УДК 004.772
Верещака М.П.[1], Лізунов С.І.[2]

[1] аспірант гр. А-172-12 НУ «Запорізька політехніка»
[2] доцент НУ «Запорізька політехніка»

МОДЕЛЮВАННЯ ПРОЦЕСІВ ВИЯВЛЕННЯ ШКІДЛИВОГО ТРАФІКУ НА ОСНОВІ ПОВЕДІНКОВОГО АНАЛІЗУ

У сучасних умовах стрімкого розвитку інформаційних технологій та глобалізації мережевих сервісів значно зростає кількість і складність кіберзагроз. Традиційні сигнатурні методи виявлення атак стають недостатньо ефективними, оскільки вони здатні ідентифікувати лише відомі шаблони атак і не враховують появу нових або модифікованих загроз. Це обмеження зумовлює необхідність використання поведінкових підходів, які базуються на моделюванні нормальної активності мережі та виявленні відхилень від неї.

Поведінковий аналіз у системах виявлення вторгнень (Intrusion Detection Systems, IDS) передбачає формування математичної моделі нормального мережевого трафіку. У такій моделі кожен мережевий сеанс описується вектором ознак:

$$X=\{x_1,x_2,...,x_n\}$$

де x_n — параметри мережевого трафіку (тривалість сесії, кількість пакетів, середній розмір пакета, кількість байтів, частота запитів тощо). Будь-яке відхилення від нормального профілю визначається як аномалія та класифікується як потенційно шкідлива активність.

У межах дослідження запропоновано модель виявлення аномального трафіку на основі поведінкового аналізу, яка включає наступні етапи:

1. Збір даних. Для навчання та тестування моделі використовується набір даних мережевого трафіку, наприклад CICIDS2017 [1], що містить реалістичні сценарії атак (DDoS, brute-force, botnet, тощо). Даний набір включає десятки параметрів, які описують поведінку мережевих потоків.

2. Попередня обробка даних. На цьому етапі виконується підготовка даних до аналізу: видалення пропущених та аномальних значень; нормалізація ознак (Min-Max або Z-score); усунення дублікатів; балансування класів за допомогою методів oversampling (SMOTE) або undersampling.

Цей етап є критично важливим, оскільки якість вхідних даних безпосередньо впливає на точність моделі.

3. Формування поведінкової моделі. Для побудови моделі використано комбінацію алгоритмів машинного навчання:

Random Forest — для класифікації трафіку на нормальний та шкідливий;

Autoencoder — для виявлення аномалій шляхом реконструкції вхідних даних;

Функція помилки автоенкодера визначається як:

$$E=||X-\dot{X}||,$$

де $\dot{X}$ — реконструйований вектор ознак. Якщо значення помилки перевищує заданий поріг, трафік класифікується як аномальний. Такий підхід дозволяє виявляти навіть невідомі типи атак без необхідності попереднього визначення сигнатур [2].

4. Виявлення аномалій. На етапі детекції застосовуються як класичні алгоритми, так і методи глибокого навчання: рекурентні нейронні мережі (LSTM) для аналізу часових залежностей; згорткові нейронні мережі (CNN) для виявлення прихованих шаблонів у трафіку.

Це дозволяє ефективно виявляти складні багатостадійні атаки, які не можуть бути визначені за допомогою простих статистичних методів [3].

5. Оцінка ефективності моделі. Для оцінки якості роботи моделі використовуються стандартні метрики:

Accuracy — загальна точність;

Precision — точність позитивних передбачень;

Recall — повнота виявлення атак;

F1-score — гармонійне середнє точності та повноти.

Матриця помилок дозволяє оцінити кількість: істинно позитивних (TP), хибно позитивних (FP), істинно негативних (TN) та хибно негативних (FN) результатів

Проведене моделювання показало, що використання поведінкового аналізу дозволяє: виявляти невідомі (zero-day) атаки; підвищити точність детекції до 90–95%; зменшити залежність від сигнатурних баз; адаптувати систему до змін у мережевому середовищі.

Моделювання процесів виявлення шкідливого трафіку на основі поведінкового аналізу є ефективним підходом до побудови сучасних систем кібербезпеки. Запропонована модель дозволяє реалізувати проактивний підхід до безпеки, при якому загрози виявляються на ранніх етапах їх розвитку.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. CICIDS2017 Dataset [Електронний ресурс]. URL: https://www.unb.ca/cic/datasets/ids-2017.html (дата звернення: 08.03.2026).

2. Evaluation of Machine Learning Techniques for Traffic Flow-Based Intrusion Detection. [Електронний ресурс]. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC9740321 (дата звернення: 08.03.2026).

3. Anomaly-based intrusion detection on benchmark datasets for network security: a comprehensive evaluation. [Електронний ресурс]. URL: https://www.nature.com/articles/s41598-026-38317-w (дата звернення: 08.03.2026).

4. Anomaly-based intrusion detection system for IoT networks through deep learning model. [Електронний ресурс]. URL: https://www.sciencedirect.com/science/article/abs/pii/S0045790622001100 (дата звернення: 08.03.2026).