

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»

Інститут інформатики та радіoeлектроніки,
Факультет комп'ютерних наук і технологій
(повне найменування інституту, факультету)

Кафедра програмних засобів
(повне найменування кафедри)

Пояснювальна записка

до дипломного проєкту (роботи)

магістр

(ступінь вищої освіти)

на тему ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ МЕТОДІВ
РОЗПІЗНАВАННЯ МОВНОГО ПАРОЛЮ ДЛЯ УПРАВЛІННЯ ДОСТУПОМ
RESEARCH AND SOFTWARE IMPLEMENTATION OF SPEECH PASSWORD
RECOGNITION METHODS FOR ACCESS CONTROL

Виконав: студент 2 курсу, групи КНТ-120м
Спеціальності 121 Інженерія програмного
забезпечення
(код і найменування спеціальності)

Освітня програма (спеціалізація)
Інженерія програмного забезпечення

Тюнін О.О.

(прізвище та ініціали)

Керівник Олійник А.О.

(прізвище та ініціали)

Рецензент Скрупський С.Ю.

(прізвище та ініціали)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»
 (повне найменування закладу вищої освіти)

Інститут, факультет ІРЕ, ФКНТ
 Кафедра програмних засобів
 Ступінь вищої освіти магістр
 Спеціальність 121 Інженерія програмного забезпечення
(код і найменування)
 Освітня програма (спеціалізація) Інженерія програмного забезпечення
(назва освітньої програми (спеціалізації))

ЗАТВЕРДЖУЮ

Завідувач кафедри ПЗ, д.т.н, проф.
С.О. Субботін
 “ ” 2021 року

З А В Д А Н Н Я

НА ДИПЛОМНИЙ ПРОЄКТ (РОБОТУ) СТУДЕНТА(КИ)

Тюніну Олегу Олександровичу
(прізвище, ім'я, по батькові)

1. Тема проєкту (роботи) Дослідження та програмна реалізація методів розпізнавання мовного паролю для управління доступом. Research and Software Implementation of Speech Password Recognition Methods for Access Control

керівник проєкту (роботи) Олійник Андрій Олександрович, д.т.н., доцент,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом закладу вищої освіти від “05” листопада 2021 року № 435

2. Строк подання студентом проєкту (роботи) 5 грудня 2021 року

3. Вихідні дані до проєкту (роботи) рекомендована література, технічне завдання

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) 1. Аналіз проблемної області та постановка завдань дослідження. 2. Матеріали та методи. 3. Розробка архітектури програми. 4. Основні рішення щодо реалізації компонентів системи. 5. Експлуатація, тестування та експериментальне дослідження програми.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень) Слайди презентації

6. Консультанти розділів проєкту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	прийняв виконане завдання
1-5 Основна частина	Олійник А.О., професор		
Нормоконтроль	Липовець М.В., асистент		

7. Дата видачі завдання “ 4 ” вересня 2021 року.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проєкту (роботи)	Строк виконання етапів проєкту (роботи)	Примітка
1	Постановка завдання роботи.	1 тиждень	Завдання, ТЗ
2	Аналіз предметної області.	2-3 тижні	Розділ 1
3	Розробка та удосконалення методів, моделей й алгоритмів вирішення задачі.	4-5 тижні	Розділ 2
4	Розробка архітектури програми.	6 тиждень	Розділ 3
5	Розробка програми.	7-8 тижні	Розділ 4
6	Тестування та експериментальне дослідження програмного забезпечення.	9-10 тиждень	Розділ 5
7	Оформлення пояснювальної записки та документів до неї.	11-13 тиждень	Додатки
8	Нормоконтроль та рецензування.	14-15 тижні	
9	Захист роботи.	16 тиждень	

Студент(ка)

(підпис) Тюнін О.О.
(прізвище та ініціали)

Керівник проєкту (роботи)

(підпис) Олійник А.О.
(прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка до дипломної кваліфікаційної роботи магістра:
101 с., 29 рис., 4 табл., 4 дод., 21 джерело.

ІДЕНТИФІКАЦІЯ, ЗАХИСТ ДАНИХ, РОЗПІЗНАВАННЯ ОБРАЗІВ,
РОЗПІЗНАВАННЯ МОВЛЕННЯ, ЛІНІЙНЕ ПРОГНОЗУВАННЯ.

Об'єкт дослідження – процес розпізнавання мовного паролю для ідентифікації користувачів з метою управління доступом.

Предмет дослідження – методи розпізнавання мовлення.

Мета роботи – дослідити, проаналізувати та реалізувати методи розпізнавання мовлення з метою контролю доступу.

Матеріали, методи та технічні засоби: структурне та об'єктно-орієнтоване програмування, технології лінійного прогнозування, пакет Matlab та інтегроване середовище розробки, персональний комп'ютер з процесором Intel Core i5-11600K під управлінням операційної системи Microsoft Windows 10.

Результати. Створено програмний застосунок, що реалізує метод розпізнавання користувача за мовним паролем на основі лінійного прогнозування.

Висновки. Розроблено програмний застосунок для розпізнавання користувача за мовним паролем.

Галузь використання – системи захисту інформації.

ABSTRACT

Explanatory note to the diploma qualifying work of the master: 101 pages, 29 figures, 4 tables, 4 appendixes, 21 sources.

IDENTIFICATION, DATA PROTECTION, PATTERN RECOGNITION, SPEECH RECOGNITION, LINEAR FORECASTING.

The object of research is the process of recognizing a speech password to identify users for access control purposes.

The subject of the research is speech recognition methods.

The aim of the work is to research, analyze and implement speech recognition methods for access control purposes.

Materials, methods and technical tools: structural and Object-Oriented Programming, linear forecasting technologies, Matlab package and integrated development environment, personal computer with Intel Core i5-11600k processor running the Microsoft Windows 10 operating system.

Results. A software application has been created that implements a method for recognizing a user by speech password based on linear forecasting.

Conclusions. A software application has been developed for recognizing a user by their speech password.

The field of use is information security systems.

ЗМІСТ

	С.
Перелік скорочень та умовних позначок.....	8
Вступ.....	9
1 Аналіз проблемної області та постановка завдань дослідження.....	11
1.1 Ідентифікація особистості	11
1.2 Базові інформаційні ресурси	13
1.3 Правові основи ведення ідентифікації	14
1.4 Машинне навчання та методи машинного навчання для розпізнавання образів.....	15
1.5 Ідентифікація на основі біометричних параметрам з використанням методів розпізнавання образів.....	24
1.6 Постановка завдання	30
1.7 Висновки за розділом 1	31
2 Матеріали та методи	32
2.1 Структура систем розпізнавання дикторів.....	32
2.2 Аналіз методів розпізнавання мовного пароля на малому словнику...	34
2.3 Метод на основа лінійного передбачення.....	39
2.4 Висновки за розділом 2	47
3 Розробка архітектури програми	49
3.1 Опис основних вимог до програми.....	49
3.2 Вибір мови програмування.....	50
3.3 Вибір формату звукових файлів.....	53
3.4 Проектування розробки	54
3.5 Вимоги до технічних засобів.....	55
3.6 Висновки за розділом 3	56
4 Основні рішення щодо реалізації компонентів системи	57
4.1 Розробка програми введення мовного сигналу	57
4.2 Розробка програми вилучення інформаційних ознак з мовного сигналу.....	59

4.3 Розробка програми розпізнавання диктора.....	62
4.4 Висновки за розділом 4	65
5 Експлуатація, тестування та експериментальне дослідження програми ..	66
5.1 Опис ПЗ	66
5.2 Верифікація ПЗ	68
5.3 Висновки за результатами тестування	71
Висновки.....	72
Перелік джерел посилання.....	74
Додаток А Технічне завдання.....	77
Додаток Б Опис програми.....	82
Додаток В Текст програми.....	86
Додаток Д Слайди презентації	93

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАК

- CBR – case based reasoning;
PCM – pulse code modulation;
WAV – waveform audio format;
AP – авторегресія;
ЕОМ – електронно-обчислювальна машина;
КПФ – короткочасне перетворення Фур'є;
КС – ковзке середнє;
ЛПМ – лінійне передбачення мови;
МН – машинне навчання;
ОС – операційна система;
ПЗ – програмний застосунок;
ПК – персональний комп'ютер;
ШНМ – штучна нейронна мережа.

ВСТУП

З розвитком науки і техніки, появою перших електронно-обчислювальних машин (ЕОМ), все більш актуальним є питання спрощення спілкування між людиною і машиною. В даний час в світі є тенденція зростання частки інформації, що передається у формі мовних повідомлень. Все більша увага приділяється інтерфейсам, які використовують мову для взаємодії між Користувачем і комп'ютером. Основна частина систем розпізнавання мови включає процес аналізу і обробки аналогового сигналу і процес розпізнавання. При аналізі аналогового сигналу з мови виділяються ознаки (властивості), які використовуються далі в процесі розпізнавання для визначення сказаного.

Системи ідентифікації особистості по голосу для користувача є найбільш зручним способом аутентифікації, що забезпечує розпізнавання на більш високому рівні і дозволяє знизити витрати за рахунок автоматизації процесу. З усіх існуючих методів, ідентифікація по голосу є найменш нав'язливою, найбільш доступною для людини, і може здійснюватися віддалено, наприклад, по телефону. Варто відзначити, що голосова ідентифікація не вимагає застосування спеціалізованого дорогого обладнання, достатньо лише мати мікрофон, який зараз є практично скрізь (в комп'ютерах, мобільних телефонах).

Актуальність роботи пояснюється тим, що системи розпізнавання можна застосовувати для вирішення завдань розпізнавання слів, верифікації та ідентифікації диктора. Останні два використовуються при банківських і ділових операціях, в управлінні доступом до пристроїв і до інформації для обраних осіб або для осіб, які спостерігають за дотримання закону, в новому обладнанні для установ.

Верифікація та ідентифікація можуть використовуватися в системах управління доступом. При верифікації система обробляє вимовлену парольну фразу і порівнює її з еталоном, отриманим від диктора. На основі порівняння

виноситься рішення про ідентичність і можливість надання диктору доступу до інформації.

Мета роботи – розробити метод та його програмну реалізацію для ідентифікації диктора на основі мовного пароля.

Для досягнення цієї мети необхідно виконати наступні завдання:

- аналіз методів розпізнавання диктора;
- розробка програми розпізнавання диктора, на основі методу лінійного передбачення мови;
- верифікація програми;

1 АНАЛІЗ ПРОБЛЕМНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАВДАНЬ ДОСЛІДЖЕННЯ

Перший розділ буде присвячено детальному аналізу проблемної області пов'язаної з ідентифікацією. Особливу увагу буде приділено використанню методів розпізнавання для розбудови автоматичних систем ідентифікації. Під час роботи буде проведено дослідження основних термінів таких, як ідентифікація, розпізнавання, методи машинного навчання.

1.1 Ідентифікація особистості

1.1.1 Визначення області ідентифікації

Особистість – це конкретна людина, носій індивідуальних біологічних і психічних властивостей, а також певного соціального і юридичного статусу, зареєстрованого в документах [1].

Ідентифікація особистості – це процес встановлення тотожності особистості людини за сукупністю ознак шляхом порівняльного їх дослідження [1].

Встановити особу людини в більшості випадків означає визначити її прізвище, ім'я, по батькові, рік народження, Місце народження та інші установчі дані, які необхідні для однозначної ідентифікації людини. Для цілей встановлення особистості розроблені і використовуються безліч методів [1], [2].

Візуальна ідентифікація. Найбільш поширеним в економічній, правоохоронній та деяких інших сферах діяльності людини на сьогоднішній день є метод встановлення особи за особистими документами людини, які покликані підтверджувати основні настановні дані людини. Основний документ, що засвідчує особу, в нашій країні – паспорт (або id-картка). Якщо людина демонструє його або при ньому виявлений паспорт (або аналогічний йому документ), то вважається, що ця людина той, чиї дані знаходяться в

паспорті. Для підтвердження цього в паспорт поміщена фотографія, яка дозволяє методом порівняння зовнішності підтвердити або спростувати особу власника паспорта [3].

Електронна ідентифікація. При ідентифікації людини в «електронному вигляді» (інтернет, пластикові карти державні послуги в електронному вигляді, електронні та мобільні платежі, тощо) необхідно використання ідентифікаційного номера, що присвоюється або людині, або його персональним даними, який може однозначно визначити відповідність персональних даних і цієї людини [1]-[3].

1.1.2 Ідентифікаційна політика

Ідентифікація людини (персональних даних людини) може проводитися різними способами [1]-[3]:

- єдиний, незмінний протягом усього життя номер;
- номер, змінюваний у зв'язку зі зміною даних людини;
- використання сукупності номерів для ідентифікації.

Питання вибору способів ідентифікації має визначальне значення для розробки будь-якої інформаційної системи, що враховує персональні дані громадян та іншу інформацію, що дає людині можливість отримувати будь-які послуги як державні, так і комерційні.

1.1.3 Типи ідентифікації

Централізована ідентифікація – передбачає збір всієї інформації на верхньому (державному) рівні [4].

Розподілена ідентифікація – зберігання інформації про суб'єкта в місці його реєстрації. При розподіленій ідентифікації може бути організований обмін ідентифікаційними даними між різними рівнями реєстрації.

Ієрархічна ідентифікація – інформація про ідентифікований об'єкт присутня на всіх рівнях ієрархії [1]-[5].

1.2 Базові інформаційні ресурси

1.2.1 Єдиний базовий інформаційний ресурс

Для організації міжвідомчої взаємодії при наданні будь-яких послуг населенню в електронному вигляді можуть використовуватися різні способи ведення ресурсів, що містять ідентифікаційні дані людини. Як ідеальний варіант наявність єдиного базового інформаційного ресурсу. Ресурс, який є первинним для будь-яких інших інформаційних ресурсів, що містять і накопичують дані при здійсненні своїх функцій [1]-[5].

У разі використання єдиного базового інформаційного ресурсу будь-яка зміна ідентифікаційних даних про громадянина викликає зміну цих даних у вторинних ресурсах, тобто інформація, що зберігається в єдиному базовому інформаційному ресурсі є єдино вірною для всіх інших. Це ідеальна ситуація [1]-[5].

1.2.2 Базові інформаційні ресурси

У перехідний період (до створення єдиного базового інформаційного ресурсу) допустимо використання декількох базових інформаційних ресурсів, які накопичили вже досить великий обсяг інформації про громадян і продовжують здійснювати це накопичення при виконанні своїх функцій (в Україні – це єдиний державний реєстр, міністерство юстиції, міграційна служба, центри надання адміністративних послуг, тощо). При використанні декількох базових інформаційних ресурсів виникає необхідність вирішення ряду досить серйозних питань [1]-[5]:

- визначення базових ідентифікаторів;
- встановлення відповідності базових ідентифікаторів один одному;

- визначення первинних (master) ресурсів, які послужать еталоном для всіх інших;
- визначення процедури виправлення помилок і введення змін в базові ресурси при виявленні таких.

При наявності великої кількості таких ресурсів робота з ними буде вельми скрутна, також як і ризик виникнення ситуації з неправильно ідентифікованими персональними даними громадян [1]-[5].

1.3 Правові основи ведення ідентифікації

1.3.1 Ідентифікація громадян

Питання присвоєння ідентифікатора персональним даним громадян в обов'язковому порядку повинен бути визначений нормативно-правовим актом, в якому має бути визначено де, коли, ким присвоюється ідентифікатор, ким він ведеться, де зберігається і як застосовується. При наявності декількох ідентифікаторів обов'язково повинні бути описані зони їх застосування і зони їх зіткнення. Без нормативно-правового закріплення цих основоположних норм процес використання базових ідентифікаторів приречений на провал [1]-[3].

1.3.2 Створення інформаційного (інформаційних) ресурсів

Після нормативно-правового закріплення основ ідентифікації необхідно визначити основні принципи розробки інформаційної системи (систем), що дозволяють проводити ідентифікацію громадян. Дана система (системи) повинні не тільки забезпечувати накопичення інформації про громадян, а й мати всю інфраструктуру. Яка дозволить користуватися даним інформаційним ресурсом, причому не тільки в програмній частині, але і в технічному забезпеченні. Використання цього ресурсу повсюдно повинно припускати наявність розгалуженої мережі передачі даних, доступ до якої

повинен бути забезпечений з будь-якої точки країни (і з інших країн також), повинні бути визначені всі принципи захисту інформації, забезпечення захисту від несанкціонованого доступу як зсередини, так і ззовні. Крім цього, також повинні бути вирішені питання забезпечення достатньої продуктивності даного інформаційного ресурсу, оскільки інформаційний обмін з базовим інформаційним ресурсом вище на кілька порядків, ніж з ресурсами, використовуваними для будь-якої однієї галузі діяльності [1]-[5].

1.4 Машинне навчання та методи машинного навчання для розпізнавання образів

Загальний термін машинне навчання (МН) позначає безліч математичних, статистичних і обчислювальних методів для розробки методів, здатних вирішити задачу не прямим способом, а на основі пошуку закономірностей в різноманітних вхідних даних. Рішення обчислюється не за чіткою формулою, а за встановленою залежністю результатів від конкретного набору ознак і їх значень. Наприклад, якщо кожен день протягом тижня земля покрита снігом і температура повітря істотно нижче нуля, то найімовірніше, настала зима. Тому МН застосовується для діагностики, прогнозування, розпізнавання і прийняття рішень в різних прикладних сферах: від медицини до банківської діяльності [6]-[9].

Виділяють 2 типи машинного навчання [6]-[9]:

- індуктивне або за прецедентами, яке засноване на виявленні емпіричних закономірностей у вхідних даних;
- дедуктивне, яке передбачає формалізацію знань експертів і їх перенесення в цифрову форму у вигляді бази знань.

Дедуктивний тип прийнято відносити до області експертних систем, тому загальний термін МН означає навчання за прецедентами. Прецеденти або навчальна вибірка – це набори вхідних об'єктів і відповідних їм результатів. При цьому не існує чіткої формули, яка аналітично описує

залежність між результатами і входами. Наприклад, яка погода буде завтра, якщо протягом тижня дні були морозні, сонячні, з низькою вологістю повітря, без вітру і опадів? При цьому слід врахувати ще безліч параметрів: географічні координати, рельєф місцевості, рух теплих і холодних фронтів повітря та ін., необхідно побудувати алгоритм, який видасть досить точний результат для будь-якого можливого входу. Точність результатів регулюється оціночним функціоналом якості. Таким чином, рішення формується емпірично, на основі аналізу накопиченого досвіду. При цьому навчається система повинна бути здатна до узагальнення-адекватного відгуку на дані, що виходять за межі наявної навчальної вибірки. На практиці вхідні дані можуть бути неповними, неточними і різнорідними. Тому існує безліч методів машинного навчання. Можна сказати, що МН реалізує підхід Case Based Reasoning (CBR) – метод вирішення проблем міркуванням за аналогією, шляхом припущення на основі подібних випадків (прецедентів) [6]-[9].

1.4.1 Методи МН

Існує безліч методів МН. Слід навести найпопулярніші, залишивши їх детальну класифікацію спеціалізованим ресурсам. Виділяють 2 види класичного МН [6]-[9].

З учителем, коли необхідно знайти функціональну залежність результатів від входів і побудувати алгоритм, на вході приймає опис об'єкта і на виході видає відповідь. Функціонал якості, як правило, визначається через середню помилку відповідей алгоритму по всіх об'єктах вибірки. До навчання з учителем відносяться завдання класифікації, регресії, ранжирування і прогнозування [6]-[9].

Без вчителя, коли відповіді не задаються, і потрібно шукати залежності між об'єктами. Сюди входять завдання кластеризації, пошуку

асоціативних правил, фільтрації викидів, побудови довірчої області, скорочення розмірності та заповнення пропущених значень.

До неklasичних, але вельми популярним методам відносять навчання з підкріпленням, зокрема, генетичні алгоритми, і штучні нейронні мережі. В якості вхідних об'єктів виступають пари: ситуація – прийняте рішення, а відповідями є значення функціоналу якості, який характеризує правильність прийнятих рішень (реакцію середовища). Ці методи успішно застосовуються для формування інвестиційних стратегій, автоматичного управління технологічними процесами, Самонавчання роботів та інших подібних завдань [6]-[9].

1.4.2 Розпізнавання образів

Розпізнавання образів – науковий напрямок, пов'язаний з розробкою принципів і побудовою систем, призначених для визначення приналежності даного об'єкта до одного із заздалегідь виділених класів об'єктів [6]-[8].

З розвитком обчислювальної техніки стало можливим вирішити ряд завдань, що виникають в процесі життєдіяльності, полегшити, прискорити, підвищити якість результату [6]-[8]. Наприклад, робота різних систем життєзабезпечення, взаємодія людини з комп'ютером, поява роботизованих систем та ін. [6]-[8]. Проте, відзначимо, що забезпечити задовільний результат в деяких завданнях (розпізнавання швидкорухливих подібних об'єктів, рукописного тексту) в даний час не вдається [6]-[8]. Таким чином, в цій статті пропонується обговорити методи і принципи, що застосовуються в обчислювальній техніці для виконання поставленого завдання [6]-[8].

Завдання пошуку зображення за зразком є частиною (підзадачею) більш загальної задачі розпізнавання образів [6]-[8]. При несистематизованому і ненаправленому пошуку схожих об'єктів з множини об'єктів, їх можна перераховувати нескінченно довго і не прийти до завершення із заданою ймовірністю [6]-[8]. У окремих випадках об'єкти

характеризуються такими ідентифікаційними параметрами (ознаками), як форма, колір, положення, рухливість, за відмітними особливостями, їх комбінації і т.п. залежно від цих факторів об'єкти піддаються класифікації [6]-[8]. Часто стоїть не глобальне завдання класифікації всіх навколишніх об'єктів, а необхідність виділити в надходить відео-потоці об'єкти певного роду. Далі розглянуті найбільш поширені класифікаційні ознаки [6]-[9].

Розпізнавання форм (об'єктів, сигналів, ситуацій, явищ або процесів) - задач ідентифікації об'єкта або визначення будь-яких його властивостей по його зображенню (оптичне розпізнавання) або аудіозапису (акустичне розпізнавання) та інших характеристик [6]-[8].

Одним з базових є поняття, що не має конкретного формулювання – множина [6]-[8]. У комп'ютерах множина представлена складом неповторних однотипних елементів [6]-[8]. Термін неповторних означає, що якийсь елемент в множині або є, або його там немає [6]-[8]. Універсальна множина включає всі керовані для розв'язуваної задачі елементи, порожнє не стримає жодного [6]-[8].

Образ – класифікаційне угруповання в системах класифікації, що об'єднує (виділяє) певну групу об'єктів за деякою ознакою [6]-[8]. Щоки мають характерним своїм, що виявляється в тому, що позначення з кінцевим числом явищ з одного і того ж множення дає можливість отримати наскільки приємно велике число його представників [6]-[8]. Щоки мають характерні об'єктивні властивості в тому сенсі, що різні люди, які навчаються різному матеріалу спостережень, великою частиною однаково і незалежні один від одного класифікують одні й ті ж об'єкти [6]-[8]. У класичній постановці завдання розпізнається універсальне множення. Кожне відображення будь-якого об'єкта на вибухають органи розпізнає системи, незалежні від його розміщення власне цих органів, прийнято викликати зображення об'єкта, а множення таких зображень, об'ємні якими-небудь загальними, являють собою власні форми [6]-[8].

Методика віднесення елемента до якого-небудь щоду називається вирішальним правилом [6]-[8]. Ще одне важливе значення – метрика, здатність визначення відстані між елементами універсального множення. Чим менша ця відстань, тим більше схожими є об'єкти (символи, звуки та ін.) [6]-[8]. Звичайні елементи задаються у вигляді складок чисел, а метрика – в широких функціях [6]-[8]. Від вибору подання форм і реалізації метрики залежить ефективність програми, один алгоритм визнання з різними метриками буде помилятися з різною частотою [6]-[8].

Навчанням зазвичай називають процес вироблення в деякій системі тієї чи іншої реакції на групи зовнішніх ідентичних сигналів через багаторазового водіння по системі зовнішньої коригування [6]-[8]. Таку зовнішню коригування в навчанні прийнято називати нагородами та штрафами [6]-[8]. Механізм генерації цього коригування практична повнота визначає метод навчання [6]-[8]. Самонавчання відрізняється від навчання тим, що тут додаткова інформація про вірність реакції системи не повідомляється [9].

Адаптація – це процес зміни параметрів і структур системи, а також керуючих команд, до основ поточної інформації з метою досягнення певного стану системи при основній невизначеності і змінюють умовах роботи [9].

Навчання – це процес, в результаті якого системи поступово набуває здатність реагувати необхідними реакціями на певні сукупності зовнішніх властивостей, а адаптація-це підстроювання параметрів і структур системи з метою досягнення необхідної якості управління в умовах неперервних змін зовнішніх умов [9].

Приклади завдань визнання освіт [6]-[9]:

- визнання букв (символи абетки або ж окремі криптографічні позначення);
- розпізнання штрихів-кодів;
- розпізнавання автомобільних номерів;
- розпізнавання осіб та інших біометричних даних;

- розпізнавання зображень;
- розпізнавання слів.

1.4.3 Методи розпізнавання образів

В цілому, можна виділити наступні методи розпізнавання образів [6]-[9].

Метод перебору. У цьому випадку проводиться порівняння з базою даних, де для кожного виду об'єктів представлені всілякі модифікації відображення [6]-[9]. Наприклад, для оптичного розпізнавання образів можна застосувати метод перебору виду об'єкта під різними кутами, масштабами, зсувами, деформаціями і т.д. для букв потрібно перебирати шрифт, властивості шрифту і т. д. у разі розпізнавання звукових образів, відповідно, відбувається порівняння з деякими відомими шаблонами (наприклад, слово, вимовлене декількома людьми) [7]-[9].

Другий підхід – проводиться більш глибокий аналіз характеристик образу [7]-[9]. У разі оптичного розпізнавання це може бути визначення різних геометричних характеристик [7]-[9]. Звуковий зразок в цьому випадку піддається частотному, амплітудному аналізу і т. д. [7]-[9].

Наступний метод – використання штучних нейронних мереж (ШНМ). Цей метод вимагає або великої кількості прикладів завдання розпізнавання при навчанні, або спеціальної структури нейронної мережі, що враховує специфіку даного завдання [7]-[9]. Тим не менш, його відрізняє більш висока ефективність і продуктивність [7]-[9].

Експертний метод, заснований на безперервному навчанні експертної системи в процесі експлуатації [7]-[9].

1.4.4 Загальна характеристика задач розпізнавання образів

Загальна структура системи розпізнавання і етапи в процесі її розробки показані на рис. 1.1.

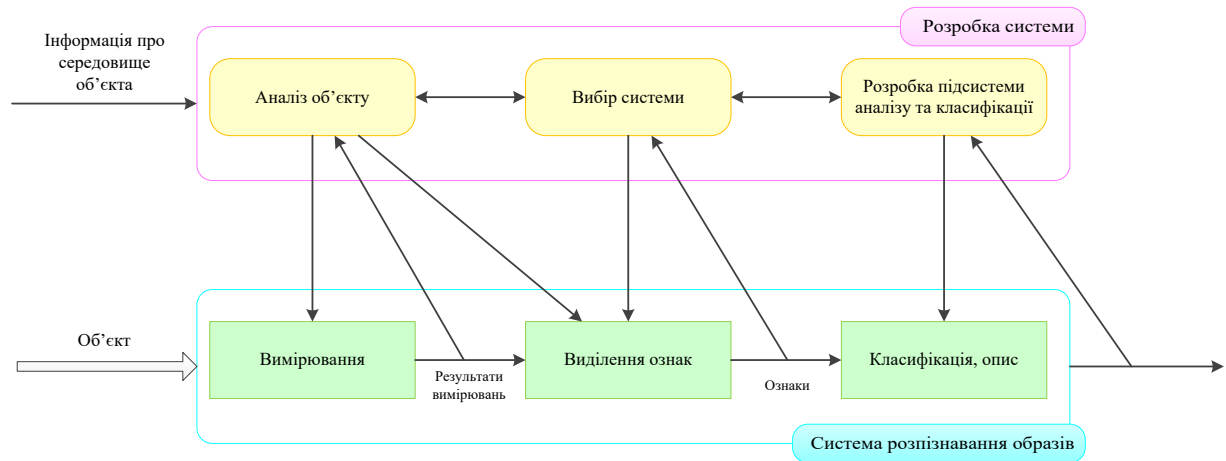


Рисунок 1.1 – Приклад структури системи розпізнавання

Завдання розпізнавання – це інформаційні завдання, що складаються з двох етапів [7]-[9]:

- перетворення вихідних даних до вигляду, зручного для розпізнавання;
- власне розпізнавання (вказівка належності об'єкта певному класу).

У цих завданнях можна вводити поняття аналогії або подібності об'єктів і формулювати правила, на підставі яких об'єкт зараховується в один і той же клас або в різні класи [7]-[9]. У цих завданнях можна оперувати набором прецедентів-прикладів, класифікація яких відома і які у вигляді формалізованих описів можуть бути пред'явлені алгоритму розпізнавання для налаштування на задачу в процесі навчання [7]-[9]. Для цих завдань важко будувати формальні теорії і застосовувати класичні математичні методи (часто недоступна інформація для точної математичної моделі або виграш від використання моделі і математичних методів непорівнянний з витратами) [7]-[9].

Виділяють наступні типи завдань розпізнавання [7]-[9]:

- завдання розпізнавання – віднесення пред'явленого об'єкта за його описом до одного із заданих класів (навчання з учителем) [7]-[9];
- завдання автоматичної класифікації – розбиття множини об'єктів, ситуацій, явищ за їх описами на систему непересічних класів (таксономія, кластерний аналіз, самонавчання) [7]-[9];
- завдання вибору інформативного набору ознак при розпізнаванні;
- завдання приведення вихідних даних до вигляду, зручного для розпізнавання [7]-[9];
- динамічне розпізнавання і динамічна класифікація – завдання 1 і 2 для динамічних об'єктів [7]-[9];
- завдання прогнозування-попередній тип, в якому рішення має відноситися до деякого моменту в майбутньому.

1.4.5 Визначення задачі розпізнавання образів

Розпізнавання образів, науковий напрямок, пов'язаний з розробкою принципів і побудовою систем, призначених для визначення приналежності даного об'єкта до одного із заздалегідь виділених класів об'єктів [7]. Під об'єктами в розпізнавання образів розуміють різні предмети, явища, процеси, ситуації, сигнали [7]. Кожен об'єкт описується сукупністю основних характеристик (ознак, властивостей) $X = (x_1, \dots, x_i, \dots, x_n)$, де i -а – координата вектора X визначає значення i -ої характеристики, і додатково характеристикою S , яка вказує на приналежність об'єкта до деякого класу (образу) [7]. Набір заздалегідь прокласифікованих об'єктів, тобто таких, у яких відомі характеристики X та S , використовується для виявлення закономірних зв'язків між значеннями цих характеристик, і тому називаються навчальною вибіркою [7]. Ті об'єкти, у яких характеристика S невідома, утворюють контрольну вибірку. Окремі об'єкти навчальної та контрольної вибірок називаються реалізаціями [7]. Одна з основних завдань розпізнавання образів-вибір правила (вирішальної функції) D , відповідно до

якого за значенням контрольної реалізації X встановлюється її приналежність до одного з образів, тобто вказуються найбільш правдоподібні значення характеристики S для даного X [7]-[9].

Успіх у вирішенні завдання розпізнавання образів залежить значною мірою від того, наскільки вдало обрані ознаки X [7]. Вихідний набір характеристик часто буває дуже великим. У той же час прийнятне правило має бути засноване на використанні невеликого числа ознак, найбільш важливих для відмінності одного образу від іншого [7]. Так, в завданнях медичної діагностики важливо визначити, які симптоми і їх поєднання (синдроми) слід використовувати при постановці діагнозу даного захворювання [7]. Тому проблема вибору інформативних ознак-важлива складова частина проблеми розпізнавання образів [7]-[9].

Проблема розпізнавання образів тісно пов'язана із завданням попередньої класифікації, або таксономією [7]-[9].

В основному завданні розпізнавання образів побудови вирішальних функцій D використовуються закономірні зв'язки між характеристиками X і S , що виявляються на навчальній вибірці, і деякі додаткові апіорні припущення, наприклад, такі гіпотези: характеристики X для реалізацій образів являють собою випадкові вибірки з генеральних сукупностей з нормальним розподілом; реалізації одного образу розташовані компактно (в деякому сенсі); ознаки в наборі X незалежні і т. д. [7].

В області розпізнавання образів істотно використовуються ідеї і результати багатьох ін. наукових напрямків – математики, кібернетики, психології і т. д. [7].

У 60-х рр. 20 ст. у зв'язку з розвитком, електронної техніки, зокрема ЕОМ, широке застосування отримали автоматичні системи розпізнавання [7]. Під системами розпізнавання зазвичай розуміють комплекси засобів, призначених для вирішення описаних вище, завдань [7]. Методи розпізнавання образів використовуються в процесі машинної діагностики різних захворювань, для прогнозування корисних копалин в геології, для

аналізу економічних і соціальних процесів, в психології, криміналістиці, лінгвістиці, океанології, хімії, ядерної та космічної фізики, в автоматизованих системах управління і т.д. Їх застосування виправдано практично всюди, де доводиться мати справу з класифікацією експериментальних даних [7].

Одним із способів вирішення завдання розпізнавання образів є використання спектральних методів [7]-[9].

1.5 Ідентифікація на основі біометричних параметрам з використанням методів розпізнавання образів

1.5.1 Ідентифікація особистості за біометричними параметрами

Біометрією називається сукупність способів і пристроїв для ідентифікації людини, які засновані на його унікальних фізіологічних або поведінкових характеристиках [10]-[13].

Цей вид ідентифікації може застосовуватися для запобігання забороненого доступу в будівлі, до комп'ютерів, банкоматів, мобільних телефонів і так далі [10]-[13].

Біометричні властивості це [10]-[13]:

- відбитки пальців;
- геометрія обличчя;
- райдужна оболонка очей;
- малюнок сітківки;
- голос;
- почерк;
- друк на клавіатурі;
- візерунок вен на руках та ін.

1.5.2 Переваги біометричної ідентифікації

Біометричний захист дає більший ефект у порівнянні, наприклад, з використанням паролів, смарт-карт, PIN-кодів, жетонів або технології інфраструктури відкритих ключів [14]-[16]. Це пояснюється можливістю біометрії ідентифікувати не пристрій, але людину [14]-[16].

Звичайні методи захисту загрожують втратою або крадіжкою інформації, яка стає відкритою для незаконних користувачів [14]-[16]. Винятковий біометричний ідентифікатор, наприклад, відбитки пальців, є ключем, що не підлягає втраті [14]-[16].

Класифікація способів біометрії [14]-[16].

За типом використовуваної інформації біометрична ідентифікація ділиться на [10]-[12].

1. Статичні способи, засновані на унікальних властивостях, даних людина від народження і невід'ємних від нього [14]-[16]. Фізіологічні показники (геометрія долоні або папілярний візерунок пальців) є незмінними для людини [14]-[16].

2. Динамічні способи, засновані на поведінкової (тобто динамічної) характеристиці особистості. Ці особливості характерні для підсвідомих рухів при відтворенні будь-яких дій (мови, підпису, динаміки клавіатурного набору) [14]-[16]. Такі поведінкові характеристики відчувають вплив керованих і не дуже керованих психічних факторів [14]-[16]. Через їх мінливість біометричні зразки повинні оновлюватися при їх використанні [14]-[16].

1.5.3 Порівняння способів ідентифікації особи за біометричними параметрами

1. Дактилоскопія. Цей метод розпізнавання є найпоширенішим [10]-[16]. Він використовує неповторність папілярних візерунків пальців для

кожної людини [10]-[16]. Спеціальним сканером отримують зображення пальцевого відбитка [10]-[16]. Воно трансформується в цифровий код і зіставляється з шаблоном, введеним раніше [10]-[16].

Процес ідентифікації триває більше декількох секунд [10]-[16]. Певний недолік, що стримує розвиток цього методу, полягає в упередженні деяких людей, які не бажають залишати дані про свої відбитки пальців [10]-[16]. Контраргумент розробників апаратури полягає в тому, що інформація про папілярному візерунку не зберігається, а зберігається тільки короткий ідентифікаційний код, збудований за відбитком пальця і не дозволяє відтворити візерунок для порівняння [10]-[16]. Перевагою методу є простота у використанні, надійність і зручність [10]-[16].

2. Ототожнення за формою руки. Цей статичний метод заснований на вимірюванні форми кисті руки [10]-[16]. Вона також є унікальним біометричним параметром людини [10]-[16]. Спеціальний пристрій дозволяє отримати тривимірний вигляд кисті [10]-[16]. В результаті отримують вимірювання для створення унікального цифрового коду, що ідентифікує людину [10]-[16].

Даний метод за своєю технологією і точності можна порівняти з методом ототожнення за відбитком пальця, хоча сам пристрій для реалізації методу займає багато місця [10]-[16]. Надзвичайно мала ймовірність наявності двох ідентичних кистей рук, що мають однакову геометрію, хоча руки з віком змінюються [10]-[16].

Сьогодні ідентифікація з геометрії руки застосовується в законодавчих органах, лікарнях, міжнародних аеропортах і т. д. [10]-[16].

3. Аутентифікація за райдужною оболонкою. Основою цього методу є винятковість візерунка на райдужній оболонці ока [10]-[16]. Для його виконання потрібна камера, щоб отримувати зображення ока з достатньою роздільною здатністю, і спеціальне програмне забезпечення для виділення з отриманого зображення малюнка на райдужній оболонці [10]-[16]. По ньому і створюється цифровий код, службовець для ідентифікації людини [10]-[16].

Перевагою сканерів є те, що від людини не вимагають зосереджуватися на цілі, оскільки зразок плям райдужної оболонки зосереджений на поверхні ока [10]-[16]. Сканування можливо на відстані менше 1 м. це зручно для використання, наприклад, в банкоматах [10]-[16].

4. Ідентифікація по сітківці ока. Сітківки сканується за допомогою низькоінтенсивного інфрачервоного світла, який направляється до кровоносних судин задньої стінки ока через зіницю [10]-[16]. Сканери сітківки широко поширені в системах доступу на секретні об'єкти, оскільки у них майже не буває неправильного дозволу доступу [10]-[16]. Помилки можуть пояснюватися відхиленням голови від еталонного положення і неправильної фокусуванням погляду на джерелі світла [10]-[16].

Навіть у близнюків різниться капілярний малюнок сітківки [10]-[16]. Ось чому цей спосіб може успішно використовуватися для ідентифікації особистості [10]-[16].

Недоліком таких систем можна віднести психологічний фактор: не кожна людина може дивитися в темний отвір, в якому в око щось світить. Крім того, ці системи чутливі до невірної орієнтації сітківки, тому треба уважно стежити за положенням ока по відношенню до отвору [10]-[16].

5. Розпізнавання по рукописному почерку [10]-[16]. Основою ідентифікації по почерку служить унікальність і стабільність цього фактора для кожної людини [10]-[16]. Характеристики вимірюються, переводяться в цифровий вигляд і піддаються комп'ютерній обробці. Тобто для порівняння вибирається не лист як продукт, а сам процес [10]-[16].

Поширені два методи обробки даних: звичайне порівняння зі зразком і динамічна верифікація [10]-[16]. Перший ненадійний, тому що підпис не завжди однаковий [10]-[16]. Такий метод призводить до великого відсотка помилок [10]-[16]. Динамічна верифікація полягає в більш складних обчисленнях [10]-[16]. Цим методом в реальному часі реєструються параметри самого процесу підпису: швидкість руху руки на різних ділянках, силу тиску і тривалість різних етапів підпису [10]-[16]. Це виключає

підробку, так як неможливо в точності скопіювати руху руки автора підпису [10]-[16].

6. Розпізнавання по клавіатурному почерку [10]-[16]. Цей метод, загалом, аналогічний описаному вище, проте підпис в ньому замінюється певним кодовим словом, а з обладнання потрібна лише звичайна клавіатура [10]-[16]. Основною ідентифікаційною характеристикою є динаміка клавіатурного набору кодового слова [10]-[16].

Згідно з сучасними дослідженнями, клавіатурний почерк володіє певною стабільністю, завдяки чому можна однозначно ідентифікувати особистість [10]-[16]. Вихідними даними є час між натисканням клавіш і їх утримання [10]-[16]. Причому час між натисканням показує темп роботи, а утримання – стиль роботи, тобто плавне натискання або різкий удар [10]-[16].

Спочатку на етапі фільтрації видаляються дані про «службових» клавішах – функціональних, управління курсором і т. д. [10]-[16].

Потім виділяються наступні характеристики користувача [10]-[16]:

- число помилок в процесі набору;
- час між натисканнями на клавіші;
- швидкість набору.
- час на утримання клавіш;
- аритмічність при наборі.

7. Розпізнавання за голосом [10]-[16]. Біометричний метод ідентифікації голосу зручний в застосуванні. Причинами його впровадження є широке поширення телефонних мереж і вбудовування мікрофонів в комп'ютери [10]-[16]. Недоліками можна вважати фактори, що впливають на розпізнавання: перешкоди в мікрофонах, навколишні шуми, помилки в процесі проголошення, різний емоційний стан людини при ідентифікації і т. п. [10]-[16].

Головне в побудові пристроїв аутентифікації по голосу – вибір параметрів, найкраще описують індивідуальність голосу [10]-[16]. Ці параметри сигналу називаються ознаками індивідуальності. Такі ознаки, крім

даних про особливості голосу, повинні мати і інші властивостей [10]-[16]. Наприклад, вони повинні легко вимірюватися, і мало залежати від шумів і перешкод [10]-[16]. Крім того, вони повинні володіти стабільністю в часі і чинити опір імітації [10]-[16].

Розроблено системи із застосуванням методу комбінованого аналізу голосу з мімікою. Виявляється, міміка мовця відрізняє тільки його і буде інший у вимовляє ті ж слова іншої людини [10]-[16].

Порівняємо способи ідентифікації за біометричними показниками, результати порівняння представимо в табличному виді – табл. 1.1.

Таблиця 1.1 – Результати порівняння способів ідентифікації за біометричними показниками

	Вимоги до додаткового обладнання	Вимоги до середовища	Надійність при великому рівні схожості	Вимоги до обробки вхідних даних
Дактилоскопія	Мінімальні (сканер)	Мінімальні	Висока	Середні
За формою руки.	Мінімальні (сканер)	Мінімальні	Середня	Середні
За райдужною оболонкою	Середні (фотосканер)	Середні	Висока	Високі
За сітківкою ока	Середні (фотосканер)	Середні	Висока	Високі
По рукописному почерку	Залежить від способу	Середні	Середня	Середні
По клавіатурному почерку	Високі	Мінімальні	Середня	Високі
За голосом	Мінімальні (мікрофон)	Середні	Висока	Середні

З порівняння можна побачити, що на сьогоднішній день більшість методів ідентифікації на основі біометричних параметрів мають прийнятний рівень надійності, навіть у ситуаціях, коли необхідно провести ідентифікацію за достатньо схожими даними (наприклад, у ситуації ідентифікації

близнюків). Проте, такий рівень забезпечується або через високі вимоги до скануючого обладнання, або до обробки вхідних даних. Так при ідентифікації по клавіатурному почерку необхідно замірювати достатньо багато параметрів, що вимагає додаткових сканерів та фіксуючих пристроїв. При ідентифікації за сітківкою ока або райдужною оболонкою знадобиться деталізоване зображення очей аби в подальшому можливо було відслідкувати велику кількість дрібних параметрів пігментного або судинного візерунку. В цьому випадку ідентифікація за голосом є більш оптимальним вибором, адже для такої процедури достатньо звичайного мікрофона смартфона, а основні дії з обробки вхідних даних направлені на усунення навколишнього шуму.

1.6 Постановка завдання

Сучасні дослідження в області безпеки показують, що зловмисник може порівняно легко заволодіти паролем (PIN-кодом, смарт-картою і т.д.), а потім, знаючи цей пароль, отримати доступ до персональних даних людини, до банківського рахунку або персонального комп'ютера. Отже, пароль повинен бути прив'язаний до персони. Такими можливостями володіють біометричні ознаки людини. До них відносяться: відбитки пальців, геометрія долоні, геометрія особи, райдужна оболонка ока, сітківка ока, голос, рукописний почерк та інші. У даній роботі розглядається голосова біометрія, так як голос невіддільний від людини, і система при цьому розпізнає саму людину, а не того хто знає пароль.

У голосової біометрії є ряд переваг, які не властиві іншим способам управління доступом. Такими є не дорогий пристрій зчитування сигналу (мікрофон) і можливість віддаленої ідентифікації користувача (по телефону), чого немає у інших систем. Так само варто зауважити, що даний спосіб ідентифікації дуже зручний: смарт-карта може бути втрачена, PIN-код – забутий, голос же завжди з нами.

Метою даної роботи є розробка програми ідентифікації диктора на основі мовного пароля. Метод ідентифікації має бути заснований на методі розпізнавання образів для підвищення надійності та пришвидшені роботі системи. Програмний застосунок (ПЗ), що реалізує запропонований метод у якості вхідних даних буде використовувати аудіофайли формату – waveform audio format (WAV) в кодуванні pulse code modulation (PCM).

1.7 Висновки за розділом 1

Під час роботи над першим розділом було розглянуто проблемну область ідентифікації особистості. Було з'ясовано, що більш надійними та актуальними методами ідентифікації є методи засновані на розпізнаванні біометричних параметрів особистості. Використання таких підходів дозволить підвищити якість роботи методу, що буде розроблено та реалізує чого його ПЗ.

2 МАТЕРІАЛИ ТА МЕТОДИ

В першому розділі було проаналізовано сучасне становище проблемної області ідентифікації особистості та ідентифікації за біометричними параметрами. Було проаналізовано сучасні практики біометричної ідентифікації.

Інтерес до систем ідентифікації обумовлений широким спектром практичних застосувань: перевірка прав доступу до системи (баз даних, каналів зв'язку, приміщень, пристроїв і механізмів, банківських рахунків і т.д.), криміналістична експертиза.

Переваги голосової ідентифікації випливають з характеристик голосу: не відчужуємо від людини, не вимагає безпосереднього контакту, не вимагає складних технічних пристроїв. Голос диктора унікальний через специфіку фізіологічної будови його артикулярного апарату і специфіки його мови. Це обумовлює інтерес до нього як до біометричного об'єкту.

Метою розділу є аналіз існуючих методів обробки мовного сигналу. З'ясувати який метод найбільш оптимальний при ідентифікації диктора на малому словнику.

У цьому розділі розглянуті основні методи ідентифікації диктора по голосу.

2.1 Структура систем розпізнавання дикторів

Робота систем розпізнавання містить два основних етапи: Реєстрація користувачів в системі і процес розпізнавання (ідентифікація або верифікація). Користувачі заздалегідь реєструються в системі, записавши свої голоси. Кожен зразок проходить етап обробки з метою вилучення ознак, які в подальшому можуть бути використані для розпізнавання [2], [13]-[16].

На основі ознак будуються моделі, які являє собою якусь структуру, що дозволяє при даних ознаках оцінити ступінь подібності або відразу прийняти рішення.

Верифікація диктора полягає у визначенні того, що фраза вимовлена певним диктором, для якого в системі зберігається еталонний зразок фрази. Користувач намагається увійти в систему, пред'являючи ідентифікатор і зразок голосу. Ознаки, витягнуті з пред'явленого зразка, порівнюються з відповідною моделлю, збереженою в базі. Результат порівнюється із заданим порогом і видається рішення про допуск [2], [13]-[16].

Ідентифікація диктора полягає у визначенні того, який диктор з групи дикторів вимовив дану фразу. Для цього треба виконати порівняння вимовленої фрази з безліччю еталонів, кожен з яких відповідає одному диктору з групи. На підставі порівняння, виноситься рішення про ідентичність фрази одному з наявних у системі еталонів.

Загальна схема системи розпізнавання реалізується за допомогою наступних основних етапів або рівнів [7]-[9].

1. Рівень обробки сигналів: на цьому рівні відбувається Витяг ознак з мовного сигналу, які істотні для завдання розпізнавання. Потім сигнал представляється у вигляді послідовності векторів витягнутих ознак [7].

2. Рівень моделей. При реєстрації користувача даний рівень використовує інформацію (вектора ознак), отриману від рівня обробки сигналів, для побудови моделі. Моделювання може полягати як в простому копіюванні векторів ознак, так і в побудові імовірнісних моделей або інших структур. Після чого стає можливим при даних ознаках обчислити ступінь подібності між ознаками і збереженою моделлю [7].

3. Рівень прийняття рішень. Функції прийняття рішень традиційно виділяють в окремий рівень, хоча він може виконувати тривіальні функції або бути відсутнім, якщо на рівні моделей обчислюються кінцеві рішення. Для прийняття рішень використовуються ступені подібності, обчислені на рівні моделей, і, якщо необхідно, задані пороги [7].

2.2 Аналіз методів розпізнавання мовного пароля на малому словнику

Існує безліч методів обробки мовного сигналу. Вибір певної методики залежить від поставленого завдання розпізнавання. У нашому випадку необхідна саме біометрична ідентифікація людини по голосу, тому будемо розглядати методи, які можна використовувати для вирішення цього завдання. Такими методами є: спектральний аналіз, кепстральний аналіз (гомоморфна обробка) і методи лінійного передбачення мови [2], [13]-[16].

Основним завданням аналізу мови при верифікації особистості по голосу є виділення акустичних параметрів, що мають відношення до характеристик голосового тракту людини. Такими характеристиками є, перш за все, форманти (формантні частоти) і основна частота (період основного тону) [2], [13]-[16].

Для розпізнавання можна також використовувати наступні ознаки: енергія сигналу, коефіцієнти лінійного передбачення, кепстральні коефіцієнти [2], [13]-[16].

2.2.1 Спектральний аналіз мовного сигналу

Для вивчення спектральних характеристик мови зручно використовувати короткочасне перетворення Фур'є (КПФ), що дозволяє аналізувати спектр на інтервалах часу 10-30 мс. Цей метод широко застосовується в різних додатках аналізу мови. За спектрами КПФ можуть бути виділені формантні частоти локалізованих звуків. Ці частоти є індивідуальними характеристиками диктора і можуть бути використані для його ідентифікації. За КПФ може бути виділений основний тон. Траєкторія основного тону задає інтонацію мови, яка може використовуватися при розпізнаванні зливої мови [2], [13]-[16].

КПФ передбачає побудову спектра для окремих фрагментів мовного сигналу заданої довжини (вікон). Чим більше точок захоплює тимчасове вікно, тим більше дозвіл спектра по частоті (вузькосмуговий спектральний аналіз). У такому спектрі добре проглядаються окремі гармоніки сигналу, але в той же час, логарифмічний спектр виходить сильно коливається, і чітко визначити місце розташування формант як сплески енергії дуже складно. При меншій довжині вікна виходить широкосмуговий спектр, що має більш згладжену форму. Незважаючи на це, однозначна ідентифікація формантних частот скрутна [2], [13]-[16].

Короткочасне перетворення Фур'є, залежне від часу, визначається за формулою [2], [13]-[16].

$$S_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(n-m) \cdot s(m) \cdot e^{-j\omega m}, \quad (2.1)$$

де $w(n-m), m = \dots -2, -1, 0, 1, 2, \dots$ – представляє собою послідовність підрахунків часового вікна;

ω – частота звукових коливань;

n – дискретний час.

В результаті аналізу виходить функція двох змінних: частоти ω і дискретного часу n . Перетворення Фур'є дискретизованого сигналу є безперервною періодичною функцією і позначається $S_n(e^{j\omega})$. Період по круговій частоті ω дорівнює 2π .

Результат короткочасного аналізу спектра залежить від вибору вікна $w(n)$. Існує значне різноманіття вікон, що застосовуються в спектральному аналізі. Метод використання всіх видів однаковий, але на спектрі КПФ відображаються особливості кожного їх вікон: розмірність частотної характеристики вікна в середній його частині; крутизна спаду (зрізу); наявність, протяжність і амплітуда бічних пелюсток вікна [2], [13]-[16].

Завдання ідентифікації диктора, як і завдання пофонемного розпізнавання мови часто вирішується на основі аналізу формантних частот мовного сигналу. При цьому для ідентифікації диктора використовують не менше чотирьох формант. Формантні частоти швидко убують по потужності. Перша форманта має високу інтенсивність, наступні форманти стають слабшими. Правильність розпізнавання фонем залежить від здатності визначити можливе більше число формант, що мають дуже низьку інтенсивність. Для цього завдання застосовується короточасний спектральний аналіз. Сигнал розкладається на базисні функції, які для дискретних сигналів мають вигляд [2], [13]-[16]:

$$\left\{ \cos\left[\frac{2\pi}{N\tau} \cdot k \cdot n\tau\right]; \sin\left[\frac{2\pi}{N\tau} \cdot k \cdot n\tau\right] \right\}_{k,n=1,2,\dots}, \quad (2.2)$$

де N – довжина послідовності відліків сигналу;

τ – період його дискретизації;

n та k – номер відліку за часом і номер відліку за частотою при КПФ.

При цьому тільки ті частоти сигналу, які збігаються з частотами базису, відображаються правильно на базисній функції. Всі інші частоти матимуть нульові проекції на будь-який з векторів базису. Це явище називається розмиванням. Якщо природний період сигналу непорівнянний з періодом спостерігача, то відбувається розмивання. Вікна, що використовуються в спектральному аналізі, призначені для зменшення ефекту розмивання [2], [13]-[16].

2.2.2 Кепстральний аналіз, гомоморфна обробка

Даний метод набув широкого поширення в області обробки мови. Його методика дозволяє отримати добре згладжений спектр, за яким форманти можуть бути виділені пошуком екстремумів. Але практика

показує, що коректна оцінка ширини смуги форманти при такому згладжуванні скрутна і часто спектр, отриманий при кепстральному аналізі, містить побічні піки, які можна помилково прийняти за форманти. Тому для достовірної оцінки формант гомоморфна обробка мови має обмежене застосування [2], [13]-[16].

Кепстр – це зворотне перетворення Фур'є логарифма спектра сигналу. Спектр $X(e^{j\omega})$ є комплексним, тому потрібно обчислювати комплексний логарифм:

$$X(e^{j\omega}) = \log|X(e^{j\omega})| + j \cdot \arg[X(e^{j\omega})]. \quad (2.3)$$

Кепстр обчислений за безперервним спектром має вигляд:

$$c(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log|X(e^{j\omega})| \cdot e^{jn\omega} d\omega. \quad (2.4)$$

При обробці мови можна не використовувати комплексний логарифм. Тоді втрачається інформація про фазу сигналу, але, як показує практика, для завдань обробки мови це не дуже важливо. Такий кепстр називається дійсним. Дійсний кепстр – це зворотне перетворення Фур'є логарифма модуля спектра сигналу [2], [13]-[16].

Кепстр сегмента мовного сигналу на інтервалі від 0 до 2 мс є відображенням імпульсної характеристики мовного тракту і дозволяє її визначити. Кепстр сегмента мовного на інтервалі від 2 до 20 мс відображає період сигналу збудження і тому досить повно характеризує сам сигнал збудження [2], [13]-[16].

Кепстральні коефіцієнти можуть бути описані математично наступним чином [2], [13]-[16]:

$$c(n) = \frac{1}{N} \sum_{k=0}^{N-1} \log|X(k)| \cdot e^{j \frac{2\pi kn}{N}}, \quad (2.5)$$

де $X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j \frac{2\pi kn}{N}}$ – перетворення Фур'є фрейму сигналу.

2.2.3 Методи лінійного передбачення мови

Методи лінійного передбачення мови (ЛПМ). Мають ряд переваг: володіють мінімальною складністю і часом обчислення, максимальною точністю оцінювання формант. При оцінці формантних частот за допомогою ЛПМ аналізується спектр моделі мовного сигналу, який представляє огинає спектру сигналу і дозволяє найбільш точно і легко визначити місце розташування формант, їх частоти, амплітуди і ширину смуг. Побічні форманти в такому спектрі практично виключені. Крім того, результатом ЛПМ є модель, що добре описує характеристики голосового тракту і дозволяє зберігати всю необхідну для аналізу і синтезу інформацію у вигляді набору параметрів моделі (коефіцієнтів ЛПМ) [4].

Завдання аналізу на основі лінійного передбачення-отримати параметри, необхідні для відтворення вихідного звуку: тип звуку, значення основної частоти, коефіцієнти фільтра. Мета кодування з лінійним прогнозом полягає в моделюванні сигналу як лінійної комбінації попередніх відліків. Даний метод є досить ефективним, оскільки мова – це високо корельований протягом короткого проміжку часу сигнал, а, значить, передбачення може бути виконано з мінімальною помилкою. У задачі розпізнавання мови даний метод застосовується для моделювання спектра сигналу як авторегресійного процесу.

Модель мовного тракту (як системи) враховує нулі і полюси всередині одиничного кола. Передавальна функція дозволяє отримати зв'язок між вхідним сигналом збудження мовного тракту $u(n)$ і виходом $s(n)$.

$$s(n) = - \sum_{k=1}^p a_k \cdot s(n-k) + G \cdot \sum_{l=1}^q b_l \cdot u(n-l) + G \cdot u(n). \quad (2.7)$$

На вхід системи подається спостережувана послідовність – сигнал збудження $u(n)$. Значення мовного сигналу на виході $s(n)$ може бути представлено у вигляді лінійної комбінації попередніх значень вхідного сигналу і його поточного значення. Така залежність називається лінійним прогнозом. У загальному випадку модель сигналу має полюси і нулі. Окремими випадками є [2], [13]-[16]:

- модель, що має тільки нулі (модель ковзного середнього (КС));
- модель, що має тільки плюси (модель авто регресії (АР)).

Найбільш широко застосовується модель авторегресії, оскільки вона враховує полюси передавальної функції. Плюси присутні завжди як наслідок резонансів, що виникають в мовному тракті. Моделі АР і КС мають властивість оборотності. Модель з кінцевим числом коефіцієнтів може бути представлена моделлю АР з нескінченним числом коефіцієнтів. Тому на практиці використовують модель АР зі збільшеним числом коефіцієнтів, що дозволяє в якійсь мірі врахувати також і нулі передавальної функції [2], [13]-[16].

2.3 Метод на основа лінійного передбачення

2.3.1 Попередня обробка мовного сигналу

Перш ніж приступити до спроб розпізнавання диктора, необхідно провести попередню обробку сигналу. Мовний сигнал є нестационарним, так як мова складається зі звукових одиниць, тривалість яких може змінюватися в інтервалах від 30 до 200 мс. Отже, аналіз і обробку мови доцільно виконувати на невеликих сегментах, на яких сигнал можна вважати локально стаціонарним. Такий підхід називається методом короткочасного аналізу [2], [13]-[16].

Виділення кордону мовного сигналу з вхідного сигналу є важливим етапом попередньої обробки. Для виділення ділянок містять тільки мова використовують кілька характеристик, в нашій програмі буде застосовуватися короткочасна енергія сигналу [2], [13]-[16].

Повна енергія сигналу E з дискретними відліками обчислюється за формулою:

$$E = \sum_{m=-\infty}^{\infty} s^2(m). \quad (2.8)$$

Короткочасна енергія сигналу може бути отримана, як енергія частини сигналу, обмеженої вікном $w(n-m)$:

$$E(n) = \sum_{m=-\infty}^{\infty} [s(m) \cdot w(n-m)]^2, \quad (2.9)$$

де функція прямокутного вікна визначена співвідношенням:

$$w(n) = \begin{cases} 1, & 0 \leq n \leq N \\ 0, & \text{else} \end{cases}. \quad (2.10)$$

Результат короткочасного аналізу залежить від ширини вікна N . Якщо N мало (порядку періоду основного тону і менше), то енергія $E(n)$ буде змінюватися дуже швидко, відповідно до тонкої структури мовного сигналу. Якщо N велике, то $E(n)$ буде змінюватися дуже повільно і не буде адекватно описувати локальні особливості мовного сигналу. Отже, не існує єдиного значення N , що задовольняє всім вимогам. Зазвичай N вибирають пропорційно частоті основного тону. Тривалість вікна вибирають близько 10-30 мс.

2.3.2 Використання лінійного передбачення

Кодування з лінійним передбаченням-це техніка кодування мовної інформації шляхом моделювання голосового тракту. Мета кодування з лінійним прогнозом полягає в моделюванні сигналу як лінійної комбінації попередніх відліків. Даний метод є досить ефективним, оскільки мова - це високо корельований протягом короткого проміжку часу сигнал, а, значить, передбачення може бути виконано з мінімальною помилкою [14], [15].

Методи лінійного передбачення можна застосовувати для аналізу не тільки в тимчасовій, але і в частотній області. Шляхом застосування даного методу можливе отримання апроксимації спектра. Для цього необхідно розглянути схему передбачення з зворотним фільтром $A(z)$. Спектр сигналу по моделі лінійного передбачення виходить при $z = e^{j\omega}$. Модель має p полюсів, а спектр має $p/2$ піків в смузі частот від 0 до $f_s/2$. Положення піків відповідає положенню формант. Якщо спектр має провал, то модель не зовсім адекватна і піки будуть зміщеними щодо їх істинного положення. Однак для більшості систем модель адекватна і застосовується в більшості систем з низькою швидкістю передачі [14], [15].

Модель лінійного передбачення ґрунтується на припущенні, що будь-який відлік мовного сигналу $s(n)$ можна наближено оцінити лінійною комбінацією деякого числа p попередніх йому відліків, що призводить до наступного співвідношення [14], [15]:

$$s(n) = \sum_{i=1}^p a_i \cdot s(n-i) + G \cdot u(n), \quad (2.11)$$

де $a_1, a_2, a_3, \dots, a_p$ – коефіцієнти передбачення;

$u(n)$ – нормалізована послідовність збуджень (помилка передбачень);

G – коефіцієнт посилення.

У z -області коефіцієнти передбачення задають передавальну функцію:

$$H(z) = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}} = \frac{G}{A(z)}. \quad (2.12)$$

Помилка передбачення $e(n)$ визначається як різниця між вихідними і наближено обчисленими (передбаченими) відліками:

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p a_k s(n-k). \quad (2.13)$$

Основне завдання методу лінійного передбачення-це обчислення набору коефіцієнтів передбачення, які мінімізують $e(n)$. Для вирішення цього завдання використовують автокореляційний і коваріаційний методи. Зазвичай автокореляційний метод дає меншу помилку передбачення і обчислювальні витрати для нього менше завдяки наявності швидкого алгоритму [14], [15]. Автокореляційний метод передбачає, що $s(n) = 0$ поза заданим сегментом сигналу $0 \leq n \leq N$, і мінімізує помилку передбачення на нескінченному інтервалі. Завдання зводиться до вирішення системи:

$$\sum_{i=1}^p a_i r(i-j) = -r(j), \quad (2.14)$$

де $j = 1, 2, \dots, p$ та $c_{ij} = \sum_{n=p}^{N-1} s(n-i)s(n-j)$.

Автокореляційний метод гарантує отримання стабільного фільтра

$H(z) = \frac{G}{A(z)}$ (тобто всі корені $A(z)$ лежать всередині одиничної окружності),

у той час як коваріаційний метод не може цього гарантувати і, як наслідок,

рідше застосовується в обробці мови, хоча може давати більш точні результати [14], [15].

Для шумового ділянки мови $U(z)$ представляє спектр білого шуму, для тонального ділянки $U(z)$ – це спектр періодичного сигналу. Вважається що $U(z)$ описує тонку структуру спектра мовного сигналу. Складова $H(z)=1/A(z)$ описує огинаючу спектра, яка краще представляє формантну структуру сигналу, і називається моделлю спектра:

$$P(\omega) = |H(e^{j\omega})|^2 = \frac{G^2}{|1 - \sum_{k=1}^p a_k e^{-j\omega k}|^2} = \frac{G^2}{|A(e^{j\omega})|^2}. \quad (2.15)$$

Сигнал помилки передбачення в z -поданні має вигляд:

$$E(z) = U(z) - A(z) \cdot S(z) = \left[1 - \sum_{k=1}^p a_k z^{-k} \right] \cdot S(z). \quad (2.16)$$

Спектр потужності сигналу визначається виразом:

$$P(\omega) = \frac{|E(e^{j\omega})|^2}{|A(e^{j\omega})|^2}. \quad (2.17)$$

З цього випливає, що істинний спектр $P(\omega)$ відрізняється від оцінки спектра за моделлю $P(\omega)$ тим, що $E(e^{j\omega})$ замінюється величиною G . Реальний спектр замінюється гладким спектром [14], [15].

Метод обчислення спектра за моделлю лінійного передбачення дозволяє виключити тонку структуру, яка не дозволяє точно визначити формантні частоти, і дає можливість більш точно локалізувати форманти [14], [15].

2.3.3 Динамічна деформація часу

Виконавши обчислення ознак на кожному послідовному фреймі сигналу, отримаємо набір векторів, які характеризують зміну характеристик мовного сигналу в часі. Цей набір ознак слід порівняти з набором ознак еталонного сигналу. Однак, еталонна і вимовлена фраза можуть мати різний часовий масштаб. Тому до виконання порівняння необхідно привести обидві фрази до одного тимчасового масштабу. Таке приведення виконується динамічно і називається динамічною деформацією часу. Динамічна деформація виконується для того, щоб точки на вимірній траєкторії збігалися найкращим чином з точками на еталонній траєкторії [14], [15].

Однопараметричний випадок. Розглянемо дискретний ряд $\{a_k, k = 0 \dots K - 1$ значень ознаки, і дискретний ряд значень образу $bl, l = 0 \dots L - 1$. На рис. 2.1 наведено 2 траєкторії ознаки для еталонної і вимовленої фрази. Узгодження траєкторій виконується за допомогою функції $l = g(k)$. На рис. 2.1. показано приклад траєкторій ознаки та вимірної фрази.

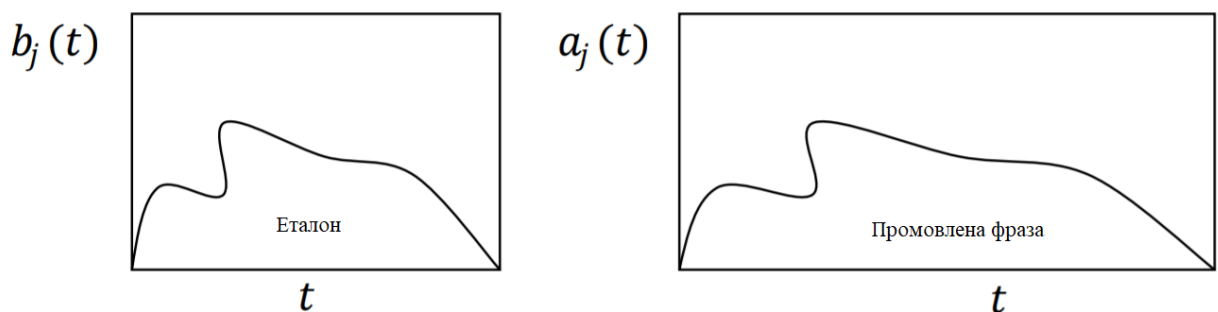


Рисунок 2.1 – Траєкторія ознаки для вимірної фрази і її образу
(однопараметричний випадок)

Як правило, еталон відповідає швидко вимовленої фразі, тому приймемо, що $L < K$. Деформуємо траєкторію розпізнаваної ознаки так, щоб вона мала мінімальне розбіжність із зразковою траєкторією. При цьому на кожному кроці порівнянь виконуємо оптимальну деформацію. Динамічне

програмування дає ефективний алгоритм пошуку оптимальної деформації часу для розпізнавання. Деформація траєкторії в середньому задана лінійною залежністю $l = g(k)$, – область, в околиці якої визначається район пошуку оптимальної деформації [14], [15].

Першим наближенням є лінійна залежність, яка визначає, яким чином в середньому виконується деформація часу:

$$l = g(k) = k \frac{L-1}{K-1}. \quad (2.18)$$

На окремих локальних ділянках допускається більш швидке або більш повільне проголошення фрази, ніж визначено в середньому залежністю (2.18). Величина r визначає, в яких межах допускається відхилення від середньої швидкості проголошення фрази. Величина r задає поле допуску, всередині якого можна виконати пошук оптимальної деформації часу для фрази [14], [15]. Величина r залежить від дискретного сигналу. Оптимальне узгодження між траєкторіями досягається, коли накопичене відстань між послідовностями мінімально:

$$D^* = \min_{g(k)} \left\{ \sum_{k=1}^K d(b(k), a[g(k)]) \right\}, \quad (2.19)$$

де d – нормалізована відстань між траєкторіями.

Оптимальне значення D^* знаходиться методом динамічного програмування. Величина накопиченої відстані між зразками нормується на кількість векторів у виміряному зразку:

$$\rho(a, b) = D^*/K. \quad (2.20)$$

Для вирішення задачі визначення оптимальної деформації часу для вимірної траєкторії ознаки доцільно враховувати обмеження Ітакури. Якщо мовний сигнал сегментований, таким чином, що один відлік відповідає мінімальній фонемі, то обмеження Ітакури має вигляд [16]:

$$g(k+1) - g(k) = \begin{cases} 0; 1; 2 & g(k) \neq g(k-1) \\ 1, 2 & g(k) = g(k-1) \end{cases} \quad (2.21)$$

Згідно з обмеженням Ітакури ознакою дозволено змінюватися на локальних ділянках вдвічі повільніше, ніж у зразку, і вдвічі швидше, ніж у зразку. Іншими словами, обмеженням допускається відхилення швидкості проголошення контрольної фрази не більше ніж в два рази порівняно з еталоном пароля [14], [15].

Багатопараметричний випадок. У багатопараметричному випадку маємо множину, що складається з J векторів ознак вимірної фрази $\{a_j(k), k = 0 \dots K-1, j = 1 \dots J\}$, і еталонної фрази $\{b_j(l), l = 0 \dots L-1, j = 1 \dots J\}$.

Виконаємо обчислення повної метрики за всіма J ознаками на основі приватних заходів розрізнення для кожної j -ої ознаки:

$$D_j = \sum_{k=1}^K \left\{ d(b_j(k), a_j[g(k)]) \right\}^2 / \sigma_{aj}^2, \quad (2.22)$$

де σ_{aj}^2 – середньоквадратичне відхилення для j -ої ознаки у вимірній фразі.

Тут використовується нормалізація ознак шляхом ділення на дисперсію [14], [15].

Для прийняття рішення про відповідність вимірної і еталонної фрази обчислюється середньозважений квадрат відстані між еталоном і вимірним сигналом по кожній з ознак:

$$D = \sum_{j=1}^J c_j D_j, \quad (2.23)$$

де c_j – вага j -ої ознаки;

D_j – квадрат відстані між контрольованою фразою і еталонною за ознакою j .

2.3.4 Алгоритм прийняття рішення у запропонованому методі

При прийнятті рішення обчислюється ймовірність або відстань від поданого на вхід системи зразка до еталонних векторів ознак, отримане значення порівнюється з порогом, який фіксується для всіх дикторів. Також можуть бути використані механізми нормалізації відстані, це необхідно для підвищення стійкості при наявності заважають факторів, якими можуть бути: варіабельність проголошення паролі фрази одного і того ж диктора, настрій диктора, різні манери і інтонації проголошення, гучність проголошення і т. д.

2.4 Висновки за розділом 2

У другому розділі проаналізовано запропонований метод розпізнавання мовного паролю на базі методу лінійного прогнозування.

Підхід до процесу обробки мовного сигналу полягає у використанні короткочасного аналізу. При цьому сигнал розбивається на часові вікна фіксованого розміру, на яких, як передбачається, параметри сигналу не змінюються. Тривалість вікна (фрейму) вибирають близько 10-30 мс. Для більш точного уявлення сигналу між вікнами роблять перекриття (наприклад, на біля вікна). Потім до кожного фрейму застосовуються алгоритми вилучення ознак (метод лінійного передбачення).

Коефіцієнти передбачення визначають частотну характеристику фільтра, що характеризує стан голосового тракту в певний момент часу. Після вилучення ознак з вхідного сигналу, застосовується метод динамічної трансформації часу, який дозволяє знайти відстань між еталонним зразком і отриманим сигналом.

Після обчислення відстані, отримані значення порівнюються зі зразками в системі, і виносяться рішення про відповідність сказаної диктором фрази з еталоном, що зберігається в системі.

3 РОЗРОБКА АРХІТЕКТУРИ ПРОГРАМИ

Другий розділ було присвячено дослідженню запропонованого методу розпізнавання мовного паролю на базі методу лінійного прогнозування.

Третій розділ буде присвячено висвітленню питань з проектування архітектури та реалізації ПЗ для дослідження запропонованого методу та можливих модифікацій. Особливу увагу буде присвячено вибору технічних засобів: мови програмування, середовища розробки, тощо.

3.1 Опис основних вимог до програми

Програма верифікації та ідентифікації диктора на основі мовного пароля призначена для контролю доступу до інформації, що міститься на персональному комп'ютері (ПК).

Таким чином серед функціональних вимог виділяються:

- виконувати попередню обробку сигналу;
- обчислювати коефіцієнти лінійного передбачення;
- порівняти вхідний сигнал з наявним в системі еталоном;
- за результатами порівняння винести рішення про ідентичність фрази одному з наявних у системі еталонів.

Забезпечити надання доступу до ПК, в разі відповідності голосу користувача з, що зберігаються в системі, еталоном голосу, записаним користувачем заздалегідь.

Передбачити захист конфіденційних даних від несанкціонованого доступу, шляхом відмови доступу в разі невідповідності мовного сигналу з зберігаються в системі.

Вхідні параметри – аудіозапис мовного пароля в форматі WAV, в кодуванні PCM, частота дискретизації 16кГц [17].

Вихідні параметри – позитивне або негативне рішення про відповідність вхідного мовного сигналу з еталонним зразком.

3.2 Вибір мови програмування

MATLAB – це платформа програмування та числових обчислень, яка використовується мільйонами інженерів та вчених для аналізу даних, розробки алгоритмів та створення моделей [18]-[21].

MATLAB поєднує в собі середовище робочого столу, налаштоване для ітеративного аналізу і процесів проектування, з мовою програмування, який безпосередньо виражає математику матриць і масивів [18]-[21]. Він включає в себе живий редактор для створення сценаріїв, які об'єднують код, висновки і форматований текст у виконуваному блокноті [18]-[21].

Професійно побудований. Набори інструментів MATLAB професійно розроблені, ретельно протестовані і повністю документовані [18]-[21].

З інтерактивними додатками. Програми MATLAB дозволяють побачити, як різні алгоритми працюють з вашими даними [18]-[21]. Можливе повторне виконання до тих пір, поки не отримаєте бажані результати, а потім автоматично згенерується програма MATLAB для відтворення або автоматизації вашої роботи [18–21].

Можливість масштабування. Масштабування власних аналізів для виконання на кластерах, графічних процесорах і хмарах з незначними змінами коду [18]-[21]. Немає необхідності переписувати код або вивчати програмування великих даних і методи роботи з нестачею пам'яті [18]-[21].

Розгортання в корпоративних застосунках. Код MATLAB готовий до роботи, тому можна перейти безпосередньо до власних хмарних і корпоративних систем і інтегруватися з джерелами даних і бізнес-системами [18]-[21].

Обладнання. Запуск на вбудованих пристроях. Автоматичне переконвертування алгоритмів MATLAB в код C/C++, HDL і CUDA для запуску на вбудованому процесорі або FPGA/ASIC [18]-[21].

Запуск на вбудованих пристроях. Інтеграція з дизайном на основі моделей. MATLAB працює з Simulink для підтримки проектування на основі

моделей, яке використовується для багатодоменого моделювання, автоматичної генерації коду, а також тестування і перевірки вбудованих систем [18]-[21].

Для MATLAB є можливість створювати спеціальні набори інструментів, що розширюють його функціональність [18]-[21]. Набори інструментів являють собою колекції функцій і об'єктів, написаних на мові MATLAB для вирішення певного класу завдань [18]-[21]. Компанія Mathworks поставляє набори інструментів, які використовуються в багатьох областях, включаючи наступні [18]-[21]:

- цифрова обробка сигналів, зображень і даних – набори функцій і об'єктів, що дозволяють вирішувати широкий спектр завдань обробки сигналів, зображень, проектування цифрових фільтрів і систем зв'язку [18]-[21];

- системи управління – набори функцій і об'єктів, що полегшують аналіз і синтез динамічних систем, проектування, моделювання та ідентифікацію систем управління, включаючи сучасні алгоритми управління, такі як робастное управління та інші [18]-[21];

- фінансовий аналіз – функцій і об'єктів, що дозволяють швидко і ефективно збирати, обробляти і передавати різну фінансову інформацію;

- аналіз і синтез географічних карт, включаючи тривимірні [18]-[21];

- збір і аналіз експериментальних даних – набори функцій і об'єктів, що дозволяють зберігати і обробляти дані, отримані в ході експериментів, в тому числі в реальному часі [18]-[21]. Підтримується широкий спектр наукового та інженерного вимірювального обладнання [18]-[21];

- візуалізація та представлення даних – дозволяє створювати інтерактивні світи і візуалізувати наукову інформацію за допомогою технологій віртуальної реальності і мови VRML [18]-[21];

- засоби розробки – інструменти, що дозволяють створювати незалежні додатки з середовища MATLAB [18]-[21];

- взаємодія із зовнішніми програмними продуктами – набори функцій, що дозволяють зберігати дані різних видів таким чином, щоб інші програми могли з ними працювати [18]-[21];

- бази даних – інструменти роботи з базами даних [18]-[21];

- наукові та математичні пакети – набори спеціалізованих математичних функцій і об'єктів, що дозволяють вирішувати широкий спектр наукових та інженерних завдань, включаючи розробку генетичних алгоритмів, рішення задач в приватних похідних, цілочисельні проблеми, оптимізацію систем і інші [18]-[21];

- нейронні мережі – інструменти для синтезу та аналізу нейронних мереж [18]-[21];

- нечітка логіка – інструменти для побудови та аналізу нечітких множин [18]-[21];

- символні обчислення – інструменти для символних обчислень з можливістю взаємодії з символним процесором програми Maple [18]-[21].

Крім перерахованих вище, існують тисячі інших наборів інструментів для MATLAB, написаних іншими компаніями і ентузіастами [18]-[21].

Проведемо порівняння мов програмування для реалізації ПЗ, результати порівняння представимо в табличній формі (табл. 3.1). У якості критеріїв будемо використовувати наступні [18]-[21]:

- швидкість роботи за стосунків – загальний критерій оцінки швидкості роботи за стосунків розроблених з використанням мови програмування на однакових апаратних платформах;

- інтегроване середовище розробки – наявність та якість існуючих середовищ розробки;

- фреймворки для методів МН – наявність, швидкість оновлення та підтримка фреймворків для реалізації методів МН;

- спільнота – обсяг активних розробників, що використовують мову програмування та кількість інформаційних ресурсів.

За результатами порівняння можна зробити висновок, що найбільш оптимальним та простим буде використання Matlab у якості мови програмування для розробки ПЗ. Так застосунки розроблені з використанням Matlab не відрізняються найшвидшою роботою, проте наявність безлічі відкритих бібліотек для МН та цілісне середовище розробки робить Matlab одним з найбільш потужних інструментів для розробки за стосунків з використанням технологій МН.

Таблиця 3.1 – Порівняння мов програмування

Критерій	Matlab	Python	Java	R
Швидкість роботи застосунків	±	±	+	+
Інтегроване середовище розробки	+	+	+	+
Бібліотеки	++	++	±	+
Фреймворки для методів МН	++	++	+	+
Спільнота	++	++	±	+

3.3 Вибір формату звукових файлів

Робота з цифровим звуком на комп'ютері передбачає, що нам, так чи інакше, доведеться мати справу зі звуковими файлами. Формат звукового файлу визначає структуру і особливості подання звукових даних при зберіганні на запам'ятовуючому пристрої ПК. Аудіофайл складається з інформації про амплітуду і частоту звуку, які зберегли для відтворення на програвачі або ПК. Виділяють наступні групи форматів звукових файлів [17]:

- аудіоформати без стиснення (AIFF, AU, CDDA, DSD, DXD, RAW, WAV і т. д.);
- аудіоформати зі стисненням звуку без втрат (FLAC, WavPack, ADX, Apple Lossless і т. д.);

– аудіоформати зі стисненням з втратами (MP2, MP3, Vorbis, WMA і т. д.).

У середовищі Matlab (функція `audioread`) підтримуються такі формати файлів, як WAV (.wav), OGG (.ogg), FLAC (.flac), AU (.au), AIFF (.aiff, .aif), AIFC (.aifc), MP3 (.mp3), MPEG-4 AAC (.m4a, .mp4) [2].

Будемо розглядати файли без стиснення, для збереження всіх параметрів і точності читання. Так як треба враховувати, що будь-яке стиснення аудіоданих вносить спотворення в форму аудіо сигналу [17].

Найвідоміший і часто використовуваний формат – WAV, розроблений корпораціями Microsoft і IBM для зберігання нестисненої, цифрової звукової (аудіо-) інформації. При подальшій обробці з 8-ми бітних або 16-ти бітних даних (звітів) WAV-формату, формуються інші формати звуку, які в більшості випадків є – стислими форматами звуку: ADPCM, MP3, OGG, Дельта-модуляція та інші [17].

WAV файл представляє з себе 2 області: заголовок файлу і область даних. У заголовку знаходиться інформація про Розмір файлу, кількість каналів, частоту дискретизації, кількість бітів у семплі (глибина звучання).

3.4 Проектування розробки

Matlab Compiler розширює функціональні можливості Matlab Compiler, дозволяючи створювати спільні бібліотеки C/C++, Microsoft. Мережеві збірки, класи Java і пакети Python з програм Matlab. Ці компоненти можуть бути інтегровані з користувацькими додатками, а потім розгорнуті в настільних, веб- і корпоративних системах [18]-[21].

Пакет SDK компілятора Matlab включає версію розробки Matlab Production Server для тестування і налагодження коду додатків і надбудов Excel перед їх розгортанням у веб-рфснйсеуrf[і корпоративних системах [18]-[21].

Програми, створені з використанням програмних компонентів з пакету Matlab Compiler SDK, можуть безкоштовно надаватися користувачам, яким Matlab не потрібен. Ці програми використовують середовище виконання Matlab, набір загальних бібліотек, який дозволяє виконувати скомпільовані програми або компоненти Matlab [18]-[21].

Використання засобів середовища Matlab для розгортання за стосунків зображено на рис. 3.1.

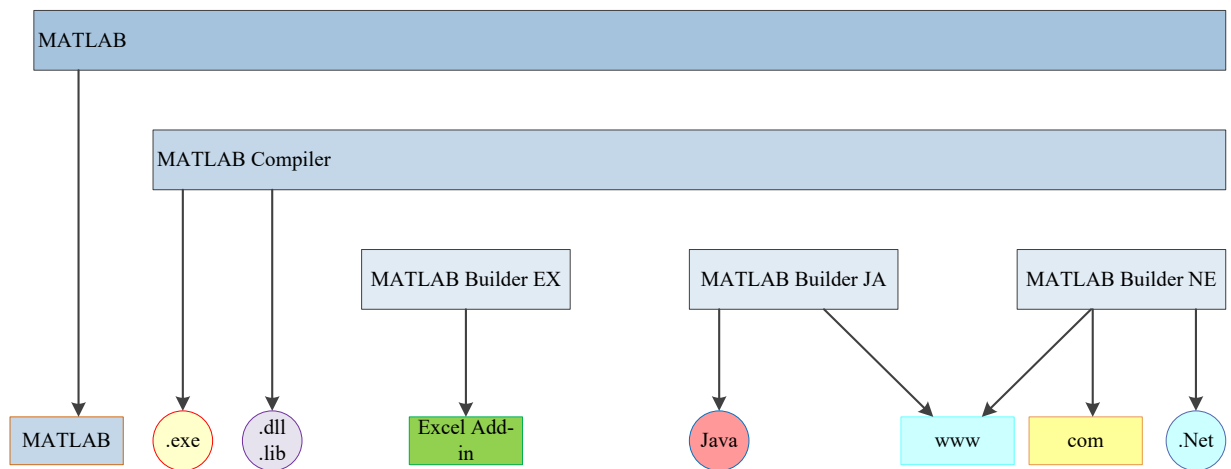


Рисунок 3.1 – Розгортання застосунку розробленого за допомогою Matlab

Математичне середовище Matlab надає зручні засоби для розробки та розгортання додатків. Завдяки цьому, на сервері середовища розробки Matlab, іменовану Mathworks, реалізуються публікації розробленого алгоритму або додатка, який має спільний доступ іншим людям, що працюють в Matlab. Не менш важливим є те, що в Matlab існує свій компілятор під назвою Matlab Compiler [18]-[21]. Він дозволяє поділитися додатком Matlab, а також дозволяє запускати додатки без встановленого середовища Matlab (рис. 3.2).

3.5 Вимоги до технічних засобів

Вимоги до складу і параметрів технічних засобів

Для нормальної роботи програми необхідні комп'ютер і мікрофон.

Технічна характеристика:

- ПК з операційною системою (ОС) Microsoft Windows 7 та старше;
- процесор – Intel(R) Core(TM) i3-2367M CPU @ 1.40 GHz, 1400

МГц, ядер: 2, логічних процесорів: 4;

- мікрофон Realtek High Definition Audio.

Вимоги до інформаційної та програмної сумісності:

- ОС ПК – Windows: Windows XP і вище, Linux, OS X.

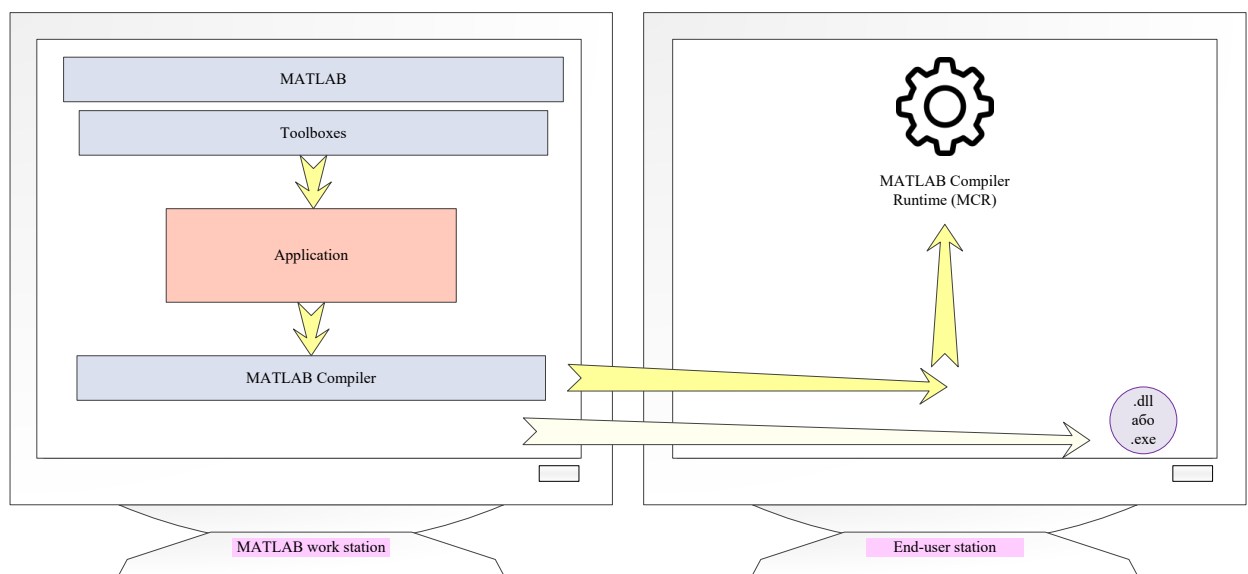


Рисунок 3.2 – Розгортання застосунку

3.6 Висновки за розділом 3

Під час роботи над розділом було досліджено технічні засоби, які будуть використовуватися під час розробки ПЗ, що реалізує запропонований метод. Так у якості мови програмування було обрано Matlab, через потужну базу вбудованих функцій, розширень та надбудов. Також важливим фактором стала наявність інтегрованого середовища розробки. Також достатню увагу було приділено вибору формату кодування звукових файлів. Були спроектовані рішення, щодо розгортання розроблюваного ПЗ.

4 ОСНОВНІ РІШЕННЯ ЩОДО РЕАЛІЗАЦІЇ КОМПОНЕНТІВ СИСТЕМИ

У третьому розділі було досліджено та вирішено ряд питань з проектування архітектури та реалізації ПЗ для дослідження запропонованого методу та можливих модифікацій. Особливу увагу було присвячено вибору технічних засобів: мови програмування та формату кодування звукових файлів.

Завдання розпізнавання диктора по голосу зводиться до виділення інформаційних ознак з вхідного сигналу, їх класифікації, після система повинна відповідним чином реагувати на людську мову, тобто приймати рішення про відповідність вхідного сигналу з зберігається в системі чином. Matlab є сильним інструментом для обробки сигналів. Це обумовлено широкими можливостями даного середовища.

Метою четвертого розділу є висвітлення рішень, щодо розробки програми розпізнавання диктора на основі мовного пароля. Метод розпізнавання заснований на методі лінійного передбачення мови, використовується динамічна деформація часу, для можливості порівняння еталонного зразка з вхідним.

У розділі описаний поетапний процес розробки програми в середовищі Matlab. Описаний спосіб введення мовного пароля в систему, обчислення ознак з вхідного сигналу на основі коефіцієнтів лінійного передбачення, динамічна деформація часу.

4.1 Розробка програми введення мовного сигналу

Введення мовного сигналу буде здійснюватися за допомогою функції пакета Matlab для зчитування інформації з аудіо файлів [18]-[21].

```
[y,Fs,bits] = wavread('filename')
```

Ця функція призначена для зчитування інформації з аудіофайлу формату WAV під ім'ям `filename`. В імені файлу слід вказати повний шлях, за винятком випадків, коли аудіофайл знаходиться в поточному каталозі або в одному з каталогів, які входять в список пошуку Matlab. У змінну `y` записується весь вміст зазначеного файлу. В результаті виходить матриця, рядки якої відповідають відлікам сигналу, а стовпці каналам. У wav-файлі може міститися один (моно-канал) або два (стереоканал). У `Fs` записується частота дискретизації в герцах, яка використовувалася при записі аудіо. У змінну `bits` записується число біт на відлік.

На вхід подається WAV файл в кодуванні PCM, з одним (моно) каналом, частота дискретизації 16кГц. Для аналізу аудіосигналу потрібно знати тільки зміст масиву `y`, тому в програмі використовується простий спосіб виклику даної функції:

```
y = wavread('filename')
```

Отриману інформацію можна вивести на графік за допомогою функції: `plot(y)`.

На рис. 4.1 показаний приклад графіка звукового сигналу, вимовленого двома різними дикторами.

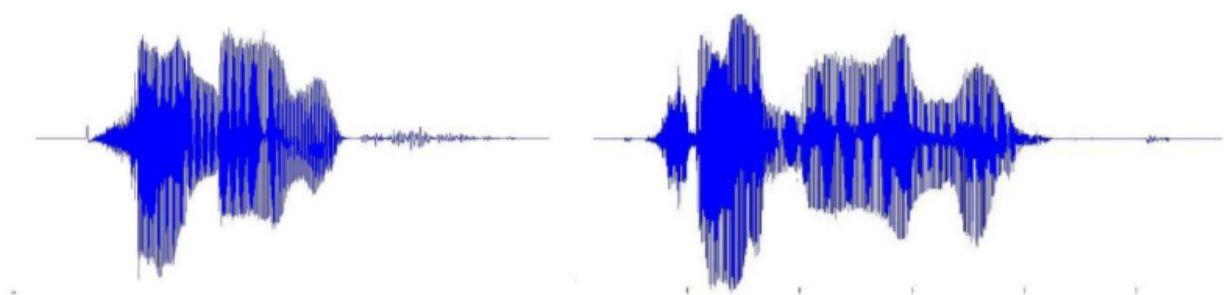


Рисунок 4.1 – Два диктори говорять одне і те ж слово

Для вилучення інформації з таких сигналів можна використовувати метод лінійного передбачення, який використовуються в розпізнавальній системі цієї роботи.

4.2 Розробка програми вилучення інформаційних ознак з мовного сигналу

Після того, як аудіофайл завантажений, з нього слід витягти ознаки, які необхідні для аналізу сигналу. Першим етапом слід розрахувати короточасну енергію мовного сигналу, що надійшов в систему. Це необхідно для визначення меж мовного вимовленого слова.

Для аналізу мовного сигналу, що є нестационарним сигналом, слід використовувати його короточасні характеристики в тимчасовій області. Тобто сигнал розділяється на сегменти (фрейми), таким чином, щоб кожен фрейм був коротким інтервалом, який представляє окремий звук. Для зменшення крайових ефектів сегменти беруть з деяким перекриттям (рис. 4.2). У програмі буде використовуватися перекриття на $1/4$ фрейма.

Короточасна енергія сигналу розраховується відповідно до пункту 2.1. справжньої роботи.

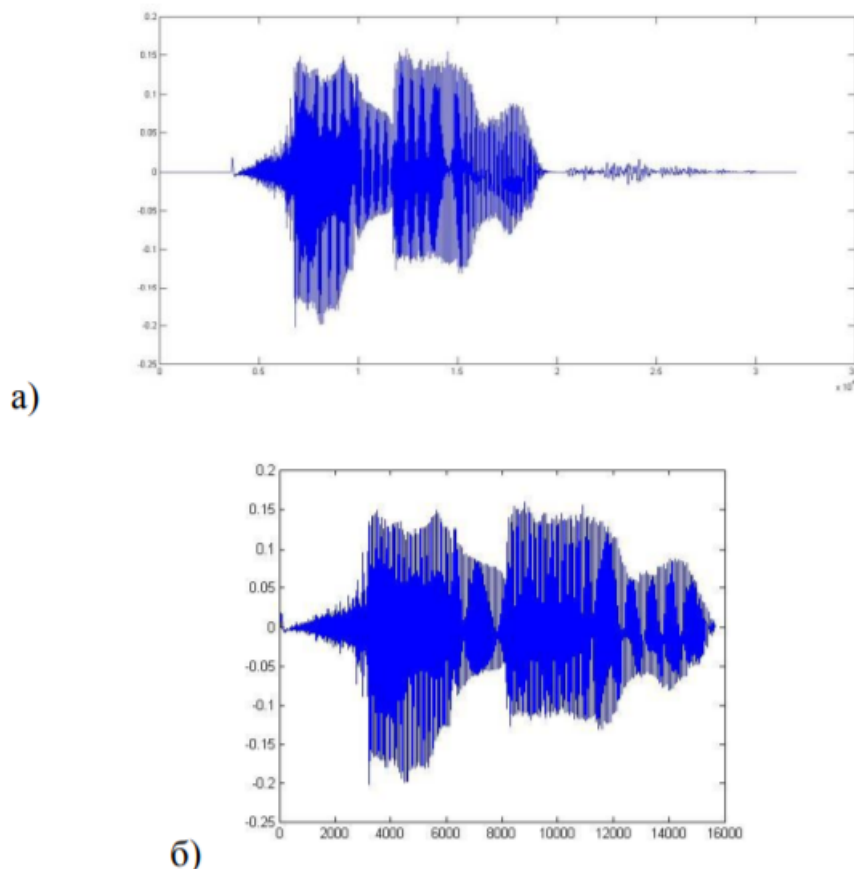


Рисунок 4.2 – Сигнал до обробки (4.2.а) і сигнал після обробки (4.2.б)

Для обчислення коефіцієнтів лінійного передбачення використовується функція Matlab: $[a,g] = \text{lpc}(x,p)$.

Ця функція розраховує коефіцієнти лінійного передбачення автокореляційним методом.

Функція lpc обчислює коефіцієнти передбачувального фільтра шляхом мінімізації середнього квадрата помилки передбачення. Знаходить коефіцієнти лінійного нерекурсивного пророкуючого фільтра p -го порядку, який найкращим чином пророкує поточне значення послідовності x відліків сигналу за її попередніми відліками:

$$\hat{x}(n) = -a(2)x(n-1) - a(3)x(n-2) - \dots - a(p-1)x(n-p), \quad (3.1)$$

Параметр p задає порядок передбачувального фільтра: $a = [1a(2)\dots a(p+1)]$. За замовчуванням функція lpc приймає $p = \text{length}(x)-1$, в разі, якщо параметр p не вказано. У вихідний параметр g записується значення рівне середньому квадрату помилки передбачення.

Параметр x є масивом, який відповідає мовному сигналу. Функція lpc повертає коефіцієнти передбачення для кожного сигналу у вихідний масив a , а середній квадрат сигналу помилки-значення g .

Для обчислення короткочасної енергії сигналу і коефіцієнтів лінійного передбачення, була створена функція в Matlab: $\text{function [energy, M, G] = Energy}(x)$.

Функція повертає значення енергії для кожного кадру в масиві energy , вектор коефіцієнтів лінійного передбачення в вихідний параметр M , і масив значень середнього квадрата помилки для кожного сегмента сигналу в G .

Після обчислення ознак на кожному послідовному сегменті сигналу, одержаний набір ознак, що характеризують зміну мовного сигналу в часі, слід порівняти з набором ознак еталонного сигналу. Однак, еталонна і вимовлена фраза можуть мати різний часовий масштаб. Отже, до виконання

порівняння обидві фрази необхідно привести до одного тимчасового масштабу за допомогою динамічної деформацією часу.

Для реалізації алгоритму динамічної деформації часу, була розроблена функція в середовищі Matlab: `function [Dist,D,k,Path]=DTW(t,r).`

Вхідними параметрами є вектора ознак t розмірністю N і r розмірністю M , еталонного і вимовленого зразків мовного пароля відповідно.

Функція повертає нормоване значення відстань між фразами в змінну `Dist`, матрицю відстаней між ознаками в D , по цій матриці здійснюється пошук оптимального шляху, який записується в вектор `Path`, k – це нормалізує фактор, в який записаний розмір шляху.

Шлях `Path` повинен задовольняє трьом умовам:

– рубіжні умови. Як правило, передбачається, що $\omega_1 = (1,1)$ і $\omega_k = (n,m)$, тобто початок і кінець шляху знаходяться на діагоналі в протилежних кутах матриці D ;

– умова безперервності. Нехай $\omega_k = (a,b)$ і $\omega_{k-1} = (p,q)$. Тоді $a - p \leq 1$, $b - q \leq 1$. Таке обмеження необхідно для того, щоб в кроці шляху `Path` брали участь тільки суміжні елементи матриці D (включаючи сусідні по діагоналі).

– умова монотонності. Нехай $\omega_k = (a,b)$ і $\omega_{k-1} = (p,q)$. Тоді $a - p \geq 1$, $b - q \geq 1$ дане обмеження потрібно, щоб точки `Path` монотонно переміщалися в часі.

Таким умовам може задовольняти безліч шляхів. Але нам потрібен такий, в якому досягається мінімум вартості шляху:

% шукаємо мінімум між суміжними елементами матриці відстаней

`[~,number]=min([D(m-1,n),D(m,n-1),D(m-1,n-1)]);`

...

% записуємо значення в масив

`Path=[m n; Path];`

```
% деформація шляху для кожного вхідного сигналу
rPath=r(Path(:,1));
tPath=t(Path(:,2));
% нормуємо відстань між векторами ознак
Dist=sum ((rPath-path).^2)/M;
Вихідний код програми наведено в додатку В.
```

4.3 Розробка програми розпізнавання диктора

Після того, як реалізована функція підрахунку відстані між вимовленою фразою і еталонної, необхідно визначити порогову відстань. Ця відстань має бути підібрана таким чином, щоб система при прийнятті рішення про відповідність вимовленої фрази зі збереженим зразком, давала негативний результат при спробі несанкціонованого доступу, і позитивний результат у разі санкціонованого доступу.

Для прийняття рішення про величину порогового значення відстані необхідно побудувати гістограми розподілів відстаней при проголошенні одного пароля різними дикторами. У випробуваннях взяли участь чотири диктора.

Диктор 1 – зареєстрований, користувач в системі, для якого зберігається еталонний зразок пароля.

Диктори 2, 3 і 4 – не зареєстровані в системі.

Для побудови гістограми кожен диктор вимовив парольне слово сорок п'ять разів. Для кожного вимовленого слова було виміряно відстань до еталону, що зберігається в системі. Нормовані відстані записуються в масив У, після чого для кожного масиву будується гістограма, для підрахунку числа влучень значень результатів в певний інтервал відстані.

На рис. 4.3 – рис. 4.6 представлені гістограми виміряних відстаней для чотирьох дикторів, які вимовляють парольне слово: «Вереск».

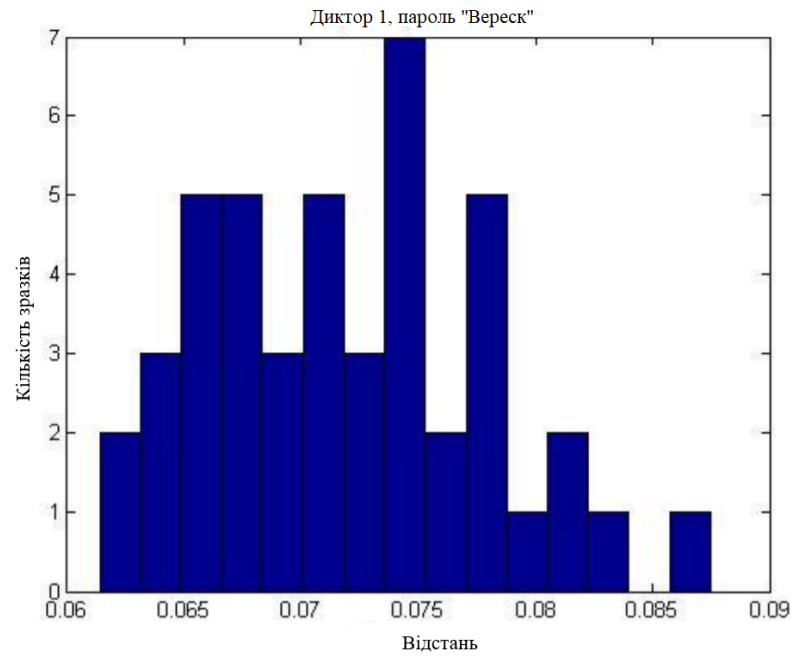


Рисунок 4.3 – Гістограма розподілів відстаней вимовлених слів для Диктора 1

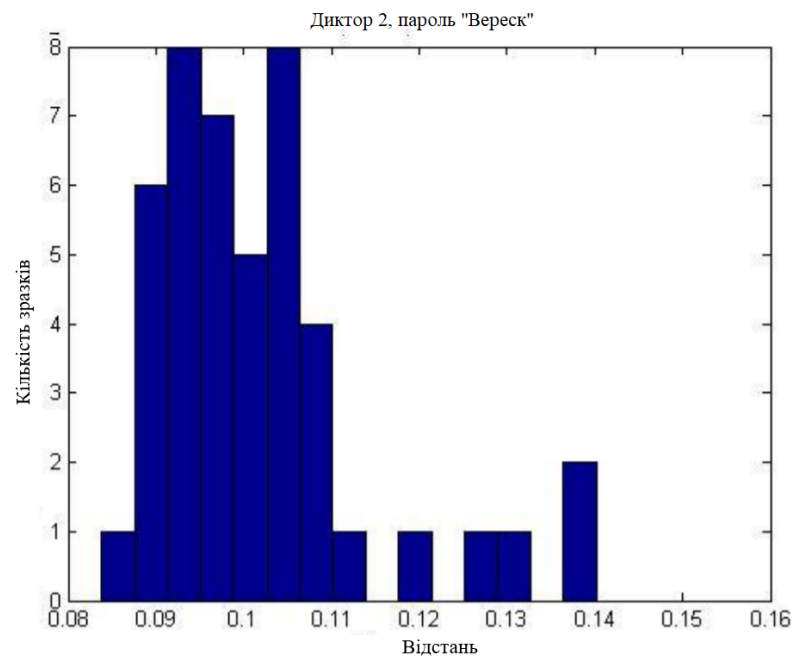


Рисунок 4.4 – Гістограма розподілів відстаней вимовлених слів для Диктора 2

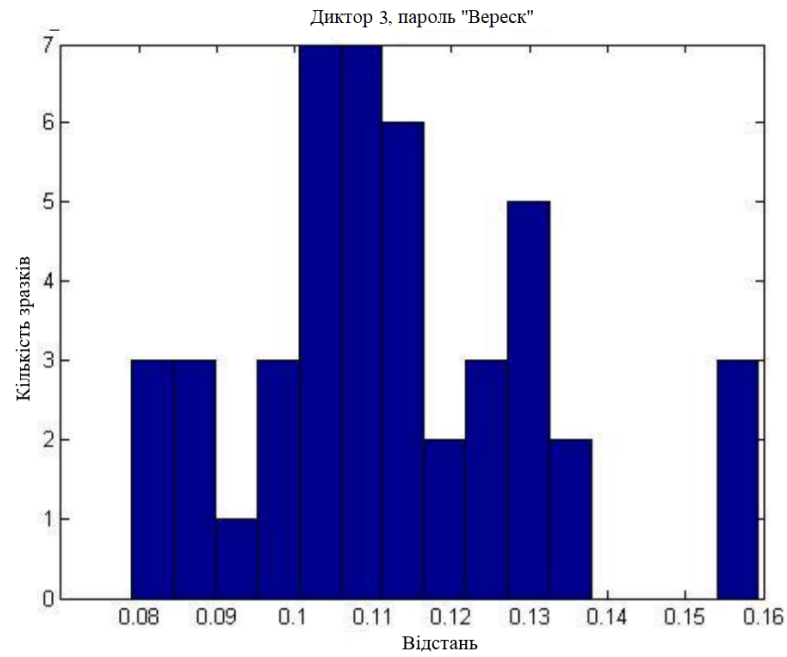


Рисунок 4.5 – Гістограма розподілів відстаней вимовлених слів для Диктора 3

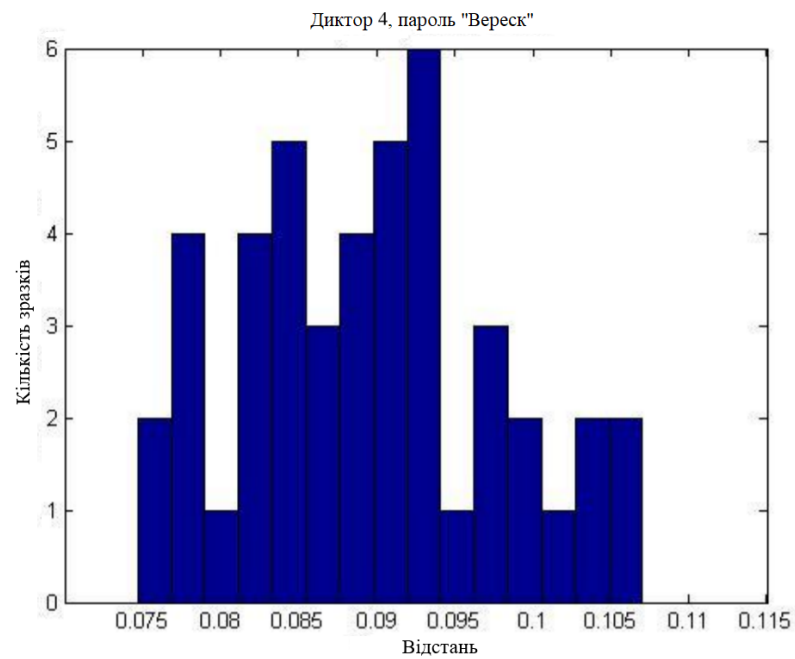


Рисунок 4.6 – Гістограма розподілів відстаней вимовлених слів для Диктора 4

З гістограм розподілів відстаней видно, що для Диктора 1 відмінності відстаней від еталонної фрази знаходяться в інтервалі від 0,060 до 0,086.

Для Диктора 2 інтервал від 0,085 до 0,140.

Для Диктора 3 інтервал від 0,080 до 0,160.

Для Диктора 4 інтервал від 0,075 до 0,106.

Виходячи з отриманих значень, можна припустити, що порогове значення відстані можна встановити на рівні 0,08. Якщо значення відстані між вимовленим словом і еталоном перевищить задане порогове значення, відповідь системи буде негативний.

4.4 Висновки за розділом 4

В четвертому розділі був описаний процес розробки програмного забезпечення для розпізнавання диктора по голосу, який складається з попередньої обробки сигналу, вилучення інформаційних ознак з вхідного сигналу, динамічної деформації часового ряду, підрахунку відстані між еталоном і вимовленим словом.

Наведено методи з середовища MaOaB, які використовуються при написанні програми, такими є `x=wavread('file')`, для зчитування аудіофайлу формату WAV, і `[a,g]=lpc(x,p)` для підрахунку коефіцієнтів лінійного передбачення.

Реалізована техніка динамічної деформації часу, для можливості підрахунку відстані між зразком і еталоном.

У цьому розділі було знайдено оптимальну порогову відстань, яка дозволяє контролювати доступ до інформації. У разі, коли відстань до зразка вище порогового, система повертає негативний результат ідентичності.

5 ЕКСПЛУАТАЦІЯ, ТЕСТУВАННЯ ТА ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ ПРОГРАМИ

Завдання розпізнавання диктора на основі мовного пароля є актуальним завданням області мовних технологій, яка об'єднує ідентифікацію і верифікацію диктора. Процес ідентифікації полягає у визначенні особистості за зразком шляхом порівняння цього зразка з шаблонами, які зберігаються в базі. Процес верифікації полягає в тому, що представлений зразок порівнюється з збереженим в базі еталоном для перевірки ідентичності.

Метою п'ятого розділу є перевірка роботи розробленої програми розпізнавання диктора по мовному паролі.

У цьому розділі описаний процес верифікації програми, наведено опис програмного забезпечення.

5.1 Опис ПЗ

Розроблена програма управління доступом на основі мовного пароля написана в середовищі Matlab. Програма є залежною від диктора, розпізнавання здійснюється для пароля, що складається з одного слова. Блок-схема алгоритму програми представлена на рис. 5.1.

На вхід програми подається аудіофайл формату WAV, в кодуванні PCM, частота дискретизації 16кГц.

З аудіофайлу зчитується інформація, для якої далі розраховуються значення короткочасної енергії і коефіцієнти лінійного передбачення в кожному фреймі. Наступним кроком є Розрахунок відстані між вимовленим словом і еталоном, який зберігається в системі. Ця відстань порівнюється з пороговим значенням, яке відповідає 0,08, дана відстань було знайдено в пункті 4.3. справжньої роботи. Якщо виміряна відстань перевищує порогове значення, то програма виносить рішення про розбіжність голосу мовця зі

збереженим зразком голосу в системі. У разі, якщо значення виміряного відстань не перевищує порогове, система повертає позитивний результат.

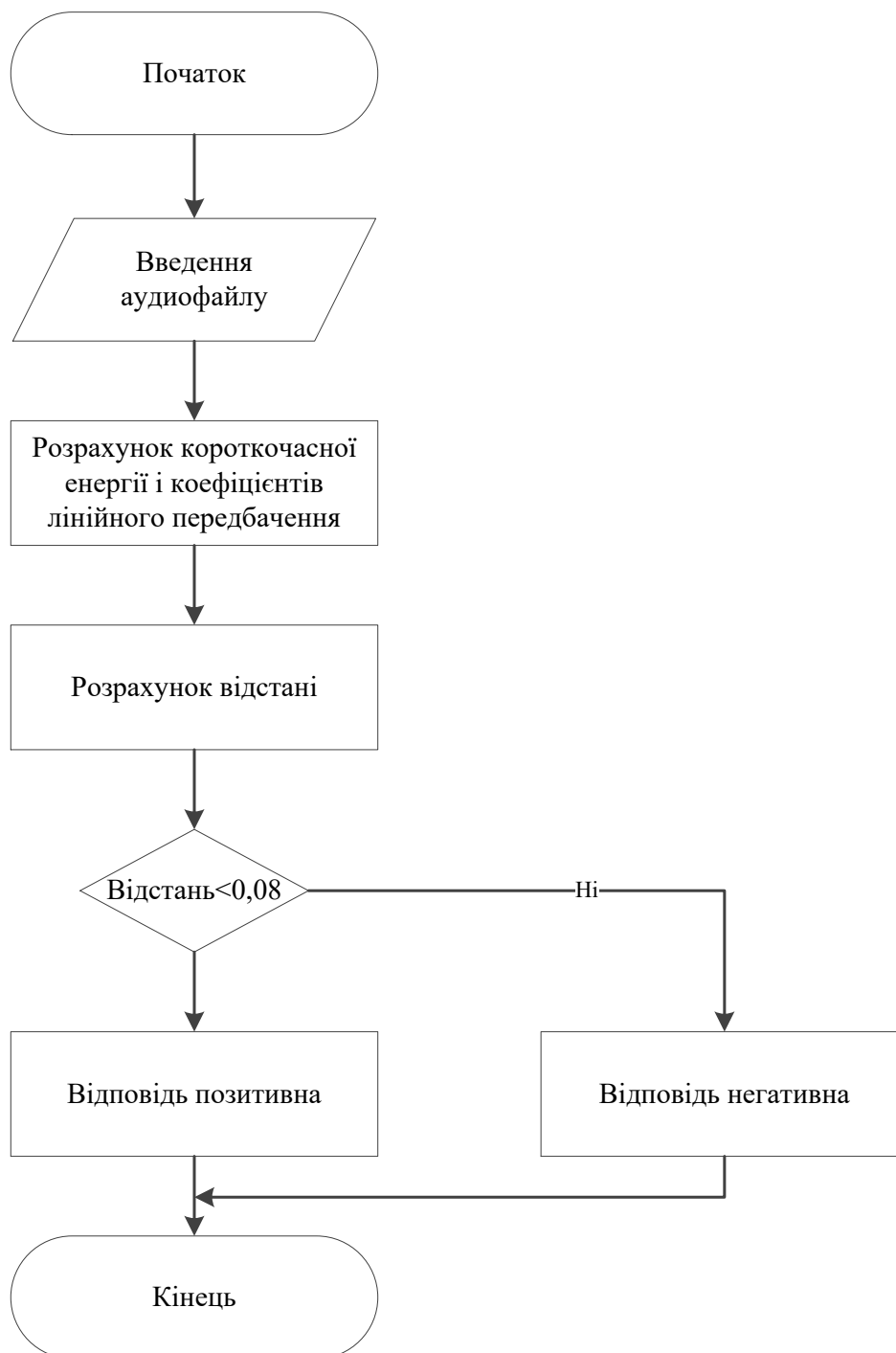


Рисунок 5.1 – Схематичне представлення роботи розробленого ПЗ

Система чутлива до шумів, запис звуку слід проводити в тихому приміщенні.

5.2 Верифікація ПЗ

При верифікації система обробляє вимовлену парольну фразу і порівнює її з еталоном, отриманим від диктора. На основі порівняння виноситься рішення про ідентичність і можливість або надання диктору доступу до інформації, або відмови в доступі.

У випробуваннях взяли участь 6 дикторів. У першому випадку перевіряємо реакцію розробленої програми на санкціонований доступ. Для кожного диктора в системі збережений еталон вимовленого парольного слова. Кожен диктор вимовив слово-пароль сім разів, в результаті перевірки програма повертала або позитивне, або негативне значення. Результати тестування продемонстровані в табл. 5.1. Символ «+» – доступ дозволений, символ «–» – відмова в доступі.

Таблиця 5.1 – Тестування програмного забезпечення: санкціонований доступ

Диктор	Результати за експериментом						
	№1	№2	№3	№4	№5	№6	№7
Диктор 1	+	+	+	+	+	+	+
Диктор 2	+	+	+	–	+	+	+
Диктор 3	–	+	+	+	+	+	+
Диктор 4	+	+	+	+	+	+	+
Диктор 5	+	–	–	+	+	+	+
Диктор 6	+	+	+	+	–	–	+

У другому випадку здійснюємо перевірку реакції програми на несанкціонований доступ. В системі зберігається еталон тільки для Диктора 1. Диктори 2-6 вимовляють парольне слово. Кожне записане слово порівнюється з еталоном. Результат верифікації програми для дикторів 2, 3, 4, 5, 6 наведено в табл. 5.2. Символ «+» – доступ дозволений, символ «–» – відмова в доступі.

Таблиця 5.2 – Тестування програмного забезпечення: несанкціонований доступ

Диктор	Результати за експериментом						
	№1	№2	№3	№4	№5	№6	№7
Диктор 1	—	—	—	—	—	—	—
Диктор 2	—	—	—	—	—	—	—
Диктор 3	+	+	—	—	—	—	—
Диктор 4	—	—	—	+	—	—	—
Диктор 5	—	—	—	—	—	—	—
Диктор 6	—	—	—	—	—	—	—

Виходячи з отриманих результатів, можна розрахувати ймовірності помилкових реакцій. Для цього введемо позначення:

- ω_1 – слово пароль вимовлено диктором;
- ω_2 – слово пароль вимовлено чужинцем;
- Ω_1 – доступ дозволений;
- Ω_2 – доступ заборонений.

На рис. 5.2 продемонстровані результати тестування програми.

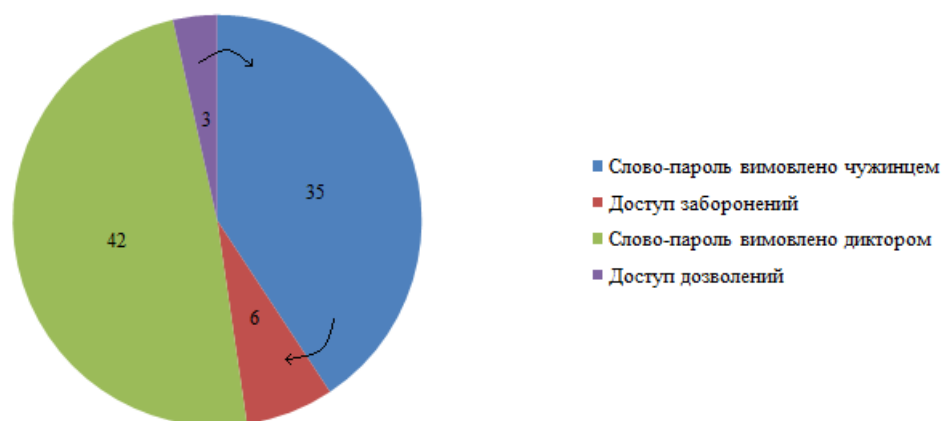


Рисунок 5.2 – Результат тестування програми

$N = 42$ – кількість перевірок слів вимовлених дикторами, для яких зберігається еталон.

$M = 35$ – кількість перевірок слів вимовлених дикторами, для яких немає еталону в системі. Загальна кількість перевірок $X = 42 + 35 = 77$. $F = 6$ – кількість дикторів, яким система невірно відмовила в доступі. $T = 3$ – кількість дикторів, помилково розпізнаних системою.

Розрахунок ймовірності помилки першого роду:

$$P\left(\Omega_2 / \omega_1\right) = \frac{F}{N} = \frac{6}{42} = 0,15 = 15\%.$$

Цей вираз описує ймовірність не допуску диктора при спробі санкціонованого доступу.

Розрахунок ймовірності помилки другого роду:

$$P\left(\Omega_1 / \omega_2\right) = \frac{T}{M} = \frac{3}{35} = 0,09 = 9\%.$$

Дана вираз описує ймовірність допуску чужинця при спробі несанкціонованого доступу.

Всього допущених дикторів $42 + 3 - 6 = 39$.

Всього не допущених $35 - 3 + 6 = 38$.

Розрахунок апостеріорної ймовірності:

$$P\left(\omega_2 / \Omega_1\right) = \frac{3}{39} = 0,08 = 8\% ,$$

$$P\left(\Omega_2 / \omega_1\right) = \frac{6}{38} = 0,16 = 16\%$$

Із загальної кількості допущених дикторів, 8% є помилково допущеними.

Із загальної кількості не допущених дикторів, 16% є помилково не допущеними.

Розрахунок достовірності:

$$D = \frac{(N - F) + (M - T)}{X} = 0,88 = 88\%$$

Вірно прийнятих рішень 88%.

В результаті проведеного тестування програмного забезпечення були розраховані помилки першого і другого роду, апостеріорна ймовірність і достовірність роботи програми. Достовірність наданих результатів програми становить 88%.

5.3 Висновки за результатами тестування

В якості вхідних параметрів даної програми слід подавати файл формату WAV з частотою дискретизації 16кГц, і одним каналом.

З аудіофайлу зчитуємо інформацію, по якій далі розраховуємо значення короткочасної енергії і коефіцієнти лінійного передбачення в кожному фреймі. Розраховуємо відстань між вхідним сигналом і еталонним зразком. Після чого виносяться рішення про ідентичність.

В ході тестування програмного забезпечення були розраховані ймовірності помилкових рішень першого і другого роду, які складають 15% і 9% відповідно. Розрахована достовірність наданих результатів рівна 88%. Однак, програма вузьконаправлена і розрахована на пароль, що складається з одного слова

ВИСНОВКИ

Під час виконання дипломної кваліфікаційної роботи магістра, було розглянуто основні методи розпізнавання диктора на малому словнику:

- спектральний аналіз;
- кепстральний аналіз;
- метод лінійного передбачення мови.

Було розглянуто загальну структуру систем розпізнавання диктора. Структура складається з трьох етапів:

- рівень обробки сигналів;
- рівень моделей;
- рівень прийняття рішень.

Був розроблений сценарій обробки мовного сигналу. Попередня обробка полягає в обчисленні короткочасної енергії сигналу, для визначення меж початку і кінця слова. Використовуваний в програмі метод вилучення ознак, заснований на лінійному передбаченні. Для кожного вхідного сигналу обчислюються коефіцієнти лінійного передбачення мови, які характеризують зміну характеристик мовного сигналу в часі.

Еталонна і вимовлена фраза можуть мати різний часовий масштаб. Для приведення сигналів до одного тимчасового масштабу в цій роботі виконаний метод динамічної деформації часу, це необхідно для того, щоб точки на вимірній траєкторії збігалися найкращим чином з точками на еталонній траєкторії.

Обчислено відстань між вхідним сигналом і еталоном. Отримане значення порівнюємо зі зразками в системі, і виносимо рішення про відповідність сказаної диктором фрази з еталоном, що зберігається в системі.

В ході тестування програмного забезпечення розраховані ймовірності помилкових рішень першого і другого роду, які складають 15% і 9% відповідно.

Розрахована достовірність наданих результатів рівна 88%.

Таким чином, в процесі підготовки дипломної кваліфікаційної роботи магістра була досягнута основна мета роботи і вирішені поставлені завдання

Під час розробки програми та аналізу розробленого програмного комплексу з'явилися ідеї з удосконалення, такі як: використання мультиагентних методів та штучних нейронних мереж з метою видалення шуму при скануванні голосу.

Нові можливості зможуть розширити коло вирішення практичних задач, адже зменшення шуму при скануванні аудіопотіку в реальному часі значно полегшить роботу систем ідентифікації.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Купін А.І. Ідентифікація та автоматизоване керування в умовах процесів збагачувальної технології на основі методів обчислювального інтелекту [Текст] : [монографія] / А. І. Купін, А. О. Сенько, Б. С. Мисько. – 2-ге вид., перероб. та допов. – Кривий Ріг : Сінельников Д. А. [вид.], 2019. – 297 с.
2. Мешков О. Ю. Аналіз голосових сигналів людини та аутентифікація особистості за голосом [Текст] : монографія / Мешков О. Ю., Новіков О. О., Новіков В. О. ; Херсон. нац. техн. ун-т. – Херсон : Вишемирський В. С. [вид.], 2018. – 167 с.
3. Тимченко О. О. Віддалена аутентифікація на основі ТOTP-алгоритму [Текст] / О. О. Тимченко // Моделювання та інформаційні технології. – 2018. – Вип. 83. – С. 186-190.
4. Комплексні системи захисту інформації в інформаційно-телекомунікаційних системах [Текст] : навч. посіб. / В. Д. Козюра [та ін.]. – Ніжин : Лук'яненко В. В. [вид.] : Орхідея, 2019. – 144 с.
5. Інформаційні технології. Криптографічний захист інформації. Алгоритм шифрування коротких повідомлень, що ґрунтується на скручених еліптичних кривих Едвардса [Текст]. – Чинний від 2020-11-01. – Київ : УкрНДНЦ, 2020. – IV, 36 с.
6. Ключин Д. А. Статистичні моделі розпізнавання образів [Текст] : навч. посіб. / Д. А. Ключин ; Київ. нац. ун-т ім. Тараса Шевченка. – Київ : Київський університет, 2019. – 85 с.
7. Методи та засоби оброблення великих даних в системах діагностування та розпізнавання образів [Текст] : монографія / [С. О. Субботін та ін.] ; під заг. ред. д-ра техн. наук, проф. С. О. Субботіна, канд. техн. наук, доц. А. О. Олійника ; Запоріж. нац. техн. ун-т. – Запоріжжя : Запоріж. нац. техн. ун-т, 2018. – 227 с.

8. Рупіч С. С. Багатокласове розпізнавання технічного стану зварного резервуару з дефектами з використанням нейромережевого класифікатора [Текст] / С. С. Рупіч, О. О. Лук'янченко // Вісник Київського політехнічного інституту. Серія : Приладобудування. – 2019. – Вип. 58(2). – С. 23-28.

9. Літвинчук А. М. Покращення моделей розпізнавання облич за допомогою згорткових нейронних мереж, навчання подібності та методів оптимізації [Текст] / А. М. Літвинчук, Л. В. Барановська // Проблемы управления и информатики. – 2021. – № 5. – С. 140-149.

10. Рибін О. І. Аутентифікація особи за динамічно введеним підписом з використанням нормального перетворення [Текст] / О. І. Рибін, Ю. Х. Ніжебецька, А. М. Луцків // Вісник Національного технічного університету України "Київський політехнічний інститут". Серія : Радіотехніка. Радіоапаратобудування. – 2010. – Вип. 40. – С. 26-30.

11. Шабала Є. Є. Біометричні методи захисту від несанкціонованого доступу на територію аеропорту [Текст] / Є. Є. Шабала, В. В. Ключова // Управління розвитком складних систем. – 2019. – Вип. 38. – С. 51-55.

12. Монастирський Л. Розпізнавання відбитків пальців у недорогій біометричній системі [Текст] / Л. Монастирський, В. Лозинський, Я. Бойко, Б. Соколовський // Електроніка та інформаційні технології. – 2018. – Вип. 9. – С. 120-124.

13. Снегірьова Є. О. Методики розпізнавання мовлення дослідження відбитку голосу [Текст] / Є. О. Снегірьова // Лінгвістика ХХІ століття. – 2017. – С. 163-170.

14. Шелепкова С. М. Аудитивний метод оцінки розпізнавання голосу і мовлення людини у різних умовах подання мовленнєвого сигналу [Текст] / С. М. Шелепкова // Молодий вчений. – 2019. – № 4(1). – С. 169-172.

15. Hutter H.-P. Comparison of classic and hybrid HMM approaches to speech recognition over telephone lines [Text] : diss. / H. Hutter ; Swiss Federal Institute of technology Zürich. – Zürich : [б.в.], 1996. – 238 p.

16. Sazhok M. M. Automatic Speech Recognition For Ukrainian Broadcast Media Transcribing [Текст] / М. М. Sazhok, R. A. Seliukh, D. Ya. Fedoryn, O. A. Yukhymenko, V. V. Robeiko // Control systems & computers. – 2019. – № 6. – С. 46-57.
17. Формат WAV [Електрон. ресурс]. – Режим доступа: <https://audiocoding.ru/formats/wav/>.
18. Стеблянюк П.О. Основи диференціальної геометрії (застосування сучасних комп'ютерних технологій, зокрема системи MatLab) [Текст] : [навч. посіб. для студ. ун-тів] / П. О. Стеблянюк, О. М. Коломієць ; Черкас. нац. ун-т ім. Б. Хмельницького. – Черкаси : [Вид. від. ЧНУ ім. Б. Хмельницького], 2011. – 203 с.
19. Xue D. Solving applied mathematical problems with Matlab® [Text] / Dingyü Xue, Yangquan Chen. – Boca Raton[etc.] : CRC press, Taylor & Francis group, 2009. – XIV, 432 p.
20. Matlab та комп'ютерні обчислення в освіті, науці та інженерії [Текст] = Matlab and computer calculations in education, science and engineering : тези загальноукр. конф., Київ, 16-18 трав. 2019 / [уклад.: В. Г. Романенко, Т. С. Соломаха, Є. О. Гаєв] ; Нац. авіац. ун-т, Ф-т аеронавігації, електроніки та телекомунікацій, ЦИТМ Експонента, офіц. дистриб'ютор MathWorks на території СНД, GlobalLogic – інновації, орієнт. на цифр. технології. – Київ : НАУ, 2019. – 67 с.
21. MATLAB Academy – MathWorks [Electronic resource]. – Access mode: <https://matlabacademy.mathworks.com/>.

ДОДАТОК А
Технічне завдання

A.1 Вступ

A.1.2 Призначення та область застосування

Програма верифікації та ідентифікації диктора на основі мовного пароля призначена для контролю доступу до інформації, що міститься на персональному комп'ютері. Існуюча програма реалізована в середовищі Matlab.

A.2 Підстави для розробки

Підставою для розробки є завдання на магістерську роботу, яке було визначене за темою «Дослідження та програмна реалізація методів розпізнавання мовного паролю для управління доступом» і затверджене наказом № 435 від 05 листопада 2021 р. за Національним університетом «Запорізька політехніка».

A.3 Призначення розробки

Програма призначена для управління доступом до персонального комп'ютера за допомогою розпізнавання диктора по голосу.

A.4 Вимоги до програми

Програма повинна вирішувати наступні завдання:

- виконувати попередню обробку сигналу;
- обчислювати коефіцієнти лінійного передбачення;
- порівняти вхідний сигнал з наявним в системі еталоном;
- за результатами порівняння винести рішення про ідентичність фрази одному з наявних у системі еталонів.

A.4.1 Вимоги до функціональних характеристик

Забезпечити надання доступу до комп'ютера, в разі відповідності голосу користувача з, що зберігаються в системі, еталоном голосу, записаним Користувачем заздалегідь.

Передбачити захист конфіденційних даних від несанкціонованого доступу, шляхом відмови доступу в разі невідповідності мовного сигналу з зберігаються в системі.

Вхідні параметри-аудіозапис мовного пароля в форматі WAV, в кодуванні РСМ, частота дискретизації 16кГц.

Вихідні параметри-позитивне або негативне рішення про відповідність вхідного мовного сигналу з еталонним зразком.

A.4.2 Вимоги до надійності

Для сталого функціонування програмного забезпечення необхідно:

- наявність безперебійних блоків живлення персонального комп'ютера.

Програма повинна видавати відмову доступу до комп'ютера при спробі несанкціонованого доступу.

A.4.3 Вимоги до технічних та програмних засобів

Вимоги до складу і параметрів технічних засобів

Для нормальної роботи програми необхідні комп'ютер і мікрофон.

- Технічна характеристика:
- ПК з ОС Microsoft Windows 7 та старше;
- процесор – Intel(R) Core(TM) i3-2367M CPU @ 1.40 GHz, 1400 МГц, ядер: 2, логічних процесорів: 4;
- мікрофон Realtek High Definition Audio.

Вимоги до інформаційної та програмної сумісності:

- ОС ПК – Windows: Windows XP і вище, Linux, OS X.

A.5 Вимоги до програмної документації

Склад програмної документації повинен включати в себе:

- технічне завдання;
- опис проектування;
- опис тестування;
- опис програми.

A.6 Стадії і етапи розробки

A.6.1 Стадії розробки

Розробка повинна бути проведена в три стадії:

- розробка технічного завдання;
- робоче проектування;
- впровадження.

A.6.2 Етапи розробки

На стадії розробки технічного завдання повинен бути виконаний етап розробки, погодження та затвердження справжнього технічного завдання [20].

На стадії робочого проектування повинні бути виконані перераховані нижче етапи робіт [20]:

- розробка програми;
- розробка програмної документації;
- випробування програми.

На стадії впровадження повинен бути виконаний етап розробки підготовка і передача програми [20].

А.6.3 Зміст робіт по етапах

На етапі розробки технічного завдання повинні бути виконані перераховані нижче роботи [20]:

- постановка задачі;
- визначення і уточнення вимог до технічних засобів;
- визначення вимог до програми;
- визначення стадій, етапів і термінів розробки програми та документації на неї;
- узгодження і затвердження технічного завдання [20].

На етапі розробки програми повинна бути виконана робота з програмування (кодування) і налагодженні програми [20].

На етапі розробки програмної документації повинна бути виконана розробка програмних документів відповідно до вимог до складу документації [20].

На етапі випробувань програми повинні бути виконані перераховані нижче види робіт [20]:

- розробка, узгодження та затвердження та методики випробувань;
- проведення приймально-здавальних випробувань;
- коректування програми і програмної документації за результатами випробувань [20].

ДОДАТОК Б
Опис програми

Б.1 Загальні свідчення

Програма верифікації та ідентифікації диктора на основі мовного пароля призначена для контролю доступу до інформації, що міститься на персональному комп'ютері. Існуюча програма реалізована в середовищі Matlab.

Б.2 Функціональне призначення

Забезпечити надання доступу до комп'ютера, в разі відповідності голосу користувача з, що зберігаються в системі, еталоном голосу, записаним Користувачем заздалегідь.

Передбачити захист конфіденційних даних від несанкціонованого доступу, шляхом відмови доступу в разі невідповідності мовного сигналу з зберігаються в системі.

Вхідні параметри-аудіозапис мовного пароля в форматі WAV, в кодуванні PCM, частота дискретизації 16кГц.

Вихідні параметри-позитивне або негативне рішення про відповідність вхідного мовного сигналу з еталонним зразком.

Б.3 Опис логічної структури

Matlab Compiler розширює функціональні можливості Matlab Compiler, дозволяючи створювати спільні бібліотеки C/C++, Microsoft. Мережеві збірки, класи Java і пакети Python з програм Matlab. Ці компоненти можуть бути інтегровані з користувацькими додатками, а потім розгорнуті в настільних, веб- і корпоративних системах [18]–[21].

Пакет SDK компілятора Matlab включає версію розробки Matlab Production Server для тестування і налагодження коду додатків і надбудов

Excel перед їх розгортанням у веб-рфснјсеурf і корпоративних системах [18]–[21].

Програми, створені з використанням програмних компонентів з пакету Matlab Compiler SDK, можуть безкоштовно надаватися користувачам, яким Matlab не потрібен. Ці програми використовують середовище виконання Matlab, набір загальних бібліотек, який дозволяє виконувати скомпільовані програми або компоненти Matlab [18]–[21].

Використання засобів середовища Matlab для розгортання за стосунків зображено на рис. Б.1.

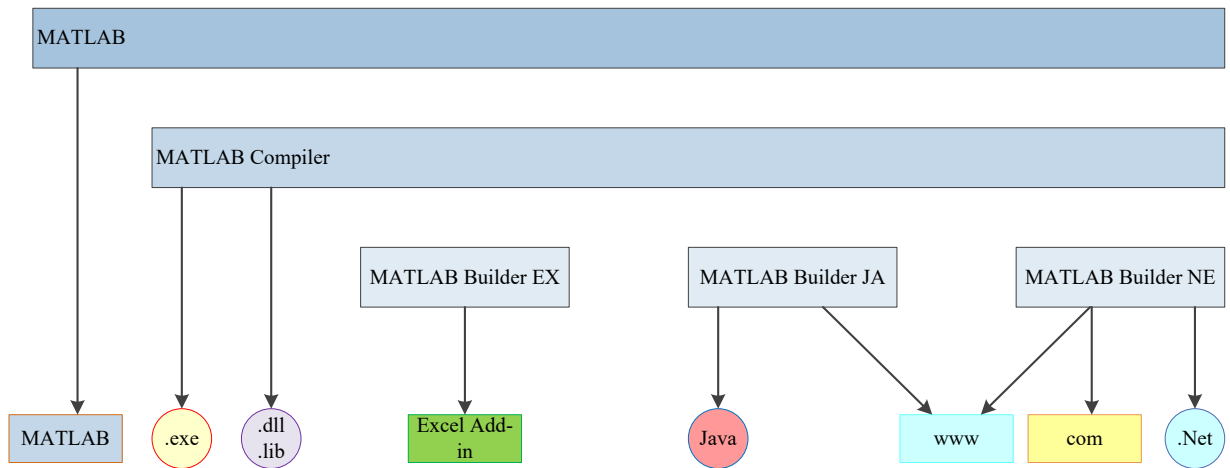


Рисунок Б.1 – Розгортання застосунку розробленого за допомогою Matlab

Математичне середовище Matlab надає зручні засоби для розробки та розгортання додатків. Завдяки цьому, на сервері середовища розробки Matlab, іменовану Mathworks, реалізуються публікації розробленого алгоритму або додатка, який має спільний доступ іншим людям, що працюють в Matlab. Не менш важливим є те, що в Matlab існує свій компілятор під назвою Matlab Compiler [18]–[21]. Він дозволяє поділитися додатком Matlab, а також дозволяє запускати додатки без встановленого середовища Matlab (рис. Б.2).

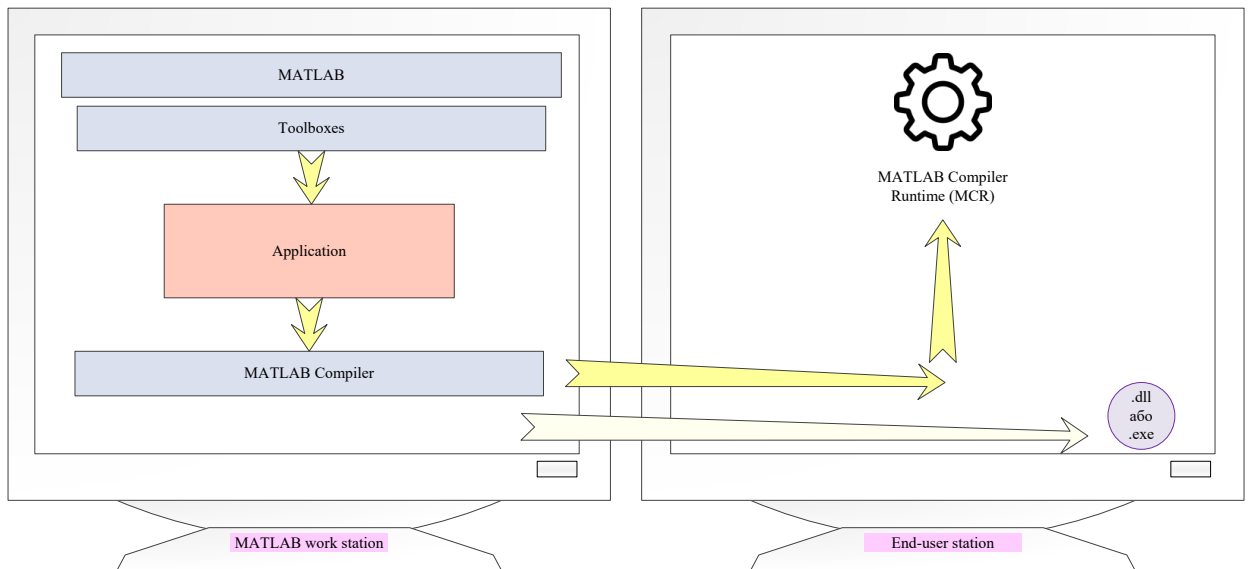


Рисунок Б.2 – Розгортання застосунку

Б.4 Використані технічні засоби

Вимоги до складу і параметрів технічних засобів

Для нормальної роботи програми необхідні комп'ютер і мікрофон.

Технічна характеристика:

- ПК з ОС Microsoft Windows 7 та старше;
- процесор – Intel(R) Core(TM) i3-2367M CPU @ 1.40 GHz, 1400 МГц, ядер: 2, логічних процесорів: 4;
- мікрофон Realtek High Definition Audio.

Вимоги до інформаційної та програмної сумісності:

- ОС ПК – Windows: Windows XP і вище, Linux, OS X.

ДОДАТОК В
Текст програми

Сирцевий код розробленого застосунку:

DTW.m

```
function [Dist,D,k,Path]=DTW(t,r)
```

```
% Dynamic Time Warping Algorithm % Dist ненормалізована відстань
між t і r % D матриця відстаней % k нормалізуючий фактор % Path
оптимальний шлях % t еталон % r ознака
```

```
% зчитування інформації з масиву ознак, обчислюється
```

```
% розмірність масивом
```

```
[row,M]=size(r); if (row > M) M=row; r=r'; end;
```

```
[row,N]=size(t); if (row > N) N=row; t=t'; end;
```

```
% Розрахунок відстані між векторами ознак d=(repmat(r',1,N)-repmat
(t,M,1)).A2;
```

```
% ініціалізації матриці відстаней D=zeros (size (d));
```

```
% визначення граничних умов D( 1,1)=d (1,1); for m=2: M
```

```
D(m,1)=d(m,1)+D(m-1,1); end
```

```
for n=2:N
```

```
D(1 ,n)=d(1,n)+D( 1,n-1); end
```

```
for m=2:M
```

```
for n=2:N
```

```
D(m,n)=d(m,n)+min(D(m-1 ,n),min(D(m-1 ,n-1 ),D(m,n-1))); end end
```

```
Dist=D(M,N);
```

```
n=N;
```

```
m=M;
```

```
k=1;
```

```
% ініціалізація шляху Path=[M n]; while ((n+m)~=2) if (n-1)==0 m=m-1;
elseif (m-1)==0 n=n-1; else
```

```
% вибирається мінімальна відстань
```

```
[~,number]=min([D(m- 1,n),D(m,n-1 ),D(m- 1,n-1)]); switch number case
1 m=m-1; case 2 n=n-1; case 3 m=m-1; n=n-1; end end
```

```
k=k+1; % підрахунок довжини шляху
```

```

Path=[m n; Path]; % запис значення мінімуму відстані
end

% деформація часу векторів ознак
rPath=r(Path(:,1));
tPath=t(Path(:,2));

% нормалізація відстані Dist=sum ((rPath-tPath)A2) / M;

Energy.m

function [energy, M, G] = Energy( x )

%UNTITLED2 Summary of this function goes here % Detailed
explanation goes here % довжина вікна frameWidth = 256;

% зміщення вікна вздовж сигналу frameSlip=frameWidth / 4;

% Розмір файлу numSamples = length (x);

% кількість вікон (фреймів)
numFrames = floor(numSamples/frameWidth);

energy = zeros (1, numFrames); % ініціалізація енергії
startSample = zeros (1, numFrames); % ініціалізація початку фрейму
endSample = zeros (1, numFrames); % ініціалізація кінця кадру

% ініціалізація коефіцієнти лінійного передсказання
M=[];
G=[];

startSample(1) = 1; endSample(1) = frameWidth;

% енергія для першої ділянки
energy(1) = sum(x(startSample(1):endSample(1)).A2);
[a,g]=lpc(x(startSample(1):endSample(1)),16);
M=a;
G=g;

% розрахунок енергії кожної ділянки, зі зміщенням на 1/4 for frame =
2:numFrames % початок фрейма зі зміщенням на 1/4 вікна startSample(frame)
= (frame-1)*frameWidth-frameSlip+1; % кінець фрейма
endSample(frame) = frame*frameWidth-frameSlip;

```

```

% індексація ділянки
frameIndex = (startSample(frame):endSample(frame));

% обчислення енергії, сума квадратів амплітуд energy(frame) = sum(x
(frameIndex).A2);

% обчислення коефіцієнтів лінійного передбачення
[a,g]=lpc(x(frameIndex),16);
M=[M a];
G=[G g]; end end
main.m
clear all; close all hidden;

% еталон
input 1 =wavread('Audio/Y arml .wav'); input2=wavread('Audio/Salbi 1
.wav'); ipnastay=wavread('Audio/Y arma2.wav');

% обчислюємо короткочасну енергію сигналу %обчислення
коефіцієнтів лінійного передбачення
[energy1, LPC1, G1]=Energy(input1); % Диктор 1 [energy3, LPC3,
G3]=Energy (ipnastay); % Диктор 2 [energy2, LPC2, G2]=Energy (input2); %
Диктор 1
ip1=wavread('file');
[energy, LPC, G]=Energy(ip1);
[Distip,Dsy3 ,ksy3 ,wsy3]=DTW(LPC2,LPC); if Distip > 0.08
answer='NO' else
answer='YES'
end
prompt = 'Ласкаво просимо, натисніть клавішу "enter", щоб увійти ';
input(prompt);

%%%%%%%% генерації звуку
pimp=5;
fs=11025; % частота дискретизації

```

```
time=0.3;
```

```
freq=196;
```

```
values=0:1/fs:time;
```

```
=pimp*sin(2*pi* freq*values);
```

```
sound(11025)
```

```
%%%%%%%%%%%% зберігання зразків для перевірки
```

```
for i=1:3
```

```
prompt = 'для того, щоб записати свій голос - натисніть "enter";
```

```
input(prompt);
```

```
recc = AUDrec(11025, 16, 1);
```

```
disp ('початок записуючі')
```

```
RCRDblock(recc, 1);
```

```
disp ('кінець записуючі');
```

```
z = getaudiodata(recc);
```

```
z=abs(fft(z));
```

```
z=z(1:3000);
```

```
[pks,locs]
```

```
=
```

```
fndpks(z,'MinPeakProminence',20,'Threshold',4,'MinPeakDistance',30);
```

```
size(locs);
```

```
if size(locs)<12
```

```
    locs=[locs; нулі(12-length(locs),1)];
```

```
else
```

```
locs=locs(1:12);
```

```
end
```

```
cmd = ['c' num2str(i) '= locs;'];
```

```
eval(cmd);
```

```
end
```

```
cf=(c1+c2+c3)/3;
```

```
c=ceil (cf); % % % вхідної мережі
```

```
%%%%%%%%%%%% створення масиву дата
```

```
d = [a1 a2 a3 a4 a5 a6 a7 a8 a9 a10 A11 a12 a13 a14 a15 a16 a17 a18 a19  
a20 a21 a22 a23 a24 a25 a26 a27 a28 a29 a30 b1 b2 b3 b4 b5 b6 b7 b8 B9 b10 U1  
u2 u3 u4 u5 u6 u7 u8 u9 u10]; %% BS Бруно,Каіо,Джейк
```

```
%%%%%%%%%%%% створення матриці відповідей
```

```
r = [нулів(1,30) ones (1,20)];
```

```
d=d';
```

```
r=r';
```

```
%%%%%%%%%%%% остаточна перевірка
```

```
f= rede_x2(c);
```

```
if f<=0.4
```

```
disp ('вхід проведений успішно! ');
```

```
else
```

```
    disp ('Доступ заборонено!');
```

```
end
```

ДОДАТОК Д
Слайди презентації

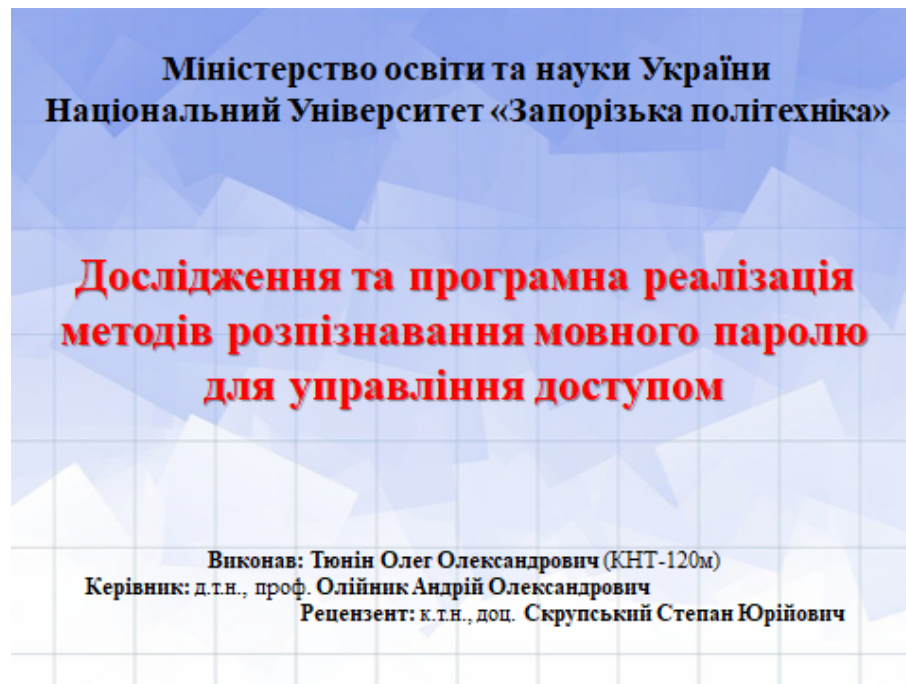


Рисунок Д.1 – Слайд №1



Рисунок Д.2 – Слайд №2

Об'єкт, предмет та мета роботи

Об'єкт дослідження – процес розпізнавання мовного паролю для ідентифікації користувачів з метою управління доступом.

Предмет дослідження – методи розпізнавання мовлення.

Мета роботи – дослідити, проаналізувати та реалізувати методи розпізнавання мовлення з метою контролю доступу.

3

Рисунок Д.3 – Слайд №3

Постановка завдання

Метою даної роботи є розробка програми ідентифікації диктора на основі мовного пароля.

Метод ідентифікації має бути заснований на методі розпізнавання образів для підвищення надійності та пришвидшені роботі системи.

Програмний застосунок (ПЗ), що реалізує запропонований метод у якості вхідних даних буде використовувати аудіофайли формату – WAV в кодуванні PCM.

4

Рисунок Д.4 – Слайд №4

Порівняння способів ідентифікації особи за біометричними				
	Вимоги до додаткової обладнання	Вимоги до середовища	Надійність при великому рівні складності	Вимоги до обробки вхідних даних
Дактилоскопія	Мінімальні (сканер)	Мінімальні	Висока	Середні
За формою руки.	Мінімальні (сканер)	Мінімальні	Середня	Середні
За райдужною оболонкою	Середні (фотосканер)	Середні	Висока	Високі
За сітківкою ока	Середні (фотосканер)	Середні	Висока	Високі
По рукописному почерку	Залежить від способу	Середні	Середня	Середні
По клавіатурному почерку	Високі	Мінімальні	Середня	Високі
За голосом	Мінімальні (мікрофон)	Середні	Висока	Середні

Рисунок Д.5 – Слайд №5



Рисунок Д.6 – Слайд №6

Порівняння мов програмування

Критерій	Matlab	Python	Java	R
Швидкість роботи застосунків	±	±	+	+
Інтегроване середовище розробки	+	+	+	+
Бібліотеки	++	++	±	+
Фреймворки для методів МН	++	++	+	+
Спільнота	++	++	±	+

Рисунок Д.7 – Слайд №7

Розгортання застосунку розробленого за допомогою Matlab

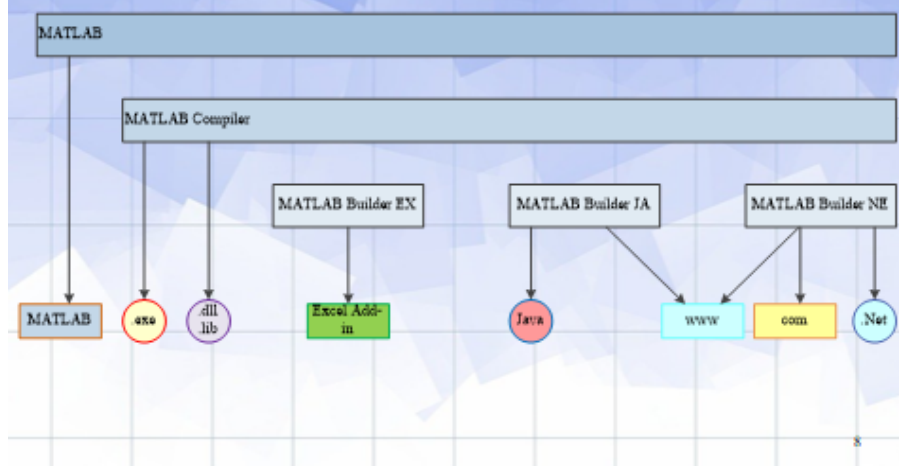


Рисунок Д.8 – Слайд №8

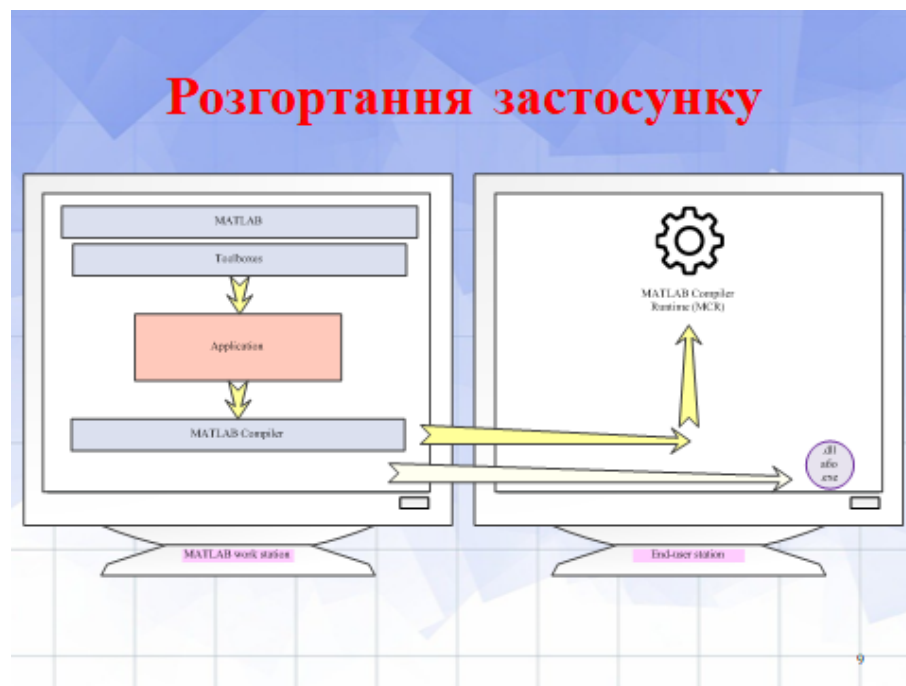


Рисунок Д.9 – Слайд №9

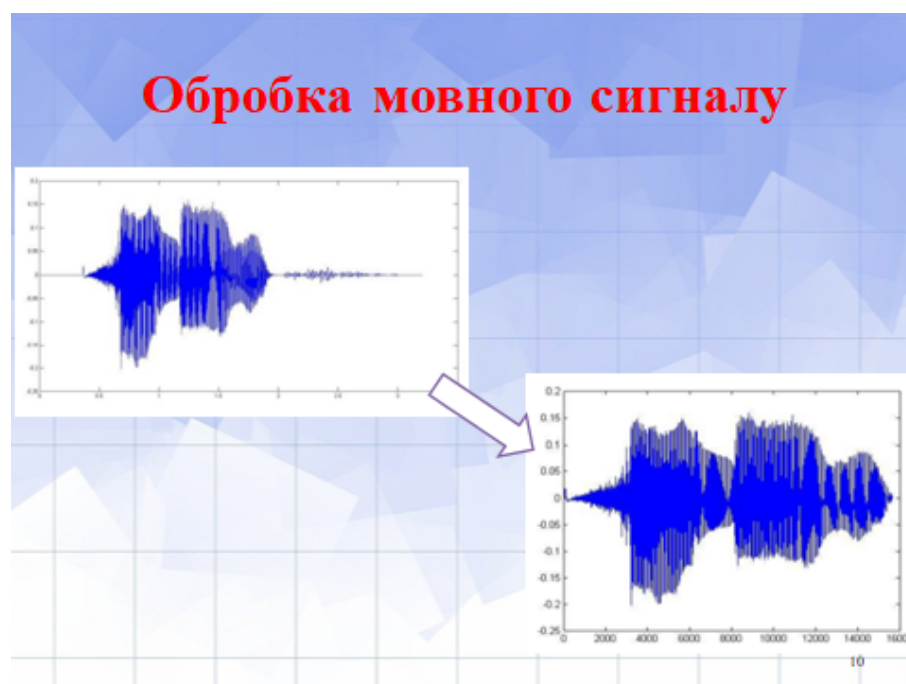


Рисунок Д.10 – Слайд №10

**Гістограми розподілів відстаней
при проголошенні одного пароля
різними дикторами**

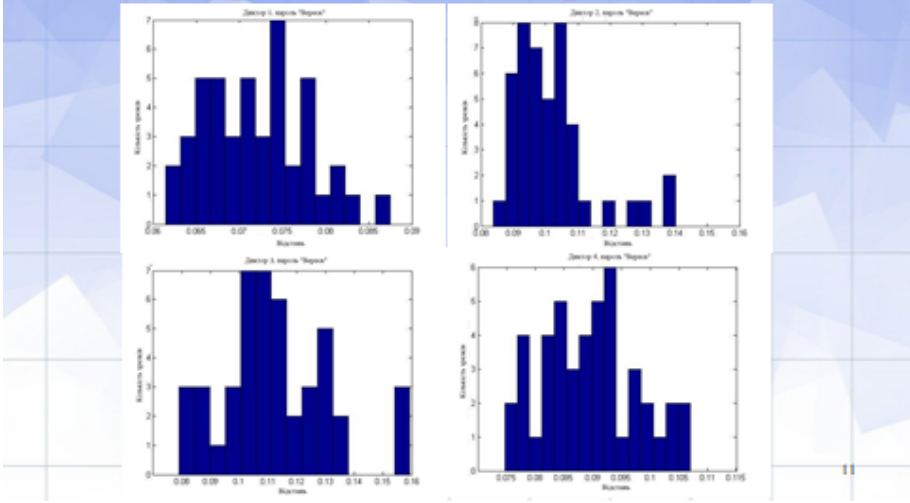


Рисунок Д.11 – Слайд №11

**Тестування програмного
забезпечення: санкціонований
доступ**

Диктор	Результати за експериментом						
	№1	№2	№3	№4	№5	№6	№7
Диктор 1	+	+	+	+	+	+	+
Диктор 2	+	+	+	–	+	+	+
Диктор 3	–	+	+	+	+	+	+
Диктор 4	+	+	+	+	+	+	+
Диктор 5	+	–	–	+	+	+	+
Диктор 6	+	+	+	+	–	–	+

Рисунок Д.12 – Слайд №12

Тестування програмного забезпечення: несанкціонований доступ							
Диктор	Результати за експериментом						
	№1	№2	№3	№4	№5	№6	№7
Диктор 1	-	-	-	-	-	-	-
Диктор 2	-	-	-	-	-	-	-
Диктор 3	+	+	-	-	-	-	-
Диктор 4	-	-	-	+	-	-	-
Диктор 5	-	-	-	-	-	-	-
Диктор 6	-	-	-	-	-	-	-

Рисунок Д.13 – Слайд №13



Рисунок Д.14 – Слайд №14

Висновки

Був розроблений сценарій обробки мовного сигналу. Попередня обробка полягає в обчисленні короткочасної енергії сигналу, для визначення меж початку і кінця слова. Використовуваний в програмі метод видлучення ознак, заснований на лінійному передбаченні. Для кожного вхідного сигналу обчислюються коефіцієнти лінійного передбачення мови, які характеризують зміну характеристик мовного сигналу в часі.

Еталонна і вимовлена фраза можуть мати різний часовий масштаб. Для приведення сигналів до одного тимчасового масштабу в цій роботі виконаний метод динамічної деформації часу, це необхідно для того, щоб точки на вимірній траєкторії збігалися найкращим чином з точками на еталонній траєкторії.

Обчислено відстань між вхідним сигналом і еталоном. Отримане значення порівнюємо зі зразками в системі, і виносимо рішення про відповідність сказаної диктором фрази з еталоном, що зберігається в системі.

В ході тестування програмного забезпечення розраховані ймовірності помилкових рішень першого і другого роду, які складають 15% і 9% відповідно.

Розрахована достовірність наданих результатів рівна 88%.

15

Рисунок Д.15 – Слайд №15