current consumption, which is manifested in the reduction of the dynamic component ΔI_{PTD} while maintaining the static component, is determined by the following formula:

$$\Delta I_{PT} = \Delta I_{PTD} = I_{DC} (d + z (n - x + 1)) / n. \quad (1)$$

The possibilities of the proposed method are governed by the minimum decrease $\Delta I_{PTD.MIN}$ in the dynamic component of the current consumption, which is achieved under the condition $x = n - d + 1$, i.e. in the case of the location of all disabled triggers in a row at the right end of the shifting register.

The minimum decrease of the total current consumption and its dynamic component is determined from the formula (1) as follows:

$$\Delta I_{PT.MIN} = I_{DC} (z + 1) d / n = I_{PTD} r. \quad (2)$$

Thus, we can expect a decrease in total current consumption by at least a fraction of the dynamic component, equal to the fraction r of the disabled triggers. We can estimate the values of r and $d = r n$, based on the inequality $I_T - I_{PT.MIN} < \Delta I_{PT.MIN}$, where I_T – is the total current consumption resulting from the simulation. Then, according to the formula (2), $r > (I_T - I_{PT.MIN}) / I_{PTD}$ and

$$d > n (I_T - I_{PT.MIN}) / I_{PTD}.$$

The method uses the total current consumption I_M , measured by the sensor [32], and the fault detection threshold $I_{PT.MIN}$. In case of $I_M < I_{PT.MIN}$, the method identifies the fault.

4 Experimental Assessment of Suggested Method

For the experimental evaluation of the proposed method in the Quartus Prime 18.1 Lite Edition CAD system [33] the VHDL-project of a 32-bit shifting register with information input SI , reset input Res , clock inputs CLK for each individual register trigger, and information SO output has been created (Fig. 1).

The project was implemented on Intel Max 10 FPGA 10M50DAF672I7G [34]. Setting the time characteristics was made in the utility TimeQuest Timing Analyzer [35].

Setting activity values A_I of the input and internal signals of the circuit, and modeling the total current consumption I_T of FPGA core, as well as its dynamic I_D and static I_S components, was performed by the PowerPlay Power Analyzer utility [36].

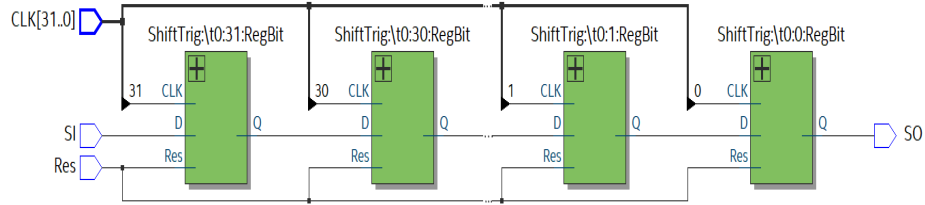


Fig. 1. An example of the RTL-scheme of a 32-bit shifting register

The simulation consisted in determining the values of the current consumption of the circuit of shifting register for all properly functioning 32 bit ($d = 0$) and then disabling the CLK inputs of triggers at the right end of the register starting with row of $d = 4$ triggers with a shutdown step of 4. The values of currents consumption for all cases determines with range of activity A_I of the input information signal SI and internal shifting register signals from 0% to 100% of the clock signal frequency CLK , which was 250 MHz.

Currents I_T , I_S , I_D modelling results are given in Table 1.

Tab. 1. Experiment results for 32-bit shifting register

A_I , %	$d = 0$			$d = 4$			$d = 8$			$d = 12$		
	I_D , mA	I_S , mA	I_T , mA	I_D , mA	I_S , mA	I_T , mA	I_D , mA	I_S , mA	I_T , mA	I_D , mA	I_S , mA	I_T , mA
0	5.77	11.81	17.59	5.22	11.80	17.02	4.23	11.79	16.02	3.63	11.78	15.41
12.5	6.48	11.82	18.30	5.75	11.80	17.55	4.68	11.79	16.47	4.00	11.78	15.78
25	7.18	11.82	19.01	6.28	11.81	18.08	5.13	11.79	16.93	4.36	11.78	16.14
37.5	7.89	11.83	19.71	6.80	11.81	18.61	5.58	11.79	17.38	4.73	11.78	16.51
50	8.59	11.83	20.42	7.33	11.81	19.14	6.04	11.80	17.83	5.09	11.78	16.88
62.5	9.30	11.84	21.13	7.86	11.81	19.67	6.49	11.80	18.29	5.46	11.79	17.24
75	10.00	11.84	21.84	8.38	11.81	20.20	6.94	11.80	18.74	5.82	11.79	17.61
87.5	10.70	11.85	22.55	8.91	11.82	20.73	7.39	11.80	19.19	6.19	11.79	17.98
100	11.41	11.85	23.26	9.44	11.82	21.26	7.84	11.80	19.65	6.55	11.79	18.34

In Table 1, value $I_{PT} = 17.59$ of current I_T for the case $d = 0$ and $A_I = 0$ is highlighted in bold. This value is used to calculate the fault detection threshold $I_{PT,MIN} = 17.41$. In the columns, where $d > 0$, the values of the current I_T is highlighted in bold when it is below the $I_{PT,MIN}$ threshold, i.e. in case of fault detection.

Table 2 compares the calculated and experimental results.

Tab. 2. Comparison of the calculated and experimental results

A_I , %	I_D , mA	I_S , mA	I_T $d=0$, mA	I_T $d=4$, mA	I_T $d=8$, mA	I_T $d=12$, mA	$\delta_{T\ d=4}$, %	$\delta_{T\ d=8}$, %	$\delta_{T\ d=12}$, %
0	5.77	11.81	17.59	16.87	16.15	15.43	0.88	-0.81	-0.13
25	7.18	11.82	19.01	18.11	17.21	16.32	0.17	-1.65	-1.12
50	8.59	11.83	20.42	19.35	18.23	17.20	-1.10	-2.47	-1.90
75	10.00	11.84	21.84	20.59	19.34	18.09	-1.93	-3.20	-2.73
100	11.41	11.85	23.26	23.26	20.41	18.98	-2.68	-3.87	-2.44

The total current consumption $I_{T\ d=4, 8, 12}$ when d triggers are disconnected is calculated by reducing the current $I_{T\ d=0}$ by the value $\Delta I_{PT\ MIN} = I_D d / 32$. This expected current value is compared with the current I_T , obtained for the corresponding d experimentally (Table 1). The error $\delta_{T\ d=4, 8, 12}$ is calculated by the formula:

$$\delta_{T\ d} = (I_{T\ d} - I_T) / I_T.$$

Fig. 2 visualize the values of calculated and experimental I_T for $d = 8$. Fig 3 visualize the values of error δ_T for $d = 4, 8, 12$.

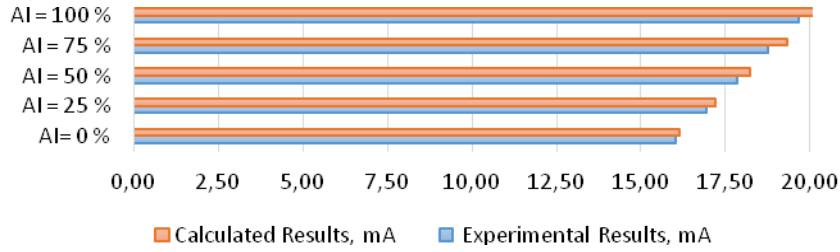


Fig. 2. Visualization of the values I_T for $d = 8$

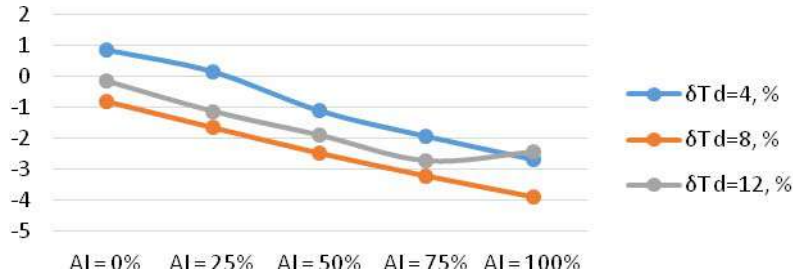


Fig. 3. Visualization of the calculated and experimental results for various d

The obtained error values indicate a good agreement between the theoretical and experimental results that characterize the suggested method for monitoring common signals.

5 Conclusions

Methods and tools for testing and on-line testing, as well as fault-tolerant solutions are the basis for ensuring the functional safety of safety-related systems. However, the possibilities of these approaches, combined by the features of logical checking, are significantly limited by the structural, structurally functional, and dual-mode checkability of digital circuits, respectively.

The lack of dual-mode checkability, the carriers of which are matrix digital circuits in the components of safety-related systems, creates in them the problem of hidden faults. These faults pose a significant threat to the safety of systems and control objects in critical applications. The use of hazardous imitation modes that recreate emergency conditions to detect hidden faults is an indisputable proof of the significance of this problem, which prevents fault-tolerant solutions from becoming fault-safe.

In conditions of limited structurally functional checkability of digital circuits in normal mode, many faults can become hidden. In this case, faults in chains of common signals can cause serious functional disturbances in the system components and manifest themselves with the transition to emergency mode. Means of online testing can be blocked by such faults or taken for the correct state of the circuit, fixed by the fault.

Logical checking of circuits in critical applications requires the involvement of other forms of checking.

The suggested method of common signals monitoring continues the development of approaches that take into account the energy impact of faults on the operation of the circuit. Known solutions that perform temperature monitoring of circuits take into account changes in power dissipation using temperature sensors. The proposed method uses the achievements of green and FPGA technologies, which allow to detect the energy trace of faults based on much more accurate estimates of current consumption.

The method is shown for bitwise pipelines that have the greatest logical checkability in critical applications, using the example of a shifting register and a fault in the synchronization of its triggers. The results of the experiments have showed high efficiency of the method and their good convergence with the theoretical estimates of the method.

References

1. Gorbenko, A., Kharchenko, V., Tarasyuk, O., Zasukha, S.: Safety of Rocket-Space Engineering and Reliability of Computer Control Systems and Software: 2000-2009 YRS. In: First Intern. Workshop "Critical Infrastructure Safety and Security", pp. 79–93, Kirovograd, Ukraine (2011).
2. Efanov, D., Lykov, A., Osadchy, G.: Testing of relay-contact circuits of railway signalling and interlocking. In: IEEE East-West Design and Test Symposium, EWDTS 2017, pp. 242–248, Novi Sad, Serbia (2017).
3. International Electrotechnical Commission, Nuclear Power Plants: Instrumentation and Control for Systems Important to Safety – General Requirements for Systems, Rep. IEC 61513, IEC, Geneva (2001).
4. IAEA NS-G-1.3. Instrumentation and control systems important to safety nuclear power plants. Safety guide, Vienna (2002).
5. Romankevich, V.: Self-testing of multiprocessor systems with regular diagnostic connections. Automation and Remote Control, vol. 78, no. 2, pp. 289–299 (2017).
6. Abramovichi, M., Stroud, C., Hamilton, C., Wijesuriya, S., Verma, V.: Using roving STARs for on-line testing and diagnosis of FPGAs in fault-tolerant applications. In: IEEE International Test Conference, pp. 973–982. Atlantic City, USA (1999).
7. Yatskiv, V., Yatskiv, N., Jun, S., Sachenko, A., Zhengbing, H.: The use of modified correction code based on residue number system in WSN. In: 7th IEEE International Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, pp. 513–516. Berlin, Germany (2013).
8. Palagin, A.V., Opanasenko V.N.: Design and application of the PLD-based reconfigurable devices. In: Adamski, M., Barkalov, A., Wegrzyn M. (eds.) Design of Digital Systems and Devices, Lecture Notes in Electrical Engineering, vol. 79, pp. 59–91. Springer Verlag, Berlin Heidelberg (2011).
9. Kharchenko, V.: Multy-version Systems: Models, Reliability, Design Technologies. In: 10th European Conference on Safety and Reliability, pp. 73–77, Munich, Germany (1999).
10. Brezhnev, E.: An approach for assessing risk of common cause failures in critical infrastructure. Information & Security, vol. 28, no. 1, 199–210 (2012).
11. Hahanov, V., Litvinova, E., Obrizan, V., Gharibi, W.: Embedded method of SoC diagnosis. Elektronika in Elektrotechn, no 8, pp. 3–8 (2008).
12. Drozd, A., Drozd, M., Martynyuk, O., Kuznietsov, M.: Improving of a Circuit Checkability and Trustworthiness of Data Processing Results in LUT-based FPGA Components of Safety-Related Systems. In: CEUR Workshop Proceedings, vol. 1844, pp. 654–661 (2017).
13. Drozd, A., Antoshchuk, S., Drozd, J., Zashcholkina, K., Drozd, M., Kuznietsov, N., Al-Dhabi, M., Nikul, V.: Checkable FPGA Design: Energy Consumption, Throughput and Trustworthiness. In book: Green IT Engineering: Social, Business and Industrial Applications, Studies in Systems, Decision and Control, V. Kharchenko, Y. Kondratenko, J. Kacprzyk (eds), vol. 171, pp. 73–94. Berlin, Heidelberg: Springer International Publishing (2018). DOI: 10.1007/978-3-030-00253-4_4.
14. Nicolaidis, M., Zorian, Y., Pradhan, D (eds): On-Line Testing for VLSI. In: Journal of Electronic Testing: Theory and Application, vol. 12, no. 1/2, pp. 7–159 (1998).

15. Drozd, A., Drozd, M., Antonyuk, V.: Features of Hidden Fault Detection in Pipeline Components of Safety-Related System, CEUR Workshop Proceedings, vol. 1356, pp. 476–485 (2015).
16. Drozd, A., Drozd, J., Antoshchuk, S., Antonyuk, V., Zashcholkina, K., Drozd, M., Titomir, O.: Green Experiments with FPGA. In book: Green IT Engineering: Components, Networks and Systems Implementation, V. Kharchenko, Y. Kondratenko, J. Kacprzyk (eds), vol. 105, pp. 219–239. Berlin, Heidelberg: Springer International Publishing (2017). DOI: 10.1007/978-3-319-55595-9_11.
17. Murugesan, S., Gangadharan, G.: Harnessing Green IT. Principles and Practices, UK: Wiley and Sons Ltd. (2012).
18. Tyurin, S., Kamenskiy, A.: Green Logic: Models, Methods, Algorithms. In: Kharchenko V., Kondratenko Y., Kacprzyk J. (eds) Green IT Engineering: Concepts, Models, Complex Systems Architectures. Studies in Systems, Decision and Control, vol 74, pp 69–86, Springer (2017). DOI: https://doi.org/10.1007/978-3-319-44162-7_4.
19. Drozd, A., Drozd, M., Kuznetsov, M.: Use of Natural LUT Redundancy to Improve Trustworthiness of FPGA Design. In: CEUR Workshop Proceedings, vol. 1614, pp. 322–331 (2016).
20. Kondratenko, Y., Gordienko, E.: Implementation of the neural networks for adaptive control system on FPGA. In: Annals of DAAAM for 2012 & Proceeding of the 23th Int. DAAAM Symposium "Intelligent Manufacturing and Automation", vol. 23, no.1, B. Katalinic (Ed.), DAAAM International, pp. 389–392, Vienna, Austria (2012).
21. Matrosova, A., Nikolaeva, E., Kudin, D., Singh, V.: PDF testability of the circuits derived by special covering ROBDDs with gates. In: IEEE East-West Design and Test Symposium, EWDTS, pp. 1–5. Rostov-on-Don, Russia (2013).
22. Goldstein, L. M., Thigen, E. L.: SCOAP: Sandia Controllability/Observability Analysis Program. In: 17th Design Automation Conference, pp. 190–196, Minneapolis, MN, USA, (1980).
23. ANSI/IEEE Std 754-1985, IEEE Standard for Binary Floating-Point Arithmetic (1985).
24. IEEE Std 754™-2008 IEEE Standard for Floating-Point Arithmetic. IEEE 3 Park Avenue New York, NY 10016–5997, USA (2008).
25. Garofalo, V.: Truncated Binary Multipliers with Minimum Mean Square Error: Analytical Characterization, Circuit Implementation and Applications. In: Ph.D. Dissertation. University of Studies of Naples "Federico II", Naples, Italy (2008).
26. Drozd, A., Lobachev, M., Hassonah, W.: Hardware Check of Arithmetic Devices with Abridged Execution of Operations. In: European Design and Test Conf, p. 611. Paris, France (1996) DOI: 10.1109/EDTC.1996.494375.
27. Drozd, A., Lobachev, M.: Efficient On-line Testing Method for Floating-Point Adder. In: Design, Automation and Test in Europe. Conference and Exhibition 2001, pp. 307–311. Munich, Germany (2001). DOI: 10.1109/DATE.2001.915042.
28. Bennets, R. G., Maunder, C. M., Robinson, G. D.: CAMELOT: a computer-aided measure for logic testability. In: IEE Proceedings E – Computers and Digital Techniques, vol. 128, no 5, pp. 177–189 (1981).
29. Kharchenko, V., Gorbenko, A., Sklyar, V., Phillips, C.: Green Computing and Communications in Critical Application Domains: Challenges and Solutions. In: 9th International Conference on Digital Technologies (DT'2013), pp. 191–197. Zhilina, Slovakia (2013).
30. Gillis, D.: The Apocalypses that Might Have Been. [Online]. Available: <http://www.popmech.ru/go.php?url=http%3A%2F%2Fwww.damninteresting.com%2F%3Fp%3D913>.
31. U.S.-Canada Power System Outage Task Force: Final Report on the August, 14, 2003 Blackout in the United States and Canada: Causes and Recommendations USA, (2004).
32. MAX 10 FPGA Development Kit User Guide (2017), <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug-ug-max10m50-fpga-dev-kit.pdf>, last accessed 2019/03/20.

33. Intel Quartus Prime Standard Edition User Guide: Getting Started, <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug-qps-getting-started.pdf>, last accessed 2019/03/20.
34. Max 10 FPGA Device Architecture (2017), https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/hb/max-10/m10_architecture.pdf, last accessed 2019/03/20.
35. Intel Quartus Prime Standard Edition User Guide: Timing Analyzer (2018), <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug-qps-timing-analyzer.pdf>, last accessed 2019/03/20.
36. Intel Quartus Prime Standard Edition User Guide: Power Analysis and Optimization (2018), <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug-qps-power.pdf>, last accessed 2019/03/20.

Information and Communication Technology: Case of Using Petri Nets for Grain Delivery Simulation at Logistics System

Natalya Shramenko^{1,2[0000-0003-4101-433X]}, Olexiy Pavlenko^{3[0000-0003-4237-4310]} and Dmitriy Muzylyov^{4[0000-0002-8540-6987]}

¹ Kharkiv Petro Vasylenko National Technical University of Agriculture, Alchevskiyh Str. 44, 61002, Kharkiv, Ukraine

² Ukrainian State University of Railway Transport, Feierbakh Square 7, 61050, Kharkiv, Ukraine
nshramenko@gmail.com

³ Kharkiv National Automobile and Highway University, Yaroslava Mudrogo Str. 25, 61002, Kharkiv, Ukraine
ttpov@ukr.net

⁴ Kharkiv Petro Vasylenko National Technical University of Agriculture, Alchevskiyh Str. 44, 61002, Kharkiv, Ukraine
murza_1@ukr.net

Abstract. The paper is showing an example of applied use of modern information and communication technologies (ICT) in logistics. It was proved application expediency of this type of mathematical tools for the solution of the logistic problem on grain delivery based on Petri Nets analysis. It is shown the relevance using of information and communication technologies at such type when choosing the best delivery variant for export cargo turnover of grain freights. The statistics for export of this agricultural products category was presented. It demonstrates the annual growth for volumes of foreign sending. This fact gives prerequisites to using ICT for the successful solution of the considered problems. Interrelations at information exchange by participants of grain delivery were presented. The two-level representation structure for logistics system was described considering this fact. The visual scheme is developed for the system of information and communication technologies when transportation variants of grain cargo turnovers were simulated based on Petri Nets. The mathematical description is executed for operators used for modeling by a network. It is also presented visually network view for an assessment of delivery terms by time indicator. The histogram for three alternatives of a supply chain is designed with the help of simulation results. It allows choosing quickly the acceptable delivery scheme considering the minimum time of cargo turnover passing and, also accounting existence of transportation capacities during the concrete time period. This ICT considerably will accelerate the adoption for correct decision on the choice of a supply chain that minimizes expenses of the logistic companies when exporting the Ukrainian grain.

Keywords: Petri Nets, System, Logistics, Grain, Data, Information, Transport, Delivery, Simulation, Technology.

1 Introduction

The different and considerable data volumes appear in modern conditions at information saturation of society and various production spheres. This fact gives prerequisites for using computer technologies and intelligent systems for fast processing of information flow. Especially it is important to arrive at correct decisions in an operational period of planning for complex systems, such as the supply chain of agricultural cargoes. In particular, the logistic system of grain cargoes delivery has many communication links, including information, and, also consists of a significant elements amount. Therefore, there are some problems with effective management and creation of "correct time-table" of delivery from a production location to the place of its consumption [1-2]. There is a necessity to investigate the behavior of the studied system and to obtain information on its most important characteristics. The systems of parallel information processing and in parallel the functioning objects allow coming to the decision of problems connected with effective interaction of separate elements for the considered delivery system. The best example is the models based on Petri Nets theory [3].

Active information technology development allowed accelerating in recent years process of representation, processing, and use of information for cargo delivery system [4]. However, the companies sustain losses because of the untimely or doubtful obtained data on an operational situation at each link of a supply chain [5]. Today the attention is more and more paid to the information flow by means of which a material flow can be planed, operated and controlled [6]. Improvement of informational content and the organization quite often can bring a bigger effect, than technical innovations [7].

Today Ukraine carries out the biggest export of corn, barley, and wheat in the world (Fig.1) [8-9].

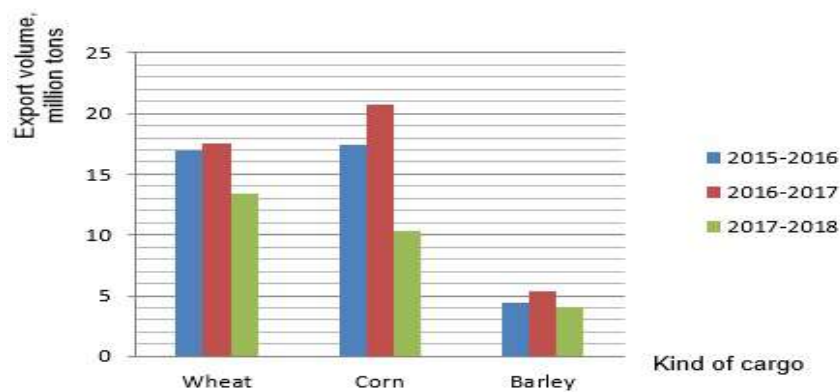


Fig. 1. Volume of the Ukrainian export grain at marketing period in 2015-2018 years

According to Fig.1 wheat was sent for export in bigger volume than others grain within the last year. Logistic costs, as a rule, include the expenses connected with transportation, warehousing, storage [10], loading and unloading operations, registration paperwork [11], cleaning, drying of cargoes, packing, safety and, also payment of any commission charges, tariffs, and taxes which connected with grain export. At the same time, delivery costs of grain cargoes exceed in Ukraine same tariffs at leading Europe countries (Fig. 2.). This explains the necessity of more flexible planning with the use of the modern systems which are based on information and communication technologies

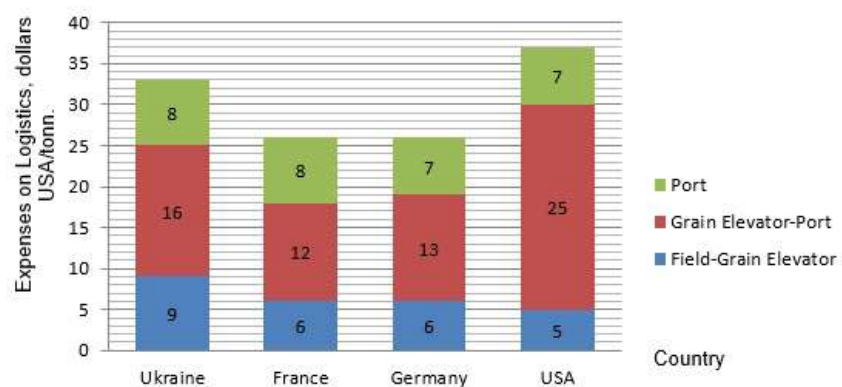


Fig. 2. Values of logistics costs per wheat delivery in different countries

The relevance of this research is also based on fact that in 2018 amount of cargo overloading in Ukrainian seaports reached 135.2 million tons. Due to the increase in cargo turnover according to four main nomenclatures which provide 71.4% of overloading: agricultural goods, ore, steel products, container transportation [12]. Thus, it is necessary to design models for logistic system of grain cargoes delivery from the producer to departure seaport and to carry out simulation for determination of impact level by obtaining information about forwarding material flow to the efficiency of decision adoption [13].

A modeling problem becomes simple when Petri Nets use. This tool can help to consider a bigger number of random factors which characterize any logistical systems during simulation. Thus, using such mathematical tools as Petri Nets reduces the probability of errors and failures emergence in a system. Therefore, the conducted modeling turns out the most adequate.

2 Literature Review

The approaches based on heuristic methods, own experience of managers and discipline of orders execution can be used now in goods transportation management in Ukraine. In practice it means emergence of the real losses connected with increase in cargo delivery period, idle times of cargo in wagons at "the abandoned trains" on

approaches to ports, boundary transitions and in warehouses of the enterprises (elevators), work arrhythmia for subsystems of transport and long investigations - "who is guilty". The sampling analysis for performance of cargoes delivery periods (based on data from the transportation document and calculated in the automated systems of normative delivery period) demonstrates what violations of term occurs when transporting (by all kinds) shipping at the level of 5%, container 32%, wagon carriage - 17% and group shipping - 8% [14]. Therefore, the logistics system of grain cargoes delivery should be considered from the viewpoint of the transport organization with delivery by trucks, railway wagons and with possibility using of containers [15] and piggyback traffic [16].

The simulation approaches quickly develops in recent years accounting active development of technologies at information processing. By the author in [17] was proved that imitating modeling of any events proceeding in active processes is relevant.

In recent years, many countries of the world made traceability compulsory procedure in a supply chain [18]. Management of cargo turnovers is represented as a difficult task for participants of a multimodal transport chain that is the main driving force in a supply chain [19].

The purpose of this work consists in offering the intelligent container tracking system for goods in the context of multimodal transportation: intelligent container tracking system for goods (i-TSCG). The tracking system of goods supply chain is proposed by several researchers because has many communications links and ensample factors. At the same time, models of the studied objects are optimized by using modern information systems and technologies [20-24].

Today neural networks are mainly used for the solution of optimization problems at logistics systems functioning as element-wise and in general [25-27]. They are applied at a stage of obtaining forecasting values to the further process organization. However, for their design it is necessary to have a bigger time series and to spend certain time for control (training) of network. At the same time tools based on Petri Nets [28, 29] allow to investigate behavior of an object in the real time mode at the expense of system productivity assessment (a time factor) [30-32] without forecast data.

3 Design of ICT for Grain Delivery Simulation (Research Methodology)

The article purpose is demonstration of applied using Petri Nets as bases of information and communication technologies for an operational decision for problems in logistical systems. In particular, for the choice of the best supply chains variant of grain cargoes from the sender to a seaport.

Creation of information systems demands system thinking. The logistics system structure of the enterprise and material flow providing logistics information systems are interconnected and interdependent. In order that information systems could provide the required efficiency of logistic processes, they should be integrated vertically and horizontally [33].

During the present time, the logistical system at grain cargoes delivery in Ukraine has the structure based on a participants' interaction: traders, as organizers of grain products export, producers, transport companies and, also the existing infrastructure of ports, railroads, highways networks (Fig.3).

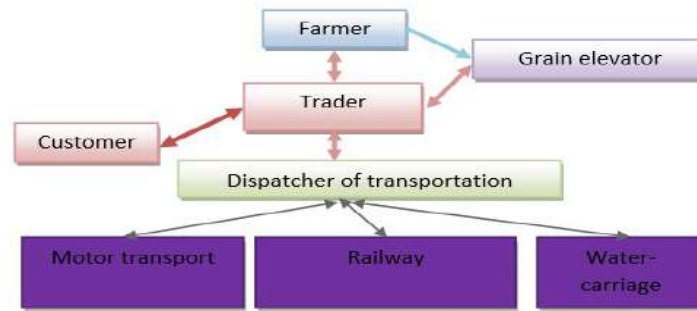


Fig. 3. The scheme of information exchange between participants in the logistical system

Traders in this scheme are the centralized system of necessary technological level which is directed to the organization of processes for participants and, also to receiving and the analysis of the data coming from a system for ensuring bigger transparency [34].

Such complete prospect is able in real time for all participants in a logistical system will allow traders to have enough opportunities for adoption of the relevant short-term and long-term decisions which will fit the organizer purpose and, also the purposes of all process. The system should be designed as two-level (Fig.4): on the top level — there are models of collecting and data processing, lower level — is presented by a simulation model about the functioning of a logistical system.

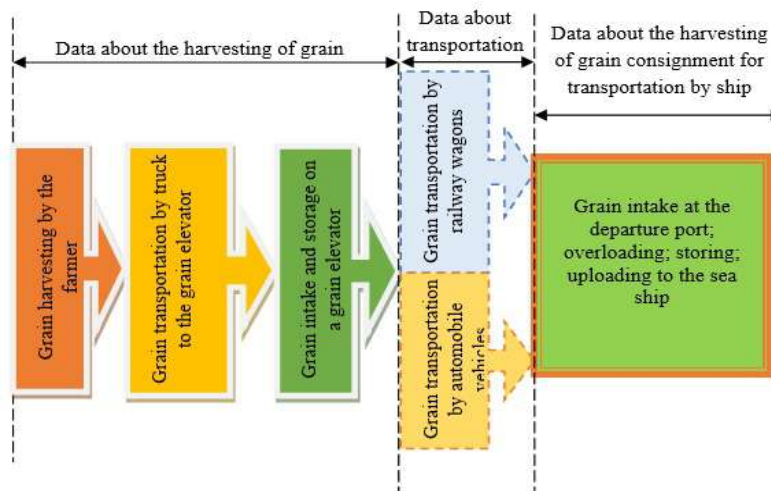


Fig. 4. Two-level representation of a logistical system at grain delivery

For ensuring work of the top level the advanced information technologies, such as technology of intelligent agents, data mining, big data, cloud computing (CC), the Internet of Things (IoT), RFID and Industry 4.0 technology which can be used for providing of data sharing and information on all supply chain can be used. [35-37].

The efficiency of functioning by a logistic system at the top level is defined by a three-element vector condition for forming the database about decision-making I_1, I_2, I_3 (Fig.4):

$$I_n = \{I_1, I_2, I_3\} \quad (1)$$

Each element is characterized by quantitative parameters N^i (mechanisms, vehicles, workers, technical means); by time which spent for performance of each technological operation, considering waiting and idle times – T^i ; by cost of work performance – C^i ; by receipt intensity at physical units (cargo, vehicles, mechanisms) – D^i ; by productivity of resources (technical means, workers) – P^i :

$$I_n = \{N_n^i, T_n^i, C_n^i, D_n^i, P_n^i\}, i = 1, \dots, k, \quad (2)$$

where each value is defined by the resources' quantity involved on i -th stage of technological operation to k -th maximum possible value.

The lower level is presented as a grain delivery process in a logistical system. We will carry out modeling of the considered real object at it. (Fig.5).

Petri Net is the evident and well-formalized behavior model of parallel systems with asynchronous interactions. It reflects the structure of the relationship between system elements and changing dynamics of its states in a compact form under the set entry conditions [38]. Thus, it acts as a static and dynamic model of the object represented with its help that in turn gives the chance to solve a rather wide range of tasks [39]. Petri Nets allow making a simulation in assembly lines [40, 41], to modeling and analysis of technological process [42-43], simulation of resources distribution at the enterprise [44-45].

It is offered to use estimated stochastic Petri Nets for model design in which tops of places have an integer of tags and are used for quantitative assessment by criteria of reliability and productivity for some parameters at systems functioning. In them, the probability of its operation per definite time is attributed to each transition [46].

This scheme (Fig.5) visually displays the formation sequence of the necessary database for grain delivery modeling. In parallel with the preparation of initial information, the choice and tuning are carried out for Petri Nets which will be used for studies process simulation.

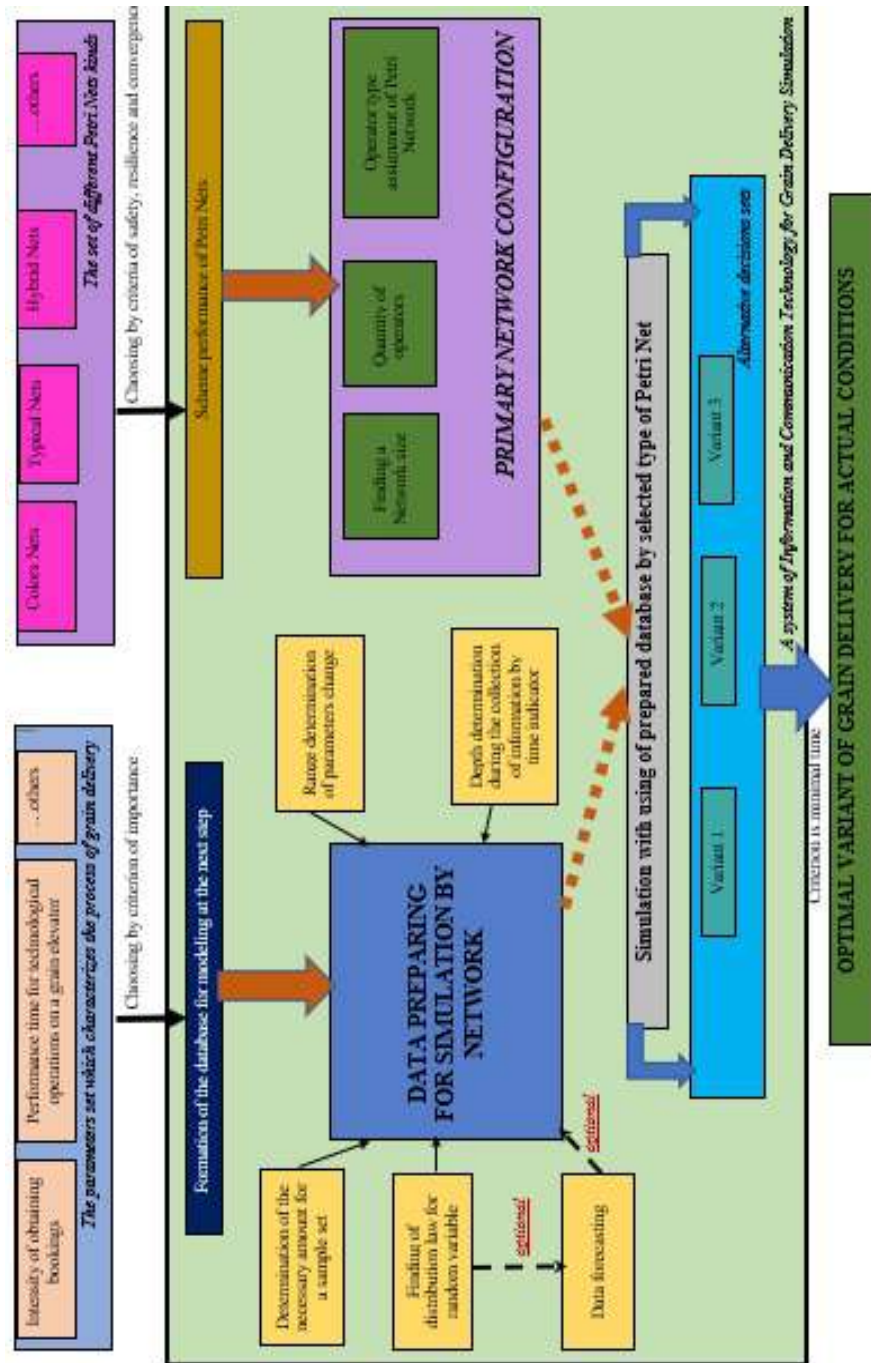


Fig. 5. View of information and communication technology system for grain delivery simulation

Considering functional communications between elements, it is possible to design the integrated model of Petri Net (Fig. 6), describing the functioning of the logistics system. This is a process of grain forwarding from the sender (farmer) to a seaport. The submitted scheme describes the process on the established elements: on transitions: T_1, T_2 – the receipt intensity of requests is defining from the trader and the farmer; T_3, T_5, T_6 – time of cargo shipping by trucks and railway wagons is defining; T_4, T_7, T_8, T_9 – distribution of cargoes, information, technical means is defined; by positions: P_1, P_2 – grain availability at the farmer and orders (technical means) at the trader is defined; $P_3 - P_{11}$ – the grain provision at each stage of operations performed is defined; $P_{12} - P_{14}$ – information availability on the performed work on each element is defined. The key places of the presented model are transitions T_4, T_7, T_8, T_9 and positions $P_{12} - P_{14}$ according to the previous thoughts. Therefore, maximum attention must be paid to these places during further modeling.

So main concentration of complex technological operations occurs on a model fragment $P_7 - P_9 - P_{10}$ for a logistical system of grain cargoes delivery [47].

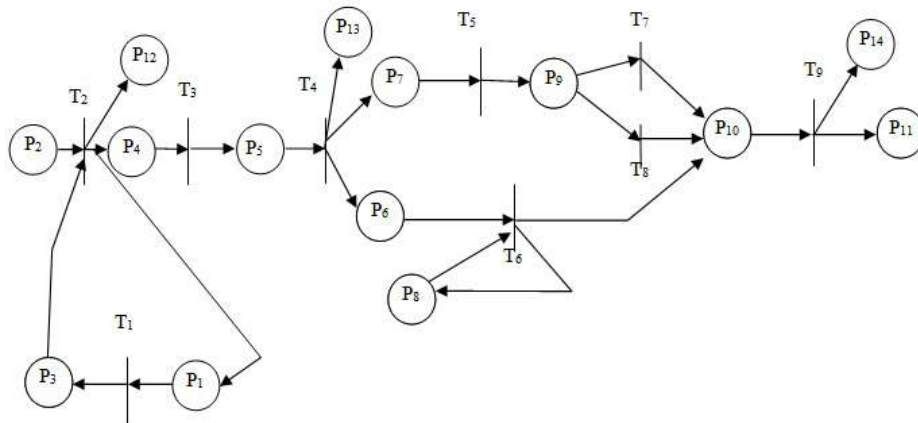


Fig. 6. The primary scheme of grain forwarding from the sender to the seaport

The designed macro-level model is universal for finding correct decisions of problems for this class (management of supply chains). At the same time adaptation of values for certain parameters is required, but the model structure does not change. Also, it is offered in this model to fix information on the performed work on each basic element. It is necessary for decision-making at the top information level. An expense of machine time is insignificant during simulation and its value makes up from two to ten seconds, depending on values of the influencing parameters (order volume). It allows making fast enough decisions when choosing the variant of the organization for cargo delivery in the logistical system.

It should be noted that the full-fledged model is quite bulky and therefore it is perceived hardly. This conditional scheme allows understanding the procedure of modeling considering formal representation of Petri Net which has the next view [38, 48]:

$$PN = \{P, T, F, W, M_0\}, \quad (3)$$

where P – final set of positions; T – final set of transitions; F – set of arcs (flow relations).

The expanded scheme of Petri Nets is provided for process visualization in the micro-level. This figure (Fig. 7) shows an example of a simulation procedure at the destination (cargo port railway station).

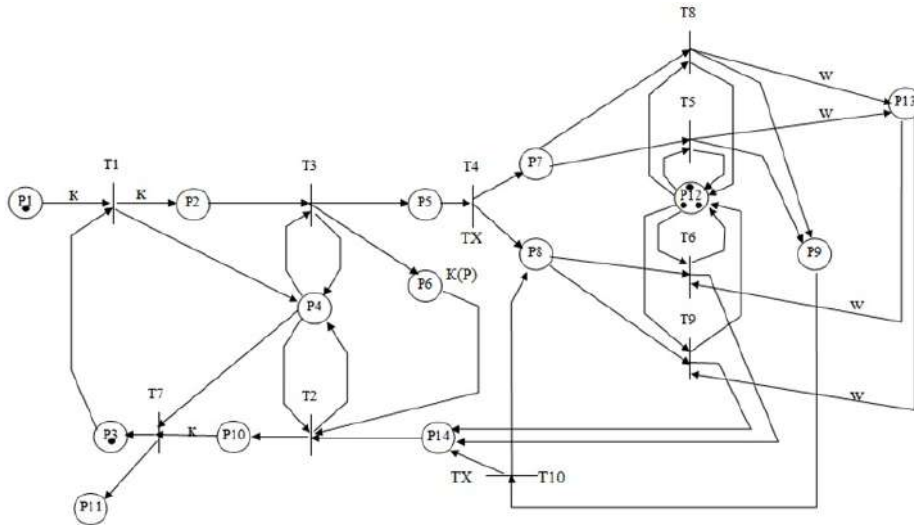


Fig. 7. The primary scheme which shows results of working by cargo-front of the port railway station

The designed networks were analyzed on safety, limitation, integrity, reliability. Assessment and analysis of complex systems by means of Petri Nets expediently to carry out by means of imitation modeling. It is confirmed also by the results of our experiment. At the same time, input flows of orders were set on the first step. They define the corresponding reaction of a system. Output parameters are calculated by means of statistical material processing which is accumulated during modeling.

4 Results and Discussion

Statistical analysis is carried out for parameters which were used for setting of initial values on transitions and positions at process modeling of logistical system operation during grain cargoes delivery.

T_i -indicators were divided into groups by characteristic criterion for statistical values assessment of used parameters which were set in transitions of model (Table.1) [49-51].

Table 1. Statistical assessment of model parameters

Name of parameters group	Volume of sampling	Diapason of parameter expected value	Distribution law
Intensity of obtaining bookings, units/hour	68	0.1-6	Exponential
Admission rate for grain from the field, t/hour	54	190-1200	Gamma-distribution
Performance time of technological operations on the motor transport, hour	28	0.3-4	Normal
Arrival intensity of trucks, unit/hour	58	1-12	Exponential
Performance time of technological operations on railway, hour	87	0.8-4	Normal
Arrival intensity of railway wagons, unit/hour	94	1-20	Exponential
Arrival intensity of railway locomotives, unit/hour	52	2.1-12.9	Exponential
Trains intensity of their arrival and departure, unit/hour	80	0.1-7.3	Exponential
Performance time of technological operations in port, hour	47	0.7-6.8	Normal
Performance time for technological operations on a grain elevator, hour	122	0.4-120	Exponential
Arrival intensity of sea ships to the port, unit/month	26	12-58	Exponential

The offered parameters set used in simulation for delivery time of grain at three alternative supply chains. At the same time, some random variables are distributed on other than the classical definition of distribution laws for these parameters. It can be explained by some formation features for this type of supply chain and forwarding specifics of grain cargoes turnover through them.

As a modeling result, data are obtained on the performance duration for main technological operations on all logistical system. Let's analyze the behavior of model depending on export order volumes of grain (Fig.8). The values amount of cargo turnover is defined proceeding from a confidence level, an error, and observations quantity. It has 68 units. At the same time, cargo turnover accepts values from 20 and to 3370 tons [52]. Three options distribution of cargo turnover are offered when using motor and railway transport: 1st variant – 100% of cargo to be transported by railway; the 2nd variant – equally on each type of transport; the 3rd supply chain – 100% of cargo is transported by the motor transport.

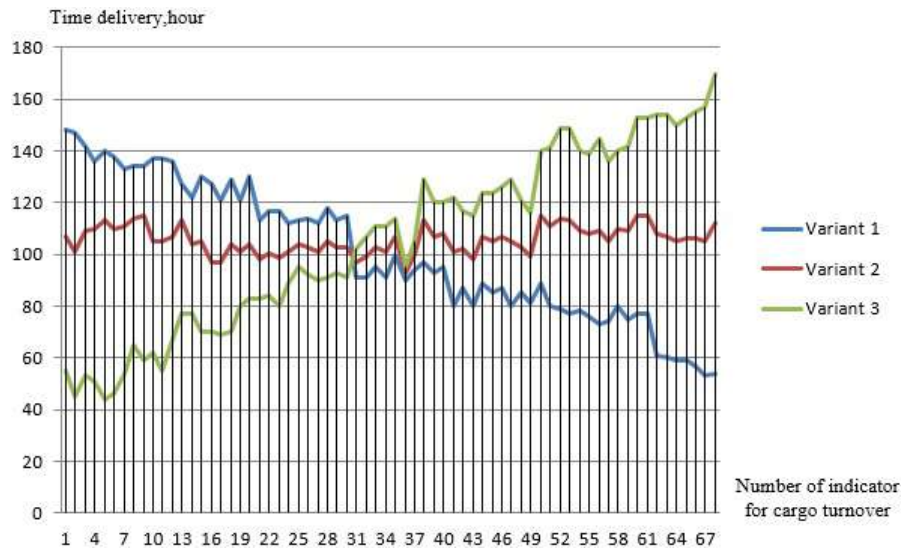


Fig. 8. Simulation results of grain cargoes delivery at logistical system by three variant of supply chains using Petri Net

Histogram (Fig. 8) allows making a fast decision on the best supply chain variant of grain cargoes from the sender to the recipient (port of departure). The choice is carried out by a criterion of the minimum delivery time. At the same time, the trend is accurately traced that to number 31 to which there corresponds cargo turnover near the 350th tons it will be favorable to use the motor transport. It is expedient to carry out delivery from a grain elevator to the port only by rail when volumes of sending grain will grow after 31st values. The alternative of trucks using is unacceptable in this case. It's explained by necessity to have a considerable quantity of trucks that are not favorable to the domestic agricultural enterprises.

It is visible according to the histogram (Fig. 8) that the average delivery time will have the second variant of the supply chain – a combination with two kinds of transport practically for all volumes of cargo turnover. This result of modeling completely shows the reality of the transport process in domestic conditions. The average time of forwarding cargo turnover is explained by the fact that motor transport is used on a small distance when the volumes of consignments are even considerable.

In this case, the application of Petri Net helps indirectly to define the necessary quantity of the rolling stock for delivery of the demanded grain volume to the port. The considered approach allows making the decision on scheme variant not only considering time but still not directly to estimate the possible costs connected with expenses per transportation. This type of expenses will depend in many respects from own carrying opportunities which enterprise has and from a percentage of leased vehicles.

Thus, application of simulation on the models developed by means of Petri Nets allows to track and analyze results at the first stage, and on the second step - to predict and to control a process for grain cargoes forwarding at the logistic system.

5 Conclusion

The successful example of Petri Nets use was shown in research in the form of ICT system for logistic problems decision. This modeling instrument allows fast receiving values of time about grain forwarding through a supply chain. It considerably will facilitate the primary choice of a supply chain.

The analysis of the theoretical basis of Petri Nets is made and its adequacy of their application for transport processes modeling is proved. It is shown that Petri Nets allow to investigate more deeply behavior of the modeled system and to obtain information on its most important characteristics. They also take into accounting that logistical systems of grain delivery must be considered as the systems of parallel information processing and in parallel operating of objects. Therefore, the offered approach is universal for the problems' decision of a similar class.

An advantage of Petri Nets application consists of small labor input when developing and an opportunity to correctly simulate process at a small amount of initial information.

The developed visual scheme for the system of information and communication technologies when modeling options of transportation of grain cargo-turnover based on Petri Nets. It allows unifying staging of the choice for a type of mathematical tools and in parallel to format the initial database.

That's why using of such ICT considerably will accelerate the adoption for correct decision on the choice of a supply chain that minimizes expenses of the logistic companies when exporting the Ukrainian grain.

References

1. Wang, F., Xu, Z., Zhen, T., Zhang, X., Zhang, M.: Grain Logistics Management Information System Based on Short Message Service Technology. *Mechatronics and Automatic Control Systems. Lecture Notes in Electrical Engineering*. vol. 237. Springer, Cham. (2014). doi: 10.1007/978-3-319-01273-5_62
2. Shramenko, N.Y., Shramenko, V.O.: Mathematical model of the logistics chain for the delivery of bulk cargo by rail transport, *Scientific Bulletin of National Mining University*, Vol. 5, No. 167, 136-141 (2018). doi: 10.29202/nvngu/2018-5/15
3. Cao, R., Hao, L., Wang, F., Gao, Q.: Modelling and analysis of hybrid stochastic timed Petri net, *Journal of Control and Decision*. Vol. 6:2, 90-110. (2019) doi: 10.1080/23307706.2017.1419079
4. Peraković, D., Periša, M., Sente, R.E.: Information and Communication Technologies Within Industry 4.0 Concept. *Advances in Design, Simulation and Manufacturing. DSMIE. Lecture Notes in Mechanical Engineering*, pp. 127-134, Springer, Cham (2018) doi: 10.1007/978-3-319-93587-4_14
5. Fernando, E., Surjandy, Harco Leslie Hendric Spits Warnars, Meyliana, Kosala, R.; Abdurachman, E.: Critical Success Factor of Information Technology Implementation in Supply Chain Management: Literature Review 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). IEEE. (2018) doi: 10.1109/ICITACEE.2018.8576979

6. Zelikov, V.A., Akopova, E.S., Pilivanova, E.K., Popova, L.K.: Model of Management of the Risk Component of Intermodal Transport: Information and Communication Technologies of Transport Logistics. Perspectives on the Use of New Information and Communication Technology (ICT) in the Modern Economy. ISC 2017. Advances in Intelligent Systems and Computing. Vol. 726. Springer, Cham. (2019). doi: 10.1007/978-3-319-90835-9_77
7. Shramenko, N.Y.: Methodology for evaluation of synergy effect in terminal cargo delivery system. Actual Problems of Economics, Vol. 8, No. 182, 439-444 (2016)
8. Ukrainian Grain Exports, <http://minagro.gov.ua/>, last accessed 2019/02/19
9. Muzylyov, D.A., Kravcov, A.G., Karnayh, N.V., Berezhnaja, N.G. & Kutiya, O.V.: Development of a Methodology for Choosing Conditions of Interaction Between Harvesting and Transport Complexes. Eastern European Journal of Enterprise Technologies, Vol. 2, No.3, 11-21 (2016). doi: 10.15587/1729-4061.2016.65670.8.
10. Shramenko, N.Y.: Effect of process-dependent parameters of the handling-and-storage facility operation on the cargo handling cost. Eastern European Journal of Enterprise Technologies, Vol. 5/3, No. 77, 43-47 (2015). doi: 10.15587/1729-4061.2015.51396.
11. Shramenko, N. Y.: The methodological aspect of the study feasibility of intermodal technology of cargo delivery in international traffic. Scientific Bulletin of National Mining University, Vol. 4, No. 160, 145-150 (2017).
12. Ukrainian Ports Turnover in 2018: Main Trends, ports.com.ua/articles/gruzooborot-portov-ukrainy-2018-glavnye-trendy, last accessed 2019/03/21
13. Shramenko, N., Muzylyov, D., Karnaukh, M.: The Principles of the Choice of Management Decisions Based on Fuzzy Logic for Cargo Delivery of Grain to the Seaport. International Journal of Engineering & Technology, Vol. 7, No. 4.3, 211-216 (2018). <https://doi.org/10.14419/ijet.v7i4.3.19789>.
14. Kirichenko, A.I. The Application of Information Technologies in the Management of Cargo Delivery Processes. Transport Problems: Collection of Scientific Papers, Vol. 9, NTU, 17-27 (2012).
15. Velykodnyi, D., Pavlenko, O.: The choice of rational technology of delivery of grain cargoes in the containers in the international traffic. International journal for traffic and transport engineering, Vol. 7, No. 2, 164-175 (2017). doi: 10.7708/ijtte.2017.7(2).02
16. Shramenko, N. Y.: Evaluation of the effectiveness of piggyback traffic in the context of creating transport and logistics clusters. Scientific Bulletin of National Mining University, Vol. 6, No. 162, 151-155 (2017)
17. Xu, X.Q.: Research on Applied Technology with Development of Modern Computer Simulation Technology, Advanced Materials Research. Vol. 908, 563-566 (2014). doi: 10.4028/www.scientific.net/AMR.908.563
18. Konovalenko, I., Ludwig, A.: Event processing in supply chain management – The status quo and research outlook. Computers in Industry. Vol. 105, 229-249 (2019). doi: 10.1016/j.compind.2018.12.009
19. Abed, M., Charfeddine, I., Benaissa, M., Starostka-Patyk, M.: Intelligent Traceability System of Containerized Goods, Applied Mechanics and Materials. Vol. 309, 241-251 (2013). doi: 10.4028/www.scientific.net/AMM.309.241
20. Li, F., Zhu, Y.P., Wu, H.R.: Modeling and Optimization of Traceability System for Agriculture Products Supply Chain, Advanced Materials Research. Vols. 605-607, 574-579 (2013). doi: 10.4028/www.scientific.net/AMR.605-607.574
21. Qu, J.H., Yao, X. S., Ying, J.L.: Agricultural Products Logistics Operational Pattern Based on Information Center, Advanced Materials Research. Vols. 361-363, 1679-1683, (2012) doi: 10.4028/www.scientific.net/AMR.361-363.1679

22. Vendrell-Herrero, F., Bustinza, O. F., Parry, G., Georgantzis, N.: Servitization, digitization and supply chain interdependency. *Ind. Mark. Manag.* Vol. 60, 69-81 (2017). doi: 10.1016/j.indmarman.2016.06.013
23. Xue, L., Zhang, C., Ling, H., Zhao X.: Risk mitigation in supply chain digitization: system modularity and information technology governance. *J. Manag. Inf. Syst.* Vol. 30, 325-352 (2013). doi: 10.2753/MIS0742-1222300110
24. Aulin, V., Lyashuk, O., Pavlenko, O., Velykodnyi, D., Hrynkyv, A., Lysenko, S., Holub, D., Vovk, Y., Dzyura, V., Sokol, M.: Realization of logistic approach in the international cargo delivery system. *Communications - Scientific Letters of the University of Zilina.* Vol. 21(2), 5-14, (2019).
25. Jin, C.F., Yang, H.M., Ling, L.: Research on Optimization and Debugging Simulation Model of Logistics Center Based on Neural Network, *Applied Mechanics and Materials.* Vols. 37-38, 1060-1063 (2010). doi: 10.4028/www.scientific.net/AMM.37-38.1060
26. He, W., Lu, T., Yu, C.Q.: A Novel Traffic Flow Forecasting Method Based on the Artificial Neural Networks and Intelligent Transportation Systems Data Mining, *Advanced Materials Research.* Vol. 842, 708-711 (2014). doi: 10.4028/www.scientific.net/AMR.842.708
27. Subbotin, S.O., Oliynyk A.O.: *Neural Networks: Teach. manual.* Zaporizhzhya, ZNTU (2014).
28. Wang, L., Zhu, X.N., Xie, Z.Y.: Object-Oriented Petri Net Modeling and Analysis of China Railway Container Freight Yard Logistic System, *Key Engineering Materials.* Vols. 467-469, 990-995 (2011). doi: 10.4028/www.scientific.net/KEM.467-469.990
29. Zhong, W.Z., Fu, X.Q., Wang, Y.P.: Petri Net Modeling: Container Terminal Production Operation Processing System Analysis, *Applied Mechanics and Materials.* Vols. 409-410, 1320-1324 (2013). doi: 10.4028/www.scientific.net/AMM.409-410.1320
30. Zhang, G.X., He, S., Zhang, Z.Y.: Study on Model Validation of Logistics System Based on Petri Nets, *Applied Mechanics and Materials.* Vols. 195-196, 853-858 (2012). doi: 10.4028/www.scientific.net/AMM.195-196.853
31. Gaied, M., M'halla, A., Lefebvre D.: Robust control for railway transport networks based on stochastic P-timed Petri net models, *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, (2019). doi: 10.1177/0959651818823583
32. Zhong, W.Z., Fu, X.Q., Wang, Y.P.: LOPN-Based BPR of Container Terminal Production Controlling System, *Applied Mechanics and Materials.* Vol. 367, 441-446 (2013). doi: 10.4028/www.scientific.net/AMM.367.441
33. Perdana, Y.R.: Logistics Information System for Supply Chain of Agricultural Commodity *Procedia - Social and Behavioral Sciences.* Vol. 65(3), 608-613 (2012). doi: 10.1016/j.sbspro.2012.11.172
34. Trzuskawska-Grzeńska, A.: Control towers in supply chain management – past and future, *Journal of Economics & Management.* Vol. 27, 114-133 (2017). doi: 10.22367/jem.2017.27.07
35. Marinagi, C., Skourlas, C., Galiotou, E.: Advanced information technology solutions for implementing information sharing across supply chains. In: *22nd Pan-Hellenic Conference on Informatics*, pp. 99-102 (2018). doi: 10.1145/3291533.3291575
36. Borgi, T., Zoghalmi, N., Abed, M., Saber Naceur, M.: Big Data for Operational Efficiency of Transport and Logistics: A Review. In: *6th IEEE International Conference on Advanced Logistics and Transport (ICALT)* (2017). doi: 10.1109/ICAdLT.2017.8547029

37. Yaqiong, L.V., Lei, T.U., Lee, C.K.M, Xin, T.: IoT based Omni-Channel Logistics Service in Industry 4.0. In: International Conference on Service Operations and Logistics, and Informatics (SOLI). IEEE. pp. 240-243 (2018). <https://doi.org/10.1109/SOLI.2018.8476708>
38. Chinnusamy, T.R., Karthikeyan, T., Krishnan, M., Murugesan, A.: A Comprehensive Survey of Flexible Manufacturing System Scheduling Using Petri Nets, *Advanced Materials Research*. Vols. 984-985, 111-117 (2014). doi: 10.4028/www.scientific.net/AMR.984-985.111
39. Zhang, Y.H., Zhang, X.F.: Modeling and Simulation Method of Component Based Petri Net, *Advanced Materials Research*. Vol. 936, 2254-2258 (2014). doi: 10.4028/www.scientific.net/AMR.936.2254
40. Dong, H., Xin, H.B., Yan J.H.: Modeling in Mixed-Model Assembly Lines Based on Object-Oriented Petri Net, *Advanced Materials Research*. Vols. 945-949, 3156-3159 (2014). doi: 10.4028/www.scientific.net/AMR.945-949.3156
41. Tian, S.L., Tian, Y., Xiao, X. H., Chen, D.X., Wang, T.Y., Zhang, L.: Defining Method Research of Technological Conflict Based on Petri Net Theory, *Key Engineering Materials*. Vol. 693, 2016-2021 (2016). doi: 10.4028/www.scientific.net/KEM.693.2016
42. Culea, G., Cristea, I., Puiu, P.G.: Modeling and Analysis of Processing Flow in a Flexible Manufacturing System Using Synchronized Petri Nets, *Applied Mechanics and Materials*. Vols. 809-810, 1474-1479 (2015). doi: 10.4028/www.scientific.net/AMM.809-810.1474
43. Prayoga, P.H.N., Shinoda, T.: A Petri Net Model and its Simulation for Straddle Carrier Direct-System Operation in a Container Terminal, *Applied Mechanics and Materials*. Vol. 862, 202-207 (2017). doi: 10.4028/www.scientific.net/AMM.862.202
44. Klosowski, G., Gola, A., Świć, A.: Human Resource Selection for Manufacturing System Using Petri Nets, *Applied Mechanics and Materials*. Vol. 791, 132-140 (2015). doi: 10.4028/www.scientific.net/AMM.791.132
45. Liu, B.W.: A New Deadlock Avoidance Policy Based on Petri Nets in Distributed System, *Applied Mechanics and Materials*. Vol. 743, 575-580 (2015). doi: 10.4028/www.scientific.net/AMM.743.575
46. Prostitenko, O.V., Khalimon, V.I., Rogov, A.Y.: Simulation of discrete systems based on Petri nets: study guide. St. Petersburg, SPbGTI (TU) (2017).
47. Lomot'ko, D.V. Arsenenko, D.V.: Creation of an effective technology for the formation of railway tracks with grain cargoes using models on the Petri network. *Information management system on the railway transport*. Vol. 6, 38-45 (2018). doi: 10.18664/iksz.voi6.151642
48. Appel, M., Konigorski, U., Walther M.: A Graph Metric for Model Predictive Control of Petri Nets. *Science Direct, IFAC Papers On-Line* 51-2. pp. 254-265 (2018). doi: 10.1016/j.ifacol.2018.03.044
49. Naumov, V.S.: Forwarding services in logistics systems: Monograph / 2nd ed. – Kharkov, HNADU (2014).
50. Murad'yan, A.O.: Coordinated management of the process of handling of goods in the gantry hubs: monograph. Odessa, Kuprienko S.V. (2017).
51. The unified technological process of the work of the State Enterprise “Odessa Commercial Sea Port” and the Odessa-Port Port of the Odessa Railway, <http://www.port.odessa.ua>, last accessed 2019/03/14
52. Kernel-Trade Company Presentation, https://www.kernel.ua/wp-content/uploads/2019/02/Kernel_FY2019_H1_Presentation.pdf, last accessed 2019/03/29

Analysis of the Verification Approaches for the Cyber-Physical Systems

Sergiy Korotunov^{1[0000-0001-5184-4455]}, Galyna Tabunshchik^{2[000-0003-1429-5180]},
Karsten Henke^{3[0000-0001-5424-9053]}, Dieter Wuttke^{4[0000-0001-8030-647X]}

¹²Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063,
Ukraine

skorotunov@yahoo.com, galina.tabunshchik@gmail.com

³⁴Ilmenau University of Technology, Max-Planck-Ring str., 14, Ilmenau, 98693, Germany
karsten.henke@tu-ilmenau.de, dieter.wuttke@tu-ilmenau.de

Abstract. In the paper the possibility of utilization of the Kripke structures for applying linear-time temporal logic in problems of verification of reactive systems is considered. The definition and main characteristics of the cyber-physical systems, finite state machines, Kripke structures and temporal logics are considered by the authors. Example of modeling on the base of GOLDi is provided.

Keywords. Verification, parallel program, reactive system, cyber-physical systems, Kripke structure, finite-state machine, temporal logic.

1 Introduction

With the high complexity of modern computer systems (millions of lines of code and billions of transistors), it is impossible to completely avoid errors. It is not only related to the application software developed in a short time in a competitive market environment, but also to critical components, such as hardware (I/O devices, microprocessors) and software (compilers, operating systems). It is obvious that the systems, in the design and implementation of which mistakes are made, can in one or another situation behave in unpredictable ways and lead to serious consequences. The literature describes many cases of this kind. As vivid examples following can be represented.

In 1996, the launch vehicle Ariane 5, developed by the European Space Agency, exploded at the 39th second after launch. As it turned out later, the onboard system partially used the software of the previous version of the rocket, the Ariane 4, in particular, the inertial navigation system control module [1]. When converting from a 64-bit floating-point number representing the inclination of the rocket to a 16-bit integer, an overflow occurred that was not handled by the module. When developing the module, it was assumed that overflow was impossible due to the physical limitations of the rocket, but new engines which were used in the Ariane 5 surpassed those limitations.

Another notable example is the case of the medical device Therac-25 [2], which was launched in 1982 by the Canadian firm Atomic Energy of Canada Limited. The device was intended for radiation therapy – irradiation of a cancer. In some situations, the irradiation intensity increased by 2 orders or more due to a number of mistakes made in the design and implementation of the built-in control system. At least two patients died, and several people were disabled during the operation of the device (the period from June 1985 to January 1987).

It is important to understand that errors in computer systems are not exceptional. According to statistics, the average number of errors per thousand lines of un-debugged code ranges in $10\text{-}50^5$ [3]. Moreover, there is a tendency to the degradation of design quality (apparently, this is a consequence of the increasing complexity of systems and optimization of the costs of their creation) [4]. At the same time, our life is increasingly dependent on computers. Therefore, ensuring the correctness and reliability of computer systems (in other words, verification) is becoming increasingly important.

2 Program verification

2.1 Verification approaches

Verification is the process of checking of the compliance of a program (its model) with the requirements placed on it. If the program meets the requirements, it is called correct. Otherwise, the program is called incorrect, and the fact of non-compliance of the program with the requirements is an error. Thus, verification is the analysis of the program for the presence or absence of errors in it. An indefinite verdict is also possible when errors are not found, but their absence has not been proven.

It is necessary to take into account many nuances. First, the requirements, as a rule, are formulated informally, in natural language, so it is not always possible to unambiguously determine whether the program corresponds to them or not. Secondly, proving the absence of errors in the program is extremely difficult from a mathematical point of view (because of the fact that this task is not fully amenable to automation).

Verification methods can be divided into three main groups [5]:

- formal methods, which utilize mathematically rigorous analysis of the program model and requirements model;
- testing methods that verify the actual behavior of the program on a certain set of scenarios;
- expertise performed by people based on their knowledge and experience directly on the design results (for example, code inspection).

Each of the specified groups of methods has its advantages and disadvantages, each has its own area of applicability. Full verification of complex systems of responsible purpose is impossible without the joint use of different approaches. This work focuses on mostly formal methods of verification programs.

2.2 Formal verification

Formal verification is based on mathematical (logical) modeling programs and requirements for them. The idea is exactly the same as when using models in other areas of knowledge:

- a model is created – an idealized description of the object or phenomenon under study;
- the model is investigated using mathematical methods;
- research results are transferred to a real object or phenomenon.

Of course, the applicability of this approach is determined by the models used – it is necessary to clearly understand the conditions for their adequacy.

The general process of formal verification [6] is shown in Fig. 1:

- a formal program model is created;
- a formal requirements model is created;
- the compliance of the program model with the requirements model is formally verified;
- based on the results of the test, it is concluded that the real program does not meet the real requirements (in other words, that there are no or no errors in the program).

For the presentation of program models and requirements models, the languages of the formal specification of programs (modeling languages) and the languages of the formal specification of requirements are used respectively.

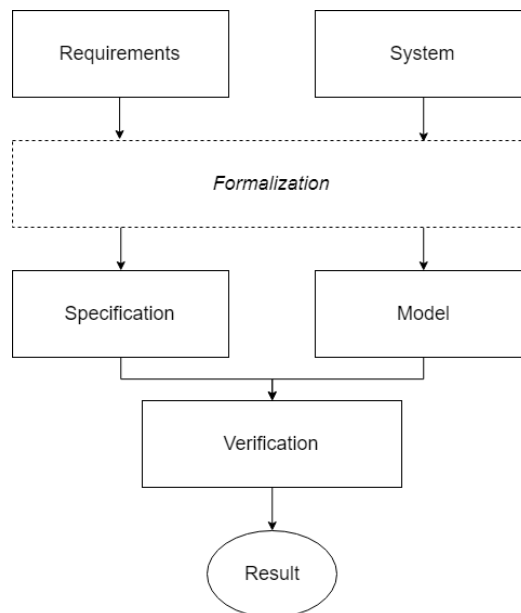


Fig. 1. Formal verification process

Formal verification methods differ in the types of program models (state machines, Petri nets, labeled transition systems), types of requirements models (software contracts, production rules, temporal statements), correspondence relations (equivalence, simulation, route inclusion) and compliance verification techniques (state space research, symbolic analysis, deductive inference). In this paper we briefly consider one of the most popular formal verification methods – model checking.

2.3 Model checking

In model checking, requirements for programs are represented by logical formulas of a certain type, and programs themselves are structures interpreting formulas of this type; the correspondence relation is the truth of a formula on a structure (a program satisfies the requirements if and only if the corresponding structure is a model of the corresponding formula).

Often, temporal logics are used to represent the requirements – logic that allows one to define the relationship of events in time, for example, Linear-time Temporal Logic or Computational Tree Logic. Accordingly, Kripke structures and related formalisms (marked transition systems) are used to represent programs.

Kripke structures are used to model reactive systems – systems operating in an “infinite loop” and interacting with their environment. As an example of such systems cyber-physical system can be seen. The behavior of the system is modeled by calculation in the Kripke structure [7].

Propositional temporal logic can be used to describe requirements, in particular LTL. The main method for establishing the correctness or incorrectness of a model is to search in the state space (bypassing the state graph).

3 Finite state machines

A finite state machine (FSM) is a computation model that can be implemented with hardware or software and can be used to simulate sequential logic (for example computer program) [8]. Finite state machine generates regular languages. Finite state machines can be used to model problems in many fields including mathematics, artificial intelligence, games, and linguistics.

3.1 Deterministic and non-deterministic finite state machines

There are two types of finite state machines: deterministic finite state machines (deterministic finite automaton) and non-deterministic finite state machines (nondeterministic finite automata).

A deterministic finite automaton (DFA) is described by a five-element tuple:

$$\langle Q, \Sigma, \delta, q_0, F \rangle. \quad (1)$$

Where Q is a finite set of states, Σ is a finite, nonempty input alphabet, δ is a series of transition functions, q_0 is the starting state, F is the set of accepting states.

There must be exactly one transition function for every input symbol in Σ from each state. DFAs can be represented by following diagrams:

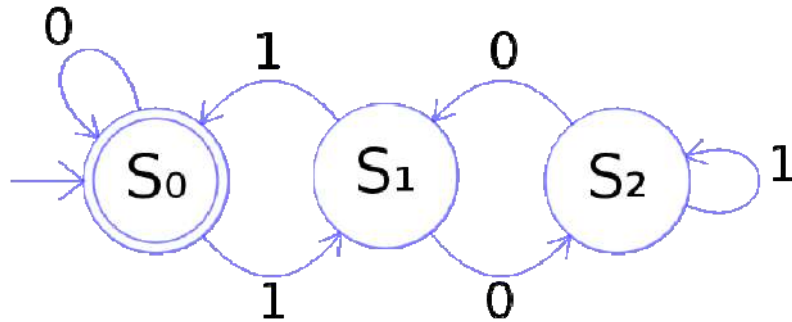


Fig. 2. Deterministic finite automaton diagram

Similar to a DFA, a nondeterministic finite automaton (NFA) is described by an above five-element tuple. Unlike DFAs, NFAs are not required to have transition functions for every symbol in Σ , and there can be multiple transition functions in the same state for the same symbol. Additionally, NFAs can use null transitions, which are indicated by ϵ . Null transitions allow the machine to step from one state to another without having to read a symbol. NFAs can be represented by diagrams of this form:

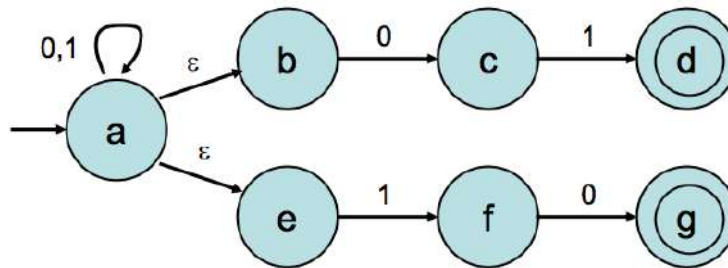


Fig. 3. Nondeterministic finite automata diagram

One might assume that NFAs can solve problems that DFAs cannot, but NFAs are just as powerful as DFAs. However, a DFA will require many more states and transitions than an NFA would take to solve the same problem. Converting from DFA to NFA and vice versa is possible which makes them equivalent.

3.2 Finite state machine example from the real world

A system where particular inputs cause particular changes in state can be represented using finite state machines. This example describes the various states of a turnstile. Inserting a coin into a turnstile will unlock it, and after the turnstile has been pushed, it locks again. Inserting a coin into an unlocked turnstile, or pushing against a locked turnstile, will not change its state. Diagram of the turnstile FSM is following:

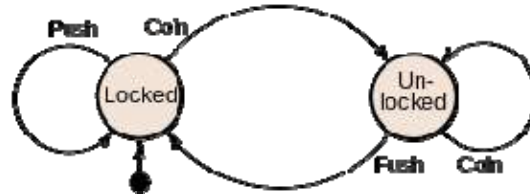


Fig. 4. Turnstile FSM diagram

4 Kripke structure

4.1 Definition of the Kripke structure

Saul Aaron Kripke introduced structures named after him to the logical and philosophical use in the late 1950s [9]. In general terms, the Kripke structure is a system of possible worlds and transitions between them: each world is static and interpreted in some traditional way. Its diagram can be seen at Fig. 5. The Kripke structure is the quadruple $\langle S, S_0, R, L \rangle$, where S is the set of states, $S_0 \subseteq S$ is the set of initial states, $R \subseteq S \times S$ is the transition relation, $L: S \rightarrow 2^{AP}$ is a labeling function which marks each state of the structure with a set of atomic propositions.

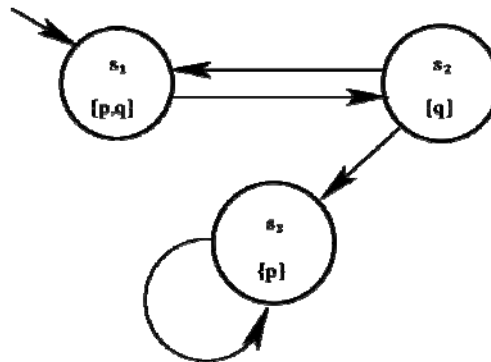


Fig. 5. Kripke structure diagram

A Kripke structure is an annotated finite-state transition graph.

4.2 Kripke structure example from the real world

Let a set of atomic propositions have the form $\{locked, lock, unlock\}$. Consider the Kripke structure $M = \langle \{S_0, S_1, S_2, S_3\}, \{S_0\}, R, L \rangle$, which simulates a door with a lock which can be seen at Fig. 6. The initial state is marked by an incoming arrow; next to each state there are statements which are true in it. The states of the structure are labeled as follows:

- $L(S_0) = \emptyset$ – the door is open; no action is taken on the lock;
- $L(S_1) = \{lock\}$ – the door is open; it is being closed;
- $L(S_2) = \{locked\}$ – the door is closed, no action is taken with the lock;
- $L(S_3) = \{locked, unlock\}$ – the door is closed; it is being opened.

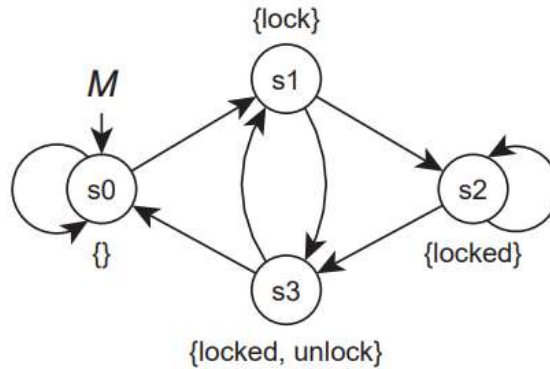


Fig. 6. Kripke structure that simulates a door lock

5 Temporal logic and its scope

5.1 Cyber-physical systems and their properties

A parallel program is a finite set of sequential programs over a common set of variables [10]. A separate program of this set is called a process. Programs running in the "infinite loop" are considered in this paper. These are the so-called reactive systems.

Such systems respond to environmental events by performing certain actions in response. This is an extensive class of programs, including cyber-physical systems, operating systems, device drivers, telecommunication environments, control systems, etc. [11].

Cyber-physical systems are the systems that provide the integration of computing, physical processes and networks, or as systems where software and physical subsystems are closely bounded, each of which works in a variety of temporal and spatial dimensions, demonstrating clear and multiple behavioral patterns, and interacts in a variety of ways [12]. Modern trends in productivity and complexity of requirements for systems use require fundamentally new design approaches in which cybernetic and physical components are integrated at different stages.

In general, the qualitative properties of cyber-physical systems can be classified into the following two broad categories:

- reachability or guarantee properties that raise the question of whether a system can achieve a configuration that satisfies a particular property;
- security properties that raise the question of whether the system can remain forever in configurations that satisfy a particular property.

The main properties of cyber-physical systems include following [13]:

- high degree of automation,
- reorganization / reconfiguration of the dynamics,
- cybernetic capabilities in each physical component,
- the ability of networks to work on multiple scales,
- integration on different time and spatial scales.

The behavior of cyber-physical systems is described in terms of sequences of events distributed in time. So-called temporal logics are often used for the specification of requirements for cyber-physical systems. Temporal logics are formal languages that allow to define the interrelationships of events in time: causal relationships, restrictions on the relative order, the magnitude of delays between events, etc. The following examples can be cited as temporal properties: the system always works without freezing; two users cannot simultaneously access shared data; a request with a higher weight will be processed before constipation with a lower weight.

5.2 Linear-time Temporal Logic

Among number of temporal logics, two have become particularly popular in computer science: Linear-time Temporal Logic (LTL) [14] and Computation Tree Logic (CTL) [15]. In this paper LTL is considered.

In LTL, two temporal operators are added to the syntax of classical propositional logic: the unary operator X (next time) and the binary operator U (until). These two operators form the LTL temporal basis.

Let a set of elementary statements AP (Atomic Propositions) be given. The LTL formula is given by the following grammar [16]:

$$\varphi := p \mid \varphi \vee \varphi \mid \neg \varphi \mid X \varphi \mid \varphi U \varphi. \quad (2)$$

Here, p is an arbitrary elementary statement from the AP set. For convenience, in LTL formulas it is possible to use:

- derived logical connectives, for example, $\wedge$ and $\rightarrow$;
- logical constants, *true* and *false*;
- derived temporal operators, for example F (in the Future), G (Globally), W (Weak until) and R (Release).

Temporal operators can be interpreted as following.

- The formula φ is true at the next point in time – $X\varphi$.
- The formula ψ is true now or will surely become true in the future, but up to this point (not inclusive) the formula φ should be true – $\varphi U \psi$.
- The formula φ is true now or will become true sometime in the future – $true U \varphi$ which is also can be displayed as $F\varphi$.
- The formula $\neg\varphi$ is false now and will never become true in the future (always, from now on, the formula φ is true) – $\neg F\neg\varphi$ which is also represented as $G\varphi$.
- The formula φ is true until the formula ψ is not become true, without requiring the formula ψ to ever become true in the future – $(\varphi U \psi) \vee G\varphi$ or $\varphi W \psi$.
- The formula ψ is true until the moment (inclusive) when φ becomes true the first time; if such a moment never comes, the formula ψ is always true (it possible to say that φ frees ψ) – $\neg(\varphi U \neg\psi)$ or $\varphi R \psi$.

Thus, the requirements for the system, which simulates a door lock and was described above, can be expressed in LTL logic:

- $G\{\neg(lock \wedge unlock)\}$ – for system it is impossible to open and close door at the same time;
- $G\{\neg(locked \wedge lock)\}$ – system should never close the closed door;
- $G\{lock \rightarrow X(locked)\}$ – if the door is closed, at the next moment in time it will become closed.

6 Experimental modeling on the base of GOLDi

Integrated Communication Systems Group at the Ilmenau University of Technology has many years of experience in integrated hard- and software systems and over 10 years of experience in dealing with Internet-supported teaching in the field of digital system design [17]. Grid of Online Lab Devices Ilmenau (GOLDi) gives the students the possibility to work on real physical systems without the need to stand in line at a lab or the need to take care of opening hours and offers the students a working environment that is as close as possible to a real world laboratory. Under real laboratory conditions disturbances can appear and lead to failures of the control algorithm that cannot be detected under virtual lab conditions.

Online laboratories offer various features like visualization and animation, which allows to observe and to test all the properties of the design. In connection with for-

mal design techniques, simulation and prototyping are used to establish a foundation for the development of a reliable system design. To check the functionality of the whole design, some special simulation and validation features are included as integral part of the GOLDI system. This offers various possibilities for the execution of simulations [18], such as:

- usage of simulation models of the physical system for visual prototyping,
- step by step and parallel execution of these prototypes,
- visualization of the simulation process with the tools also used for specification,
- features for test pattern generation and
- code generation for hardware and software synthesis.

GOLDI offers a Web-based environment supporting the above mentioned features to generate and execute a design by using simulation models.

As an example of modeling it was decided to create Kripke structure of the elevator which is located in the GOLDI. This elevator has ability to move upwards and downwards from floor to floor and open or close its door.

The atomic propositions for the Kripke structure representing the elevator are as follows:

- 1st – elevator is located at the 1st floor;
- 2nd – elevator is located at the 2nd floor;
- DO – door is open;
- MU – elevator is moving in the upward direction;
- MD – elevator is moving in the downward direction.

For clarity, each state is labeled with both the atomic propositions that are true in the state and the negations of the propositions that are false in the state. The labels on the arcs indicate the actions that cause transitions and are not part of the Kripke structure. Kripke structure of the elevator can be seen at Fig.7.

This model can be used for further formal verification. For example one might want to determine that “door of the elevator is closed and it is moving upwards”. S_0 , $p = \neg 1st \wedge \neg 2nd \wedge DO \wedge MU \wedge \neg MD$. Using Kripke structure this can be determined.

7 Conclusions

In the paper the authors considered the possibility of utilization of the Kripke structures for applying linear-time temporal logic in problems of verification of reactive systems. Review of the main characteristics of the cyber-physical systems, finite state machines, Kripke structures and temporal logics was carried out by the authors. Example of modeling on the base of GOLDI was provided in the paper.

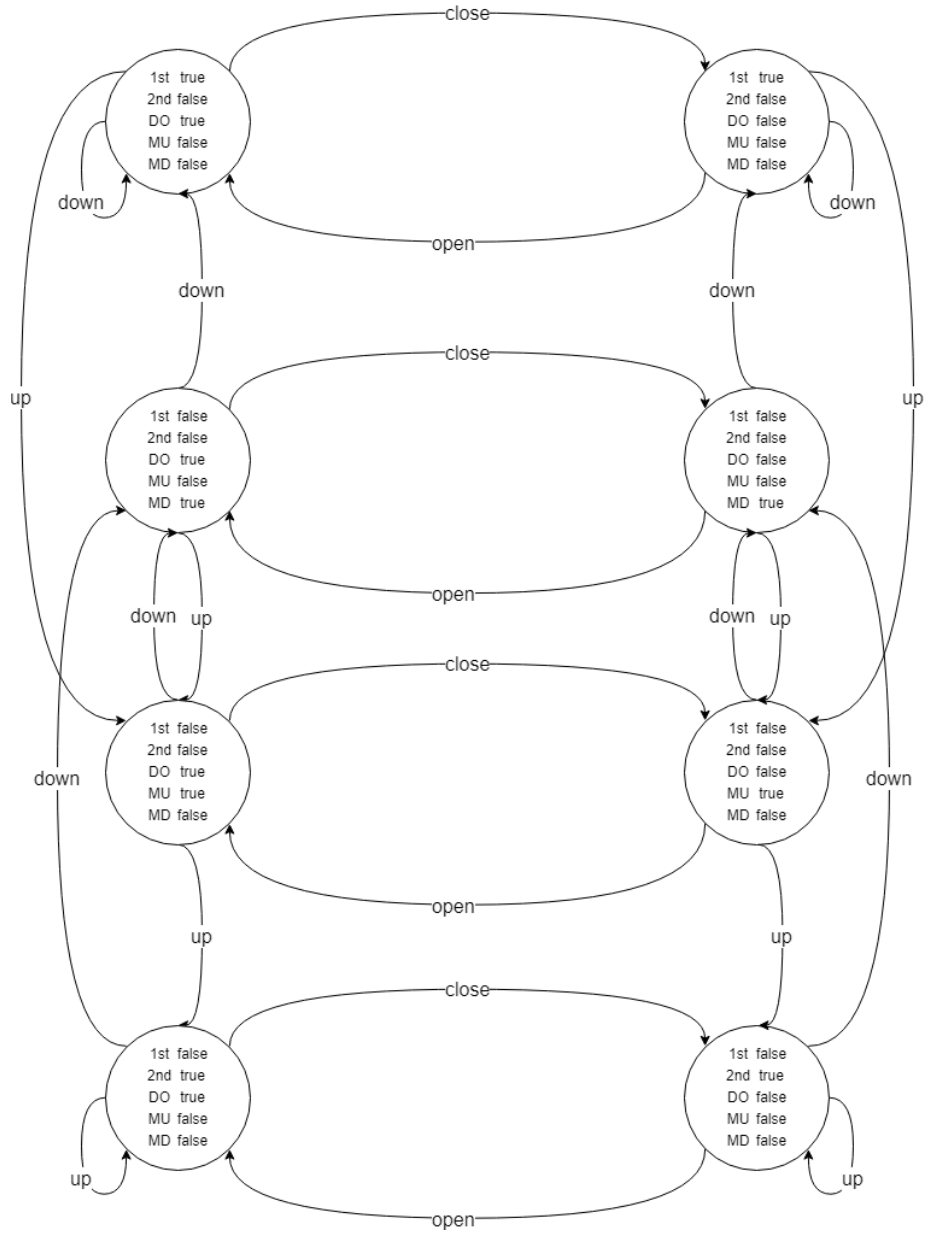


Fig. 7. Kripke structure of the elevator in GOLDi

8 References

1. Garfinkel, S., Goode, L., Barrett, B., Pardes, A.: History's Worst Software Bugs, <https://www.wired.com/2005/11/historys-worst-software-bugs/?currentPage=2>.

2. Leveson, N., Turner, C.: An investigation of the Therac-25 accidents. *Computer*. 26, 18-41 (1993).
3. McConnell, S.: Code complete, second edition. Microsoft Press, Redmond (Washington) (2004).
4. Lloyd, S.: Programming the universe. Knopf, New York (2006).
5. Habra, N., Abran, A., Lopez, M., Sellami, A.: A framework for the design and verification of software measurement methods. *Journal of Systems and Software*. 81, 633-648 (2008).
6. Tuch, H., Klein, G., Heiser, G.: OS Verification -- Now!, https://www.usenix.org/legacy/event/hotos05/final_papers_backup/tuch/tuch_html/index.html.
7. Subbotin, S.: Methods of data sample metrics evaluation based on fractal dimension for computational intelligence model building. 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). (2017).
8. Grant, P.: Elementary Computability, Formal Languages and Automata. *Software & Microsystems*. 1, 171 (1982).
9. Gabbay, D.: Kripke Saul A.. Semantical considerations for modal logics. Proceedings of a Colloquium on Modal and Many-valued Logics, Helsinki, 23-26 August, 1962, *Acta Philosophica Fennica* 1963, pp. 83-94. *The Journal of Symbolic Logic*. 34, 501 (1969).
10. Oliinyk, A., Skrupsky, S., Subbotin, S.: Parallel Computer System Resource Planning for Synthesis of Neuro-Fuzzy Networks. *Recent Advances in Systems, Control and Information Technology*. 88-96 (2016).
11. Müller-Olm, M., Schmidt, D., Steffen, B.: Model-Checking. *Static Analysis*. 330-354 (1999).
12. Korotunov, S., Tabunshchik, G., Wolff, C.: Cyber-Physical Systems Architectures and Modeling Methods Analysis for Smart Grids. 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). (2018).
13. Miclea, L., Sanislav, T.: About dependability in cyber-physical systems. 2011 9th East-West Design & Test Symposium (EWDTS). (2011).
14. Pnueli, A.: The temporal logic of programs. 18th Annual Symposium on Foundations of Computer Science (sfcs 1977). (1977).
15. Clarke, E., Emerson, E., Sistla, A.: Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Transactions on Programming Languages and Systems*. 8, 244-263 (1986).
16. Tonetta, S.: Linear-time Temporal Logic with Event Freezing Functions. *Electronic Proceedings in Theoretical Computer Science*. 256, 195-209 (2017).
17. Henke, K., Fäth, T., Hutschenreuter, R., Wuttke, H.: Gift – An Integrated Development and Training System for Finite State Machine Based Approaches. *International Journal of Online Engineering (iJOE)*. 13, 147 (2017).
18. Henke, K., Vietzke, T., Wuttke, H., Ostendorff, S.: Safety in Interactive Hybrid Online Labs. *International Journal of Online Engineering (iJOE)*. 11, 56 (2015).

A Random Forest Model Building Using A priori Information for Diagnosis

Sergey Subbotin¹[0000-0001-5814-8268]

¹Zaporizhzhia National Technical University, Zhukovsky str., 64, Zaporizhzhia, 69063, Ukraine
subbotin@zntu.edu.ua

Abstract. The problem of inductive model building on precedents for biomedical applications is considered. The model paradigm is a random forest as a set of decision tree classifiers working as ensemble. The apriori information taken from training data set is used in proposed method of random forest model building provide more accurate model saving general random character of a method. The resulting random forest provide more accurate model in comparison with a single decision tree, but its comparison with known methods of random forest model building proposed method is more accurate.

Keywords: medical diagnosis, decision tree, random forest, training, inductive learning

1 Introduction

The decision trees [1-2], which are hierarchical tree-like models for obtaining a decision on assigning a recognized instance to one of the possible classes, are useful for problems of biomedical diagnosis by precedents. However, the known methods of constructing models based on decision trees [1-24] are not always able to provide the required level of classification accuracy. The paradigm of a random forest [25-31] is used to improve the accuracy of models based on decision trees.

Random forest paradigm [25-31] is committee (ensemble) of decision trees and combines two main ideas: the Bagging method of L. Breiman [32, 33] and the method of random subspaces of T. Ho [26]. Generally, the idea of the method is to use a large ensemble of decision trees, each of which in itself gives not a lot of classification quality, but together joined they provide a good result due to the large number of predictors [25]. Decision trees are a good family of basic classifiers for bagging [32], since they are quite complex and can reach zero error on any sample. The method of random subspaces [26] allows to reduce the correlation between trees and avoid re-training. Basic methods are trained on different subsets of the features, which are also randomly distinguished.

The random forest paradigm has such advantages [25-31] as: the ability to efficiently process data with a large number of features and classes, insensitivity to scaling (and in general to any monotonic transformations) of feature values, resistance to

the inclusion of irrelevant (useless) features, the ability to process both continuous and discrete features, the ability to work with missing data, possibility to be used for feature significance estimation, insensitivity to emissions in data due to random sampling, possibility of evaluating the resulting model ability to generalize (test for out-of-bag - unselected instances), high parallelizability and scalability, high prediction accuracy, the ability to balance the weight of each class on the entire sample, or on a subsample of each tree, low tendency to retraining (in practice trees almost always only improves composition, but upon validation, after reaching a certain number of trees, the learning curve goes asymptote), the random forests by modifying their definitions can be represented as nuclear methods, which are more convenient and interpretable for analysis, random forest can be converted into the k -nearest neighbors model.

However, a random forest has such disadvantages as [25-31]: not very high accuracy of models, lower interpretability of the model compared to a single tree, lack of formal conclusions (p -values) available for assessing the feature importance, poor quality of work for samples containing a large number of sparse features, a tendency to retraining on noisy data, inability to extrapolate, model bias for data including categorical variables with different amount levels, in favour of features with a large number of levels (when a feature has many levels, the tree will be more adaptable to these features, since they can get a higher value of the optimized functional such as information growth), tendency to give preference to small groups of correlated features that have similar significance for tags to large groups, a large size of the resulting models - the spatial complexity is estimated as $O(ST)$ for storing the model, where S is the volume of the initial sample, T is the number of trees, the large time spent on building the model compared to single decision tree.

The aim of this paper is to improve random forest building method preserving its random character, but concentrating the method on such solutions, which seems to be more convenient to increase model accuracy using the a priori information extracted from the training sample.

2 Formal problem statement

Let we have a training set of S precedents (observations, instances, cases) $\langle x, y \rangle$, where $x = \{x^s\}$, $y = \{y^s\}$, $s = 1, 2, \dots, S$, x^s is an s -th instance of the sample, y^s is an output feature value associated with the s -th instance, $x^s = \{x_j^s\}$, $j = 1, 2, \dots, N$, x_j^s is a value of the j -th input feature of the s -th instance, N is a number of input features.

Then the general model constructing task for the dependence $y=f(w, x)$ is to find such a model structure f and such values of model parameters w for which the model quality criterion F is satisfied. As a rule, for the problems of approximation the model quality criterion is determined as a function of the model error (1):

$$E = \sum_{s=1}^S \left(y^s - f(w, x^s) \right)^2 \rightarrow 0. \quad (1)$$

The decision tree model structure consists of nodes and links (connections between nodes). The decision tree model parameters are the numbers of features used in tree nodes, as well as their boundary values for splitting the ranges of feature values.

A random forest is a collection of trees as structures with parameters, as well as a transformation that combines the results of trees work and the weight of trees in making the final decision.

3 Literature review

The methods of decision tree constructing [1-24] hierarchically divide the initial space into regions, in which they evaluate the output average value for the instances hit in the region (cluster). This value is assigned to the output feature of all instances hitting in this cluster. The advantage of this group of methods is the simplicity and interpretability of the resulting models, as well as the possibility of passing the cluster analysis tasks and selecting informative features. The disadvantages of the methods of this group are the low accuracy of the obtained models.

The random forest method aims to improve classification accuracy joining set of separately generated random models. This makes possible to approximate the inter-class boundary more accurately in comparison with single decision tree.

Let the training sample consist of S instances, the dimension of the feature space is N , and the parameter m (the number of selected features) is specified. Set the number of individual models (trees) in the ensemble (forest) T as well as the maximum acceptable number of instances in the tree node or the maximum allowable tree height as a stop criterion Z .

The most common way to build a committee's tree is called bagging (short for bootstrap aggregation):

To build the next t -th model in a form of decision tree ($t = 1, 2, \dots, T$):

- set the number of features for the t -th model $m_t < N$ (usually, only one value is used for all models $m_t = m$);

- generate a random subsample with S repetitions from the training set (thus, some instances will hit into it two or more times – on average $S(1 - S^{-1})^S$, and about S/e instances (here e is an Euler's number) will not be included into it at all). Those instances that are not hit in the sample are called out-of-bag (unselected).

- build a decision tree based on the generated subsample, and during the creation of the next tree node we will randomly select m_t features from the N initial features, then from the m_t randomly selected features we will choose the best feature, based on which we will split instances. The choice of the best of these features can be done in various ways, for example, on the basis of the Gini criterion or the criterion of information gain. The tree is built until the subsample is completely exhausted, or until the tree reaches the maximum given height Z , or until no more than Z instances are found in each leaf of the tree, and it is not subjected to the pruning procedure.

The optimal number of trees is chosen in such a way as to minimize the classifier error on the test set. In case of its absence, the out-of-bag error estimate is minimized, defined as a recognition error for those instances that did not fall into the training sub-sample due to repetitions (their number is approximately S/e).

It is recommended to set $m = \sqrt{N}$ in classification problems, and $m=N/3$ in regression problems. It is also recommended in the classification problems to build each tree until there is one instance in each leaf, and in regression tasks to build each tree until there are five objects in each leaf.

Recognition of sample instances by a trained model in the form of a forest of decision trees is performed by voting: each forest's tree assigns a recognized instance to one of the classes, and the class for which the most of trees voted wins. More formally, the total forest model for the recognized instance x' , submitted to its inputs, will determine the output value using the formula:

– in classification problems (2):

$$y(x') = \arg \max_{k=1,2,\dots,K} \left\{ \frac{1}{T} \sum_{t=1}^T \{1 \mid y_t(x') = k\} \right\}; \quad (2)$$

– in the evaluation problems (3):

$$y(x') = \frac{1}{T} \sum_{t=1}^T y_t(x'). \quad (3)$$

The disadvantage of the known methods for constructing a random forest of decision trees is that the resulting forest turns out to be more precise than a single tree, but at times or even more complicated because of the large number of constructed models in the form of trees of the corresponding forest. This essentially not only increases the time expenditures for the decision making and the requirements for memory resources of the computer, but also leads to a significant reduction in the generalizing properties and interpretability of the forest model in comparison with the model of the single tree.

Therefore, an urgent task is to develop methods that allow to synthesize models in the form of a random forest, free from the disadvantages noted above, or characterized by them to a lesser extent.

4 The method of forming a random forest of decision trees based on a priori information

To solve the problem of the development of a method free from disadvantages mentioned above, it is proposed in the process of building a forest of decision trees, along with a casual approach to their construction to use deterministic component - take into account the a priori information, which will allow on-directs the formation of models so as to concentrate on the most promising directions, preserving the overall stochastic nature of the model building process.

At the selecting a subsample to build a partial tree model, it is proposed to select samples for inclusion in the model randomly, but taking into account their individual informativeness, minimizing their similarity (increasing diversity) – maximizing the distance to the nearest instance.

At the current tree constructing, the choice of the root node is proposed to make randomly, but taking into account individual estimates of the informativeness of the features, increasing the chances of those features to be used in the root node, which are individually the most significant, and have not yet used in the roots of previously formed trees.

At the current non-root node forming, it is proposed to take into account the individual informativeness of the features their relationships with each other with respect to the output parameter.

1. Initialization: set the training sample $\langle x, y \rangle$.
2. Estimate the a priori information:
 - determine individual estimates of the informativeness of each j -th input feature [34-46] respectively to the output feature $I_j, j = 1, 2, \dots, N$;
 - determine individual estimates of the informativeness of each i -th input feature [34-46] respectively to the j -th input feature $I_{i,j}, j = 1, 2, \dots, N$;
 - estimate the individual informativeness of each sample instance I^s [35, 44], $s = 1, 2, \dots, S$;
 - determine the measure of the distance between the s -th and p -th instances of the sample in the feature space $d^{s,p}$. For small size samples evaluate by (4):

$$d^{s,p} = d^{p,s} = \sum_{j=1}^N (x_j^s - x_j^p)^2, s = 1, 2, \dots, S, p = 1, 2, \dots, s. \quad (4)$$

For large samples, instead of the distance between instances, it is possible to use the distances between their locally sensitive hashes [47] evaluated as (5):

$$d^{s,p} = H^s - H^p, s = 1, 2, \dots, S, p = 1, 2, \dots, s, \quad (5)$$

where H^s is a locally sensitive hash for s -th instance.

Calculation of hashes, unlike calculation of distances, will not require loading all instances into computer memory: hashes can be calculated in just one sample pass, operating with fragments in memory.

3. Building a forest of T decision trees. To build the t -th model as a decision tree ($t = 1, 2, \dots, T$):

- set the number of attributes for the t -th model $m_t < N$ (usually, only one value is used for all models $m_t = m$);

- generate a random subsample of size S with repetitions from the training sample. When selecting a subsample to build a partial model, it is proposed to select the samples for inclusion in it randomly, but taking into account their individual informative-

ness, minimizing their similarity (improving diversity) – maximizing the distance to the nearest instance using the modified rule of roulette [48] determined by (6):

$$V^s = \begin{cases} 1, r^s \leq \frac{I^s}{\sum_{p=1}^S I^p} \left(\frac{\max_{p=1,2,\dots,S} \{d^{s,p} \mid x^p \in \Theta\}}{\frac{1}{S-|\Theta|} \sum_{g=1}^S \max_{p=1,2,\dots,S} \{d^{g,p} \mid x^g \notin \Theta, x^p \in \Theta\}} \right); \\ 0, \text{otherwise,} \end{cases} \quad (6)$$

where V^s is the binary decision to use the s -th instance at subsample forming (1 is use, 0 is not to use), Θ is a set of already selected instances for inclusion in the subsample, r^s is a random number in the range $[0, 1]$, generated to select s -th instance.

This rule should be applied consistently to the original data sample, as long as $|\Theta| \leq S$, where S is a desired subsample volume.

The proposed rule will preserve the random nature of the selection, providing each instance with a chance to be selected for inclusion in the subsample, but will increase the chances of being selected for the individually most informative instances, as well as for those instances that are most different in the feature space from the already selected instances;

– for a given basic method of decision tree building, determine for each feature the value of the partitioning criterion I_{jM} used by the method (it is necessary to provide $I_{jM} \in [0,1]$), on the basis of which to determine the randomized selection criterion of the feature for splitting the tree in the root node by (7)-(9):

$$I_{jRand} = I_{jM} I_{jt} \hat{I}_j r, \quad (7)$$

$$I_{jt} = \frac{1}{1 + \sum_{p=1}^{t-1} \{1 \mid \text{root}(p) = j\}}, \quad (8)$$

$$\hat{I}_j = \frac{I_j}{\max_{i=1,2,\dots,N} \{I_i\}}, \quad (9)$$

where $\text{root}(t)$ is a number of the feature used in the root node of the t -th tree from already constructed forest of trees, r is a random number in the range $[0, 1]$ generated for the new tree.

The proposed criterion I_{jRand} , preserving the criterion I_{jM} of the chosen method as a whole, will give it randomized properties, as well as increase the chances of those features to be placed in the root node that have not yet placed in the roots of previously formed trees, as well as individually are the most significant.

- select the best feature as the root node of the created t -th tree using the criterion I_{jRand} , and then split the sample according to the selected feature;
- build a decision tree based on the generated subsample using the resulting root node, and during the creation of the next node of the tree from the N initial features, randomly select the m_t features, taking into account their individual informativeness and interrelation with each other using the criterion (10):

$$I_{j*Rand} = I_{jM} \max(\hat{I}_j, I_{j,temp})^r, \quad (10)$$

where is $temp$ a number for the current node of the feature, located in the parent node.

After that, from the randomly selected m_t features, using the base method, select the best feature, on which basis the partition will be performed. The tree is built until the subsample is completely exhausted or until the tree reaches the maximum given height Z , or until no more than Z instances are found in each leaf of the tree, and it does not undergo the pruning of branches.

The proposed method preserving the generally random nature of the selection of subsample instances in the process of building decision trees forest will increase the chances to be selected for constructing of the particular models of those instances that are individually more informative, and will also strive to provide a variety of instances used to construct particular models, also preserving the random principle of selection of a subset of candidate features for the formation of the current node of the synthesized decision tree will increase the chances to becoming the root of the tree of the individually most informative features that not yet used as root, and at the next node forming it will increase the chances to being used of those features that are individually the most significant and also most closely associated with the feature already used in the parent node.

5 Experiments and results

The proposed method were implemented as software and experimentally investigated in solving the problem of diagnosing the development of recurrent respiratory infections in young children.

Diseases of the respiratory system occupy a leading place among the entire pathology of young children of early age. It is problematic to examine children from whom infectious episodes acquire a protracted, severe, recurrent course. To create rehabilitation and prevention programs, it is necessary to determine the most significant risk factors, as well as their combination, which was solved within the framework of the task of diagnosing the development of recurrent respiratory infections in young children [48].

The data sample for the experiments was obtained in [48] on the basis of a survey of 108 children. The observations were characterized by 42 features reflecting the presence or absence of chronic diseases in the parents and the child: x_1 is child gender,

x_2 is a father's illness – allergy, x_3 is a father's illness – other chronic diseases, x_4 is a father's illness – healthy, x_5 is a mother's illness – allergy, x_6 is a maternal diseases – other chronic diseases, x_7 is a maternal diseases – healthy, x_8 is a normal pregnancy, x_9 is a mother sick during pregnancy, x_{10} is a normal delivery, x_{11} is a pregnancy was complete, x_{12} – whether the child was in the anaesthesiology department, x_{13} – whether the child was in the neonatal pathology department, x_{14} is a breastfeeding, x_{15} is a breastfeeding for up to one month, x_{16} is a breastfeeding for up to three months, x_{17} is a breastfeeding for up to one year, x_{18} is a breastfeeding feeding up to the present moment, x_{19} is a breastfeeding with a term of over a year, x_{20} is a diagnosis – cytomegalovirus infection, x_{21} is a diagnosis – anemia, x_{22} is a diagnosis – abnormality of the urinary system, x_{23} is a diagnosis – obstructive bronchitis, x_{24} is a diagnosis – recurrent obstructive bronchitis, x_{25} is a diagnosis – congenital heart disease, x_{26} is a diagnosis – urinary tract infection, x_{27} is a diagnosis – otitis, x_{28} is a diagnosis – acute respiratory viral infection, x_{29} is a diagnosis – acute stenotic laryngotracheobronchitis, x_{30} is a diagnosis – pneumonia, x_{31} is a diagnosis – rickets Ca-norm, x_{32} is a diagnosis – enterobiosis, x_{33} is a diagnosis – hypotrophy, x_{34} is a diagnosis –allergy, x_{35} is a diagnosis – central nervous system damage, x_{36} is a parents smoke, x_{37} is a older children in the family, x_{38} is a child attends care facility, x_{39} is a unfavourable living conditions, x_{40} is a child use antibiotics of 1-2 groups, x_{41} is a child use antibiotics of 3rd group, x_{42} is a child episodically ailing. The output feature y take a value "0" or "1" depending on the child's exposure to the development of recurrent respiratory infections. The fragment of data sample is presented in the Table 1.

Table 1. Fragment of data sample

S	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	...	x_{42}	y
1	0	0	0	0	0	0	1	0	1	1	...	0	1
2	1	1	0	0	0	0	1	1	0	1	...	0	1
3	1	0	0	1	0	0	1	1	0	1	...	0	1
4	0	0	1	0	1	0	0	0	1	1	...	0	1
5	0	0	1	0	0	0	1	1	0	1	...	1	1
6	0	1	0	0	0	0	1	1	0	1	...	0	1
7	1	0	0	1	0	1	0	1	0	1	...	0	1
8	1	0	0	1	0	1	0	0	1	1	...	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...
12	0	0	0	1	0	0	1	1	0	1	...	0	1
13	1	0	0	1	0	0	1	1	0	1	...	0	1
14	1	0	0	1	0	0	1	1	0	0	...	0	1
15	1	0	0	1	0	0	1	0	0	1	...	0	1
16	0	0	0	1	0	1	0	0	1	1	...	0	1
17	1	0	0	1	0	0	1	1	0	1	...	0	0
18	0	0	0	1	0	0	1	0	0	1	...	0	1
19	0	0	0	1	0	0	1	1	1	1	...	0	1
20	0	0	0	1	0	0	1	0	0	1	...	0	1
21	0	0	0	1	0	0	1	0	0	1	...	1	0
22	0	1	1	0	1	1	0	0	1	0	...	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
108	0	0	0	1	0	0	1	1	1	0	...	0	0

In the experiments, the original sample was randomly divided into training and test samples of the same size. Wherein it was ensured the preservation of the frequencies of the classes in the training sample relative to the original sample. On the basis of the training sample, models were built in the form of decision trees. For each model, the errors were determined on the basis of training and test data, as well as the time of work creation.

The results of the experiments are presented in the Table 2. Here E_{tr} is a model error for training sample, t_{tr} is a time of building, t_{trr} is a time of recognition of training sample, E_{tst} is a model error for test sample, t_{tst} is a time of recognition of test sample.

Table 2. Results of experiments

Model architecture and training method	E_{tr}	t_{tr} , sec	t_{trr} , sec	E_{tst}	t_{tst} , sec
Single decision tree, Breiman's method [2]	0	0.73	0.09	3	0.09
Random forest, method [25]	0	5.78	0.65	1	0.65
Random forest, proposed method	0	5.92	0.65	0	0.65

The results of the experiments presented in the Table 1 show that the proposed method allows to obtain models that is better in accuracy (of lower error) in comparison with existent tree building methods [2, 25], but requires a little more time to build a model, the speed of the constructed model for proposed method does not essentially differ from the speed of a random forest model constructed using the method [25], but, as was to be expected, random forest-based models require significantly more time for calculations compared to the single-tree model [2].

Increasing the accuracy of problem solving by a model built on the basis of the proposed method is provided, on the one hand, by increasing the diversity of decisions for root feature selection, and, on the other hand, by taking into account the individual informativeness of features at forming the decision-making hierarchy. Whereas the proposed method preserves the generally random nature of the formation of the forest of the decisive trees.

6 Conclusion

The problem of inductive model building on precedents for biomedical applications is considered in the paper.

The model paradigm is a random forest as a set of decision tree classifiers working as ensemble.

A random forest method is proposed. It for a given training sample builds a set of trees for hierarchical clustering of instances. The a priori information taken from training data set is used in proposed method of random forest model building to provide more accurate model saving general random character of a method.

The software implementing the proposed methods has been developed and studied at the diagnosis problem solving. The conducted experiments have confirmed the performance of the developed software and allow recommending it for use in practice.

The resulting random forest provide more accurate model in comparison with a single decision tree, but it comparison with known methods of random forest model building proposed method is more accurate.

The prospects for further research are to test the proposed methods on a wider set of applied problems, to study the dependence of the speed and accuracy (error) of method's work on the sample volume and the feature number in the original sample.

References

1. Amit, Y., Geman, D., Wilder, K.: Joint induction of shape features and tree classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11), pp. 1300–1305 (1997)
2. Breiman, L., Friedman, J. H., Stone, C. J., Olshen, R. A.: *Classification and regression trees*. Chapman & Hall / CRC, Boca Raton (1984)
3. Rabcan, J., Rusnak, P., Subbotin, S.: Classification by fuzzy decision trees induced based on Cumulative Mutual Information. In: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018 - Proceedings, Slavske, 20-24 February 2018, pp. 208-212 (2018)
4. Dietterich, T. G., Kong, E. B.: Machine learning bias, statistical bias, and statistical variance of decision tree algorithms. *Machine Learning*. 255 (1995).
5. Friedman, J.H.: A recursive partitioning decision rule for nonparametric classification. *IEEE Transactions on Computers*. 100(4), pp. 404–408 (1977)
6. Geurts, P., Irrthum, A., Wehenkel, L.: Supervised learning with decision tree-based methods in computational and systems biology. *Molecular Biosystems*. 5(12), pp. 1593–1605 (2009).
7. Geurts, P.: Contributions to decision tree induction: bias/variance tradeoff and time series classification. PhD Thesis. University of Liège (2002)
8. Murthy, K. V. S., Kasif, S., Salzberg, S.: A system for induction of oblique decision trees. *Journal of Artificial Intelligence Research archive*. 2 (1) (1994)
9. Hothorn, T., Hornik, K., Zeileis, A.: Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*. 15(3), pp. 651–674 (2006)
10. Kim, H., Loh, W.-Y.: Classification trees with unbiased multiway splits. *Journal of the American Statistical Association*. 96 (454) (2001)
11. Kufirin, R.: Decision trees on parallel processors. *Machine Intelligence and Pattern Recognition*. 20 (1997)
12. Kwok, S. W., Carter, C.: Multiple decision trees. In: *Uncertainty in Artificial Intelligence*. 9 (4), pp. 327–338 (1990)
13. Mingers, J.: An empirical comparison of pruning methods for decision tree induction. *Machine learning*. 4(2), pp. 227–243 (1989)
14. Murthy, K. V. S.: On growing better decision trees from data. PhD thesis. The Johns Hopkins University (1995)
15. Quinlan, J.R.: *C4.5 Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann. (1992)

16. Oliinyk, A.A., Subbotin, S.A.: The decision tree construction based on a stochastic search for the neuro-fuzzy network synthesis. *Optical Memory and Neural Networks (Information Optics)*, 24 (1), pp. 18-27 (2015). doi: 10.3103/S1060992X15010038
17. Quinlan, J. R.: Induction of decision trees. *Machine learning*. 1(1), pp. 81– 106 (1986)
18. Quinlan, J. R.: Simplifying decision trees. *International Journal of Man-Machine Studies*. 27 (1987). doi:10.1016/S0020-7373(87)80053-6
19. Quinlan, R. : Learning efficient classification procedures. In: *Machine Learning: an artificial intelligence approach*, Michalski, Carbonell & Mitchell (eds.), Morgan Kaufmann, pp. 463-482 (1983). doi:10.1007/978-3-662-12405-5_15
20. Subbotin, S., Kirsanova, E.: The regression tree model building based on a cluster-regression approximation for data-driven medicine. *CEUR Workshop Proceedings*. 2255, pp. 155-169 (2018)
21. Strobl, C., Boulesteix A.-L., Augustin T.: Unbiased split selection for classification trees based on the gini index. *Computational Statistics & Data Analysis*, 52(1), pp. 483–501 (2007)
22. Utgoff, P. E.: Incremental induction of decision trees. *Machine learning*. 4(2), pp. 161-186 (1989). doi:10.1023/A:1022699900025
23. Wehenkel, L.: On uncertainty measures used for decision tree induction. In: *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. pp. 413-418 (1996)
24. White, A. P., Liu, W. Z. Technical note: Bias in information-based measures in decision tree induction. *Machine Learning*. 15(3), pp. 321–329 (1994)
25. Breiman, L.: Random Forests. *Machine Learning*. 45 (1), pp. 5–32 (2001). doi:10.1023/A:1010933404324.
26. Ho, T. K.: The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 20(8), pp. 832–844 (1998)
27. Botta, V.: A walk into random forests: adaptation and application to genome-wide association studies. PhD Thesis. Université de Liège (2013)
28. Denisko, D., Hoffman, M.: Classification and interaction in random forests. *Proceedings of the National Academy of Sciences of the United States of America*. 115 (8), pp 1690–1692 (2018). doi:10.1073/pnas.1800256115. PMC 5828645. PMID 29440440.
29. Louppe, G.: Understanding Random Forests: From Theory to Practice. PhD Thesis, University of Liege (2014)
30. Boulesteix, A.-L., Janitza, S., Kruppa, J., König, I. R.: Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6), pp. 493–507 (2012)
31. Criminisi, A., Shotton, J.: *Decision Forests for Computer Vision and Medical Image Analysis*. Springer (2013)
32. Panovand, P., Džeroski, S.: Combining bagging and random subspaces to create better ensembles. *Advances in intelligent data analysis*. VII, pp. 118–129 (2007)
33. Wolpert, D. H., Macready, W. G.: An efficient method to estimate bagging's generalization error. *Machine Learning*. 35(1), pp. 41–55 (1999)
34. De Mántaras, R. L.: A distance-based attribute selection measure for decision tree induction. *Machine learning*. 6(1), pp. 81–92 (1991).
35. Subbotin, S.: The instance and feature selection for neural network based diagnosis of chronic obstructive bronchitis. *Studies in Computational Intelligence*, vol. 606, pp. 215-228 (2015)

36. Miyakawa, M.: Criteria for selecting a variable in the construction of efficient decision trees. *IEEE Transactions on Computers*. 38(1), pp. 130–141 (1989).
37. Subbotin, S.A.: Methods of sampling based on exhaustive and evolutionary search *Automatic Control and Computer Sciences*. 47(3), pp 113-121 (2013)
38. Altmann, A., Toloşi, L., Sander, O., Lengauer, T.: Permutation importance: a corrected feature importance measure. *Bioinformatics*. 26 (2010). doi:10.1093/bioinformatics/btq134. PMID 20385727.
39. Subbotin S.: Quasi-relief method of informative features selection for classification. In: 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 - Proceedings, 11–14 September 2018, pp. 318–321 (2018). doi: 10.1109/STC-CSIT.2018.8526627
40. Oliinyk, A., Subbotin, S., Lovkin, V., Leoshchenko, S., Zaiko, T.: Feature selection based on parallel stochastic computing 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 - Proceedings, 1, 11 September 2018 through 14 September 2018, pp. 347-351. (2018) DOI: 10.1109/STC-CSIT.2018.8526729
41. Painsky, A., Rosset, S. Cross-Validated Variable Selection in Tree-Based Methods Improves Predictive Performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 39 (11), pp. 2142–2153. (2017). doi:10.1109/tpami.2016.2636831. PMID 28114007.
42. Oliinyk, A., Subbotin, S., Lovkin, V., Leoshchenko, S., Zaiko, T.: Development of the indicator set of the features informativeness estimation for recognition and diagnostic model synthesis. In: 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018 - Proceedings, 20-24 February 2018, pp. 903-908 (2018)
43. Mingers, J.: An empirical comparison of selection measures for decision-tree induction. *Machine learning*. 3(4), pp. 319–342 (1989).
44. Subbotin, S.: The instance and feature selection for neural network based diagnosis of chronic obstructive bronchitis. In: *Studies in Computational Intelligence*. 606, pp. 215-228 (2015). doi: 10.1007/978-3-319-19147-8_13
45. Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A.: Conditional variable importance or random forests. *BMCbioinformatics*. 9 (2008)
46. Subbotin, S., Oleynik, A.: Entropy based evolutionary search for feature selection. In: *The Experience of Designing and Application of CAD Systems in Microelectronics - Proceedings of the 9th International Conference, CADSM 2007, Lviv-Polyana, 20-24 February 2007*, pp. 442-443 (2007). doi: 10.1109/CADSM.2007.4297612
47. Pauleve, L.; Jegou, H.; Amsaleg, L.: Locality sensitive hashing: A comparison of hash function types and querying mechanisms. *Pattern Recognition Letters*. 31 (11), pp. 1348–1358 (2010). doi:10.1016/j.patrec.2010.04.004.
48. Gerasimchuk, T., Zaitsev, S., Subbotin, S.: The use of artificial immune systems to predict the risk of recurrent respiratory infections in young children. In: *Diahnostyka ta likuvannya infektsiyno oposeredkovanykh somatychnykh zakhvoryuvan' u ditey: mizhrehional'na naukovo-praktychna konferentsiya, Donetsk, 10–11 February 2011*. pp, 27–29 (2011)

INFORMATION AND COMMUNICATION SYSTEMS BASED ON SIGNAL SYSTEMS WITH IMPROVED PROPERTIES BUILDING CONCEPT

Gorbenko Ivan¹[0000-0003-4616-3449], Zamula Alexander²[0000-0002-8973-6190], Morozov Vladyslav³[0000-0003-0427-0107]

¹V.N. Karazin Kharkiv National University, Svobodu square 4, Kharkiv, 61022, Ukraine
gorbenkoi@iit.kharkov.ua

²V.N. Karazin Kharkiv National University, Svobodu square 4, Kharkiv, 61022, Ukraine
zamylyaaa@gmail.com

³V.N. Karazin Kharkiv National University, Svobodu square 4, Kharkiv, 61022, Ukraine
ilissar@hotmail.com

Abstract. The paper presents the results of solving the actual problem of improving the performance indicators of information and communication systems (ICS), in particular, information security, noise immunity, secrecy, in terms of external and internal disturbing influences, based on the models and methods development for information exchange in ICS, synthesis new classes of nonlinear discrete cryptographic signals (CS) with the necessary properties. For the first time, complex non-linear discrete CS synthesis problem was formulated and solved, based on the use of random (pseudo-random) processes. Taking into account the requirements of cryptographic robustness and the complexity of generating a cryptographic signal as a source of random (pseudo-random) processes, the choice of a symmetric block encryption algorithm with a counter is grounded. The synthesis method obtained by the authors makes it possible to form large ensembles of discrete sequences of practically any period with given, but physically realizable, side-lobe values of auto, cross and butt correlation functions, as well as statistical characteristics of correlation functions that are not inferior to those of the best linear classes of signals, in terms of correlation functions. A method for synthesizing derived signal systems based on nonlinear discrete cryptographic signals is proposed. Improved (as compared with the known linear signal classes) ensemble, correlation and structural properties of the synthesized signal system allow improving the indicators of information security, noise immunity and secrecy of the operation of ICS.

Keywords: information security, secrecy, correlation function, complex signal, cryptographic signal, ensemble signals.

1 Introduction

In the context of intensive information counteraction by the parties, the interests and competition of which can be manifested in various fields, including, as recent events have shown, in the field of information and hybrid wars, the availability and use of

secure information and communication systems (ICS) is of particular importance. To a significant extent, such systems are based on the use of secure radio channels. In this case, under system protection we will understand, in a broad sense, their ability to provide the necessary indicators for noise immunity, information, energy and structural secrecy. Increased requirements for the effectiveness of the ICS in the conditions of internal and external influences are largely not taken into account by existing information technologies. There is a contradiction between the stringent requirements for ensuring the reliability, secrecy, confidentiality, integrity, authenticity of data transmitted over wired and wireless ICS communication lines, and existing models, methods and technologies for managing telecommunications networks, information security, services and quality of service.

New models, methods and technologies creation for managing telecommunication networks, information security, services and quality of service, in order to improve system performance indicators, under external and internal impacts (threats), through the creation of various interferences (retransmitted, structural, impulse, broadband, narrowband etc.), breaches of confidentiality, integrity and authenticity of information exchange, are, in our opinion, relevant research areas.

A significant number of ICS are multi-user systems. In such systems, multiple channels are located within a common time-frequency resource. One of the most promising ways of multiple access to services and system resources is code division multiple access in the common frequency band (CDMA). This access method is the most promising in many aspects: high noise immunity of the channels and ensuring the confidentiality of the transmitted data; high transmission rate and bandwidth efficiency; high energy efficiency and network subscriber capacity. Since the code division of ICS channels is based on the difference in signals provided to system subscribers, the construction of such systems and their characteristics are determined by the choice of signals and their properties. A promising direction for ensuring the security of information resources is the use of distributed spectrum technology (broadband noise-like signals).

The main ways to solve the above contradiction is to increase the noise immunity (in particular, energy and structural secrecy, noise immunity of signal reception) and information security of ICS by improving the methodological basis for constructing ICS by developing information exchange methods, methods for synthesizing new classes of nonlinear complex discrete signal-data carriers with the necessary ensemble, correlation and structural properties.

2 Statement and problem solution for nonlinear discrete complex cryptographic signals system synthesis

The process of choosing complex signals structures that are rational for some criteria is identical to the synthesis of the corresponding manipulating discrete sequences (DS). As a criterion for choosing a discrete signal class, as a rule, they are guided by the criterion of minimum mutual interference (minimax criterion). Such a criterion implies the ensemble signal construction with volume M , manipulated by DS, which differ as much as possible from each other during possible cyclic shifts. The quantitative measure of the differences between manipulating DSs are the maximum

ensembles side-lobe levels of the periodic autocorrelation function (ACPF) and the side lobe level of the periodic cross-correlation function (CCPF). On this basis, the broadband signals (BBS) used in ICS should have such correlation properties, when the lateral peaks of the correlation functions of the BBS are as low as possible, i.e. ideally should aim for zero. However, the requirement of ideality (zero values of side peaks) of auto- and cross- correlation functions between all cyclic shifts of K sequences and various signal system isomorphisms with period N is not feasible, since the values of side peaks cannot fall below $1/2\sqrt{B}$ (B – signal base) [1].

Information exchange methods used in ICS, as well as classes of wideband signals, used as a physical carrier of data (sets of linear recurrent sequences (M-sequence), Kasami, Golda, Kamaletdinov and others), possessing relatively small values of side lobes of auto- and cross-correlation functions that do not support necessary (for critical ICS applications) information security and noise immunity indicators [1]. Thus, in the process of information exchange in ICS for a long time, the correspondence: the message-signal bit is fixed, and the above signals have low structural secrecy, limited ensemble properties, and also exist only for a limited number of signal period values. In the case of period truncation (increase) of such signals, their correlation properties deteriorate.

The analysis showed that currently there are no regular methods for the discrete sequences (DS) synthesis that are optimal by the minimax criterion. The problem of DS synthesis turns out to be even more complex if the requirements are imposed on signal system dimension (volume), structural properties and the number of DS elements. Since, for distributed spectrum technology, data carrier signal properties are completely determined by DS properties, which manipulate the information bits of system users [1-2], the search for efficient methods for the discrete signals (sequences) synthesis of corresponding to the potentially achievable boundary characteristics of correlation functions (minimax properties) or the border of “dense packing”) and possessing the required correlation, structural, ensemble properties.

The quality of services provided by ICS is proposed to assess, including the level of information security. At the same time, information security means the ability of ICS to provide protection against the destruction, modification, blocking of information, its unauthorized leakage or against violation of the established procedure for its routing. Also, information security should be understood as the state of security of data processing and storage systems, which ensures information confidentiality, integrity and availability, as well as other information properties and services: authenticity, traceability, irrefutability and reliability. [3].

We will formulate in general terms the problem of signal synthesis with given correlation ensemble and structural properties that provide the required values of noise immunity, information security and information transmission system secrecy. We require that such signal systems have the property of “blurring” in the correlation properties. This property means that increasing or decreasing the length of a discrete sequence does not change the correlation properties of the original discrete sequence.

The need to use secure radio channels forces researchers to take a fresh look at both protected radio channels operation modes and formation and application of complex signals aspects. Therefore, in our opinion, new approaches and new views on the complex signals application processes and functions are needed. In our opinion, new fundamental understanding of methods to ensure information secrecy, that is,

functions that in traditional ICS are assigned to cryptographic systems and means. Therefore, a productive step, from the point of view of a new direction of using complex signal systems, is the synthesis of so-called cryptographic signal (CS) systems. The synthesis of such signals is based on the application of key data, and at the same time, the signals must have: absolute structural secrecy regarding the laws of their formation; improved ensemble properties (exist for almost any period value, have a significant amount of signal system); necessary to ensure the required value of noise immunity, the correlation properties. For protected radio channels, considered signal systems are defined by the applications in which they applied. In particular, these can be either individual signals or pairs of signals, as well as large sets of discrete sequences with necessary, but objectively limited values of «dense packing», cross-correlation and ensemble properties.

Under cryptographic discrete signals, it is proposed to understand symbol sequence sets (vectors) of a certain alphabet that possess the necessary (specified) structural, ensemble and correlation properties, temporal and spatial formation complexity and are formed on the basis of random (pseudo-random) processes, including using cryptographic algorithms [4-6]. Such signals must meet the requirements of chance, irreversibility, unpredictability [7], have correlation and ensemble properties determined by the owner (user) ICS. The use of CS, will improve the performance of ICS, in particular: noise immunity (noise immunity of signal reception in the face of structural, barrage, retransmitted and other types of interference, secrecy of operation) and information security. With this approach, the signal's structural secrecy is provided by randomness or pseudo-randomness of the symbol sequence, noise immunity is provided by the synthesized signal system correlation properties, ICS information security is provided based on statistical properties of CS close to the properties of random sequences, and the use of cryptographic keys, and the key length can be significantly less than the period (length) of the signal itself. It is necessary to note the special property of cryptographic signal systems: the possibility of their recovery in space and in time using keys and a number of other parameters that are used in the synthesis of this signal system.

3 Discrete cryptographic signal models

Under task of constructing (synthesizing) CS, we will understand the task of constructing subsets of discrete sequences. $(W_l^q), q = \overline{1, N}, l = \overline{1, L}$, the combination of which forms a discrete signal system with given dimension alphabet $M_k = N \times L$, such that in each of the subsets (vocabulary) the conditions that are imposed on the subset of CS in terms of structural, ensemble, correlation properties, spatial and temporal complexity of their generation are met.

CS construction is based on the analysis and use of periodic and aperiodic correlation functions and is based on the following theoretical principles.

1. Ensuring conditions for fulfilling the requirements for structural and ensemble properties, the possibilities of forming a CS subset with permissible temporal and spatial complexity, including using keys.

2. CS formation W^q , periodic autocorrelation function (ACPF) of each of which satisfies the system of nonlinear parametric inequalities (NPI):

$$R_{a_1}^q(l) \leq \sum_{i=1}^L W_i^q (W_{i+l}^q)^* \leq R_{a_2}^q(l), \quad l=\overline{1, L-1}, \quad q=\overline{1, N}, \quad (1)$$

where $R_{a_1}^q(l)$ и $R_{a_2}^q(l)$ – ACPF implementation setpoints, and indices are calculated modulo $(i+l) \bmod L$.

If $l=L$ for all $q=\overline{1, N}$ (1) convolution by value L

$$\sum_{i=0}^L W_i^q W_i^q = L, \quad q=\overline{1, N}. \quad (2)$$

3. Pair formation CS W^q and W^p , cross-correlation functions (CCF) which meet the requirements that are determined by a set of NPI systems (3), and also meet the requirements for butt cross-correlation functions (CCBF) pair CS W^q and W^p with butt discrete signals W^{qp} and W^{pq} (3 – 7):

$$R_{b_{1,1}}^{qp}(l) \leq \sum_{i=0}^{L-K} W_i^q \times (W_{i+l}^p)^* + \sum_{i=L-K+1}^{L-1} W_i^q \times (W_{i-l+K}^p)^* \leq R_{b_{2,1}}^{qp}(l) \quad (3)$$

$$R_{b_{1,2}}^{qp}(l) \leq \sum_{i=0}^{L-K} W_i^q \times (W_{i+l}^q)^* + \sum_{i=L-K+1}^{L-1} W_i^q \times (W_{i-l+K}^p)^* \leq R_{b_{2,2}}^{qp}(l); \quad (4)$$

$$R_{b_{1,3}}^{qp}(l) \leq \sum_{i=0}^{L-K} W_i^q \times (W_{i+l}^p)^* + \sum_{i=L-K+1}^{L-1} W_i^q \times (W_{i-l+K}^q)^* \leq R_{b_{2,3}}^{qp}(l); \quad (5)$$

$$R_{b_{1,4}}^{qp}(l) \leq \sum_{i=0}^{L-K} W_i^p \times (W_{i+l}^p)^* + \sum_{i=L-K+1}^{L-1} W_i^p \times (W_{i-l+K}^q)^* \leq R_{b_{2,4}}^{qp}(l); \quad (6)$$

$$R_{b_{1,5}}^{qp}(l) \leq \sum_{i=0}^{L-K} W_i^p \times (W_{i+l}^q)^* + \sum_{i=L-K+1}^{L-1} W_i^p \times (W_{i-l+K}^p)^* \leq R_{b_{2,5}}^{qp}(l); \quad (7)$$

where $l=\overline{1, L-1}$ for all kinds of connections q and p , $q=\overline{1, N}$, $p=\overline{1, N}$, $q \neq p$,

where $R_{b_{1,j}}^{qp}(l)$ and $R_{b_{2,j}}^{qp}(l)$, CCPF and CCBF implementations are set respectively, $j=\overline{1, 5}$.

In NPI systems (1) - (2) and (3) - (7) W_i^q and W_i^p are unknown values of random or pseudo-random characters CS W^q and W^p , $q=\overline{1, N}$, to be determined in the process of their construction.

Let us analyze the systems of nonlinear parametric quadratic inequalities (further - systems) (1) - (2) and (3) - (7), using the introduced model.

Systems (4) and (6) if $l=L$ for all $q=\overline{1, N}$ must give a complete convolution with the value L , i.e (4):

$$\sum_{i=1}^L W_i^q W_i^q = L, q = \overline{1, N} \quad (8)$$

and (6) gives

$$\sum_{i=1}^L W_i^p W_i^p = L, p = \overline{1, N} . \quad (9)$$

Systems (3), (5) and (7) with $l=L$ for all pairs W^q and W^p give the values of the cross-correlation function at the zero value of the shift of the corresponding type:

$$\sum_{i=1}^L W_i^q W_i^p = R^{qp}(0); q, p = \overline{1, N}, \quad (10)$$

$$\sum_{i=1}^L W_i^q W_i^p = R^{qp}(0), q, p = \overline{1, N}, \quad (11)$$

$$\sum_{i=1}^L W_i^p W_i^q = R^{pq}(0), p, q = \overline{1, N} . \quad (12)$$

In the following, systems (1–2), (3–7) and the quadratic equation (10) will be called the subset (dictionary) model CS.

We will analyze the systems (1), (2) for the existence of solutions and independence. Directly from (1) it follows that for each of q CS there L unknown - $W_1^q, W_2^q \dots W_L^q$. To find them according to (1), you can make a system of $L-1$ independent NPI. Further, using (2), we obtain one more expression, but already the equality. A feature of system (1) is that it gives a convolution of each of q CS with value L . Based on (1) and (2), each N subset of CS can be constructed to form N independent systems of quadratic NPI, each of which will contain $L-1$ quadratic inequalities of the form (1) and formally one equality, so that all of them will be L .

We will also analyze the totality of systems of parametric inequalities (3-7), taking into account (8) – (12), for solutions existence, systems independence and individual equalities. Systems (3-7) determine the permissible cross-correlation properties with respect to the CCPF and CCBF of each pair CS – W^q and W^p . They define the requirements for CCPF and butt cross-correlation functions (CCBF) specifically only two CS - W^q and W^p . When building three CDSs, we'll have 3!/2 systems like (3-7), and for N CS respectively $-N!/2$ such systems. Thus, with increasing N , the number of systems of the form (3-7) increases exponentially (in terms of factorial).

For $N = 2$, among (8) – (12) NPI systems are redundant nonlinear quadratic equalities. Equality (2) coincides with (8) and (9), therefore, the last two are already included in system (2), are dependent, and therefore cannot be used. Further, equalities (10) and (11) coincide, and equality (12) is symmetric, in terms of the correlation function, with respect to equalities (10) and (11). Therefore, for each pair p and q independent is (10).

Based on a detailed analysis, we find that all (3-7) NPI systems define different implementations of CCPF and CCBF specifically only two CS W^q and W^p . Therefore, a mathematical model for constructing two CS W^q and W^p uniquely determined by the five systems NPI like (3 -7), and, as already stated, by the equation (10).

The above analysis results allow us to determine model complexity and, on its basis, N CS subset constructing.

1. When building a CS, it is necessary, depending on the allowed values. $R_{a_1}^q(l)$ and $R_{a_2}^q(l)$, which are determined by the boundaries of «dense packing», consider $v \geq k$ systems such as (1 – 2).

2. When building two CS you need to consider $v_2 \geq k_2$ systems like (3 – 7), where k_2 determined by $R_{b_{1,j}}^{qp}(l)$ and $R_{b_{2,j}}^{qp}(l)$.

3. When building N CS you need to consider $v_N \geq k_N$ systems like (3-7), where k_N determined by $R_{a_1}^q(l)$, $R_{a_2}^q(l)$ and $R_{b_{1,j}}^{qp}(l)$, $R_{b_{2,j}}^{qp}(l)$ valid values.

Thus, on the basis of taking into account the boundaries of CS subset [1,6] physical packing, there are possibilities for constructing subsets of CS in accordance with (1 -2) and (3 -7). Similarly (1 -2) and (3- 7) the model of the CS subset (dictionary) is specified through aperiodic autocorrelation functions (ACAF). In this case, simplifications are possible. So, the system (1 -2) by analogy can be represented as the NPI system based on aperiodic correlation functions, i.e.

$$r_{a_1}^q(l) \leq \sum_{i=1}^{L-m} W_i^q \left(W_{i+1}^q \right)^* \leq r_{a_2}^q(l), \quad l = \overline{1, L}, \quad m = \overline{1, L}, \quad (13)$$

where $r_{a_1}^q(l)$ and $r_{a_2}^q(l)$ – given but admissible implementations in terms of «dense packing». Systems (1 - 2) and (3- 7) can also be represented through aperiodic cross-correlation functions (CCAF) as a system NPI

$$r_{b_{1,1}}^{qp}(l) \leq \frac{1}{L-m} \sum_{i=0}^{L-m} W_i^q \left(W_{i+1}^q \right)^* \leq r_{b_{1,2}}^{qp}(l); \quad (14)$$

$$l = \overline{1, L}, \quad m = \overline{1, L},$$

$$r_{b_{2,1}}^{qp}(l) \leq \frac{1}{L-m} \sum_{i=0}^{L-m} W_i^p \left(W_{i+1}^q \right)^* \leq r_{b_{2,2}}^{qp}(l); \quad (15)$$

$$l = \overline{1, L}, \quad m = \overline{1, L},$$

where $r_{b_{1,1}}^{qp}$, $r_{b_{1,2}}^{qp}$, $r_{b_{2,1}}^{qp}$, $r_{b_{2,2}}^{qp}$ - permissible in terms of «dense packing», ACAF and CCAF.

4 Solution for CS subset constructing (synthesis) problem

The authors first obtained complex non-linear cryptographic signal systems synthesis method. The construction (synthesis) of CS sets is based on the application of key data, as well as random or pseudo-random processes. The method includes the following steps. [6].

1. Discrete sequences formation using key data and a source of random (pseudo-random) processes.
2. Potential CS statistical properties evaluation [7].
3. Build the required number of potential CS W^q using the system (1) and key data.
4. Finding pairs or subsets CS W^q and W^p , that meet the requirements (3 – 7).
5. Obtaining a cross-correlation functions state matrix of all possible pairs of potential CS, selected by the results of the previous step, and have all the necessary properties.
6. Processing the matrix, consisting in the fact that the formation of the required number of subsets or pairs CS according to (1) and (2) and selection into a subset of sequences satisfying the boundaries «dense packing» (maximum achievable values) for the corresponding correlation functions.

In order to ensure cryptographic security and structural secrecy (complexity) of cryptographic signals, the choice of a symmetric block encryption algorithm with a counter is justified as a source of pseudo-random sequences of characters (1st stage of the method): National cryptographic standard of the block symmetric transformation DSTU 7624:2014, defining the cipher «Kalina» and its modes of operation to ensure the confidentiality and integrity of information [8]. Alternatively, such a source may be offered a source based on the AES algorithm (international standard ISO/IEC 18033). Preference is given to choosing DSTU 7624:2014, taking into account the following factors. Block symmetric ciphers are one of the most common cryptographic primitives. In addition to ensuring the confidentiality (encryption) of the main volumes of information transmitted over the network or stored locally, they are used as a constructive element of other primitives (hashing functions, message authentication codes, pseudo-random sequence generators, etc.). The significance of this cryptographic transformation is underlined by a number of international contests, such as AES, NESSIE, CRYPTTRACK, which were focused on the development of block cipher (as the main goal or as part of a set of promising solutions).

The national standard supports the block size and encryption key length of 128, 256, and 512 bits (the key length is equal to the block size or twice its size), providing a normal, high, and extremely high level of resilience (now it is the only block encryption standard in the world that supports 512 bit symmetric keys). Different standard versions provide flexibility in the choice of parameters for developers of cryptographic protection systems, which makes it possible to obtain both the highest level of performance and the greatest margin of transformation durability. High-level design uses well researched Square-like SPN-structure, used in algorithms AES/Rijndael, Whirlpool, «Stribog» and many others. The cycle transformation is built on the basis of lookup tables (S-blocks) and multiplication by an MDR-matrix over a finite field, providing the necessary cryptographic properties. The use of such design allows us to provide provable durability with respect to differential, linear and other types of cryptographic analysis, while simultaneously providing an efficient implementation for a wide range of software and hardware-software platforms. When choosing the size of the MDR-matrix, the size of the L1 cache of modern and promising processors was taken into account, which made it possible to optimize the speed of the software implementation of the cipher. Ukraine standard provides the greatest non-linearity of Boolean functions, which gives an additional margin of

stability with respect to linear cryptanalysis. In addition, in our opinion, the standard of the block symmetric transformation DSTU 7624:2014 refers to post-quantum algorithms, i.e. - it will provide (at the choice of the corresponding parameters) cryptographic resistance against attacks with the use of quantum computers.

Finding discrete sequences with the necessary correlation functions characteristics is reduced, in fact, to the enumeration of all possible variants of the sequences belonging to a certain set and the selection of those sequences that satisfy the known estimates. Moreover, the computational complexity of such methods turns out to be quite significant. It is known that there is a large group of improved brute force methods, united by the common name «branch and bound method» [9]. The main idea of such methods is to use the finiteness of the set of solutions and to replace their complete search with a reduced (directed) search. Thus, the essence of improved brute force methods is to find optimal solutions for various optimization problems, in particular, discrete and combinatorial optimization. For method implementation procedure for finding estimates (boundaries) being used. The procedure for finding estimates is to establish the boundaries for solving the problem of finding valid values. If the estimate of the subset (parameter) is greater than the boundary of the values of the function of the subset, then the value is excluded from further consideration. The general principle implementation described above has certain difficulties determined by the specifics of the optimization problem being solved. As applied to the problem of synthesizing systems of nonlinear discrete cryptographic signals formulated above, this optimization method is aimed at implementing the procedure «branching», consisting in splitting the set of permissible values of a variable X (scan steps) on subregions (subsets) of smaller sizes. The resulting subregions form a tree called the search tree or the tree of branches and borders. The nodes of this tree are constructed subregions (subsets of the set of values of the variable X). In the course of the research, the authors obtained an improved method for the synthesis of nonlinear cryptographic sequences based on the use of an abbreviated (directed) search based on the application of the method «branches and borders», by excluding from further consideration subregions (implementations of discrete sequences that have side lobes of the correlation function that exceed the boundary established by the researcher), as well as using the properties and mutual relation of ACAF and ACPF, as well as «big» and «small» steps. Simulation modeling of the above sequence synthesis method was carried out and the performance (speed) of which was estimated. As the source of nonlinear cryptographic signals, the data encryption standard of Ukraine «Kalina» was used. In the modeling process, using the above method, sequences were selected with different symbol repetition periods (from 256 to 1024), whose autocorrelation function meets «dense packing» boundary for specified periods. Analysis of the research results showed that this method provides a performance gain in nonlinear discrete cryptographic sequence systems synthesis with given correlation properties from 40 to 60 percent with respect to the method of synthesizing a signal system based on the enumeration of all possible sequence variants. When implementing this method, there may be gaps (losses) in finding the best signals in terms of signal correlation properties. But, as studies have shown, the percentage of such losses is insignificant, and for these periods it is no more than 8 percent.

Table 1. Ensemble properties of complex signals various systems

Signal classes	Sequence period	«Dense packing» boundaries	The number of pairs of sequences that satisfy the boundary
M-sequence	31	9	3
Sequences with 3-level CCPF	31	9	495
Cryptographic signals	31	9	1465137
M-sequence	127	27	36
Sequences with 3-level CCPF	127	17	11610
Cryptographic signals	127	23	47 053
M-sequence	255	36	28
Sequences with 3-level CCPF	–	–	–
Cryptographic signals	255	36	17599
M-sequence	511	63	276
Sequences with 3-level CCPF	511	33	147500
Cryptographic signals	511	63	2666671
M-sequence	1023	100	435
Sequences with 3-level CCPF	1023	65	338000
Cryptographic signals	1023	100	5293538

We will evaluate the ensemble properties of this signal system. It should be noted that nonlinear discrete cryptographic signals, in contrast to the known classes of signals used in various ICS applications, can be synthesized for any discrete signal period values. The synthesis of this signal class is based on the limitations associated with the boundary values of the functions of auto - and cross-correlation signals in periodic and aperiodic modes of information transfer. The volume of nonlinear cryptographic signals system (coding power) is determined, firstly, by the requirements resulting from the use of this signal class (detection and measurement of signal parameters, the data transfer mode of users, etc.), and secondly, by the requirements imposed on from the point of view of such indicators of the effectiveness of the functioning of ICS as the noise immunity of signal reception, information secrecy of the system. The user (owner) of the system, on the basis of these restrictions, must make a compromise decision on the choice of an ensemble of non-linear cryptographic signals with the necessary properties.

For most ICS applications, large signal sets with good cross-correlation properties are of interest. That is why the problem of the signal system synthesis was considered as a complex problem, including signal synthesis with the necessary (for certain conditions) ensemble, correlation, and structural properties.

Table 1 shows the data characterizing the ensemble properties of CS in comparison with M-sequences and sequences with 3-level CCPF.

Analysis of the data presented in Table 1 shows that CS have significantly better ensemble properties compared to M - sequences and sequences with 3-level CCPF.

Table 2 shows the data characterizing CS correlation properties for various periods. In particular, the following are given: the dimension (period) of the cryptographic symbol sequence studied; the boundary values of the side peaks of the autocorrelation functions and the signals number satisfying a given boundary in the CS class for various correlation functions; the smallest side peaks of various correlation functions and their number; signal system volume (including the number of CS pairs satisfying boundary values for the corresponding sequence period), etc.

Table 2. Cryptographic discrete signals correlation properties

№		1	2	3	4	5
Segment	dimension	31	63	127	511	1 023
CS						
Uncertainty	function	9	17	23	59	100
	boundary values					
ACPF	CSs satisfying boundary	7 743	10 868	3482	3819	8 513
	Min R_{smax}	5	9	17	45	77
	CS with min R_{smax}	155	14	51	6	9
ACAF	CSs satisfying boundary	3 622	7 166	1302	1951	6 194
CCPF	Total number of pairs	29 977 024	59 056 712	6 062 162	7 292 380	36 235 584
	Pairs satisfying boundary	1 465 137	12 214 869	47 053	122 835	5 293 538
	Min R_{smax}	5	11	19	51	79
CCAF	Pairs satisfying boundary	14 537 423	54 822 445	1 619 780	3 466 713	35 083 491

Analysis of the data in the table shows the following. So for the sequence period $N = 63$, the number of cryptographic discrete sequences pairs that satisfy the specified boundary value of the maximum side lobes of the CCPF - 17 is 12 214 869. For linear sequences class representative - sequences with a three-level cross-correlation function (Gold sets, which is optimal from the point of view of the cross-correlation functions of signals [1]), the number of pairs of signals satisfying a given boundary is - 975. The excess of CS over an ensemble composed of M-sequences is more than

10^7 . For the sequence period $N = 1023$, the number of cryptographic discrete sequences pairs that satisfy the set boundary value for the cross-correlation functions side lobes (CCF) – 100, is 5 293 538, whereas for linear sequences class representative M-sequences, the pairs number satisfying a given boundary is - 435, i.e. signal system excess is more than 10^5 .

With a slight decrease in the requirements for the limiting value of the maximum lateral peak of the CCF (in fact, reducing the noise immunity of reception), in accordance with which the selection of signals is carried out, the information security performance of the ICS can be significantly improved. So, for the period of the sequence $N = 127$, increasing the border value by 1.2 dB will allow increasing the ensemble volume with $M=11610$ (with limit value $R_{smax}=17$), up to 9 006 648 signals (with limit value $R_{smax}=27$), i.e. more than 700 times.

Thus, by varying the boundary values of the side-lobe level of the corresponding correlation function (taking into account the requirements for a telecommunications system signal reception noise immunity and systems information security point of view), the tasks of achieving the necessary noise immunity values of signal reception and telecommunications signals information secrecy.

The ICS immunity, and one of its components - the structural secrecy of the system, is largely determined by the structural or statistical properties of the signals - data carriers in the system. Statistical properties studies are carried out in the methodology of statistical tests framework based on statistical tests. In our opinion, the most acceptable (for practical use) testing methods are: FIPS PUB 140-1, AIS 20 и AIS 31, NIST 800-90b, NIST 800-22. In the process of research using the methodology NIST SP 800-22 [10] the implementation of a cryptographic sequence of characters has been tested.

The simulation results showed that nonlinear CS statistical properties (in terms of estimated probabilities values) are within the acceptable values. And this, in turn, means that the CSs satisfy the requirements for pseudo-random sequences: unpredictability of following characters, irreversibility, chance, equiprobability, independence, unpredictability, indistinguishability, etc. Essentially, CSs are indistinguishable from random sequences. Thus, using them as a physical data carrier will increase the structural and information secrecy (cryptographic resistance) of ICS.

5 Derived signal systems synthesis based on cryptographic discrete symbol sequences

Among the systems of phase-shifted signals, many are based on Walsh systems. [1-2]. It is known that the auto- and cross-correlation functions of the Walsh sequences have large side peaks. To improve signals correlation properties, derived signal systems (DSS) are formed by multiplying Walsh sequences (source sequences) by a signal that has certain properties (producing a signal), in particular, have small side peaks of the autocorrelation function.

The authors formulated a hypothesis about the possibility of using nonlinear cryptographic sequences as generating ones, the theoretical foundations of which are given in [11].

The method of synthesizing derived signal systems based on the use of CS includes the following steps.

1. The selection of M cryptographic sequences of a fixed period N, with the minimum values of the maximum side lobes ($R_{\max}$) ACPF.
2. A set of Walsh codes (matrix N·N) is formed, in which each row corresponds to a separate code.
3. Perform the multiplication of sequences (each of the lines of the Walsh code of the original sequences) on the cryptographic signal, forming N derived orthogonal signals.
4. Carry out a study of the correlation properties of the obtained derived orthogonal signals (in particular, ACPF, ACAF). To study the functions of cross-correlation, they form a matrix of dimension N·N. The number of such matrices: L·N.

Table 3 shows cryptographic sequences (M = 14), selected from a set of sequences, by the criterion of the minimum values of the maximum side lobes ACPF ($R_{\max} < 10$).

Table 3. CS with minimal ACPF side lobes

1	1110001111101000011111011100110011000101000110101101001001100101
2	1000010010000100101110011010000000110010010000010111001110011101
3	0000100100001001011100110100000001100100100000101110011100111011
4	0000100100001001011100110100000001100100100000101110011100111011
5	0001001000010010111001101000000011001001000001011100111001110110
6	0100100001001011100110100000001100100100000101110011100111011000
7	0000100101110011010000000110010010000010111001110011101100010110
8	0001001011100110100000001100100100000101110011100111011000101101
9	0010010111001101000000011001001000001011100111001110110001011010
10	0100101110011010000000110010010000010111001110011101100010110100
11	0000000010100010011000001111100001101101110001101000010111100101
12	0000000101000100110000011111000011011011100011010000101111001010
13	0000001010001001100000111110000110110111000110100001011110010100
14	010001111000110000010011001000000001101111011100101011000010110

The results of the CCPF DSS study based on cryptographic sequences show that the number of pairs of signals for a period of sequences is 64 characters, for which the values $R_{\max}$ do not exceed 17 (this is the so-called «dense packing» border, achieved in the class of the best, from the point of view of CCF, sequences with a three-level CCPF), 604 pairs (about 30% of the total number of possible combinations of pairs of signals). The number of signals pairs for which values $R_{\max}$ do not exceed 20 – 1577, which is 77% of the total number of signal pairs. With limit $R_{\max} < 25$ the maximum number of selected signal pairs is 1984 (96,8%). Such values $R_{\max}$ have a place for sequences that are most prevalent in modern telecommunications systems M-sequence.

Table 4 shows the results of studies of the correlation functions statistical characteristics of various signal classes, including DSS, when used as generating cryptographic signals. As correlation function statistical characteristics were used:

largest lateral emission values $\frac{R_{\text{max}}}{\sqrt{N}}$; expected value of the emission module $\frac{m_{|R|}}{\sqrt{N}}$; the

value of the standard deviation of emissions $\frac{D_{(R)}^{1/2}}{\sqrt{N}}$ and emission module - $\frac{D_{|R|}^{1/2}}{\sqrt{N}}$. Calculations were carried out for different values of sequence periods (from 30 to 2052).

Analysis of the data given in table. 4, indicates that the values of the maximum lateral emissions of CS, as well as the statistical characteristics of this class of signals are not inferior to the corresponding characteristics of signals based on the use of M-sequences. As follows from the data presented in the table, the statistical characteristics of the DSS are close to the corresponding characteristics for linear and nonlinear signal classes. The values of the maximum lateral peaks of the DSS cross-correlation functions are smaller than those of the linear M-sequences widely used in modern ICS.

Table 4. Statistical characteristics of the correlation functions of various classes of signals

Signal types	Specifications	$\frac{R_{\text{max}}}{\sqrt{N}}$	$\frac{m_{ R }}{\sqrt{N}}$	$\frac{D_{ R }^{1/2}}{\sqrt{N}}$	$\frac{D_{(R)}^{1/2}}{\sqrt{N}}$
Nonlinear characteristic sequences	ACAF	1,6 - 2,4	0,3 - 3,4	1,4 - 7,7	1,9 - 10,8
	ACPF	0,02 - 0,5	0,02 - 0,3	0,03 - 0,3	0,06 - 0,5
	CCAF	1,3 - 3,3	0,5 - 0,7	2,4 - 18,2	3,6 - 27
	CCPF	0,8 - 3,3	0,7 - 0,8	5,8 - 45,3	5,9 - 45,3
DSS	ACAF	0,8 - 2,4	0,4 - 0,5	0,9 - 1	1 - 1,1
	ACPF	0,7 - 2,5	0,2 - 0,7	0,2 - 0,5	0,3 - 0,9
	CCAF	1 - 2,5	0,2 - 0,7	0,2 - 0,5	0,3 - 0,7
	CCPF	1,4 - 2,8	0,2 - 0,7	0,4 - 0,5	0,6 - 0,9
Nonlinear cryptographic characteristic sequences	ACAF	0,7 - 2,5	0,4 - 0,5	0,9 - 1	0,9 - 1,2
	ACPF	0,9 - 2,5	0,3 - 0,7	0,2 - 0,5	0,3 - 0,9
	CCAF	1,2 - 2,7	0,4 - 0,7	0,3 - 0,5	0,5 - 0,7
	CCPF	1,5 - 2,8	0,5 - 0,7	0,3 - 0,5	0,8 - 0,9
Linear M-sequences	ACAF	0,7 - 1,25	0,32	0,26	0,41
	ACPF	$1/\sqrt{N}$	$1/\sqrt{N}$	0	0
	CCAF	1,4 - 5,0	0,54	0,48	0,73
	CCPF	1,9 - 6,0	0,8	0,62	1

Due to the fact that CSs have ensemble properties that are improved compared to other classes of signals, ICS protection indicators against imposing (entering) spurious messages can be improved. At the same time, it should be noted that the use of CS ensures the noise immunity of receiving signals not lower than when applying the above signals based on the linear formation laws.

Perform an assessment of ICS security against the imposition of false messages when used as a manipulating (extending the range) CS. At the same time, we will assume that the system implements a dynamic mode of operation, which implies, among other things, a change in the correspondence of m message bits - 2^m complex signals. Change of compliance is carried out at set time intervals and using a control sequence that meets the requirements of randomness. To provide the necessary noise immunity of signal reception, we will use CS systems that have good correlation properties [1-2,4-5].

The probability of imposing a false message is determined by the ability of the reaction station to determine the law of conformity: m bits of the message - 2^m complex signals, or, in other words, determine the structure (law of formation) of the control sequence establishing the specified correspondence, and is determined from the ratio:

$$P_{\text{imp./message}} = (2^{-k})^n, \quad (16)$$

where: 2^{-k} - the source control sequence possible states number; n - message length expressed in bits.

Note that the number of possible states of the source control sequence (2^{-k}) is determined by an ensemble of discrete sequences, by which the phase of the high-frequency carrier is manipulated to form a phase-shift keyed broadband discrete signal. Table 5 shows the values. $P_{\text{imp./message}}$ for various systems of discrete signals obtained on the basis of M - sequences, sequences with a three-level periodic cross-correlation function (CCPFT) and non-linear cryptographic sequences (NLCS). The message dimension is set to $n = 32$. As the sequence period N were selected: 31, 63, 127, 1023. In the calculations were used data on the ensemble NLCS, are given in table 2.

It must be emphasized that in the calculations $P_{\text{imp./message}}$ (for use in the system NLCS), sequences were selected whose correlation characteristics are close to optimal boundary values («dense packing») from CCPF point of view. Such boundary values are achieved in a class of sequences with a three-level cross-correlation function and constitute $R_{\text{ок. max}} \leq 1,5\sqrt{N}$.

Table 5. The message imposing probability values for different discrete signal systems

Sequence period (N)	$P_{\text{imp./message}}$ values for systems:		
	M-sequence	Sequences with CCPFT	NLCS
31	2^{-96}	2^{-288}	2^{-672}
63	2^{-96}	2^{-320}	2^{-768}
127	2^{-160}	2^{-448}	2^{-640}
1023	2^{-192}	2^{-608}	2^{-704}

As can be seen from the data table, the values $P_{\text{imp./message}}$ for NLCS significantly smaller than in the case of using the most widely used in practice linear classes of signals (M-sequences and sequences with three-level CCPF). In the case when the system does not have strict requirements for the $P_{\text{imp./message}}$ (imitation resistance), but it is necessary to provide increased requirements for signal reception noise immunity and to fulfill high requirements in terms of the structural secrecy of the complex signals used, the values of the maximum lateral peaks of the CCPF can be selected as the limiting values «dense packing». In this case, the volume of the system of signals satisfying this boundary will be less, and accordingly $P_{\text{imp./message}}$ will be higher, but at the same time, the noise immunity of receiving signals will be improved. Thus, when using systems of nonlinear signals, it becomes possible to vary (taking into account the requirements for ICS) indicators of the noise immunity of signal reception — the system is protected from unauthorized data modification (imitation resistance).

The above estimates suggest that in ICS, in which as a method of information exchange, the dynamic mode of changing the correspondence is implemented, m bits of the message - 2^m complex signals and apply nonlinear cryptographic signals, provide high levels of system security from unauthorized data modification and the imposition of false information.

In essence, the presented system of providing imitability is [12-13] a cryptographic system, because it contains all the attributes of such a system: an algorithm for protecting against imposing a false message and hiding the semantic content of a message, based on the implementation of the dynamic mode of ICS functioning and using cryptographic discrete signals as information carriers; algorithm for deciding the truth of the information received; a key system that implements the functions of generating a control sequence for a change of correspondence: the message bit is a complex signal, as well as the generation of cryptographic discrete signals.

6 Conclusions

Information exchange methods used in the information and communication system based on a fixed correspondence: message bit (m bit) signal (2^m signals) in the information channel and use (for a long time) in the synchronization channel of the same broadband signal (and the signals used are constructed using linear laws), do not allow to achieve the necessary values of noise immunity and information security of operation ICS. A comprehensive solution to the problem of ensuring noise immunity and information security of ICS operation can be achieved, including, based on the implementation of a dynamic information transmission mode, in which compliance: message bit - the signal changes over time according to the law, which can be predicted with a probability not exceeding the allowable in the system the values and applications of signals with the necessary correlation, ensemble, structural properties. At the same time, signal systems should be based on nonlinear rules for constructing. For the first time, a method for synthesizing nonlinear cryptographic discrete complex signals was obtained, which uses key data as well as random (pseudo-random) processes and allows you to create signals with the necessary ensemble, structural and correlation properties, which makes it possible to improve ICS performance indicators

impacts. Improvement of these efficiency indicators is achieved, in particular, due to the possibility of forming, using the obtained CS synthesis method, large discrete sequences ensembles of almost any period with the necessary side lobes of the auto, cross and butt correlation functions (for various ICS applications) in periodic and aperiodic modes of operation, as well as the statistical characteristics of the correlation functions (CF), which are not inferior to those of the best, from the CF point of view, linear classes of signals.

The volume of the system of nonlinear CS (coding power) is determined, firstly, by the requirements resulting from the use of this class of signals (detection and measurement of signal parameters, the mode of data transmission of users, etc.) and, secondly, by the requirements imposed on the system from the point of view ICS performance indicators such as noise immunity of signal reception, information security of the system. The user (owner) of the system, on the basis of these limitations, needs to make compromise decisions on the choice of a particular ensemble of nonlinear CS with the necessary properties. In general, the problem of synthesizing nonlinear discrete complex CSs is formulated and solved, the ensemble correlation properties of which can be selected depending on the requirements for noise immunity and information security of ICS. A method for optimizing the synthesis of complex nonlinear CS based on the use of abbreviated (directional) enumeration based on the application of the «branches and borders» method is proposed.

Studies of the DSS properties formed on the basis of nonlinear cryptographic sequences show that more than 30% of the total number of combinations of such DSS pairs have side peaks of the correlation function equal to potentially achievable values, for the remaining pairs the maximum side peaks of CF are less than in widely used linear M sequences. In addition, the resulting DSS have improved structural and ensemble properties compared with orthogonal signals.

Currently, mathematical model and software have been developed that implements methods for synthesizing and studying the properties of non-linear cryptographic signal systems, which is almost ready for possible use as part of prototypes and elements of modern digital communication tools, and allows: to generate non-linear cryptographic signals for almost any period; determine the values of the minimum and maximum lateral emissions of various correlation functions; compare the obtained values with known, potentially achievable boundaries for the corresponding correlation functions; assign to the implementations of the synthesized sequences, as well as the parameters used for the synthesis of signals, unique identifiers that are necessary for optimal signal processing; calculate the statistical characteristics of the various correlation functions of the synthesized signals; carry out studies of the ensemble characteristics of the synthesized signals.

References

1. Sarwate, D., Pursley, M.: Correction to "Crosscorrelation properties of pseudorandom and related sequences." Proceedings of the IEEE. 68, 1554-1554 (1980).
2. Ipatov, V.: Spread Spectrum and CDMA. John Wiley & Sons, Ltd., Hoboken (2006).
3. Joye, M., Tunstall, M.: Fault Analysis in Cryptography. Springer-Verlag, Heidelberg (2012).
4. Chen, J., Ling, B., Feng, P., Lei, R.: Computer cryptography through performing chaotic modulation on intrinsic mode functions with non-dyadic number of encrypted signals. IET Signal Processing. 13, 7-13 (2019).
5. Kuznetsov, A., Kavun, S., Panchenko, V., Prokopovych-Tkachenko, D., Kurinniy, F., Shoiko, V.: Periodic Properties of Cryptographically Strong Pseudorandom Sequences. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 129-134 (2018)
6. Gorbenko, I., Zamula, A., Semenko, A., Morozov, V.: Method for synthesis of performed signals systems based on cryptographic discrete sequences of symbols. Telecommunications and Radio Engineering. 76, 1523-1533 (2017).
7. Goubin, L.: Cryptographic hardware and embedded systems. Springer, Berlin (2006).
8. DSTU 7624:2014. National Standard of Ukraine. Information technologies. Cryptographic Data Security. Symmetric block transformation algorithm (2014).
9. Land, A., Doig, A.: An Automatic Method of Solving Discrete Programming Problems. Econometrica. 28, 497 (1960).
10. NIST 800-22 A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications (2000).
11. Gorbenko, I., Zamula, A., Semenko, A., Morozov, V.: Method for complex improvement of characteristics of orthogonal ensembles based on multiplicative combining of signals of different classes. Telecommunications and Radio Engineering. 76, 1581-1594 (2017).
12. Gorbenko, I., Zamula, A., Morozov, V.: Information security and noise immunity of telecommunication systems under conditions of various internal and external impacts. Telecommunications and Radio Engineering. 76, 1705-1717 (2017).
13. Hameed, A.: High data rate of a novel modulation scheme based on orthogonal chaotic signals. Telecommunications and Radio Engineering. 75, 1657-1663 (2016).

Protection of Information from Leakage by Technical Channels for Sources with Non Range Distribution of Probability

Igor Korobiichuk¹[0000-0002-5865-7668], Serhii Ivanchenko²[0000-0003-1850-9596], Oleksii Havrylenko³[0000-0002-9552-5832], Anatolij Golishevsky⁴[0000-0001-9981-7771], Serhii Hnatiuk⁴[0000-0002-1541-7058], Ruslan Hryshchuk⁵[0000-0001-9985-8477]

¹Warsaw University of Technology, Institute of Automatic Control and Robotics, Warsaw, 02-525, Poland

i.korobiichuk@mchtr.pw.edu.pl

²National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Institute of special communication and information protection, Kyiv, 03056, Ukraine

soivanch@ukr.net,

³National Aviation University, Kyiv, 03058, Ukraine

gavrylav@gmail.com

⁴National scientific research institute of special communications and information protection of Ukraine, Kyiv, 03142, Ukraine

tolyan207@ukr.net, gnatyk-2@i.ua

⁵Sergey Korolyov Zhytomyr Military Institute, Cybersecurity Department of the Research Center, Zhytomyr, 10004, Ukraine

dr.hry@i.ua

Abstract. Made reasoning of the information protection from leakage through technical channels for uneven distribution of source probabilities and set the negative sequences in case of its non-consideration. Shown relations of probability information security risk conditions with power at the input of intercept receiver and given requires analytical values, which allow to assess the risk relative to the current signal/noise ratio and for the given probability of the safety risk of the required values of the maximum allowable indicators, namely the throughput of the technical channel of leakage, the probability of correct reception with possible interception and the signal/noise ratio. These ratios differ from previously known by those that take into account the uneven distribution of the signs probability on output of the source of leakage. Ratios allow automated analysis of information security risk in real time with the use of modern information systems and technologies.

Keywords: information, security, protection, risk, information leak, technical channel of leakage.

1. Introduction

As is well known, functioning of almost modern technical means and systems of information processing and transfer related with quantized electromagnetic energy which circulating by its circuit implementations using electromagnetic fields, electric currents, optical energy, etc. Its – energy, which is the carrier of discrete data or information in a continuous environment and, actually, within its own properties as energy conservation law, almost always accompanied by a list of adverse effects [1-7]. These effects are: unwanted radiation of electromagnetic fields to the environment, which can spread over long distances from hardware of information processing; the impact of these fields to technical means, which have galvanic and other connections with the world; leakage of information to conductive conductors of electric currents: power supply network, grounding system, connecting lines or subscriber network of free access [8].

During processing of information with restricted access by technical means mentioned above is a certain class of threats, which is implemented by using of special means of interception. These special means may have different degrees of complexity and efficiency depending on the value of the information and the interest in outsiders knowing. The value of information is a determining factor for possible efficiency involved in interception technical means and the degree of protection of information from leakage.

The balance of protection and threats determine the security of information, and a measure of balance – its safety guarantee, which as a result of the natural factors influence may be characterized by a probability of safety risk [9-11]. This probability is an integral part, which is used by the Bayesian criterion for direct risk assessment [12-15]. A separate case of this criterion is the Kotelnikov criterion, which uses the construction of an optimal receiver and the mathematical justification of its decisive scheme as the best of all possible receiver circuits [16-19]. The decisive scheme provides an opportunity to substantiate the relationship of probabilistic indicators of the channel with the energy indices of carriers (the signal/noise ratio) at the input of the idealized receiver. Accordingly, the application of this approach to protecting information from leakage through technical channels may allow the justification of signal/noise ratio limitations that are permissible in places of possible interception and provide the desired probability of a security risk.

However, the existing assessments of receivers quality, as rule, have orientation on communication channels. It is hypothetically considered, that the best channels and channel sources are the sources and channels that respectively produce and transmit the largest amount of information [16, 17]. It is obvious that under these circumstances, the assumption concerning the equivalence of source symbols, which uses the existing substantiation of the solving schemes, is quite feasible and permissible.

For receivers that can be considered as means of interception, such an assumption is inadmissible, since for information leakage channels in comparison with communication channels is set the opposite task, namely, the impossibility of leakage of information [16, 17]. The real sources may have a different probability distribution. There-

fore, the security of information should be based not only on the equivalence of source symbols, but also on uneven distribution.

So, the actual task is analyzing the protection of information from leakage by technical channels for uneven distribution of source probabilities, which should provide reliable protection of modern technical means and systems of information processing and transmission.

2. Materials and methods of research

Let the technical channel of leakage be given as a discrete-continuous one (Fig. 1). Let for simplicity as a source of information leakage is used a binary discrete source X , which produces signs x_r , where $r = 1$ and 2 . Each of the signs in the technical processing means is represented by the intermediation of some continuous-time signals implementation $s_1(t)$ and $s_2(t)$. The implementation of each other is different in that, so that the technical means can distinguish them and, accordingly, identify them with signs x_1 and x_2 . On the channel scheme on Fig.1 an element of this identification appears by "modulator".

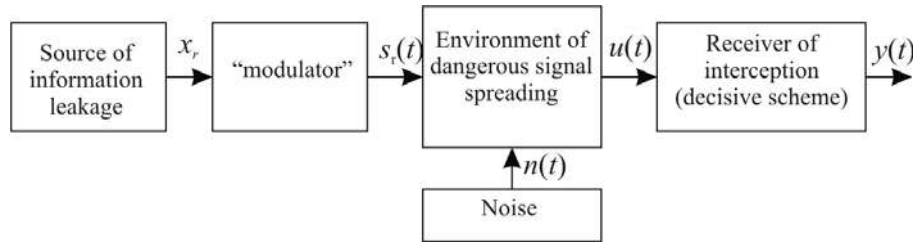


Fig. 1. Discrete-continuous channel as a technical channel of information leakage

Let the implementation $s_1(t)$ та $s_2(t)$ are finite in spectrum and have the same duration T . In the environment of the spreading a dangerous signal, forming a technical channel of leakage, there is a noise that distorts it and prevents interception of information. Noises can have a variety of characters of origin and influence and different degrees of chance and the associated masking properties. The analysis of the literature shows that most distorting effect has additive noise with normal distribution of probabilities [1, 2, 6-8, 14-17]. Unlike the others, this type of noises is quite common in natural environments of signal propagation, has a mathematical description of the density of probability distribution and allows the justification of the security with a proven guarantee of reliability [6, 7, 16, 17].

Other noises also can distort signals, revealing masking property. However, the effect of masking by them is still determined by the proximity with normal distribution law. Otherwise, their presence is defined as a protection margin. Including this, the masking efficiency of these noises is not taken in calculations.

Let the normally distributed noise has additive action to a dangerous signal in the channel, in ideally – white noise $n(t)$. At the output of the channel is formed a mixture which enters the receiver input:

$$u(t) = \mu s_r(t - \tau) + n(t) = c_r(t) + n(t), \quad (1)$$

where $c_r(t) = \mu s_r(t - \tau)$ is attenuated signal with time delay; μ is channel transmission ratio; τ is signal delay time in the channel.

The task of the receiver of interception is to analyze $u(t)$ and making a decision y_r about the sign x_r at the output of the leakage source. Obviously, that the more true this solution, then interception will be more effective and the source of information leakage will be less protected.

Probability of security risk $p_{risk.}$ can match the fate of information, that, without violating of information security, is permissible can flow through by technical channels. In accordance, this part can be providing by bandwidth - the ability of the technical channel of leakage, expressed in relative units:

$$p_{risk.} \leq \frac{C_{TCL}}{C_{TCLmax}}, \quad (2)$$

where for a discrete symmetric channel with transmission reliability q bandwidth is equal to [16, 17, 20, 21]:

$$C_{TCL} = 1 + q \log_2 q + (1 - q) \log_2 (1 - q) \text{ [bit]}, \quad (3)$$

At that, the maximum bandwidth per expression (3) is achieved on condition $q = 1$ or $q = 0$ and equals, minimum, which is desirable for the technical channel of leakage, provided that the probability of the false and correct character transfer is equal $1 - q = q$ (Fig. 2).

Compliance with minimum transmission reliability $q = 0$ and the maximum of bandwidth can be explained by the fact that on the receiving side can be used inverter. It will result in the redistribution of arrows on Fig. 2 and will cause compliance $x_1 \rightarrow y_2$ and $x_2 \rightarrow y_1$. At that q must be replaced $1 - q$ and become equal to 1. Equality same $p = q$ makes the output of the channel completely uncertain, similar to the random character generator.

As follows, having the maximum permissible probability of information security risk, it is possible to determine the maximum allowable bandwidth $C_{TCL max.per.}$ the technical channel of leakage and by solving the equation (3) – the maximum allowed probability of correct interception $q_{max.per.}$ of one binary sign x_r , $r = 1$ and 2.

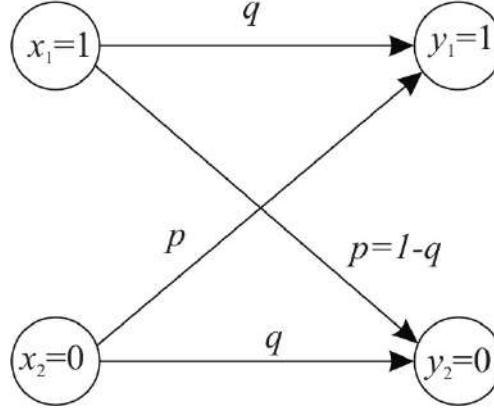


Fig. 2. A graph of states of a discrete symmetric channel without memory

Relative to the decisive scheme of the ideal receiver, by the theory of potential noise immunity for a discrete-continuous channel, the probability of an error, which is determined by the formula:

$$p = F\left(-\frac{1}{2}\sqrt{\frac{P_{\Delta}T}{N_0}}\right), \quad (4)$$

where $F(\dots)$ – Laplace integral:

$$F(\zeta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\zeta} \exp\left\{-\frac{\eta^2}{2}\right\} d\eta, \quad (5)$$

P_{Δ} is the power of the difference signal $c_{\Delta}(t) = c_1(t) - c_2(t)$:

$$P_{\Delta} = \frac{1}{T} \int_0^T c_{\Delta}^2(t) dt, \quad (6)$$

where N_0 – spectral density of noise in the technical channel of leakage.

In accordance, the authenticity of receiving a binary sign will be based on the formula:

$$q = 1 - p = F\left(\frac{1}{2}\sqrt{\frac{P_{\Delta}T}{N_0}}\right), \quad (7)$$

But the formula of estimation of probability of error (4) predicts the equivalence of source symbols $p(x_1) = p(x_2)$. As already noted above, for the modern means of

information processing and transmission, such an assumption can not always be performed, and therefore founded by the formula (4) probability is not always correct.

Provided $p(x_1) \neq p(x_2)$ the decisive scheme must be built by implementing inequality:

$$p(x_l)\lambda_{l/0}(u) > p(x_r)\lambda_{r/0}(u), \quad (8)$$

where $\lambda_{kr/0}(u) = \lambda_{r/0}(u_1, u_2, \dots, u_k)$ – k -dimensional relation of plausibility:

$$\lambda_{kr/0}(u) = \frac{\omega_k(u / x_r)}{\omega_k(u / x_0)}, \quad (9)$$

where l is the index of the correct decision, x_0 is mark of no sign on the output of the source, $\omega_k(u/x_r) = \omega_k(u_1, u_2, u_3, \dots, u_k/x_r)$ is conditional k -dimensional density of distribution of counts receiver reading.

If the frequency spectrum of signs realizations is finite in time and is completely concentrated in the band of frequencies F , then by the Kotelnikov theorem, the number of readings counts can be limited to the number $k = 2FT$, therefore:

$$\lambda_{r/0}(u_1, u_2, \dots, u_{2FT}) = \frac{\omega(u_1, u_2, \dots, u_{2FT} / x_r)}{\omega(u_1, u_2, \dots, u_{2FT} / x_0)}. \quad (10)$$

If in the "modulator" for sign x_0 is associated with so-called zero implementation $c_0(t)$, for example, input channel gets nothing, in accordance with the formula (1) at the output there will be only a noise process:

$$u(t) = c_0(t) + n(t) = n(t). \quad (11)$$

Due to the properties of white noise on the statistical independence of its conditional density of its reading counts, the denominator of the ratio (10) can be expressed in the form of a product of one-dimensional densities of the normal distribution law:

$$\omega(u_1, u_2, \dots, u_{2FT} / x_0) = \prod_{i=1}^{2FT} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u_i^2}{2\sigma^2}} = \frac{1}{(\sigma\sqrt{2\pi})^{2FT}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{2FT} u_i^2\right\}, \quad (12)$$

For r -th implementation using additivity properties of noise in the channel:

$$\begin{aligned} \omega(u_1, u_2, \dots, u_{2FT} / x_r) &= \omega(u_1 - c_{r1}, u_2 - c_{r2}, \dots, u_{2FT} - c_{r2FT} / x_0) = \\ &= \frac{1}{(\sigma\sqrt{2\pi})^{2FT}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{2FT} (u_i - c_{ri})^2\right\}. \end{aligned} \quad (13)$$

Substituting ratio (12) and (13) into the formula (10), we will get:

$$\lambda_{r/0}(u_1, u_2, \dots, u_{2FT}) = \exp \left\{ \frac{1}{2\sigma^2} \sum_{i=1}^{2FT} u_i^2 \right\} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{2FT} (u_i - c_{ri})^2 \right\}. \quad (14)$$

Taking advantage of the property the noise ergodicity, its dispersion can be replaced by the power expressed in terms of spectral density N_0 in the bandwidth F :

$$\sigma^2 = P_z = N_0 F. \quad (15)$$

After inclusion in the ratio (14) replacement (15) and direction $\Delta t \rightarrow 0$ the plausibility ratio will look:

$$\begin{aligned} \lambda_{r/0}(u) &= \lim_{\Delta t \rightarrow 0} \exp \left\{ \frac{1}{N_0} \sum_{i=1}^{2FT} u_i^2 \Delta t \right\} \exp \left\{ -\frac{1}{N_0} \sum_{i=1}^{2FT} (u_i - c_{ri})^2 \Delta t \right\} = \\ &= \exp \left\{ \frac{1}{N_0} \left[\int_0^T u^2(t) dt - \int_0^T (u(t) - c_r(t))^2 dt \right] \right\}. \end{aligned} \quad (16)$$

Substituting (16) into inequality (8) and applying a logarithm for the right and left parts of the inequality by a natural logarithm, we get:

$$p(x_l) \exp \left\{ \frac{1}{N_0} \left[\int_0^T u^2(t) dt - \int_0^T (u(t) - c_l(t))^2 dt \right] \right\} > p(x_r) \exp \left\{ \frac{1}{N_0} \left[\int_0^T u^2(t) dt - \int_0^T (u(t) - c_r(t))^2 dt \right] \right\}$$

or

$$\frac{1}{T} \int_0^T u(t) c_l(t) dt - \frac{P_l}{2} + \frac{N_0}{2T} \ln p(x_l) > \frac{1}{T} \int_0^T u(t) c_r(t) dt - \frac{P_r}{2} + \frac{N_0}{2T} \ln p(x_r). \quad (17)$$

In accordance with the inequality (17) the decisive scheme of an ideal receiver for all r must calculate and compare the right and left parts and for the maximum of found result make a decision y_l about the sign x_l produced by the source. Since the right and left sides of the inequality contain random components, there are probabilities of the existence of this and the opposite inequalities. Obviously, the probability of inequality (17) is the probability of a correct solution, and the probability of the opposite inequality is the probability of error.

The probability of a correct solution on average can be found as a expected value for all realizations r :

$$q = p(x_1) p(y_1/x_1) + p(x_2) p(y_2/x_2). \quad (18)$$

For a binary source, if the correct solution is x_1 , inequality (17) can be converted to the form:

$$\frac{1}{T} \int_0^T c_{\Delta}(t) n(t) dt > -\frac{P_{\Delta}}{2} - \frac{N_0}{2T} \ln \frac{p(x_1)}{p(x_2)}$$

or

$$p(y_1 / x_1) = p \left\{ \xi > -\frac{P_{\Delta}}{2} - \frac{N_0}{2T} \ln \frac{p(x_1)}{p(x_2)} \right\}, \quad (19)$$

where $\xi = \frac{1}{T} \int_0^T c_{\Delta}(t) n(t) dt$ is a normally distributed random variable with expected

value $M[\xi] = 0$ and dispersion $D[\xi] = \frac{N_0 P_{\Delta}}{T}$ [13, 14, 17, 18].

3. The results of research

Let's analyze the nature of throughput for the general case and justify communication of maximum throughput and needed for ensure probability of errors. In this case will suppose, that data, which simultaneously can be processed by technical means and flow by technical channels, can have very different origins, that is generated from Q different information sources with a different syntax and different semantics (Fig. 2).

Expressed with the help of Laplace integrals [19, 20]:

$$p(y_1 / x_1) = F \left(\frac{1}{2} \sqrt{\frac{P_{\Delta} T}{N_0}} + \frac{1}{2} \sqrt{\frac{N_0}{P_{\Delta} T}} \ln \frac{p(x_1)}{p(x_2)} \right) \quad (20)$$

and

$$p(y_2 / x_2) = F \left(\frac{1}{2} \sqrt{\frac{P_{\Delta} T}{N_0}} - \frac{1}{2} \sqrt{\frac{N_0}{P_{\Delta} T}} \ln \frac{p(x_1)}{p(x_2)} \right). \quad (21)$$

It should be noted that if $p(x_1) = p(x_2)$ the formula (18) in conjunction with (20) and (21) will match the formula (7).

In the case of $p(x_1) \neq p(x_2)$ were obtained the quantitative values of the probability of correct reception for multiply-polar, orthogonal signals and signals with zero realization. The character of the probability dependence of the correct receiving from the attitude $\frac{p(x_1)}{p(x_2)}$ presented on Fig. 3.

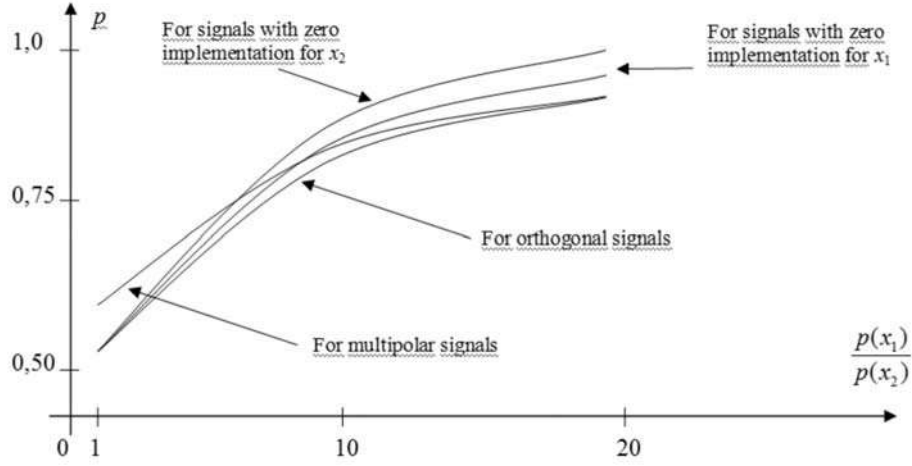


Fig. 3. A graph of the probability of correct receiving of signs related probabilities output source

As can be seen from the graphs, with the same energy conditions on the channel input, provided equal probability symbols on the source output $\frac{p(x_1)}{p(x_2)} = 1$ the security

of the multipolar signals is less (the probability of correct receiving is greater) than for orthogonal signals and signals with zero realization. Provided unequal probabilities of signs $\frac{p(x_1)}{p(x_2)} > 1$ there is a tendency to decrease security (increases of probability of correct receiving).

At that a large variation of the signs probability, for example $\frac{p(x_1)}{p(x_2)} = 20$ for signals with zero realization, leads to increased probability of correct

reception compared to orthogonal and different polar signals. Probabilities is receiving for the last two with this distortion are same.

It should also be noted, that with an increase in the distortion of probabilities of signs of the probability of correct receiving approaching the greater probability of signs. This is not difficult to explain using the formula (18), taking $p(y_1/x_1) \approx 1$ and $p(y_2/x_2) \approx 0$. So, according to the formula (18):

$$q \approx p(x_1) \times 1 + p(x_2) \times 0 = p(x_1).$$

An analysis of the dependence of the probability of a correct reception on the signal/noise ratio under the condition of strong distortion of the signs probabilities (Fig. 4).

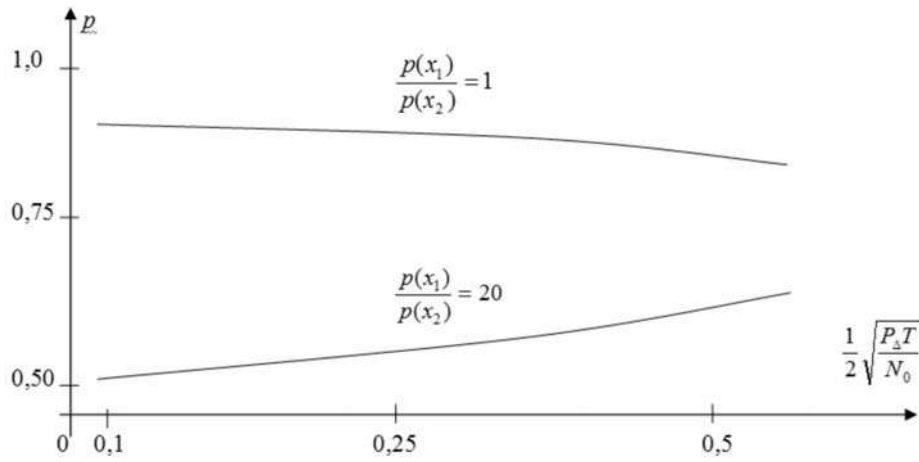


Fig. 4. Graph of the dependence of the probability of correct reception on the signal / noise ratio at the receiver input

It is shown that with increasing the level of noise in relation to the signal under the condition of equal probabilities of signs at the output of the source $\frac{p(x_1)}{p(x_2)} = 1$ probability of correct receiving usually decreases. In the case of strong distortion of signs probabilities, for example, when $\frac{p(x_1)}{p(x_2)} = 20$, under the same conditions with an in-

crease in the level of noise in relation to the signal there is an opposite effect – increasing of correct receiving probability. That is, with a strong distortion of the signs probability increasing the noise over the signal does not improve the masking of a dangerous signal, but on the contrary - improves the conditions for interception.

Thus, made reasoning of the information protection from leakage through technical channels for uneven distribution of source probabilities and set the negative sequences in case of its non-consideration. Shown relations of probability information security risk conditions with power at the input of intercept receiver and given requires analytical values. These resulting formulas provide for the use of intermediate indicators: the throughput of the technical channel of information leakage and the probability of correct reception of messages relative to the ideal receiver.

The resulting ratios allow estimating the information security risk related on the current signal/noise ratio for the given probability of risk the required values of the maximum allowable indicators, for example, the throughput of the technical channel of leakage, the probability of correct receiving with possible interception and the signal/noise ratio. These ratios differ from previously known by those that take into account the uneven distribution of the signs probability on output of the source of leakage. This unequal signs probability can take place in practice, and therefore should be

taken into account in the calculation of the protection of real hardware, information processing and transmission systems.

Conclusions

Ratios allow automated analysis of information security risk in real time with the use of modern information systems and technologies. Their use has obtained practical results that show the dependence of the probability of true interception from ratio of probabilities of signs at the source output and the signal/noise ratio at the input of the ideal receiver. It is shown that in the case of a strong distortion of the signs probabilities, the masking influence of the noises changes to the opposite than for the equal probabilities of these signs. That is, increasing the level of noise does not increase, but reduces security.

The above is advisable to take into account in the calculation of security information from its leaks.

References

1. Buzov, G.A., Kalinin, S.V., Kondratev, A.V.: Protection of information from leaks through technical channels. Goryachaya liniya, Telecom: Moskva (2005).
2. Parshutkin, A.V., Levin, D.V., Zaytsev, S.A., Egin, A.V.: Application of structural interference for data protection from information Leakage in the stray electromagnetic radiations channel. SPIIRAS Proceedings, 3 (58), pp. 160-181 (2018) doi: 10.15622/sp.58.7
3. Kuhn G. Compromising emanations: eavesdropping risks of computer displays. This technical report is based on a dissertation submitted June 2002 by the author for the degree of Doctor of Philosophy to the University of Cambridge, Wolfson College, (2002) <http://www.cl.cam.ac.uk/techreports>.
4. Kuhn, M.G.: Electromagnetic eavesdropping risks of flat-panel displays. Lecture Notes in Computer Science, 3424, pp. 88-107 (2005)
5. Kuhn, M.G.: Optical time-domain eavesdropping risks of CRT displays. Proceedings - IEEE Symposium on Security and Privacy, 2002-January, art. no. 1004358, pp. 3-18 (2002) doi: 10.1109/SECPRI.2002.1004358
6. Lenkov, S.V., Peregudov, D.A., Horoshko, V.A.: Methods and means of information protection. Tom I. Unauthorized receipt of information. Ariy: Kyiv (2008)
7. Qiu, J., Li, H., Zhao, C.: Cancelable palmprint templates based on random measurement and noise data for security and privacy-preserving authentication. Computers and Security, 82, pp. 1-14 (2019) doi: 10.1016/j.cose.2018.12.003
8. Korobiichuk, I., Dobrzahansky, O., Kachniarz, M.: Remote control of nonlinear motion for mechatronic machine by means of CoDeSys compatible industrial controller. Tehnički vjesnik/Technical Gazette, Vol. 24/No. 6, pp. 1661-1667 (2017) doi: 10.17559/TV-20151110164217
9. Information technology. Security techniques. Information security management systems. Requirements [ISO/IEC 27001:2013].
10. Korobiichuk I.V., Hryshchuk R.V., Horoshko V.O., Hokhlacheva Yu.E. Self-diagnostics of complex systems with a software-configurable structure. Informatics and Mathematical Methods in Simulation, vol 8, No. 1, pp. 36-47 (2018)

11. Korobiichuk, I., Hryshchuk, R., Mamarev, V., Okhrimchuk, V., Kachniarz, M.: Cyberattack Classifier Verification. International Conference on Diagnostics of Processes and Systems DPS 2017: Advanced Solutions in Diagnostics and Fault Tolerant Control, pp. 402-41 (2018) doi: 10.1007/978-3-319-64474-5_34
12. Ivanovsky, R.I.: Theory of probability and mathematical statistics. BHV: Petersburg (2008)
13. Niyato, D., Hossain, E.: A queuing-theoretic and optimization-based model for radio resource management in IEEE 802.16 broadband wireless networks. IEEE Transactions on Computers, 55 (11), pp. 1473-1488 (2006) doi: 10.1109/TC.2006.172
14. Isakov, V.N.: Statistical theory of radio engineering systems. ARI: Moskva (2007)
15. Couillet, R., Debbah, M.: Random matrix methods for wireless communications. Random Matrix Methods for Wireless Communications, 9781107011632, pp. 1-539 (2011) doi: 10.4324/CBO9780511994746
16. Burachenko, D.L., Zavaryn, H.D., Kliuev, N.Y., et. al.: General Theory of communication. VAS: Leningrad (1970).
17. Lee, M., Neifeld, M.A., Ashok, A.: Capacity of electromagnetic communication modes in a noise-limited optical system. Applied Optics, 55 (6), pp. 1333-1342 (2016) doi: 10.1364/AO.55.001333
18. Ivanchenko, S.O.: Justification safety risk information about its security from leaking by technical channels. Scientific and technical digest "Legal, regulatory and metrological support of information security in Ukraine", NTUU "KPI" SRC "Tezis": Kyiv, № 1 (31), pp. 9 – 13 (2016)
19. Duc, A., Faust, S., Standaert, F.-X.: Making masking security proofs concrete: Or how to evaluate the security of any leaking device. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9056, pp. 401-429 (2015) doi: 10.1007/978-3-662-46800-5_16
20. Fink, L. M.: The theory of transfer of discrete messages. [2-d edition], Sov. Radio: Moskva (1970)
21. Kulkarni, A.N., Bukate, R.R., Nanaware, S.D.: Study of Various Attacks and Routing Protocols in MANETS. 2018 International Conference on Information, Communication, Engineering and Technology, ICICET 2018, art. no. 8533696 (2018) doi: 10.1109/ICICET.2018.8533696
22. Bronshtein, Y.N., Semendiaev, K.A.: Handbook on mathematics for engineers and students of high schools. Nauka: Moskva, Ch. ed. Phys-Math. Lit. (1986)
23. Ram, M., Davim, J.P.: Mathematics applied to engineering. Mathematics Applied to Engineering, pp. 1-210 (2017) doi: 10.1016/C2015-0-06715-6

Evaluation of testing assignment for system level self-diagnosis

Viktor Mashkov¹[0000-0001-9817-3388], Jiri Fiser¹[0000-0002-5404-1727],
Volodymyr Lytvynenko²[0000-0002-1536-5542], Maria Voronenko²[0000-0002-5392-5125]

¹University J.E. Purkyne, Usti nad Labem, Department of Informatics, Usti nad Labem,
Czech Republic,
viktor.mashkov@ujep.cz, jf@jf.cz

²Kherson National Technical University, Department of Informatics & Computing Technology, Kherson, Ukraine
immun56@gmail.com, mary_voronenko@i.ua

Abstract: The research concerns a task of comparing and assessing testing graphs in context of system level self-diagnosis. System level self-diagnosis aims at diagnosing systems composed of units, with the requirement that they are able to test each other by exchanging information through available links. A set of tests performed in a system can be represented as a testing graph. The obtained testing graph can be assessed based on the quality of diagnosis that it allows to achieve. In the research, we suggest the method allowing different testing graphs to be assessed and compared. The method uses the characteristic numbers. We have shown how such numbers can be computed.

Keywords: complex systems, self-diagnosis, probabilistic algorithm, decision rule

1 Introduction

Research in the area of system level self-diagnosis started in the late 60s of the last century. Since that time there has been done a great amount of research. Achievements in theoretical research gave impulse to practical implementation and enabled a broadening of the application domains of system level self-diagnosis. Initially, system level self-diagnosis was applied in complex multiprocessor systems and then it gradually spread to distributed systems, different types of networks, multi-agent systems [1, 2, 3, 4] etc. There are four main issues, which form the context of system level self-diagnosis [5].

The first issue is the system testing assignment that defines the possible set of tests among the system units. The second issue is the assumptions underlying the diagnosis algorithms. Particularly, these assumptions tackle the possible faulty sets and the test results. These assumptions have direct impact on achievable quality of diagnosis (i.e., correctness and completeness).

The third issue is the organization of test performance. It concerns the tasks of test scheduling, test repetition, test performing (random or deterministic), etc. The fourth issue relates to the problem of determining the unit(s), which will perform a diagnosis algorithm and/or will provide environment with the information about system state.

Usually, the first and second issues are considered together, and they can be viewed as theoretical ground of system level self-diagnosis. In the theory of system level self-diagnosis, three main problems were formulated. They are characterization, diagnosability and diagnosis problems [6].

The characterization problem (i.e., the problem of testing assignment) was the first one which was considered in context of system level self-diagnosis. The work of Preparata et al. [7] studied the requirements to the system testing assignment. In this research, the authors have introduced the diagnosability measure which, to a great extent, depends on the system testing assignment. Particularly, in this research, they proved that for correct diagnosis of at most t faulty units there must be satisfied the following two requirements: a system must have n units, where $n \geq 2t + 1$; each system unit must be tested by at least t distinct other units. When these requirements are not satisfied correct diagnosis is not guaranteed. Nevertheless, there is a probability (sometimes great enough) that in this case it is also possible to achieve correct diagnosis. In this research, we suggest new diagnosability measure which allows comparing system testing assignments and evaluating the formed testing graphs obtained after performing a set of tests.

2 Formal problem statement

System level self-diagnosis is based on the results of tests performed by system units. System units test each other either according to predefined schedule or randomly.

One of the ways of how to express a testing assignment consists in presenting it as a testing graph. Testing graph is a basis for a model of test performance. This model can be used for simulation of tests execution and obtaining a syndrome. Syndrome is an input for a diagnosis algorithm (see Fig. 1).

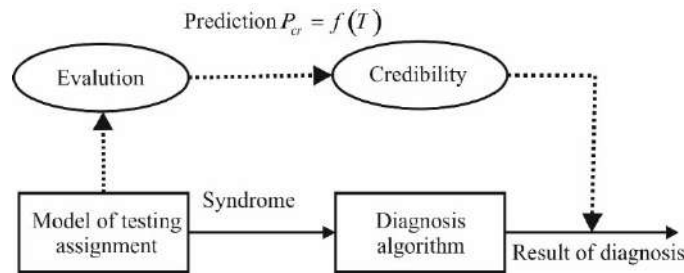


Fig.1. The role of task of evaluation of testing assignment

Different testing assignments, T (i.e., testing graphs) allow obtaining different credibility of diagnosis result (i.e., probabilities of correct diagnosis result, P_{CT}). To

predict the credibility of possible diagnosis results it is needed to evaluate the available testing graph. It can be expressed as $T \Rightarrow P_{CT}$. The problem of computing a credibility of diagnosis results consists in determining the function f , where $P_{cr} = f(T)$. Characteristic numbers C_k can be used for computing f .

3 Literature review

System level self-diagnosis (SLSD) was introduced by Preparata et al. [7] and has been deeply investigated in literature. There are four main issues that form the context of SLSD [8]. The first issue is the testing assignment that defines the possible set of tests among the system units. The second issue is the assumptions underlying the diagnosis algorithms. The third issue is the organization of tests performance. The fourth issue relates to the problem of choosing the diagnostic nucleus. The first and second issues can be considered as theoretical ground of SLSD. In the theory of SLSD, three basic problems were formulated [6, 9]. These are characterization problem, diagnosability problem and diagnosis problem. The problem of testing assignment (i.e., characterization problem) was the first task that was considered in context of SLSD. The work of Preparata et al. [7] determined the requirements to the testing assignment of a system. Diagnosis problem concerns the task of determining a fault set from an allowable family, for a given testing assignment, fault model, and syndrome [8].

There have been developed many algorithms allowing to identify a fault set uniquely (when some assumptions are made about the faulty units). The main task while developing these algorithms is to reduce their complexity. Among the most efficient algorithms there could be named algorithm proposed by Dahbura and Masson [10] which has $O(N^2.5)$ time complexity and $O(t^3 + |E|)$ algorithm suggested by Sullivan [11]. Both algorithms were developed for t -diagnosable systems under the symmetric invalidation model and when permanent fault are allowable only. There are many special classes of t -diagnosable systems that support more efficient diagnosis techniques than above mentioned ones, and this is reason to believe that an $O(|E|)$ diagnosis solution exists for all t -diagnosable systems. Preparata et al. defined the $D\delta.t$ structure in which unit u_i tests u_j if and only if $j - i = \delta m \pmod{n}$, where $m = 1, 2, \dots, t$. Meyer and Masson [12] gave $O(nt)$ solution to the case of $\delta = 1$.

In real complex systems, units are not necessarily homogeneous and can operate under different conditions. Therefore, units can have different levels of their reliability. This fact can be accounted for by assigning probabilities to the unit states. Probabilistic approach to system level self-diagnosis doesn't deal with such problems as t -diagnosability and testing assignment. Among the first who investigated the probabilistic algorithms were H. Fujiwara and K. Kinoshita [13]. Probabilistic algorithms are based on the computing of the posterior probabilities of system unit states, upon which the decision about the system state is made.

4 Testing graph

Tests in a system can be performed:

- either in accordance with a preset schedule (i.e., defined a priori)
- or in an adapted manner when, at the beginning, the tests are performed in accordance with defined a priori testing assignment.

Once a unit is diagnosed as fault free, the tests it performs are considered reliable, and therefore, any other units should only be tested ones by this fault-free unit to correctly determine its status. Thus, the testing assignment is adapted such that units diagnosed as fault-free perform all the testing in the system [14]:

- or entirely randomly (i.e., from the beginning to the end of testing);
- or adaptively randomly.

At the beginning, all units are engaged in tests performing. Tests are performed randomly. Once a test reset takes the value of 1, the units participated in this test (so-called suspected pair) should only be tested by other system units (i.e., should not perform tests on other units). The choice of each pair of units for testing is performed randomly.

Testing graph is a convenient form for presenting the tests performed in a system.

Testing graph is a directed graph $G = (V, E)$. Each vertex $v_i \in V$ of G represents unit u_i and edge $e_{ij} \in E$ represents the test which is performed by unit u_i on unit u_j .

Having obtained the testing graph, we can investigate the diagnosis properties which this graph possesses. Based on these diagnosis properties, it is possible to compare different testing graphs and find the best one from a certain criteria.

It is easy to show that diagnosis result depends on number of faulty units. Let t be the total number of faulty units that are allowable in the system. For some values of t , it is always possible to construct the testing graph, which will ensure the correct and full diagnosis whichever syndrome is obtained.

Syndrome is a set of test results. Test result is represented by binary variable r_{ij} , such that $r_{ij} = 1$ if unit u_j has not passed the test, and $r_{ij} = 0$ otherwise.

Correct diagnosis means that the detected faulty units are indeed faulty units. Full diagnosis means that every faulty unit is identified.

For the testing graphs that ensure unique diagnosis [5] of at most t faulty units (so called t -diagnosability) the following conditions should be satisfied [7]: each vertex of the graph should have at least t incoming edges and there should not be multiple edges.

For example, a testing graph may satisfy the requirements for t -diagnosability for $t = 1$ and for $t = 2$, but fails to be t -diagnosable for $t = 3$. The maximum value of t for which a testing graph is t -diagnosable is denoted as t_{max} . Many t -diagnosable testing graphs (for $t \leq t_{max}$) can exist. Thus, the task arises to determine those testing graphs which have the least number of edges for providing t -diagnosability.

Definition: Testing graph is t -optimal if it contains minimal number of edges and at the same time it ensures t -diagnosability.

For the value $t = t_{max}$, t -optimal testing graph could be simply named as optimal. The number of edges of t -optimal testing graph can be easily computed as follows:

$$l = tN \quad (1)$$

Thus, the number of edges of optimal testing graph is equal to:

$$l = t_{max}N = \left\lfloor \frac{N-1}{2} \right\rfloor N \quad (2)$$

Whichever testing graph that contains more than l edges is considered as redundant testing graph. It is also possible for the given value t to construct instances of testing graphs which provide certain diagnosis properties, e.g., capability to detect certain number of faulty units. For example, testing graph (see Fig. 1.a) is t -optimal for $t = 1$ since it provides unique diagnosis of any faulty unit, and it has minimal total number of edges which is needed for such diagnosis.

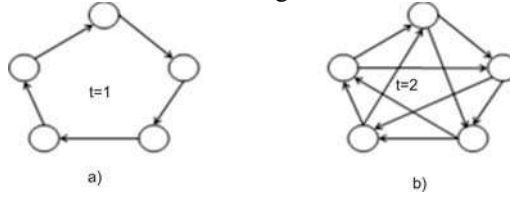


Fig. 2. Optimal testing graphs

Testing graph in Fig. 1.b has $t = 2$ and provides unique diagnosis of any two faulty units. Moreover, this graph is optimal since $t = t_{max}$.

All testing graphs depicted in Fig. 2 provide unique diagnosis of the same number of faulty units (particularly, $t = 1$) but have different total number of edges.

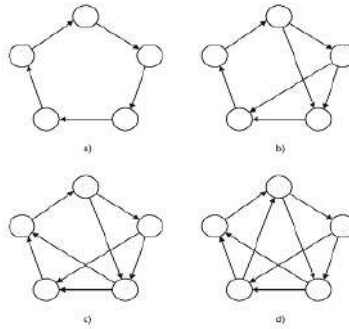


Fig. 3. Examples of the testing graphs with $t = 1$

5 Assessment of testing graphs

The method suggested by Preparata et al. [7] for assessing testing graphs does not allow comparing the testing graphs and determining more precisely their diagnosis properties. Only parameter t and its maximum value are taken into account. Besides parameter t , there also exist further criteria for evaluating diagnosis properties of testing graphs, which make it possible to assess and compare the testing graphs.

For example, to each testing graph there could be assigned the value of probability, which will reflex the fault detection property of testing graph [15, 16, 17]. It can be a probability that a syndrome (obtained after performing all tests, which are depicted in testing graph) is sufficient to detect all possible faulty units, PFD. It is only assumed that a fault-free unit is always able to detect a faulty unit (i.e., test covering is 100%). This probability can be also interpreted as probability that a system is fault-free when obtained syndrome contains only zero test results.

The probability PFD can be calculated as a sum of probabilities of the events when a fault-free units test all the other units. In the testing graph, it means that vertices which correspond to the fault-free units have edges directed to all remaining vertices.

The computation of these probabilities can be explained with help of a simple example. Let the testing graph be the one as shown in Figure 3.

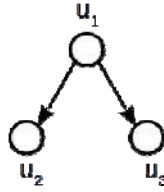


Fig. 4. Exemplary testing graph

For this testing graph there is seven $(= 2^3 - 1)$ possible combinations of faulty-free units.

The combination without faulty-free unit occurring with probability P_U^3 , where P_U is probability of unit faulty state (it is assumed that all system units have the same probability), is irrelevant because empty set of fault-free units cannot test the other system units.

More useful are three situations with one fault-free unit. In Fig. 4, this faulty-free unit is depicted by white vertex:

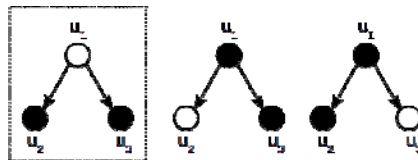


Fig. 5. Exemplary situations with one faulty free unit

Each of these situations occurs with probability $(1 - P_U)P_U^2$, but only one of them corresponds the condition that fault-free units test all the other system units. Two other situations do not allow to perform correct diagnosis of system units.

The probability of all satisfactory situations with one fault-free unit is:

$$P(A_1) = (1 - P_U)P_U^2 \cdot 1 \quad (3)$$

The probability of each situation with two faulty-free units is $(1 - P_U)^2 P_U$. There are three such situations (see Fig. 5):

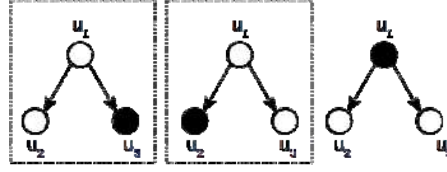


Fig. 6. Situations with two fault-free units

In the first situation, fault-free units u_1 and u_2 test the remaining unit (i.e., unit u_3). Therefore its probability $(1 - P_U)^2 P_U$ should be considered as one of the summands for computing probability P_{FD} . In a similar way, the probability of the second situations (units u_1, u_3 test unit u_2) should be taken into account. The third situation (units u_2, u_3 are fault-free) does not guarantee correct diagnosis of the remaining units (and, therefore, it is not a summand for computing probability P_{FD}).

The probability of all satisfactory situations with two faulty free units is:

$$P(A_2) = (1 - P_U)^2 P_U \cdot 2 \quad (4)$$

Now, there is only one situation left. Particularly, the situation when all units are faulty-free. This case also leads to correct diagnosis. The probability of this situations is $(1 - P_U)^3$ and it is always included in computing probability P_{FD} regardless of testing assignment:

$$P(A_3) = (1 - P_U)^3 \quad (5)$$

The total probability PFD can be generally expressed by the following sum:

$$P_{FD} = \sum_{k=1}^N P(A_k) = \sum_{k=1}^N (1 - P_U)^k P_U^{N-k} C_k \quad (6)$$

where C_k is the number of options to choose the subgraph with k vertices from which all the remaining vertices are achievable (C_N is always 1).

For the considered testing graph $C_1 = 1$ (only one subgraph is satisfactory for correct diagnosis) and $C_2 = 2$ (two subgraphs are satisfactory). Therefore, the total probability is:

$$P_{FD} = (1 - P_U) P_U^2 \cdot 1 + (1 - P_U)^2 P \cdot 2 + (1 - P_U)^2 \quad (7)$$

For given $P_U = 0.1$, we receive $P_{FD} = 0.9$.

Let's make some changes in this testing graph. Particularly, edges from u_3 to u_2 and u_1 are added, which makes testing graph 1-diagnosable (see Fig. 6).

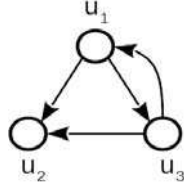


Fig. 7. Exemplary testing graph with $t=1$

For this graph, $C_1 = 1$ (both units u_1 and u_3 test all remaining units) and $C_2 = 3$ (all sets with two units test remaining unit). Therefore, for $P_U = 0.1$ we receive probability $P_{FD} = 0.9 \cdot 0.1^2 + 0.9^2 \cdot 0.1 \cdot 3 + 0.9^3 = 0.99$.

The graph in Figure 7 has six edges (every unit tests all remaining units).

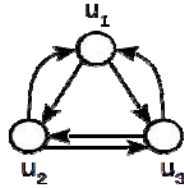


Fig. 8. Testing graph with six edges

In this graph, number C_1 is increased up to 3 (from each vertex of the testing graph now it is possible to directly reach all the remaining vertices of the graph). Clearly, number C_2 is also equal to 3. The value of P_{FD} is $0.9 \cdot 0.1^2 \cdot 3 + 0.9^2 \cdot 0.1 \cdot 3 + 0.9^3 = 0.999$.

This graph has the same value t (particularly $t=1$) and provides the same 1-diagnosability. However, from the point of probability PFD, it is evident that the graph in Figure 6 has better diagnosis properties ($P_{FD} = 0.999$) than the graph in Figure 5 ($P_{FD} = 0.99$).

For comparing the probabilities P_{FD} , we need to have the numbers C_k , $k=1,2,\dots,n$ to be computed in advance. These numbers are called as characteristic numbers.

Definition: Characteristic numbers C_k , $k=1,2,\dots,n$ are the numbers of choices of k vertices (resp. sub-graphs) from the testing graph so that all the remaining vertices of the graph are directly reachable.

For more complex graphs the characteristic numbers can be computed from the modified adjacency matrix of the testing graph.

Adjacency matrix of a testing graph $M = (a_{ij})$ has the following entries:

$$a_{ij} = \begin{cases} 1 & \text{if unit } u_i \text{ tests unit } u_j \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Modified adjacency matrix is derived from the adjacency matrix by way of setting the values of 1 alongside the main diagonal. An algorithm for computing the characteristics numbers for particular testing graph have to account all the combinations of choices a unit(s), from which all the remaining ones are directly reachable, according to the following expression:

$$C_k = \sum_{c \in C(k,n)} \left[\prod_{j=1}^n (a_{c_1j} \vee a_{c_2j} \vee \dots \vee a_{c_kj}) \right] \quad (9)$$

where $C_k, C(k,n)$ is set of all k -combination i.e. all subsets of set $\{1, 2, \dots, n\}$ with k distinct element.

6 Computation of characteristic numbers

The effectiveness of calculation of probability P_{FD} for a testing assignment strongly depends on implementation of computation of characteristic number C_k for this testing assignment.

Modern programming languages make possible straightforward computation of equation 3.2. For example in Julia programming language [18], the program is almost identical to its mathematical form (where parameter a is a modified adjacency matrix).

```
function C(a::Matrix{Int}, k::Int)
    n = size(a, 1) # size in the first = second dimension
    return sum(prod(reduce(|, a[c,j]) for j = 1:n)
               for c in combinations(1:n, k))
end
```

The outer summation iterates over sequences of every k -combinations of set $1, 2, \dots, n$ (set is generated by range $1:n$) produced by function combinations. The individual combination (denoted as c) is vector of integers. The computation of inner logical summation $(a_{c_1j} \vee a_{c_2j} \vee \dots \vee a_{c_kj})$ of selected items in columns is provided by discontinuous indexing (index is vector of integers) and by reduction (folding) using bitwise or operation.

This initial and naïve implementation is depicted in Figure 9 (example of computation of characteristic number C_2 for system with tree units).

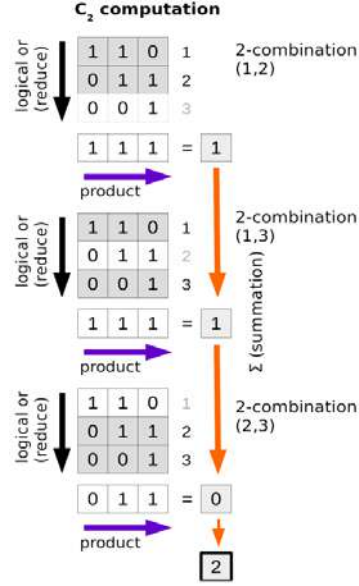


Fig. 9. Straightforward computation of character number

This implementation is compact but it has some shortcomings such as:

- effective iterator over all k -combination is not available in most programming languages (e.g. *Matlab* or *Java*);
- it does not use any form of parallelism (including bit-wise parallelism which is supported by all processors);
- the computation does not utilize already calculated values (outputs).

The first two shortcomings can be eliminated by iteration over k -combinations which are represented by bit patterns. These patterns are formed by string of bits containing k 1's and $n-k$ 0's. Sequences of all k -combinations are special cases of combinatorial Gray codes [19] (generalization of commonly used and well-known Gray code) and they are realizable by elementary bitwise operations at processor level (e.g. bit complement, bitwise OR, AND and shifts).

In our implementation we use simple generator [20] which utilizes only basic bit-wise operations as well as operation of counting trailing zeros. This operation is often directly and efficiently supported by current CPUs (for example, `trailing_zeros` function is directly translated to BSF instruction in x86-64 platform [21] by Julia programming language).

The generator of k -combinations Gray codes in Julia has a form of iterator function. For pattern v it returns pair (v, w) where w is next patterns.

function Base.next(c:: Combinations {Int}, v:: Int)
 $t = v \mid (v - 1)$

```

w = (t + 1) | (((~t & ~t) - 1) >> (trailing_zeros(v) + 1))
return (v,w)

```

end

At every step of iteration over bit patterns, only relatively simple operations are performed. The parameter *pa* is a modified adjacency matrix packed as array of integer values each of which corresponds to one column of the original matrix. For example sample matrix from Figure 5 is represented as array [6, 8, 5] i.e. [100, 110, 011] in binary notation. Parameter *gcode* is again bit pattern (integer) representing a selection of modules of diagnostic graph (the so-called subgraph).

```

function cover_test (pa:: Vector {Int64}, gcode :: Int64)
  for column in pa #iteration over packed column
    if column & gcode == 0
      return false
    end
  end
  return true #OK subgraph covers all remaining vertices
end

```

Bitwise AND operator in this code replaces slow discontinuous indexing of the original code. The reduction by logical OR is substituted by one instruction of integer equality (operator `==`). Product is performed by for each-loop which realizes short circuit evaluation (the first occurrence of zero exits loop).

This new implementation is illustrated on Figure 10 (only one of summation is depicted).

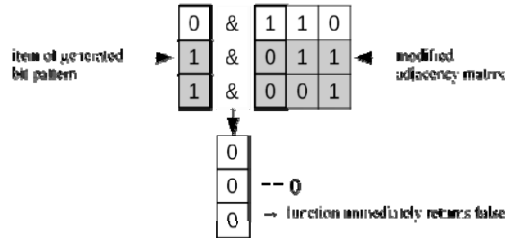


Fig. 10. Fragment of suggested computation of characteristic number C_2

Further acceleration of algorithm is possible by utilization of trivial assertion: if vertices of subgraph of a testing graph cover all the remaining vertices of this testing graph then the vertices of any graph which include this subgraph (the so-called super-graph) also cover the remaining vertices.

The improved algorithm computes characteristic numbers all together concurrently i.e. it returns vector of characteristic number $[c_1, c_2, \dots, c_n]$.

The first tested subgraph in this algorithm contains one unit and it is represented by bit pattern with 1 on the leftmost position (e.g. bit pattern 1000 for system with

four unit). If this subgraph covers the remaining vertices (i.e. function `cover_test` returns value true) then `c1` is incremented and its supergraphs containing 2 to n vertices (e.g. 1100, 1101, 1111) are taken in account in increment of values $c_2, \dots, c_n$ (number of subgraphs is determined by appropriate binomial coefficient). These supergraphs need not be consequently iterated and tested in the following phases of computation. Otherwise, the next subgraphs with two units represented by 1's on the leftmost position is tested (e.g. pattern 1100).

When all supergraphs of the first one-vertex subgraph are considered (i.e. directly tested or automatically included in bulk) the next one-vertex subgraph is tested (eg. 0100).

This implementation requires additional memory for generation of sequences of bit patterns but generated combinatorial Gray codes take into consideration partial ordering of combination by supergraph relation i.e. codes are hierarchical in some sense.

Optionally the precomputed table of binomial coefficients can be used. The asymptotic space complexity of algorithm is $O(k)$ which is acceptable because k is relatively small (time complexity is still exponential). The code is more complex but it is still usable in small devices.

```
function Cvec(a:: Matrix {Int}, sorting :: Bool)
    pa = pack(a) # packing of matrix to vector of bit patterns (i.e integer values)
    if sorting # sort by 1's count (i.e. by outdegree of given vertex), see next paragraph
        sort!(pa , by= count_ones)
    end
    n = length (pa) # number of vertices
    binom = binomials(n) # populating of matrix of binomials coefficients
    p = zeros0(Int , n + 2) #vector of indexes of the leftmost positions of ones (indexed
    from zero)
    gc = zeros0(Int64 , n + 1) # vector of subgraph bit patterns (zero based indexing)
    c = zeros(Int , n) # vector of characteristic numbers (one based indexing)
    p[0] = n + 1 # p[0] is only formal stop position
    k = 1 # number of vertices in subgraph (= number of 1's in patterns)
    while k > 0
        p[k] += 1 # shift leftmost 1-bit position to the right
        if p[k] == p[k - 1] # if it is not possible
            p[k] = 0
            k -= 1 # return to iteration over (k-1)-vertex subgraph
        else
            gc[k] = gc[k - 1] | (1 << (p[k] - 1)) # bit pattern of k-vertex subgraph
        end
        if cover_test (pa , gc[k])
            c[k] += 1 # it counts this subgraph
            for j=1:p[k]-1 # and all its super-graphs
                c[k+j] += binom[p[k] - 1, j]
            end
        end
    end
```

```

else
    k += 1 # or test next subgraph
end
end
end
return c
end

```

According to the above mentioned assertion the algorithm is rather asymmetrical. The subgraphs which are represented by 1-bits positioned on the left side of a bit pattern take advantage of the assertion on greater extent (the generator fills bit strings from the leftmost bit). The (one vertex) subgraph represented by the least significant bit does not utilize assertion anymore.

The partial improvement consists of the sorting of vertices by out-degree values because the vertex with greater out-degree is better candidate for participation in subgraph covering all remaining vertices (i.e. should be placed on more significant bit of packed column).

The efficiency of algorithms is evaluated by the time required for application of algorithms on random diagnostics graphs (see Figure 11). The random testing graphs contain approximately $n/2$ edges (close to density of edges in optimal testing graphs).

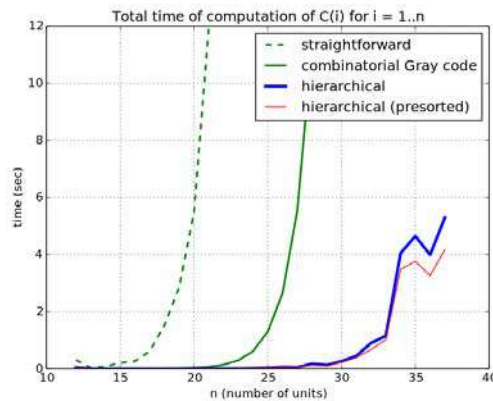


Fig. 11. Time of computation of characteristic number for different number of units

All algorithms have exponential time complexity but their application areas are different. The straightforward implementation is usable only for relatively small number of units (<17 for subseconds execution times on PC range CPU). The algorithm using bit-level parallelism is usable to greater number of units (approximately 25). The algorithm with hierarchical bit patterns makes possible computation with subseconds delays for system with 34 unit. The pre-sorting of vertices leads to some speed-up but it is in most real cases practically negligible.

The hierarchical algorithm also depends on specific structures of testing graph (details are subject to further research).

Figure 12 shows effectiveness of algorithms for almost t_{max} -optimal testing graphs i.e. or graphs with a large number of tests. For t-optimal testing graphs ($t > t_{max}$) the situation may be very different. Plot in Figure depicts the time of computation for simple t-optimal graphs for system with 25 units (t_{max} is 12) and for both algorithms based on bit patterns (straightforward implementation is useless for system with 25 units).

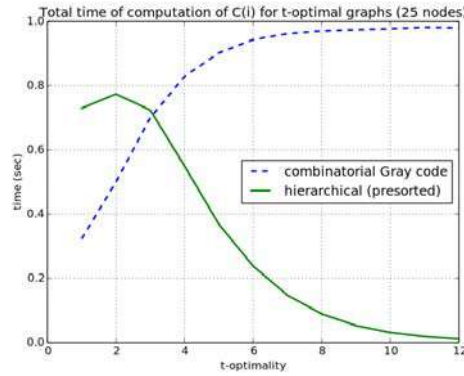


Fig. 12. Computational time for t-optimal testing graphs

For testing graphs with a small number of edges (tests) the simple (non-hierarchical) algorithm is better compared to t_{max} optimal testing graphs because iteration is simpler and short circuit evaluation of product operation is performed almost immediately (only two direct instruction are executed per subgraph). On other hand the hierarchical approach is more complicated (tens of direct instruction per subgraph) and it cannot utilize acceleration of bulk increments (small subgraphs do not cover its remaining vertices). But this acceleration outperform short-circuit for 3-optimal graph and for nearly t_{max} -optimal graphs the difference is in order of tens.

The computation of probability PFD using equation 3.1 requires floating point unit because it uses multiplication of rational numbers (computation of characteristic number is limited to integer arithmetic with simple operation without multiplication). Fortunately, the relative comparison of probabilities (i.e. evaluation of testing assignment) does not depend on probability of a unit faulty state P_U (if it is assumed that all system units have the same probability P_U). Therefore we can choose any probability PU from interval $[0,1]$. For probability 0.5 (although it is unrealistic) the equation 3.1 is simplified to form:

$$\sum_{i=1}^k \frac{C_k}{2^N} \quad (10)$$

Comparison of probabilities for concrete system requires only (integer) summation of characteristic number (value 2^N is constant). When we need to compare testing graphs with differed number of units the computation can be performed by way of bit shift operation on fixed point representation of probability values.

7 Conclusions

Tests performed in a system can be represented as a testing graph. Analysis of the obtained testing graph aims at checking whether all system units have been tested or whether the formed testing graph belongs to predefined subset of testing graphs. It depends on the value of required credibility of system self-diagnosis. Testing graph can be also used as input data for diagnosis algorithm. For different testing graphs the obtained diagnosis results will have different credibility.

In the research, we have considered the probability that all system units can be correctly diagnosed after performing all tests. This probability can be used as a diagnosability measure which will allow comparing system testing assignments and evaluating obtained testing graphs. This probability is computed by using characteristic numbers. In the research, we suggested relatively effective method for computing characteristic numbers and developed the algorithm (based on bitwise operations with integer values). Efficiency of the developed algorithm was also evaluated for different scenarios of testing assignments.

References

1. Mashkov, V.: Task allocation among agents of restricted alliance. *Proc. of IASTED ISC'2005 conference*, Cambridge, MA, USA, 2005, pp.13-18 (2005)
2. Mashkov, V.: Restricted Alliance and Coalitions Formation. *Proc. of IEEE/WIC/ACM International Conference on Intelligent Agent Technology/ Beijing, China, 2004*, pp.329-332 (2004)
3. Qin, L., He, X., Zhou, D.: A survey of fault diagnosis for swarm systems. *Systems Science and Control Engineering*. Vol. 2, 2014, pp. 13-23 (2014)
4. Mahapatro, A., Khilar, P.: Fault diagnosis in wireless sensor networks: a survey *IEEE Commun Surv Tutor*. Vol. 15, 2013, pp. 2000-2026 (2013)
5. Mashkov, V., Mashkov, O.: Interpretation of diagnosis problem of system level self-diagnosis. *Int. Journal „Mathematical Modeling and Computing“*, Vol.2, No.1, 2015, pp.71-76 (2015)
6. Barborak, M., Malek, M., Dahbura, A.: The consensus problem in fault-tolerant computing. *ACM Computing Surveys*. Vol.25, No.2, 1663, pp.171-220 (1993)
7. Preparata, T., Metze, G., Chien, R.: On the connection assignment problem of diagnosable system. *IEEE Transactions on Electronic Computers*. Vol.EC-16, No.12, 1967. pp. 848-854(1967)
8. Mashkov, V.: Selected problems of system level self-diagnosis. Lviv: Ukrainian Academic Press, 2011, 184 pages (2011)
9. Somani, A.: System Level Diagnosis: A Review. (1997) [online]. Available from [www: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.9488>](http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.9488).

10. Dahbura, A., Masson, G.: An $O(n^{2.5})$ fault identification algorithm for diagnosable systems. *IEEE Trans. Comput.*, Vol.C-33, 1984. pp.486-492 (1984)
11. Sullivan, G.: An $O(t^3 + |E|)$ fault identification algorithm for diagnosable systems. *IEEE Trans. Comput.*, Vol.C-37, 1988. pp.388-397 (1988)
12. Meyer, G., Masson, G.: An efficient fault diagnosis algorithm for symmetric multiple processor architectures. *IEEE Trans. Comput.*, Vol.C-27, 1978. pp.1059-1063 (1978)
13. Fujiwara, H., Kinoshita, K.: Some existence theorems for probabilistically diagnosable systems. *IEEE Trans. on Comp.* Vol.C-27, No.4, 1981. pp.297-303 (1981)
14. Bianchini, R., Buskens R.: An adaptive distributed system-level diagnosis algorithm and its implementation. In the *21st International IEEE Symposium on Fault-tolerant Computing*. New York (USA), 1991, pp.222-229 (1991)
15. Mashkov, V., Barabash, O.: Self-checking and self-diagnosis of module systems on the principle of walking diagnostic kernel. *Engineering Simulation*, Vol.15, 1998, pp. 43-51 (1998)
16. Jarrah, H., Sarkar, N., Gutierrez, J.: Comparison-based system-level fault diagnosis protocols for obile ad-hoc networks: A survey. *Journal of Network and Computer Applications*. Vol. 60, 2016, pp. 68-81 (2016)
17. Weber, A., Kutzke, A., Chessa, S.: Energy-aware test connection assignment for the self-diagnosis. *J Braz Comput Soc.* Vol. 18, 2012, pp. 19-27 (2012)
18. Bezanson, J., Karpinski, S., Shah, V., Edelman, A.: Julia: A Fast Dynamic Language for Technical Computing. 2012, ARXIV (2012)
19. Savage, C.: A Survey of Combinatorial Gray Codes. *SIAM Review* [online]. 1997, 39(4), 605-629 (1997) [cit. 2016-04-16].
20. Anderson, S.: Bit Twiddling Hacks [online] (2005)
<https://graphics.stanford.edu/~seander/bithacks.html>
21. *AMD64 Architecture Programmer's Manual Volume 3: General Purpose and System Instructions3* . AMD. 2011. pp. 204–205, (2011)

Microprocessor Means for Technical Diagnostics of Complex Systems

Igor Korobiichuk¹[0000-0002-5865-7668], Ruslan Hryshchuk²[0000-0001-9985-8477], Vladimir Horoshko³[0000-0001-6213-7086], Yulia Khokhlacheva³[0000-0002-1883-8704]

¹Warsaw University of Technology, Institute of Automatic Control and Robotics, Warsaw, 02-525, Poland
i.korobiichuk@mchtr.pw.edu.pl

²Sergey Korolyov Zhytomyr Military Institute, Cybersecurity Department of the Research Center, Zhytomyr, 10004, Ukraine
dr.hry@i.ua

³National Aviation University, Kyiv, 03058, Ukraine
professor_va@ukr.net, hohlachova@gmail.com

Abstract. The construction of algorithms for assessing the operability and the recognition of technical states of complex technical systems can be carried out by preliminary modeling of these problems on large computers. This work is considering the expediency of using multi-level modern microprocessor means technical diagnostics for these purposes.

Keywords: complex technical systems, technical diagnostics, technical states, algorithm, management of object maintenance.

1. Introduction

Providing high reliability of functioning of modern complex technical systems (CTS) includes a wide range of problems, the solution of which is aimed at the efforts of many specialists of different profiles in the design, manufacture, testing and operation of such systems. The issues of technical diagnostics with the subsequent organization of effective maintenance, the main task of which is to maintain the required level of reliability of complex systems during operation.

At the same time, it should be borne in mind that the characteristic features of modern CTS are the high degree of integration of the unit and the nodes, the developed logistics structure and high saturation of computer technology, the wide application of integrated circuit technology and multi-layer printed editing [1]. Therefore, maintenance of such systems is characterized by great laboriousness and a richness of control and diagnostic operations, an increase in the nomenclature of the measured CTS parameters, an increase in the requirements for accuracy, reliability and level of processing of control results [2]. Therefore, a special place in the complex of technological operations for CTS maintenance is concurrently granted to operational control. Qualitatively new approach in the development and use of operational and control-diagnostic devices leads to a change in the organizational and technological structure

of operational control and maintenance systems. Application in the control and diagnostic systems of microprocessor means allows to solve a wide range of problems in the diagnosis of CTS [2-6].

Among these problems an important place is occupied by the issues of technical diagnostics.

The architecture of constructing modern microprocessor means (MPM) for the technical diagnosis of CTS is largely determined by the constructive and functional complexity of the monitored objects. If, for example, for CTS of ordinary complexity, the automation of workability control is reduced mainly to the allowable evaluation of individual controlled parameters or their disunited summation, then for complex CTS, which are characterized by a multitude of such parameters, a peculiar hierarchical structure and unequal connection of individual aggregates a specific complex approach to the solution of the task of assessing the level of operability of objects [3]. In this case, the problem arises of recognizing the classes of technical conditions of RES, which are formed on the basis of a generalized set of performance indicators and controlled parameters. This work is considering the expediency of using multi-level MPM technical diagnostics for these purposes.

It is known that the main task of autocontrol systems is to identify one of the conditions in the monitored objects - workable or inoperable (failure, halting or malfunction). The criterion of efficiency is usually reduced to the description of the external behavior of the CTS, which is characterized by changes in its characteristics, internal or external parameters. Taking into account the working conditions of objects, which are described by the relationships between the monitored parameters, it is possible to form classes of technical states defining the region of the workable and inoperative CTS states. Obviously, the corresponding values of the internal and external parameters of the CTS determine the boundary points of these states. The task of autocontrol systems is to determine the working conditions that are most relevant to the given criterion. Direct control of the indicators describing this criterion and the search for boundary points of workability is hampered by a number of difficulties associated with the diversity of these indicators, the complexity of the representation and interrelation of the technical implementation. The development of MPM technical diagnostics, performed in a hierarchical structure, avoids these difficulties.

2. Materials and methods of research.

During operation, modern CTS can be in the following states: normal working capacity (S1), described by a complex of nominal values of characteristics and parameters; anticipatory tolerances for controllable parameters (S2); operational nominal tolerances for controlled parameters, the excess of which is classified as a fault (S3) or failure (S4); reserve condition (S5).

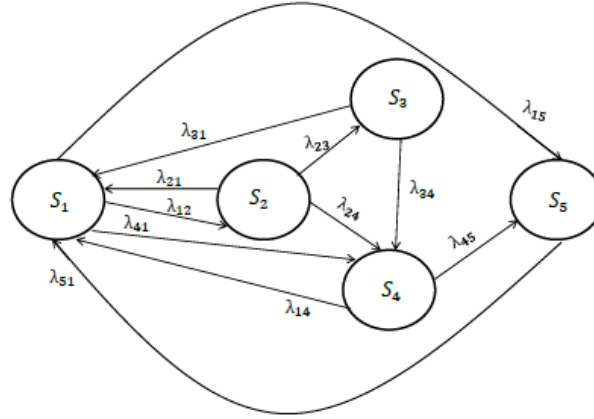


Fig. 1. The Markov model, of graph of discrete states of the object

To assess the possibility of monitoring these states, we will depict the maintenance process of CTS in the form of a Markov model, which is a graph of discrete states of the object (Fig. 1). The transition of CTS from one state to another occurs under the influence of certain events - failures, malfunctions, regulations, performance of regulated maintenance, etc. The arrows on the graph indicate possible transitions of the system from one state to another. For clarity of the presented model, we characterize the intensity of the transitions noted above the corresponding arrows of the graph: when performing scheduled maintenance in the standby state (λ_{15}); from standby to operable condition after elimination of failures or maintenance (λ_{51}); gradual failures, associated with the acquisition of controlled parameters of the values of anticipatory tolerances (λ_{12}); when conducting regulating and adjusting works (λ_{21}); sudden failures (λ_{14}); gradual failures associated with the obtaining of pre-emptive values of the monitored parameters of the states of malfunction (λ_{23}) or failure (λ_{24}); gradual failures associated with a succession of malfunctions and failures (λ_{34}).

Thus, from the state S1, the monitored CTS can be transferred to the state S5 when performing the scheduled maintenance (λ_{15}) or when gradual ($\lambda_{12} - \lambda_{23} - \lambda_{34}$) and sudden (λ_{14}) failures occur. When the REN have reached the S2 state, it is necessary to perform regulating and adjusting operations in order to transfer the object to the Si state. If this is not possible, the object goes into the conditions of malfunction (S3) or failure (S4). In the future, it is advisable to transfer this object to a standby state (λ_{45}) for restoration work.

3. The results of research

In accordance with the methodological recommendations given in [4, 5], we will compose a system of differential equations directly from the given graph (Fig. 1), describing the Markov model of discrete states of the object:

$$\begin{aligned}
\frac{dP_1}{dt} &= -P_1(\lambda_{12} - \lambda_{14} - \lambda_{15}) + P_2\lambda_{21} + P_3\lambda_{31} + P_4\lambda_{41} + P_5\lambda_{51}; \\
\frac{dP_2}{dt} &= -P_1\lambda_{12} + P_2(\lambda_{21} + \lambda_{23} + \lambda_{24}); \\
\frac{dP_3}{dt} &= -P_2\lambda_{23} + P_3(\lambda_{31} + \lambda_{34}); \\
\frac{dP_4}{dt} &= -P_1\lambda_{14} + P_2\lambda_{24} + P_3\lambda_{34} + P_4(\lambda_{41} + \lambda_{45}); \\
\frac{dP_5}{dt} &= -P_1\lambda_{15} + P_4\lambda_{45} + P_5\lambda_{51}.
\end{aligned} \tag{1}$$

Taking into account the use of the limiting probabilities of the investigated CTS states P_i (S_i), we transform a system of equations to an algebraic form (1), the left-hand side of which is equal to zero. Then $\sum_{i=1}^5 P_i = 1$. In the future, given the intensities

of the transitions λ_{ij} , the numerical values of the CTS residence probabilities in S_i states are determined, which to a large extent predetermines the architecture of constructing the MPM of technical diagnostics.

It is most expedient to build the technical diagnostics MPM on a distributed architecture, based on the functional decentralization, i.e., the monitoring and diagnostic system must be performed in multi-level with a hierarchical structure. In this case, the control and diagnostic process is subdivided into sub-processes with well-defined interconnection methods for information transfer.

Figure 2 shows a two-level MPM of technical diagnostics, which consists of two lower-level subsystems and a coordinating microcomputer. Lower-level subsystems A and B should conduct a diagnostic control of measuring parameters with different combinations of sensors and perform preliminary processing of the received information. In this case, the admission control of the incoming information determines the position of the controllable parameter in the field of system operability. Thus, with the help of lower-level subsystems, it is possible to fixate the moments of the CTS transition from the state of normal working capacity S_1 to the states $S_2 - S_5$ and to obtain quantitative information about these states. The tasks of recognition of technical states and control by measuring parameters are solved by the coordinating upper-level microcomputer. Taking into account the quantitative characteristics of the Markov model of CTS states, it is possible to preliminarily describe the problems solved by both computers and determine the architecture of their construction by defining the corresponding amount of memory, organizing the interaction of functional systems, etc. Thus, the RAM memory size of the coordinating microcomputer depends mainly on the mathematical model and the algorithm for recognizing the technical states of the Markov model.

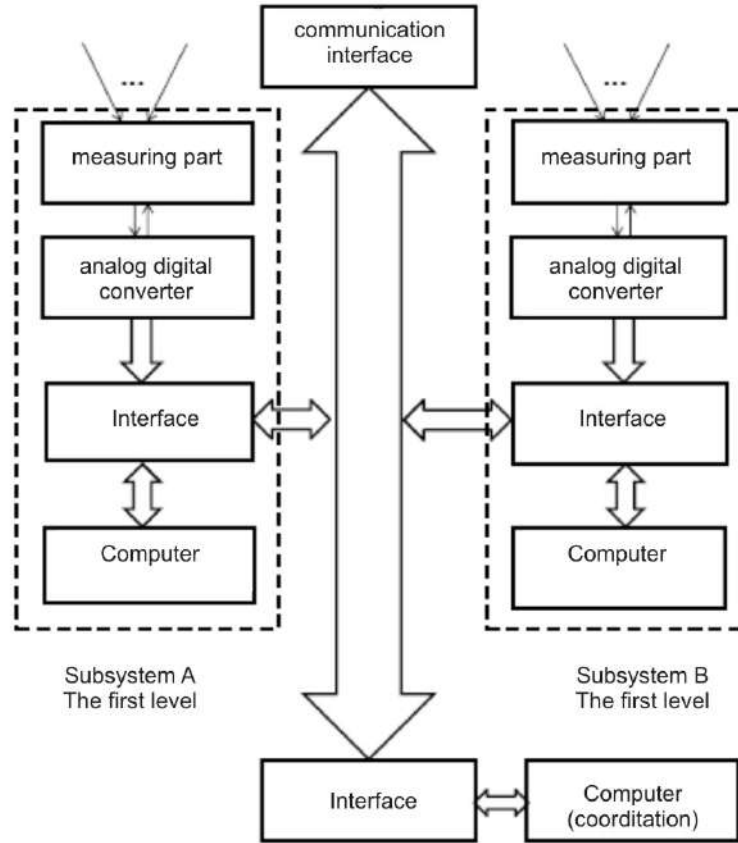


Fig. 2. The two-level MPM of technical diagnostics

The mathematical basis for the recognition of technical states of CTS can be laid by Bayesian criterion [6-8] or the maximum likelihood principle for equiprobable states P_i

$$\Lambda_k = \frac{\varphi(K_1/x_1, x_2, \dots, x_n)}{\varphi(K_2/x_1, x_2, \dots, x_n)} \geq 1, \quad (2)$$

Where Λ_k is the criterion for recognition (classification) of technical states of CTS; K_1, K_2 - the investigated classes of technical states of CTS; $x_1, x_2, \dots, x_n$ - indicators of technical condition or monitored parameters of the object.

For complex CTS, it is advisable to install the following classes of technical states [9-12]:

- class K_1 - a set of CTS status indicators requiring continuous monitoring, the deviation of which can cause a malfunction or complete failure of the object;

- class K2 - a set of indicators of the state of CTS, requiring periodic monitoring, the deviation of which can cause partial failures or malfunctions of the object;
- class K3 - a set of indicators of the state of CTS, which determine the conduction of regulated maintenance.

In the general case, the number of classes considered exceeds the number given in the Markov model (Fig. 1).

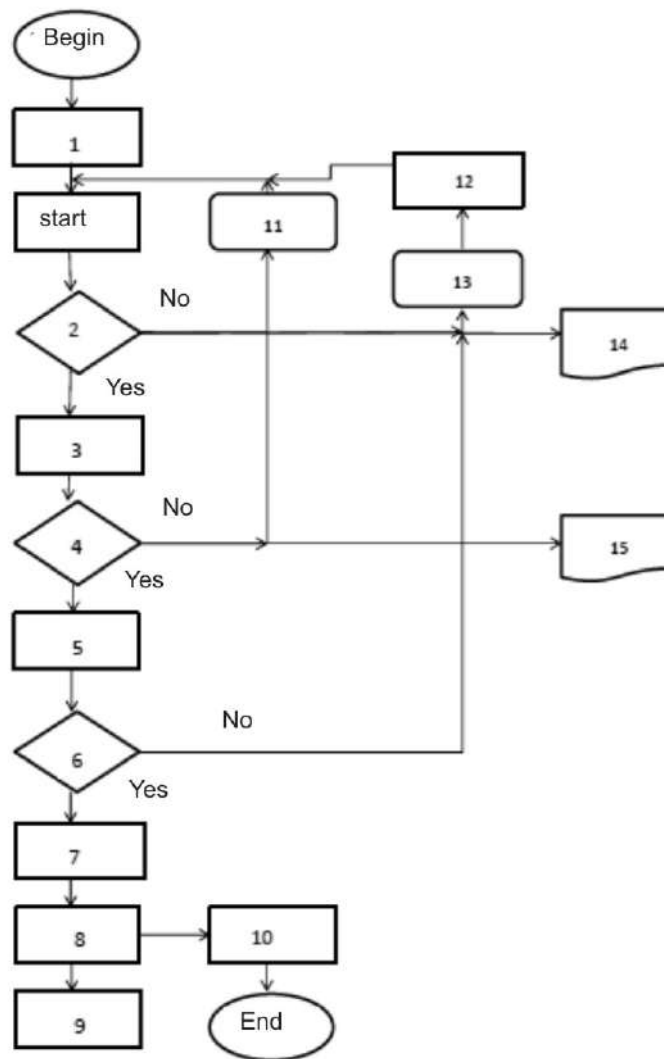


Fig. 3. The scheme of the algorithm of operation of the subsystem of the lower-level MPM

The scheme of the algorithm of operation of the subsystem of the lower-level MPM is shown in Fig. 3. The presented algorithm solves the problem of estimating the monitored parameter for finding it in the field of operability. The values of the monitored parameter are compared with the values of its admissible values, which are stored in the memory of the microcomputer (procedure 1). In addition, a ping sequence of sensors installed in the CTS and an algorithm for calculating the parameters determined by indirect methods are also stored in memory.

Procedure 2 is organizing a cycle of MPM selecting the technical diagnostics of the corresponding controlled parameter. If the MPM is not correctly connected, the system is dynamically stopped by the processor (procedure 13) and the command is repeated to connect the desired sensor to the monitoring system (procedure 12). Simultaneously, for a high-level microcomputer, a signal is generated about the absence of such a connection at a given time (procedure 14). When connecting the corresponding object sensor to the MPM in the microcomputer from the measuring part (Fig. 2), information is received on the value of the parameter controlled at a given time (procedure 3). Correspondence of the CTS parameters of the working capacity field can be carried out with the help of the admission control according to the pre-compiled algorithm stored in the micro-computer memory (procedure 4). When the monitored parameter exits the boundaries of the working capacity field, the lower-level system is dynamically stopped (procedure 11) and a message "out of tolerance" is generated for the upper-level computer (procedure 15). Then the upper-level computer is assuming direct control of the process of measuring and recognizing the technical state of the object. If the complex evaluation shows that the output of one parameter outside of the working capacity field does not affect the overall performance of the CTS, then the control of the process of measuring parameters and analyzing the state of the object is transferred again to the lower-level computer. The values of the CTS parameters that have passed the admission control are entered in the RAM of the lower-level micro-computer (procedure 5). Further, it is checked whether all the parameters are monitored (procedure 6). If the answer is negative, the system stops dynamically (procedure 13) and procedures 2-6, 12 are repeated. With full control of CTS, a conclusion is made about the state of the object (procedure 7). The result of this evaluation is stored in the computer's RAM instead of the information recorded during the procedure 5 (procedure 8). In the future, this information can be used for digital indication (procedure 9) or registration on a digital printer and output to a upper-level computer (procedure 10).

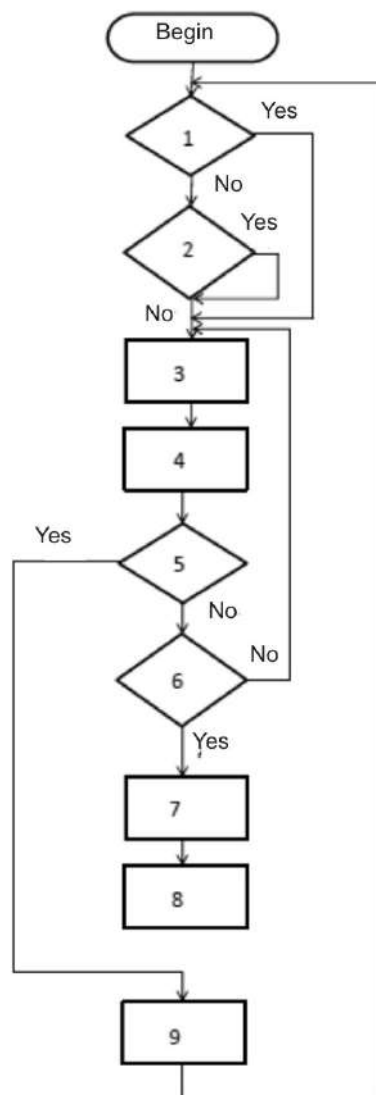


Fig. 4. The algorithm of operation of the lower-level subsystem, the upper-level computer interacts directly with the subsystem

As can be seen from the algorithm of operation of the lower-level subsystem, the upper-level computer interacts directly with the subsystem. This interaction clearly illustrates the algorithm of the coordinating microcomputer (Fig. 4). The described algorithm, in addition to the described functions for coordinating the operation of the entire technical diagnostic system, determines the collection of information from lower-level subsystems (A and B), by processing which the CTS technical condition is rec-

ognized.

The coordinating microcomputer performs a sequential ping of lower-level subsystems A (procedure 1) and B (procedure 2) for the presence of requests (procedure 15 in Fig. 3). If there are no requests, the computer reads the CTS status information (procedure 3) and processes it (procedure 4). If there is such a request, the computer switches to procedure 3 (reading data from the information buses of lower-level subsystems A and B) and the technical condition of CTS is recognized by formula (2) (procedure 4). On the basis of the received data, a decision is taken on whether the object is in operable or faulty state (procedure 5). If the technical condition of the CTS does not meet the specified requirements or the controlled parameters approach the permissible values, procedures 3-6 are repeated. If the technical condition of the object is classified as inoperative (procedure 6), the coordinating computer goes to the malfunction search routine (procedure 7), after which the faulty CTS unit is indicated (procedure 8). If the departure of parameters beyond the tolerance limits does not affect the overall operability of the object or all of its parameters are normal, the information on the current state of the REN is updated on the indicator board (procedure 9) and MPM proceeds to re-execute all the I procedures.

Conclusions

The construction of algorithms for assessing the operability and the recognition of technical states of CTS can be carried out by preliminary modeling of these problems on large computers.

Functional versatility, programmability, a large degree of integration, high reliability, uncomplicated modification of the performed functions and the low cost of microprocessor means determine the possibility of their wide application in the systems of technical diagnostics and management of object maintenance.

References

1. Korobiichuk, I.: Automatic control devices. Ed. PIAP: Warsaw (2018)
2. Korobiichuk I., Hryshchuk R., Mamarev V., Okhrimchuk V., Kachniarz M.: Cyberattack Classifier Verification. International Conference on Diagnostics of Processes and Systems DPS 2017: Advanced Solutions in Diagnostics and Fault Tolerant Control, pp. 402-41 (2018) doi: 10.1007/978-3-319-64474-5_34
3. Korobiichuk, I., Ladanyuk, A., Shumyhai, D., Boyko, R., Reshetiuk, V., Kamiński, M.: How to Increase Efficiency of Automatic Control of Complex Plants by Development and Implementation of Coordination Control System. Recent Advances in Systems, Control and Information Technology. Advances in Intelligent Systems and Computing, vol 543, p 189-195 (2017) doi: 10.1007/978-3-319-48923-0_23
4. Wentzel, E.S.: Investigation of operations. Nauka: Moscow (1980)
5. Conly, C. H.: Organizing for Computer Crime Investigation and Prosecution. National Institute of Justice (NIJ): Washington (1989)
<http://www.ncjrs.gov/App/publications/abstract.aspx?ID=118216>
6. Kredentzer, B.P.: Maintenance and reliability of systems with temporary redundancy. Phoenix: Kiev (2016)

7. Volochiy, B., Mulyak, O., Ozirkovskiy, L., Kharchenko, V.: Automation of quantitative requirements determination to software reliability of safety critical NPP I&C systems. Proceedings - 2nd International Symposium on Stochastic Models in Reliability Engineering, Life Science, and Operations Management, SMRLO 2016, art. no. 7433137, pp. 337-346 (2016) doi: 10.1109/SMRLO.2016.62
8. White, J.M., Darrah, C.: Reliability, redundancy, & long-term operations: Designing a flexible integrated membrane plant. AWWA/AMTA 2014 Membrane Technology Conference and Exposition, Code 104790 (2014)
9. Willkomm, D., Gross, J., Wolisz, A.: Reliable Link Maintenance in Cognitive Radio Systems. First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 371-378 (2005)
10. Zahmati, A.S., Fernando, X., Kojori, H.: Emerging wireless applications in aerospace: Benefits, challenges, and existing methods. 4th Annual Caneus Fly-By-Wireless Workshop, FBW 11, art. no. 5965561, pp. 87-90 (2011) doi: 10.1109/FBW.2011.5965561
11. Shi, Q., Taubenheim, D., Kyperountas, S., Gorday, P., Correal, N.: Link maintenance protocol for cognitive radio system with OFDM PHY. 2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, art. no. 4221525, pp. 440-443 (2007) doi: 10.1109/DYSPAN.2007.63
12. Petrov, A., Khoroshko, V., Scherbak, L., Petrov, A., Aleksander, M.: Reliability of Information Systems. AGH: Krakow (2016)

Information Technologies based on Wavelet Transform for Soldered Joints Diagnostic of Printed Circuit Boards

Galina Shcherbakova¹[0000-0003-0475-3854], Viktor Krylov¹[0000-0003-1950-4690],
Anatoliy Sachenko²[0000-0002-0907-3682], Pavlo Bykovyy³[0000-0002-5705-5702],
Diana Zahorodnia³[0000-0002-9764-3672], Myroslav Komar³[0000-0001-6541-0359],
Ivan Kit³, Dmytro Lendiuk³, Andriy Kaniovskyi³, Mykola Dacko³

¹ Odessa National Polytechnic University, Odessa, Ukraine

² Kazimierz Pulaski University of Technology and Humanities in Radom, Radom, Poland

³ Ternopil National Economic University, Ternopil, Ukraine
galina_onpu@ukr.net

Abstract. In modern instrumentation, the number of soldered joints in printed circuit boards can reach several thousand. Diagnostics of the soldered joints defects within the optical wave length range is carried out using automated diagnostic systems. A number of stages of the existing information technologies for such systems are implemented on the basis of target functions extremum search using the gradient estimation. In the large batches of products production, the use of expensive automated diagnostic systems within lighting subsystems and high cost positioning are justified. These subsystems can provide improved noise immunity. However, in conditions when small batches of products are produced, and at some stages (for example, when positioning by comparison with a standard/prototype images) in general, such objective functions can be noisy and can be multi-extremes. For such cases, information technologies based on methods of enhanced noise immunity are required. Such an increase in noise immunity can be provided by methods using wavelet transformation. For this purpose, information technologies were proposed using wavelet transformation-based procedures that improve noise immunity and reduce the error of procedures in the diagnostic systems of printed circuit boards and their soldered joints.

Keywords: Automated Diagnostic Systems, Solder Joint, Printed Circuit Boards, Wavelet Transformation

1 Problem description

In modern instrument production, the number of solder joints (SJ) in the printed circuit boards (PCB) can reach several thousand [1-14]. Diagnostics of solder joint defects in the optical wavelength range is carried out in automated diagnostic systems (ADS) of the solder joints (Fig. 1) [15-26]. In the existing ADS, the detection of PCB

assembly defects is carried out by comparing their images with the images without the defects (standard/prototype images). Such a comparison in the ADS is carried out using systems of the precision mechanics and high cost optics.

Reducing the requirements for such systems in order to reduce the cost of small batches of products production leads to uneven illumination of objects and the appearance of multiplicative and additive interference in images. In conditions with interferences and signal-to-noise ratio (power) for additive interference ranging from 8 to 11 and from 7 to 10 for multiplicative interference, such automated diagnostic systems are not operational. In the existing automated diagnostic systems, the combination of images for comparison is carried out on the basis of gradient methods searching for the minimum of the multimodal function and this way to determine the error in combining with the standard/prototype image. Such methods are not functional within the signal / interference ratio being less than 15. Therefore, for such a combination of images, information positioning technologies are needed because it can provide necessary reliability in the conditions of interference and a priori uncertainty due to small sets of training data. To this end, it is proposed to implement these IT on the basis of the developed multi-start optimization methods of the wavelet transformation [27-32].

Often the reduction of the resource consumption in automated diagnostic systems and reduction the operational speed are increased by conducting diagnostics on the conveyor for PCB panels (not separate PCB's) (Fig. 2). In this case, several automated diagnostic systems are present at different stages of the PSB assembly. For example, when assembling a PCB during mounting components on a surface, these are the stages of pattern printing, resistors and integrated circuits (ICs) placement (with a small output range) (Table 1).

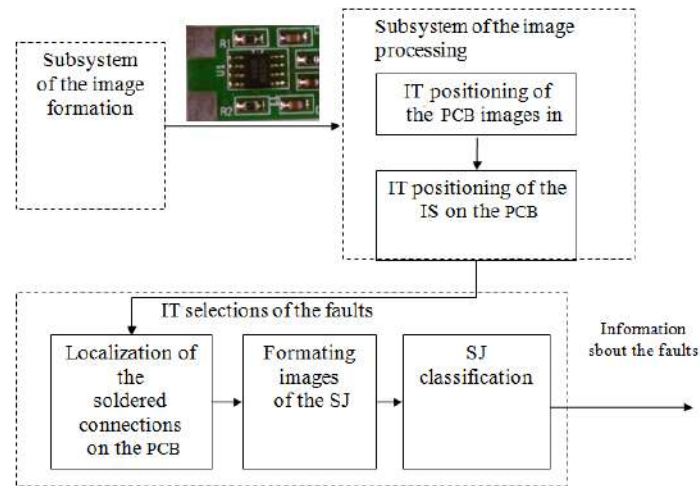


Fig. 1. Fragment of the automated diagnostic system (ADS) block diagram for soldered joints within the optical range

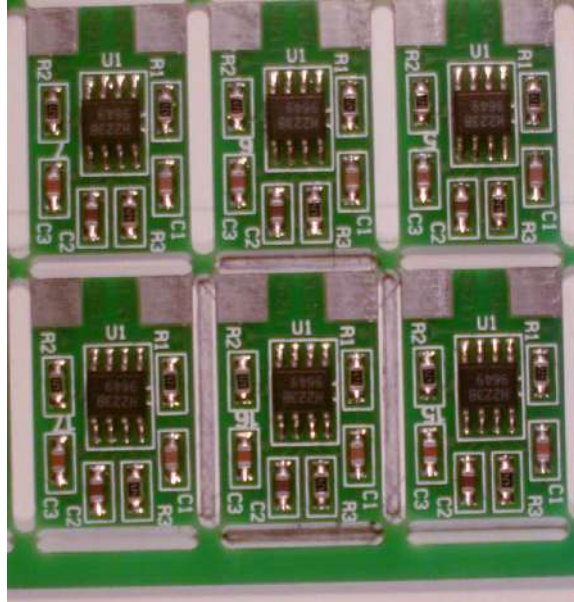


Fig. 2. PCB Panel

A generalized structure of the automated diagnostic system (ASD) in printed circuit board assemblies has been developed within the optical range (Fig. 3).

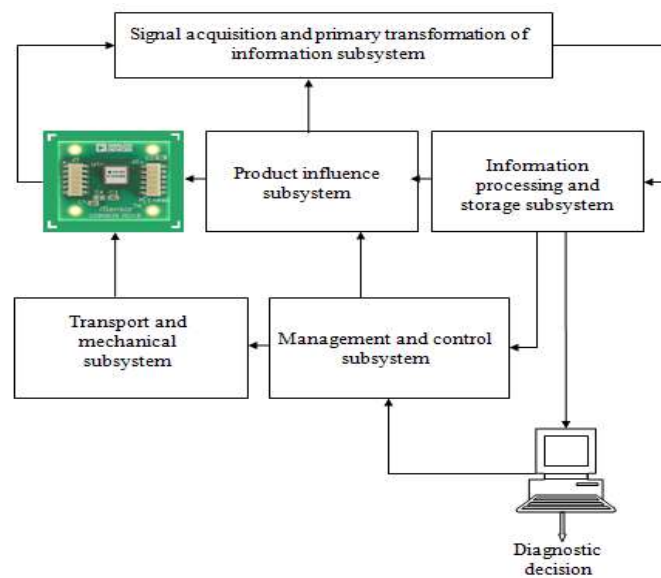


Fig. 3. Generalized structure of the automated diagnostic system for the printed circuit assemblies diagnostics within the optical range

Such an ASD may contain the following subsystems:

- subsystem influencing the object, which contains device impact on the object sources of illumination and / or penetrating fields (with fluoroscopy) and scanning devices;
- transportation mechanical subsystem, which contains the object delivery devices (loading and unloading elevators, conveyor), coordinate table, positioning devices;
- subsystems of signal acquisition and information conversion, which contain receiver-transducers of penetrating fields, devices for a signal pickup from receiver-transducers, devices for primary conversion and normalization of signals, image sensors, video processing devices (amplification, quantization and normalization of the video signal);
- subsystems of the computer-based processing and information storage, that are based on applied IT implemented signal and image processing.

Table 1. Types of defects at various PCB assembly stages

Stage	Type of defects detected during optical inspection
Screen printing soldering paste	Skip of the paste, improper application, lintels
Component positioning	The absence/presence of components, bevel, shear, wrong orientation and/or polarity. Printed circuit board (PCB) warping
Melting solder	Presence/absence, shear, bevel, polarity of components, correctness of soldered joints, defects in paste failure, solder bridges, raised connectors.

Practically, it is necessary to control the soldered joints for all PCB after the solder melting step. Such diagnostics in the existing integrated circuits are carried out by comparing the solder joints on the PCB with the reference images of the solder joints, so it is important to carry out the exact combination of such images.

When choosing the number of automated diagnostic systems, the cost of repairs after performing this operation, the type of equipment for which these launchers are assembled for are taken into consideration. It is also taken into account that up to half of all defects of the PCB is concentrated in soldered joints of integrated circuits, as well as the fact that the cost of the repairs after soldering averages up to 60% of the repair cost of the PCB. Therefore, for medical equipment and security systems (for example, air bags in cars), it is recommended to perform diagnostics at the stage of placement of the integrated circuits. Such diagnostics is carried out by comparing with reference images of PCB, therefore the exact combination of images of such type is important. Therefore it is necessary to carry exact positioning of such images (in order to reduce costs of the automated diagnostic systems, the cost of PCB and to increase the performance).

2 Main stages of the developed information technologies of the automated system for printed circuit boards and soldered joints diagnostics

The developed information technology for PCB automated diagnostic systems includes the following steps.

Stage 1. Positioning of the PCB panel image along the contours of the reference marks (fig 4).

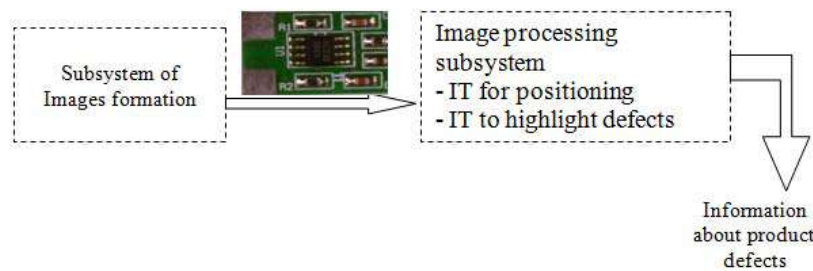


Fig.4. Processing sequence when information technology for positioning and defects selection on the print nodes are being used

This IT allows for automated positioning to use images of reference marks (RM) - marks of a familiar shape on the surface of printed circuit assemblies PCBs (Fig. 5) and has two modes: training mode and operating mode.

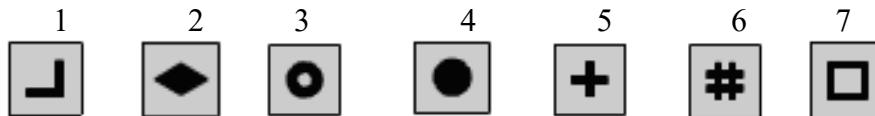


Fig. 5. Classes of reference marks

The training mode includes the following main stages (Fig. 6).

Stage 1. Pre-processing of the image of the product obtained with the help of a video camera in order to reduce the level of noise.

Stage 2. Binarization of the product image.

Stage 3. Localization of the reference marks. With the localization of reference marks, their coordinates on the image are determined, and their contours are selected and traced. The result of this stage is the coordinates of the contours of the reference signs.

Stage 4. Determination of the identification vector of the reference marks. As a result, a training sample is formed to classify reference marks.

Stage 5. Training of a multilayered neural network (MNN), as a result of which the surface is built, dividing the reference signs in the feature space. This dividing surface is used on the next stage in the operating mode during the classification.

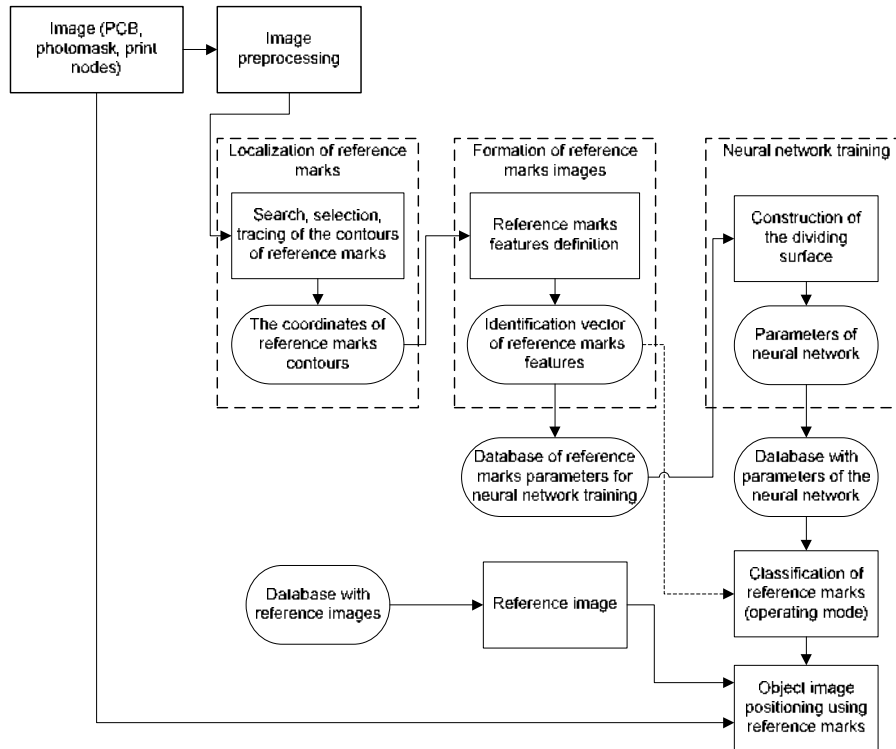


Fig. 6. Basic IT of the reference marks positioning procedures

In operating mode, the following steps are implemented.

Stage 1. Classification of reference marks. The operator of the ASD receives information about the belonging of the reference marks to a certain class - that is, the result of recognition - a classification decision.

Stage 2. Positioning of the object image on the reference marks is relative to the standard, coming from the database (Fig. 7).



Fig. 7. Fragment of the PCB image and the result of localization of reference marks

The implementation specifics of the IT procedures for the positioning of printed circuit assemblies of control stations by the reference marks are described below:

a) image pre-processing. In order to reduce the level of additive interference arising in a television sensor at the stage of forming a digital image of an object, is proposed to be conducted with the application of the median filtering. It implements a non-linear noise removal procedure. When the procedure is implemented, a window is created that moves through the image, covering an odd number of its pixels. The value of the current pixel is replaced by the median of the pixel values in the window. The intensity drops with such processing in contrast to the smoothing filters. In order to reduce the level of multiplicative noise caused by uneven illumination, homomorphic filtering was applied.

b) binarization. To highlight the defects of the printed circuit assemblies of the PCB in the ASD, binarization of their images and references is performed. This procedure should ensure invariance to intensity transformations and reduce the amount of information processed by the system. In this case, the error of binarization should not exceed the 50% of the track width of printed circuit assemblies of the PCB. During binarization, the following decision is made: is there a track (useful signal) or substrate (background) among the pixels of the image? Depending on this, the pixel is assigned the value "0" or "1".

c) localization. For the localization of reference marks, a method of localization of small-sized objects in the image using the wavelet transform was chosen. As a result of the localization procedure, a traced contour of the reference marks of PCB and the coordinates of their position (centers of gravity) are obtained (Fig. E). This reduces the amount of information for the identification procedure.

d) identification. Based on the calculation of geometric moment features using information about the characteristic points - the most informative part of the contour of the reference marks. The first and second coefficients of the discrete cosine transform of the moment-signs, the geometric sign of compactness, and the Euler number are included in the feature set. The selected features provide: scale, shift, and rotation invariances sufficient for the requirements of the practice.

e) reference marks classification. It is carried out with the help of MNN (perceptron). The result of MNN learning is stored in the database with the parameters of the MNN (see Fig. 6). Then, at the recognition stage, a set of reference marks obtained under the noisy conditions is supplied to the input of the perceptron. As a result of the calculations, the class of the relay is determined in the network output.

IT for defect allocation involves the following steps.

Stage 1. Binarization of the reference images and controlled products.

Stage 2. Median filtering of the product image. In order to exclude false defects from consideration (relative shift results by 1 pixel on images, for example, current PCB and standard one), median filtering is used to ensure the noise immunity of the decision on the presence of a defect. By varying the length of the median filter, you can identify defects with a given accuracy.

Stage 3. Determination of the resistors and capacitors solder joints defect locations on the image of the PCB.

Selection of such defects is performed by combining the binarized images of the PCB relative to the standard ones based on their pixel-by-pixel comparison. In Fig. 8 a fragment of the PCB image is presented with identified defects.

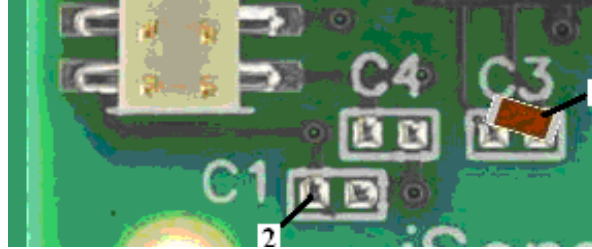


Fig.8. Fragment of the PCB assembly with defects in the capacitors soldered joints

To assess the recognition accuracy, a database was used which consisted of 100 photographs of printed circuit assemblies of various sizes, rotation and tilt angles, and various levels of illumination. As a result of testing such an ASD, the recognition accuracy was 0.85. The noise immunity of the system was increased on average by up to 20% compared to the existing systems through the use of developed information technology that meets the requirements of practice. The reliability of diagnostics in terms of interference on the signal-to-noise ratio range 10 20 (in power) increased up to 1.2 times in comparison with the existing ASD.

The next step is to determine the defects of components with a small lead pitch (Fig. 9). Information technology for components positioning in automated diagnostic systems for soldered joints with a small range of component conclusions involves the implementation of two main stages.

The first stage is the positioning of the PCB by the reference marks, the second stage is the positioning of the resistors, capacitors and integrated circuits on the PCB image (see Fig. 1).

At the second stage, in order to determine the defects of the PCB of the integrated circuit, it is necessary to identify the region of these integrated circuits location on the image of the PCB. To solve this problem, the following methods are used, the ones that use information about the entire halftone image, and methods that use information about specific areas of the image - characteristic features.

For the methods of the first group, the computational costs are high, since the parameters of the soldered joints positioning are calculated from the information about the entire half-tone image of the PCB. The characteristic features method is easy to implement and has higher speed. Therefore, this method was chosen as the baseline for the positioning on the second stage (fig. 3). When searching for the position of the integrated circuits on the PCB, the minimum of the multimodal function should be found, which determines the error of combining the integrated circuit component and a sample (fig 3.a), with a signal-to-noise ratio below 6. Gradient methods of the minimum search that are traditionally used at this stage are not functional. Signal / interference ratio is less than 15. To search for a minimum in such conditions, a multi-start optimization method in the wavelet transformation space has been developed [27]. This method was used in the work to search for the location of the integrated circuits and to align the reference image and controlled integrated circuits components in the ADS with the processing of information on soldered joints obtained in the optical range [27-29].

Within this the binarization of the PCB image by converting a halftone image (with a pixel intensity ranging from 1 to 256) into a binary one according to the intensity threshold is carried out. For this, two thresholds of binarization were accepted in the paper T_1 and T_2 . Pixel of the binarized image $B(x, y)$ takes the value of 1 if $A(x, y) < T_1$ or $A(x, y) > T_2$ and the value of 0 if $T_1 < A(x, y) < T_2$. Here $A(x, y)$ is a halftone image $T_1 = 77$ and $T_2 = 156$ - binarization thresholds. The result of binarization is shown in Fig. 9, c.

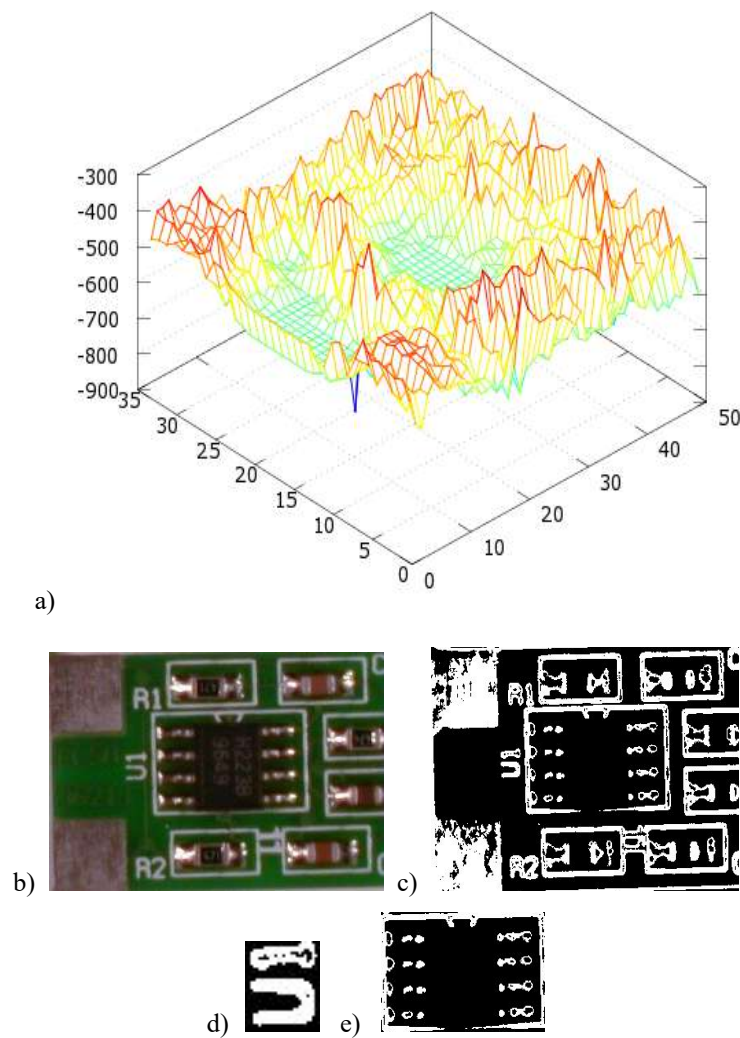


Fig. 9. Error of PCB combining function and a standard (a); a fragment of the PCB (b); binarized image (c); masks with the designation PCB (U1) (d) and with soldered joints (e).

Then the image of the PCB is scanned using a mask with the marking of U1 (the label of this IS on the PCB) (see Fig. 9, d). The minimum of the combining image error (see Fig. 9, a) during such a scanning is found with the help of the multi-start optimization method with WT. The carrier length of the wavelet function (WF) in the search for a minimum in the coordinate x is 17, in y it is 8, the discretization step of the WF is 1, and in the implementation of the iterative search with the Haar's WF $\gamma_1 = 0,1$.

The minimum of the image alignment error was found in 3 iterations (at the start of the search in $[1; 1]$, the coordinates of the minimum (25; 14), (19; 13), (17,13) were found). At the second stage of the search in the region of this minimum, scanning of the image PCB is performed with an image mask of the SJ of this IS (see Fig. 9, e). The minimum of this error function is also determined by using the optimization method with the WT for three iterations.

Then the positioning of the integrated circuit on the image PCB is made.

The developed information technology allows a fast search of soldered joints while the components are being mounted to the surface with a signal-to-noise ratio in amplitude starting from 5.

3 Conclusions

Thus, it can be concluded that during testing of the developed information technologies for soldered joints and automated diagnostic systems of the PCB based on procedures using wavelet transformation, the recognition accuracy was 0.85. In terms of interference in the signal-to-noise ratio range 10 to 20 (in power) the reliability of diagnostics increased up to 1.2 times compared to the existing automated diagnostic systems. The noise immunity was increased on average by up to 20% compared to the existing systems due to the use of developed information technologies that meet the requirements of practice.

References

1. Gu, Jian, Lei, Yongping, Lin, Jian, Fu, Hanguang, Wu, Zhongwei: Design of the printed circuit board for board level drop impact base on the JEDEC standard. 17th International Conference on Electronic Packaging Technology (ICEPT). pp. 588 – 591 (2016).
2. Krylov, V., Shcherbakova, G.: The combination of images in the optical control systems of the printed circuit boards. Technology and design in electronic equipment, № 1, pp. 23–25 (2000).
3. Panchak, R., Kolesnyk, K., Lobur, M.: Topology editing strategies in the subsystem of printed circuit boards manufacturability improvement. 11th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), pp. 246 – 247 (2011).
4. Hu, Chaohui, Zhou, Lina: Failure analysis for micro-short circuit between two pins in printed circuit board assembly. 11th International Conference on Reliability, Maintainability and Safety (ICRMS), pp. 1-4 (2016).

5. Moganti, M., Ercal, F., Dagli, C., Tsunekawa, S.: Automatic PCB inspection algorithms: a survey. *Computer vision and image understanding*. Vol. 63, No.2. pp. 287–313 (1996).
6. Zhang, Yong, Yang, Peng, Ji, Mingjiang, Liu, Guanjin: Mechanism of Solder Joint Intermittent Faults and Its Detection. In: *IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, pp. 1 – 5 (2018).
7. Weibo, Huang, Peng, Wei: A PCB Dataset for Defects Detection and Classification. *Journal of Latex Class Files*, Vol. 14, No. 8, August 2018. pp.1-9 (2018).
8. Bangshu, Xiong, Qiaofeng, Ou: A measurement system for printed circuit board parameters based on image. *4th IEEE Conference on Industrial Electronics and Applications*, pp. 2396 - 2400 (2009).
9. Borba, J.F., Facon, J.: A printed circuit board automated inspection system. *Proceedings of the 38th Midwest Symposium on Circuits and Systems*. Vol.1, pp. 69 – 72 (1995).
10. Pedro, M.A., Vitoriano, Tito, G.: Amaral: Improved Pattern Matching Applied to Surface Mounting Device Components Localized in Automated Optical Inspection. *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering*, vol:11, no:4, pp. 435-439 (2017).
11. Yao, Bin, Lu, Yudong, Luo, Daojun: Key failure modes of solder joints on HASL PCBs and root cause analysis. *14th International Conference on Electronic Packaging Technology*, Dalian (China), 11-14 Aug. 2013, pp 742-745 (2013).
12. Švecová, O., Kosina, P., Šandera, J., Szendiuch, I.: Reliability model for assessment of lifetime of lead-free solder joints. *18th European Microelectronics & Packaging Conference*, Brighton (UK), 12-15 Sept. 2011, pp 1-5 (2011).
13. Lou, Haohuan, Qu, Xin, Chen, Zhaoyi, Wang, Jiaji, Taekoo, Lee, Hui, Wang: Lifetime Assessment of Solder Joints of BGA Package in Board Level Drop Test. *6th International Conference on Electronic Packaging Technology*, Shenzhen (China), 30 Aug.-2 Sept. 2005, pp 1- 5 (2005).
14. Carroll, D., Bates, C., Zampino, M., Jones, K.: A novel technique for modeling solder joint failure during system level drop simulations. *Thermal and Thermomechanical Proceedings 10th Intersociety Conference on Phenomena in Electronics Systems (ITHERM)*, San Diego (CA, USA), 30 May-2 June 2006, pp. 861-868 (2006).
15. Belbachir, A.N., Lera, M., Fanni, A., Montisci, A.: An automatic optical inspection system for the diagnosis of printed circuit based on neural networks. *IEEE 40th Industry Applications Society Annual Meeting*. – Hong-Kong : China, 2005. pp. 680–684 (2005).
16. Garcia, H.C., Villalobos, J.R., Runger, G.C.: An Automated Feature Selection Method for Visual Inspection Systems. *IEEE Transactions on Automation Science and Engineering*. Vol. 3, Iss. 4, pp. 394 - 406 (2006).
17. Garcia, H.C., Villalobos, J.R.: Automated Refinement of Automated Visual Inspection Algorithms. *IEEE Transactions on Automation Science and Engineering*, Vol. 6 , issue 3, pp. 514 – 524 (2009).
18. Zahorodnia, D., Pigovsky, Y. Bykovyy, P., Krylov, V., Sachenko, A., Molga, A.: Automated Video Surveillance System based on Hierarchical Object Identification. *14th International Conference on Development and Application Systems (DAS)*, Suceava (Romania), 24-26 May 2018, pp. 194-199 (2018).
19. Garcia, H.C., Villalobos, J.R., Pan, R., Runger, G.C.: A Novel Feature Selection Methodology for Automated Inspection Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 31, Issue 7, pp. 1338 – 1344 (2009).
20. Park, J.-S., Tou, J.T.: A solder joint inspection system for automated printed circuit board manufacturing. *Proceedings of IEEE International Conference on Robotics and Automation*. Vol.2, pp. 1290 – 1295 (1990).

21. Park, J.-S., Tou, J.T.: Automatic inspection of assembled PC board via highlight separation and dual channel processing. *Proceedings of IEEE International Conference on Robotics and Automation*. vol.3, pp. 2702 – 2707 (1991).
22. Pierce, B.L., Shelton, D.J., Longbotham, H.G., Baddipudi, S., Yan, P.: Automated inspection of through hole solder joints utilizing X-ray imaging. *IEEE Aerospace and Electronic Systems Magazine*, vol. 9, issue 2, pp. 28 – 32 (1994).
23. Asaad, F. Said, Bonnie, L., Bennett, L.J., Karam, Pettinato, J.S.: Automated Detection and Classification of Non-Wet Solder Joints. *IEEE Transactions on Automation Science and Engineering*, vol. 8, issue 1, pp. 67 – 80 (2011).
24. Kuk, Won Ko, Young, Jun Roh, Hyung, Suck Cho, Hyung, Cheol Kimn: A neural network approach to the inspection of ball grid array solder joints on printed circuit boards. *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. (IJCNN 2000). Neural Computing: New Challenges and Perspectives for the New Millennium*, vol. 5, pp. 233-238 (2000).
25. Ko, K.W., Cho, H. S.: Solder joint inspection using a neural network and fuzzy rule-based classification method. *IEEE Transaction on electronics packaging manufacturing*. No. 23 (2). pp. 93–103 (2000).
26. Ibrahim, Z., Aspar, Z., Al-Attas, S.A.R., Mokji, M.M.: Coarse resolution defect localization algorithm for an automated visual printed circuit board inspection. *IEEE 2002 28th Annual Conference of the Industrial Electronics Society. (IECON 02)*, Vol. 4, pp. 2629 - 2634 (2002).
27. Shcherbakova, G., Gerganov, M., Antoshchuk, S., Polyakova, M., Sachenko, A., Krylov, V.: Areal Multistart Method of Optimization for Image Recognition. *IEEE Second International Conference on Data Stream Mining & Processing, Lviv (Ukraine)*, 21-25 August 2018, pp.605-608 (2018).
28. Krylov, V., Antoshchuk, S., Shcherbakova, G.: Photomasks Diagnostic in the Automated Optical Inspection System. *Modern problems of radio engineering, telecommunication and computer science: Int. conf. TCSET'2008, Lviv – Slavske, Ukraine*, 19–23 February 2008, pp. 533–534 (2008).
29. Shcherbakova G., Krylov V., Abakumov V., Brovko V., Kozina I.: Sub Gradient Iterative Method for Neural Networks Training. In: *The 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications - Proceeding of Conference, Prague*, 15-17 September 2011, pp.361 – 364 (2011).
30. Paliy, I., Dovgan, V., Boumbarov, O., Panev, S., Sachenko, A., Kurylyak, Y., Zagorodnya, D.: Fast and Robust Face Detection and Tracking Framework. *Proceedings of the IEEE 6th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2011)*, Prague (Czech Republic), 15-17 September 2011, vol.1, pp. 430-433 (2011).
31. Zahorodnia, D., Pigovsky, Y., Bykovyy, P.: Canny-based Method of Image Contour Segmentation. *International Journal of Computing*, Vol. 15(3), pp. 200-205 (2016).
32. Zahorodnia, D., Pigovsky, Y., Bykovyy, P., Krylov, V., Sachenko, A.: Information technology for structural and statistical identification of hierarchical objects. *Proceedings of the 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske (Ukraine), 20-24 February 2018, pp. 272-275 (2018).