Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра комп'ютерних наук
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр
(назва освітнього ступеня)

на тему: Методи та моделі машинного навчання для виявлення аномалій у мережевому трафіку інформаційних систем

Виконав: студент VI курсу, групи СНнм-61
спеціальності 122 Комп'ютерні науки
(шифр і назва спеціальності)

_____ Ясінський О.О.
(підпис) (прізвище та ініціали)

Керівник _____ Липак Г.І.
(підпис) (прізвище та ініціали)

Нормоконтроль _____ Никитюк В.В.
(підпис) (прізвище та ініціали)

Завідувач кафедри _____ Боднарчук І.О.
(підпис) (прізвище та ініціали)

Рецензент _____ Осухівська Г.М.
(підпис) (прізвище та ініціали)

Тернопіль
2026

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних наук
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Боднарчук І.О.
(підпис) (прізвище та ініціали)
«_13_» _квітня_ 2026 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 122 Комп'ютерні науки
(шифр і назва спеціальності)

Студенту Ясінському Олександру Олеговичу
(прізвище, ім'я, по батькові)

1. Тема роботи Методи та моделі машинного навчання для виявлення аномалій у мережевому трафіку інформаційних систем

Керівник роботи Липак Галина Ігорівна, к. соц. ком., доцент кафедри КН
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «10» березня 2026 року № 4/9-150

2. Термін подання студентом завершеної роботи 28 травня 2026 р.

3. Вихідні дані до роботи Наукові публікації про методи та моделі наукового навчання для виявлення аномалій у мережевому трафіку інформаційних систем

4. Зміст роботи (перелік питань, які потрібно розробити)
Вступ. 1 Аналіз предметної області. 1.1 Обґрунтування актуальності проблеми. 1.2 Огляд наукової літератури. 1.3 Аналіз нормативної документації. 1.4 Визначення та характеристики мережевого трафіку. 1.5 Поняття та структурна класифікація аномалій. 1.6 Загрози інформаційної безпеки у вигляді аномалій. 1.7 Мережевий трафік як джерело даних для аналізу 1.8 Сучасні підходи із використанням машинного навчання. 1.9 Висновок до першого розділу. 2 Аналіз методів, систем та моделей виявлення аномалій. 2.1 Аналіз архітектури систем виявлення вторгнень. 2.2 Характеристика ансамблевих алгоритмів машинного навчання. 2.3 Порівняльна характеристика наборів даних. 2.4 Математична формалізація процесів обробки, нормалізації та балансування класів. 2.5 Стратегії вибору ознак, сегментації аномалій Та вибір метрик ефективності. 2.6 Висновок до другого розділу. 3 Створення моделі та оцінка ефективності. 3.1 Підготовка та попередня обробка набору даних. 3.2 Відбір ознак. 3.3 Модель XGBoost. 3.4 Модель Random Forest. 3.5 Висновок до третього розділу. 4 Охорона праці та безпека в надзвичайних ситуаціях. 4.1 Організація безпечного робочого місця користувача ПК. 4.2 Захист від статичної електрики в серверних приміщеннях. 4.3 Дії персоналу при виникненні пожежі в комп'ютерному залі. 4.4 Висновки до четвертого розділу. Висновки. Перелік джерел. Додатки.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
1.Титульний слайд. 2. Актуальність теми. 3. Мета та завдання дослідження. 4. Класифікація мережевих аномалій. 5. Таксономія аномалій мережевого трафіку. 6. Види архітектур IDS. 7. Види машинного навчання. 8. Методологія обробки даних. 9. Вибір метрик та стратегія оцінки. 10. XGBoost. 11. Random Forest. 12. Аналіз результатів моделювання. 13. Висновки. 14. Завершальний слайд.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Гурик О.Я., доцент кафедри МТ		
Безпека в надзвичайних ситуаціях	Теслюк В.М., проректор з адміністративно-господарської роботи та будівництва		

7. Дата видачі завдання 13 квітня 2026 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	13.04.2026	*Виконано*
2.	Підбір та опрацювання наукових публікацій, збір	13.04.2026-20.04.2026	*Виконано*
	даних по темі роботи		
3.	Виконання дослідження згідно теми кваліфікаційної	21.04.2026-03.05.2026	*Виконано*
	роботи		
4.	Оформлення розділу «Аналіз предметної області»	04.05.2026-10.05.2026	*Виконано*
5.	Оформлення розділу «Аналіз методів, систем та	04.05.2026-10.05.2026	*Виконано*
	моделей виявлення аномалій»		
6.	Оформлення розділу «Створення моделі та оцінка	04.05.2026-10.05.2026	*Виконано*
	ефективності»		
7.	Виконання завдання до підрозділу «Охорона праці»	27.04.2026-10.05.2026	*Виконано*
8.	Виконання завдання до підрозділу «Безпека в	27.04.2026-10.05.2026	*Виконано*
	надзвичайних ситуаціях»		
9.	Оформлення кваліфікаційної роботи	11.05.2026-13.05.2026	*Виконано*
10.	Нормоконтроль	14.05.2026	*Виконано*
11.	Перевірка на плагіат	15.05.2026	*Виконано*
12.	Попередній захист кваліфікаційної роботи	18.05.2026	*Виконано*
13.	Захист кваліфікаційної роботи	28.05.2026	

Студент _____ Ясінський О.О.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Липак Г.І.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Методи та моделі машинного навчання для виявлення аномалій у мережевому трафіку інформаційних систем // Кваліфікаційна робота освітнього рівня «Магістр» // Ясінський Олександр Олегович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра комп'ютерних наук, група СНнм-61 // Тернопіль, 2026 // С. 77, рис. – 22, табл. – 5, кресл. –, додат. – 1, бібліогр. – 67.

Ключові слова: мережевий трафік, виявлення аномалій, машинне навчання, інформаційна безпека, класифікація, кіберзагрози, ансамблеві методи, інтелектуальний аналіз даних.

Кваліфікаційна робота присвячена методам виявлення аномалій у мережевому трафіку за допомогою машинного навчання. У першому розділі описано таксономію аномалій та проаналізовано загрози інформаційній безпеці. Висвітлено характеристики трафіку та розглянуто інтелектуальні підходи до аналізу даних.

У другому розділі досліджено архітектуру систем IDS, SIEM та EDR. Подано математичну формалізацію підготовки даних та обґрунтовано вибір ансамблевих алгоритмів XGBoost і Random Forest.

У третьому розділі описано створення моделей та вибір значущих ознак на основі датасету CSE-CIC-IDS2018. Проведено експериментальну оцінку ефективності моделей за допомогою метрик точності.

ANNOTATION

Methods and Machine Learning Models for Anomaly Detection in Network Traffic of Information Systems // The educational level "Master" qualification work // Yasinskyi Oleksandr // Ternopil Ivan Pulyuy National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Computer Science, SNnm-61 group // Ternopil, 2026 // P. 77, fig. – 22, tables – 5, posters –, annexes – 1, ref. – 67.

Key words: network traffic, anomaly detection, machine learning, information security, classification, cyber threats, ensemble methods, intelligent data analysis.

This qualification work is dedicated to developing methods for detecting network traffic anomalies using machine learning. The first chapter describes the taxonomy of anomalies and analyzes information security threats. It highlights network traffic characteristics and reviews intelligent data analysis approaches.

The second chapter investigates the architecture of IDS, SIEM, and EDR systems. It provides a mathematical formalization of data preparation and justifies the selection of ensemble algorithms, specifically XGBoost and Random Forest.

The third chapter describes the model creation process and the selection of significant features based on the CSE-CIC-IDS2018 dataset. An experimental evaluation of model effectiveness was conducted using accuracy metrics.

ПЕРЕЛІК СКОРОЧЕНЬ

APT (англ. Advanced Persistent Threat) – Розвинена стала загроза.

DDoS (англ. Distributed Denial of Service) – Розподілена відмова в обслуговуванні.

IDS (англ. Intrusion Detection System) – Система виявлення вторгнень.

IP (англ. Internet Protocol) – Інтернет-протокол.

IPS (англ. Intrusion Prevention System) – Система запобігання вторгненням.

ML (англ. Machine Learning) – Машинне навчання.

NIDS (англ. Network Intrusion Detection Systems) – Мережева система виявлення вторгнень.

SIEM (англ. Security Information and Event Management) – Система керування інформацією та подіями безпеки.

ЗМІСТ

ВСТУП

Актуальність теми. У міру стрімкого розвитку мережевих технологій та зростання складності кіберзагроз, традиційні методи захисту, що базуються на сигнатурному аналізі, стають дедалі менш ефективними проти новітніх атак та вразливостей «нульового дня». Сучасні системи виявлення вторгнень потребують впровадження інтелектуальних підходів, здатних аналізувати великі масиви трафіку в реальному часі та ідентифікувати приховані аномалії. Більшість комерційних рішень у сфері кібербезпеки є закритими продуктами інтелектуальної власності, що обмежує можливість глибокого розуміння та адаптації їхніх алгоритмів під специфічні потреби інформаційних систем. Тому розроблення методів та моделей для автоматизованого виявлення аномалій у мережевому трафіку є актуальним напрямом сучасних наукових досліджень у галузі захисту інформації.

Мета і задачі дослідження. Метою даної кваліфікаційної роботи освітнього рівня «Магістр» є підвищення ефективності та точності виявлення різних типів кібератак у сучасних інформаційних системах. Для досягнення поставленої мети потрібно виконати ряд завдань, зокрема:

- Проаналізувати сучасний стан досліджень у сфері кібербезпеки та класифікацію мережевих аномалій.
- Дослідити архітектуру та функціональні можливості систем IDS, SIEM та засобів захисту кінцевих точок.
- Виконати порівняння еталонних наборів даних та обґрунтувати вибір датасету для моделювання.
- Проаналізувати математичні засади роботи алгоритмів XGBoost та Random Forest у задачах класифікації.
- Розробити методику попередньої обробки, нормалізації та балансування даних мережевого трафіку.
- Створити та експериментально оцінити моделі машинного навчання за допомогою метрик точності.

Об’єкт дослідження. Процеси виявлення аномальної активності в мережевому трафіку інформаційних систем на основі методів машинного навчання.

Предмет дослідження. Методи та моделі машинного навчання для виявлення та аналітичного опрацювання аномалій у мережевому трафіку інформаційних систем.

Наукова новизна одержаних результатів. Отримано подальший розвиток методів виявлення аномальної активності у мережевому трафіку на основі ансамблевих алгоритмів машинного навчання, що, на відміну від існуючих підходів, дозволяє підвищити точність ідентифікації прихованих загроз шляхом оптимізації процесів попередньої обробки та відбору значущих ознак даних.

Практичне значення одержаних результатів. Розроблено та програмно реалізовано моделі машинного навчання на основі алгоритмів XGBoost та Random Forest для автоматизованого виявлення мережевих вторгнень. Запропоновано методику оптимізації підготовки даних, що підвищує точність і швидкість аналізу трафіку в інформаційних системах. Результати дослідження можуть бути впроваджені в сучасні системи виявлення вторгнень (IDS) для посилення кіберзахисту.

Апробація результатів магістерської роботи. Основні результати проведених досліджень обговорювались на XIV міжнародній науково-практичній конференція молодих учених та студентів «Актуальні задачі сучасних технологій» та IX міжнародній студентській науково-технічній конференції «Природничі та гуманітарні науки. Актуальні питання».

Публікації. Основні результати кваліфікаційної роботи опубліковано у двох працях конференції (Див. додаток А).

Структура й обсяг кваліфікаційної роботи. Кваліфікаційна робота складається зі вступу, чотирьох розділів, висновків, списку літератури з 67 найменувань та 1 додатка. Загальний обсяг кваліфікаційної роботи складає 87 сторінки, з них 77 сторінки основного тексту, який містить 22 рисунки та 5 таблиць.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Обгрунтування актуальності проблеми

Сучасний етап розвитку глобального інформаційного простору характеризується безпрецедентною динамікою цифрової трансформації, що водночас призводить до критичного розширення поверхні потенційних кібератак. Глобальний стан кібербезпеки демонструє стійку тенденцію до професіоналізації зловмисників, які все частіше використовують автоматизовані інструменти та штучний інтелект для пошуку вразливостей [1]. В контексті України ця проблема набуває особливої гостроти через постійні гібридні загрози та цілеспрямовані атаки на об'єкти критичної інформаційної інфраструктури, що потребує впровадження найбільш прогресивних засобів моніторингу та захисту даних [2].

Головною деструктивною особливістю сучасних кіберзагроз є стрімке зростання кількості атак нульового дня, які за своєю природою не мають заздалегідь відомих сигнатур. Традиційні антивірусні засоби та брандмауери, що функціонують на основі порівняння вхідних даних із базами відомих шкідливих шаблонів, виявляються неспроможними протидіяти таким втручанням [3]. Оскільки розробка сигнатури завжди відбувається вже після факту виявлення та аналізу нової атаки, виникає часовий розрив, протягом якого інформаційна система залишається повністю вразливою. Це зумовлює перехід наукової та практичної спільноти до концепції проактивного захисту, де основна увага приділяється не пошуку відомих вразливостей, а аналізу самої структури мережевої поведінки.

Саме виявлення аномалій у мережевому трафіку стає ключовим інструментом для ідентифікації прихованих загроз у таких складних умовах [4]. Будь-яка несанкціонована активність, незалежно від її новизни, неминуче спричиняє відхилення від нормального профілю функціонування мережі, будь то нетиповий обсяг переданих пакетів, специфічні часові затримки або звернення до нехарактерних портів. Аналіз цих відхилень дозволяє фіксувати присутність

зловмисника на ранніх етапах інфільтрації, ще до моменту нанесення прямої шкоди системі. Таким чином, перехід від статичних правил до динамічного аналізу аномалій на основі машинного навчання є єдиним дієвим шляхом для забезпечення надійної безпеки в умовах постійної еволюції кіберзлочинності [5].

1.2 Огляд наукової літератури

Сучасний стан наукових досліджень у сфері виявлення мережевих вторгнень базується на фундаментальних працях, що заклали основу для переходу від простих фільтрів до складних аналітичних систем. Вагомий внесок у розвиток цієї галузі зробив Верн Паксон, чиї дослідження щодо архітектури систем виявлення вторгнень та розробка системи Bro продемонстрували важливість глибокого аналізу протоколів та поведінкових характеристик трафіку [6]. Паралельно з цим, група дослідників з Канадського інституту кібербезпеки здійснила прорив у забезпеченні наукової спільноти якісними даними, створивши серію датасетів CICIDS [7]. Ці набори даних стали стандартом для тестування алгоритмів завдяки детальній розмітці різноманітних типів атак, що дозволило іншим вченим об'єктивно порівнювати ефективність нових методів детектування.

Наукова дискусія останніх років зосереджена на порівнянні класичних статистичних підходів та методів машинного навчання. Традиційні статистичні моделі базуються на побудові профілю нормальної поведінки та фіксації відхилень від середньоквадратичних показників, що забезпечує високу швидкість обробки, проте часто призводить до великої кількості помилкових спрацьовувань у динамічних мережах [8]. На противагу цьому, підходи на основі машинного навчання, зокрема градієнтний бустинг та глибокі нейронні мережі, здатні виявляти складні нелінійні залежності між ознаками трафіку. Це дозволяє моделям краще адаптуватися до мінливого середовища та розпізнавати завуальовані загрози, які залишаються непоміченими для класичних статистичних фільтрів.

Незважаючи на значні успіхи, у сучасній науці залишаються певні «білі плями», що стримують масове впровадження інтелектуальних систем. Ключовою проблемою є суттєвий розрив між теоретичними показниками точності на статичних датасетах та реальною ефективністю систем при аналізі живого трафіку в реальному часі. Більшість академічних моделей демонструють ідеальні результати в лабораторних умовах, проте стикаються з проблемою невідповідності ознак та обчислювальною складністю при обробці високошвидкісних потоків даних [9]. Питання трансформації складних статистичних агрегатів, на яких навчаються моделі, у прості ознаки окремих пакетів для миттєвого реагування залишається однією з найбільш актуальних та найменш досліджених тем у сучасній кібербезпеці.

1.3 Аналіз нормативної документації

Аналіз нормативно-правового поля та міжнародних технічних регламентів є критично важливим етапом, оскільки він визначає легітимність та необхідність впровадження інтелектуальних систем захисту в сучасну IT-інфраструктуру. Фундаментальним документом у вітчизняному правовому просторі є Закон України «Про основні засади забезпечення кібербезпеки України», який покладає на суб'єктів критичної інфраструктури обов'язок щодо створення систем виявлення кібератак та своєчасного реагування на кіберінциденти [10].

На міжнародному рівні ключовим орієнтиром є оновлений стандарт ISO/IEC 27001:2022, який у переліку заходів безпеки чітко визначає необхідність безперервного моніторингу мережевої активності [11]. Ця норма вимагає від організацій не просто фіксації подій, а глибокого аналізу поведінкових характеристик систем для виявлення аномалій, що не піддаються сигнатурному розпізнаванню. Додаткову методологічну базу забезпечують рекомендації NIST SP 800-94 Rev. 1, які деталізують архітектурні принципи побудови систем виявлення вторгнень (IDS), наголошуючи на важливості автоматизації аналізу великих потоків даних для мінімізації часу реакції на загрозу.

Відповідність розробки моделі відповідає вимогам стандарту ISO/IEC 27005:2022 щодо управління ризиками та дозволяє інтегрувати її в загальну систему менеджменту безпеки підприємства [12]. Такий підхід забезпечує не лише технічну ефективність виявлення аномалій, а й гарантує дотримання принципів конфіденційності та цілісності даних згідно зі світовими стандартами. При проєктуванні систем збору та моніторингу даних велике значення має впровадження концепцій захищених інформаційних систем, які забезпечують надійне збереження, конфіденційність та безпечну консолідацію накопичених масивів інформації від потенційних кіберзагроз і несанкціонованого доступу [13]. Отже, розробка інтелектуальної моделі для аналізу мережевого трафіку є логічним та нормативно обґрунтованим кроком, що відповідає актуальним векторам розвитку глобальної та національної кібербезпеки в умовах еволюції цифрових загроз.

1.4 Визначення та характеристики мережевого трафіку

Мережевий трафік являє собою потік даних, що передається через комп'ютерну мережу протягом певного періоду часу. Кількісно трафік зазвичай вимірюється в пакетах або байтах, переданих за одиницю часу.

В контексті кібербезпеки, зокрема для задач систем виявлення вторгнень, мережевий трафік є основним джерелом даних для аналізу. Ефективність сучасних NIDS значною мірою залежить від якості та повноти даних, на яких вони навчаються. Однак аналіз трафіку ускладнюється через його фундаментальні характеристики [14]:

- Великий обсяг. Сучасні мережі генерують величезні обсяги даних, що висуває високі вимоги до продуктивності інструментів для їх збору, зберігання та обробки в реальному часі.
- Різноманітність типів. Трафік складається з безлічі протоколів та різноманітних форматів даних. Кожен протокол має унікальні патерни поведінки та потенційні вектори атак.

- Динамічність. Патерни трафіку постійно та швидко змінюються. Ця мінливість залежить від часу доби, поведінки користувачів, бізнес-процесів та появи нових мережевих додатків. Це значно ускладнює визначення статичної «норми» для систем поведінкового аналізу.
- Гетерогенність. Розподіл трафіку між вузлами та каналами мережі є вкрай нерівномірним. Це обумовлює необхідність локалізованого або розподіленого аналізу замість єдиної точки моніторингу.

Сукупність цих характеристик створює значні виклики для систем безпеки, оскільки шкідлива активність може бути прихована у великих обсягах легітимних даних.

1.5 Поняття та структурна класифікація аномалій

У загальному сенсі, аномалія визначається як будь-яке суттєве відхилення від очікуваної або нормальної поведінки. У контексті мережевої безпеки, виявлення аномалій є процесом ідентифікації подій або патернів, які не відповідають встановленій «нормі». Для побудови ефективної системи захисту критично важливо розуміти як існуючі підходи до виявлення, так і саму природу аномалій [15]. На рисунку 1.1 зображено вигляд мережевої аномалії.

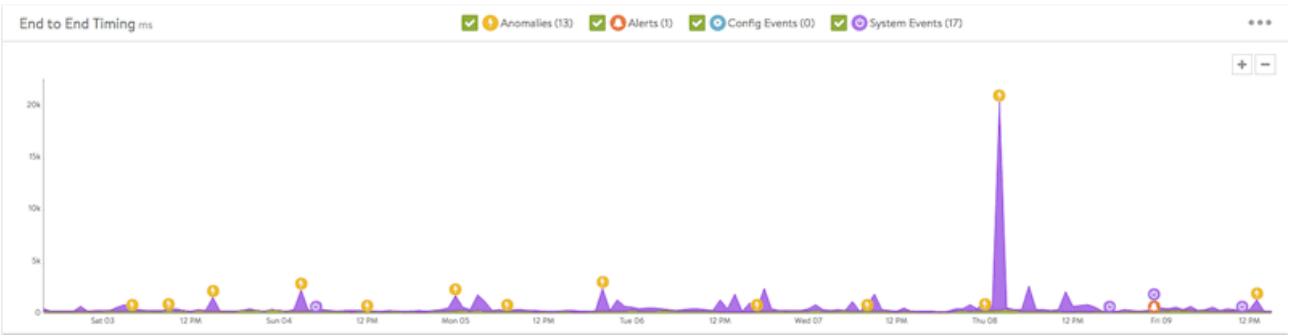

Рисунок 1.1 – Мережева аномалія

Як ми бачимо із рисунку 1.1, аномалія це сплеск певної активності в певний момент часу, що суттєво контрастує з трафіком в інші часові проміжки [16].

У сфері виявлення вторгнень історично склалися два основні підходи. Перший – це виявлення зловживань, або сигнатурний метод. Він працює за

простим принципом: шукає у трафіку відомі патерни атак, які називають сигнатурами. Майже всі комерційні IDS використовують саме його. Його плюс – висока точність при виявленні відомих атак і низький рівень помилкових спрацьовувань. Але є і величезний мінус: такі системи вразливі до нових, невідомих чи модифікованих атак, для яких сигнатур ще просто не існує. На рисунку 1.2 зображено схему роботи сигнатурного методу [17].

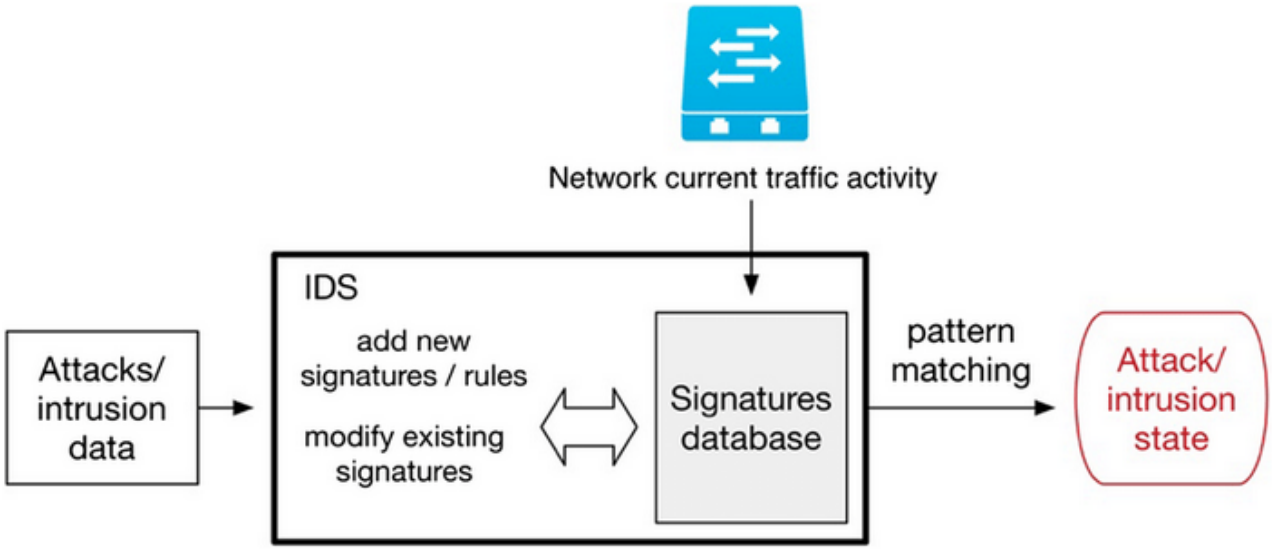

Рисунок 1.2 – Сигнатурний метод

Другий підхід – це виявлення аномалій, який є фокусом нашої роботи. Цей метод спочатку будує модель, або профіль, нормальної поведінки системи. А вже потім будь-яка активність, що суттєво від цієї моделі відхиляється, позначається як аномалія і, отже, як потенційне вторгнення. Головна перевага цього методу очевидна – він здатний виявляти нові та невідомі атаки [18]. На рисунку 1.3 зображено метод виявлення аномалій.

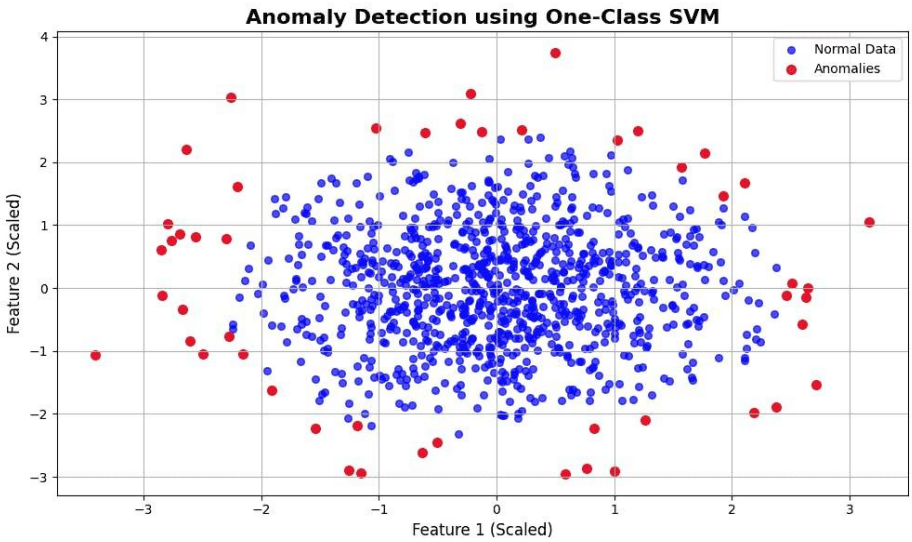

Рисунок 1.3 – Метод виявлення аномалій

Із рисунку 1.3 можна побачити що профіль нормальної поведінки зображено синім кольором, а точки що відхиляються від «норми» зображено червоним.

Однак і тут є свої складнощі. По-перше, буває важко точно визначити саму межу «норми». По-друге, це часто призводить до вищого рівня помилкових спрацьовувань, адже не кожна аномалія автоматично є атакою – це може бути просто рідкісна, але цілком легітимна поведінка.

Оскільки ми фокусуємося на виявленні аномалій, нам важливо їх правильно класифікувати. Наприклад, українські дослідники пропонують досить широку таксономію з 11 категорій, куди входять аномалії поведінки користувачів, мережевого трафіку, системних журналів, даних, ресурсів, ПЗ, архітектури, ідентифікації, захисту від вразливостей, а також хмарні та часові аномалії [19].

Ця таксономія дуже детальна, але вона охоплює всю інформаційну систему. Для наших завдань, а саме для NIDS вона занадто широка. NIDS аналізує виключно мережевий трафік. Вона не може «зазирнути» в програмне забезпечення на хості чи оцінити архітектуру; вона бачить лише прояви цих проблем у мережі.

Тому для цілей даного дослідження необхідно сфокусувати цю таксономію. Релевантними для NIDS є лише ті категорії, які можна спостерігати в трафіку:

- аномалії в мережевому трафіку;
- аномалії в ресурсах;
- аномалії в ідентифікації;
- аномалії, пов’язані з часом.

Для цілей NIDS таксономія фокусується виключно на тих аномаліях, які мають безпосередній прояв у структурі, ідентифікації або динаміці мережевого трафіку.

Окрім того, де шукати аномалії, важливо розуміти їхню структуру з точки зору аналізу даних . Це потрібно для вибору правильних математичних моделей. Виділяють три основні типи:

Точкові аномалії – це найпростіший тип, окремі екземпляри даних, які сильно вибиваються із загальної маси. Це може бути раптовий, короткий сплеск трафіку, пакет з аномально великим розміром або з'єднання з дивними прапорами TCP [20]. На рисунку 1.4 зображено точкову аномалію.

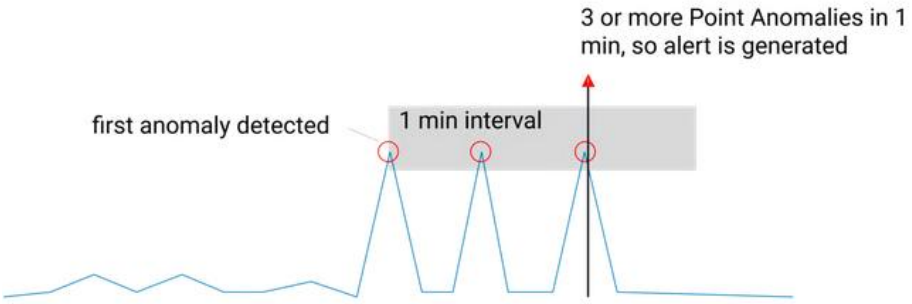

Рисунок 1.4 – Точкова аномалія

З рисунку 1.4 можна зробити висновок що на протязі однієї хвилини було помічено 3 точкових аномалії що спричинило тривогу.

Контекстуальні аномалії – це дані, які самі по собі є нормальними, але стають аномальними в певному контексті.

Наприклад, вхід адміністратора в систему о 3 годині ночі – сам по собі вхід легітимний, але часовий контекст робить його підозрілим. Інший

приклад – інтенсивний FTP-трафік на веб-сервері, де зазвичай має бути лише HTTP. На рисунку 1.5 зображено контекстуальну аномалію.

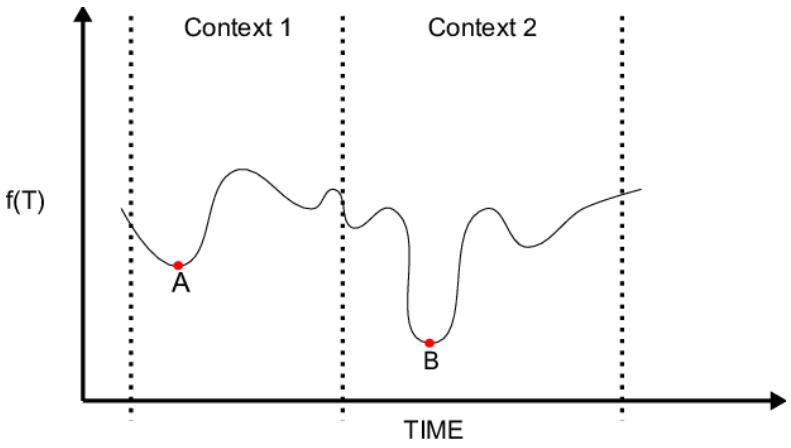

Рисунок 1.5 – Контекстуальна аномалія

Колективні аномалії – це найскладніший тип. Тут ціла сукупність даних є аномальною, хоча кожен окремий елемент у цій групі може виглядати абсолютно нормальним. Класичний приклад – повільне сканування портів. Кожен окремий запит до порту виглядає легітимним, але їх сукупність, розтягнута в часі, формує чіткий патерн розвідки [21]. На рисунку 1.6 зображено колективну аномалію.

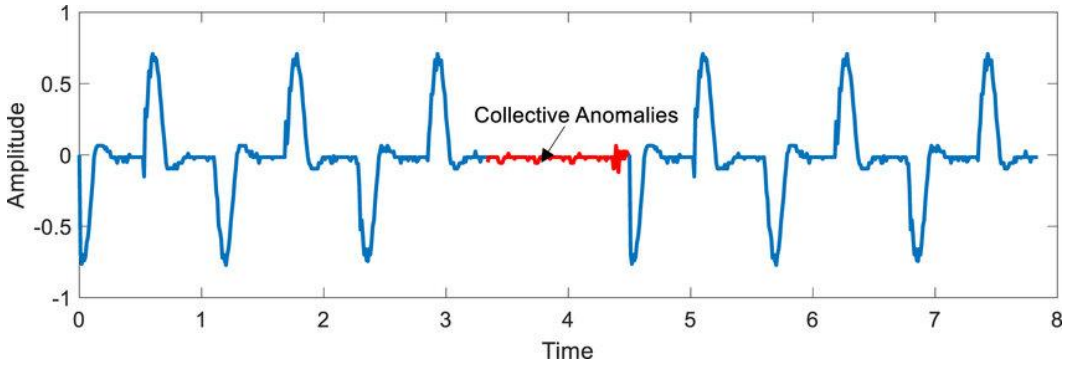

Рисунок 1.6 – Колективна аномалія

З рисунку 1.6 видно що колективна аномалія ідентифікується як сегмент сигналу (виділений червоним кольором), де звична періодична послідовність піків переривається аномально стабільним плато з незначними шумоподібними коливаннями. На відміну від одиничних викидів, дана аномалія визначається саме тривалістю нетипової поведінки часового ряду, що свідчить про системне

порушення нормального ритму процесу в інтервалі між третьою та п’ятою одиницями часу.

1.6 Загрози інформаційної безпеки у вигляді аномалій

Сучасні інформаційні системи не можуть функціонувати без мережевої інфраструктури. Проте, саме вона, на жаль, і є головним вектором для різноманітних кіберзагроз.

Серед основних загроз, що проявляються у вигляді аномалій, варто виділити наступні:

- Розподілені атаки «відмова в обслуговуванні». Вони характеризуються аномально великим обсягом трафіку, роблячи її недоступною.
- Сканування портів. Це дії зловмисника, який намагається знайти відкриті порти та вразливі сервіси на сервері. У трафіку це виглядає як велика кількість спроб підключення з однієї IP-адреси до багатьох портів (або навпаки).
- Діяльність ботнетів. Це скоординовані дії цілої групи заражених машин, якими керують з єдиного центру. Вони проявляються через нетипову, часто приховану, комунікацію з командно-контрольними серверами.
- Атаки Brute Force. Це, по суті, спроби «вгадати» облікові дані (наприклад, логін та пароль до FTP чи SSH) шляхом простого перебору великої кількості варіантів.
- Веб-атаки. Сюди відносять SQL-ін'єкції, Cross-Site Scripting та інші атаки, що використовують вразливості у веб-додатках.
- Інфільтрація та ексфільтрація. Несанкціоноване проникнення в мережу, за яким слідує викрадення чутливих даних.

Систематизація знань про природу мережевих аномалій дозволяє побудувати ефективну багаторівневу систему захисту, оскільки кожен вид деструктивного впливу залишає свій унікальний цифровий відпечаток у трафіку. Точкові аномалії зазвичай відповідають різким сплескам активності, характерним для експлуатації вразливостей або спроб несанкціонованого доступу, тоді як контекстні аномалії вимагають глибшого аналізу часових чи

просторових залежностей для виявлення прихованої розвідки. Найскладнішими для ідентифікації залишаються колективні аномалії, які можуть свідчити про розподілені атаки на кшталт DDoS або складні стійкі загрози, де кожен окремий пакет виглядає легітимним, але загальна послідовність дій вказує на атаку.

Для глибшого розуміння механізмів виявлення варто детальніше розглянути таксономію, наведену в таблиці 1.1, яка класифікує відхилення за їхньою структурою та логікою виникнення.

Таблиця 1.1 – Таксономія аномалій мережевого трафіку

Тип аномалії	**Контекст кібербезпеки**	**Прояв у трафіку**	**Опис**	**Приклад атаки**
Точкова	Аномалії в мережевому трафіку; Аномалії в ресурсах	Аномалії обсягу	Одиничний екземпляр даних значно відхиляється від норми.	DDoS-атака, сканування портів
Контекстуальна	Аномалії, пов’язані з часом; Аномалії в поведінці користувачів	Поведінкові аномалії	Дані є нормальними за значенням, але аномальними у певному контексті (час доби, геолокація, попередня поведінка).	Вхід адміністратора в систему з нетипової країни; мережева активність у неробочий час.
Колективна	Аномалії в мережевому трафіку	Поведінкові аномалії	Група пов'язаних екземплярів даних є аномальною в сукупності, хоча кожен окремо може виглядати нормальним.	Скоординована комунікаційна поведінка ботнету

Наприклад, DDoS та DoS-атаки найчастіше проявляються як точкова або колективна аномалія у вигляді екстремального сплеску обсягу трафіку.

Сканування портів – це типова колективна аномалія, яка характеризується великою кількістю спроб з'єднань до різних портів з одного джерела.

Складніші атаки, як-от APT, намагаються бути непомітними [22]. Вони генерують ледь помітні контекстуальні (нетиповий час входу) та колективні

аномалії (повільна передача даних назовні, нехарактерні з'єднання між серверами всередині мережі).

Навіть програми-вимагачі залишають сліди [23]. На етапі поширення це може бути нетипова активність SMB, а під час шифрування файлів на мережевих дисках – це генерує колективну аномалію у вигляді масових операцій читання/запису файлів за короткий час.

1.7 Мережевий трафік як джерело даних для аналізу

Існують два основні підходи до збору мережевих даних, що різняться рівнем деталізації, вимогами до ресурсів та сферою застосування.

Аналіз на рівні пакетів, також відомий як повне захоплення пакетів, передбачає захоплення та збереження повних мережевих пакетів, включно з усіма заголовками та корисним навантаженням. Дані зазвичай зберігаються у форматі pcap [24].

До переваг необхідно віднести максимальний рівень деталізації для аналізу мережевими аналітиками. Це дозволяє досліджувати безпосередній вміст переданих даних, що є критичним для виявлення атак, сигнатури яких знаходяться саме в корисному навантаженні.

До недоліків слід віднести те що метод є вкрай ресурсоємним та вимагає значних обчислювальних потужностей та дискового простору для зберігання й обробки. Ключовою проблемою сучасності є широке розповсюдження шифрування. Аналіз зашифрованого корисного навантаження неможливий без складних процедур розшифрування, що часто є непрактичним, неможливим або небажаним з точки зору приватності.

Аналіз на рівні потоків агрегує пакети, що мають спільні характеристики, в єдиний запис, який називається «потоком». Стандартне визначення потоку базується на комбінації п'яти полів:

1. IP-адреса джерела.
2. Порт джерела.
3. IP-адреса призначення.

4. Порт призначення.

5. Протокол транспортного рівня.

До переваг необхідно віднести те що аналіз потоків значно зменшує обсяг даних, що зберігаються, порівняно з повним захопленням пакетів. Це робить його ідеальним для довгострокового моніторингу та аналізу високошвидкісних мереж. Аналіз статистичних характеристик потоків дозволяє ефективно виявляти багато типів мережевих аномалій, таких як DDoS-атаки, сканування портів або мережева розвідка. У процесі дослідження систем безпеки «розумного міста» особлива увага приділяється технологіям логічної консолідації та попередньої обробки розподілених інформаційних ресурсів для забезпечення їх подальшого аналізу алгоритмами машинного навчання [25].

До недоліків слід віднести відсутність доступу до корисного навантаження унеможливлює виявлення атак, які не змінюють статистичні патерни метаданих потоку.

1.8 Сучасні підходи із використанням машинного навчання

Машинне навчання (machine learning, ML) дає набагато потужніший і гнучкіший підхід для виявлення аномалій, ніж традиційні статистичні методи. Головна перевага ML-моделей у тому, що вони здатні «знаходити» складні, нелінійні зв'язки у величезних масивах даних та адаптуватися до змін.

Коли йдеться про виявлення мережевих аномалій, найчастіше говорять про кероване навчання (Supervised Learning, SL). Це, по суті, «навчання з вчителем», коли модель тренують на даних, де кожен потік вже має готову мітку: «Normal» або «Malicious». Задача зводиться до звичайної класифікації, де модель вчиться розпізнавати патерни атак [26]. Головна вимога для цього – мати якісний, детально розмічений набір даних.

Протилежний підхід – некероване навчання (Unsupervised Learning, UL), або «навчання без вчителя». Тут модель отримує дані без будь-яких міток, що набагато ближче до реальних умов, де готових розміток атак майже ніколи немає. Замість класифікації, модель сама шукає приховану структуру в даних,

наприклад, через кластеризацію (де нормальні дані групуються разом, а аномалії – це викиди) або через пряме виявлення викидів [27].

Навчання з підкріпленням (Reinforcement Learning, RL) – це коли модель вчиться методом «проб і помилок», отримуючи «винагороду» чи «штраф» за свої рішення [28]. Цей підхід, щоправда, рідше використовують для самого виявлення, а частіше – у системах, які мають активно реагувати.

У контексті систем виявлення вторгнень існує такий собі фундаментальний парадокс між тим, що роблять в академічних дослідженнях, і тим, що потрібно на практиці.

Академічні дослідження переважно використовують кероване навчання [29]. Причина проста: є готові, повністю розмічені датасети. На них дуже легко порахувати об'єктивні метрики, а саме: Accuracy, Precision, Recall, F1-score і чітко порівняти, який алгоритм спрацював краще. Для підвищення надійності функціонування інтелектуальних систем детекції станів загроз доцільно використовувати передові архітектурні рішення у сфері інтеграції та оптимізації моделей штучного інтелекту, що забезпечують стабільні показники точності під час класифікації складних багатофакторних процесів [30]. При використанні методів штучного інтелекту для детекції аномалій необхідно враховувати ризики безпеки самих моделей, зокрема загрози впровадження бекдор-атак на етапі навчання, які здатні цілеспрямовано викривляти можливості розпізнавання нейронних мереж та знижувати точність класифікації загроз [31].

1.9 Висновок до першого розділу

В першому розділі кваліфікаційної роботи освітнього рівня «Магістр» проаналізовано сучасний стан предметної області, визначено ключові характеристики мережевого трафіку та класифіковано основні типи аномалій і кіберзагроз, що дозволило обґрунтувати необхідність впровадження інтелектуальних методів аналізу для підвищення рівня безпеки інформаційних систем.

2 АНАЛІЗ МЕТОДІВ, СИСТЕМ ТА МОДЕЛЕЙ ВИЯВЛЕННЯ АНОМАЛІЙ

2.1 Аналіз архітектури систем виявлення вторгнень

Системи виявлення та запобігання вторгненням є критично важливими компонентами ешелонованого захисту інформаційних систем. Основними завданнями цих засобів є ідентифікація, реєстрація та нейтралізація несанкціонованої активності в мережі або на окремих вузлах.

Мережеві системи (Network-based IDS, NIDS) орієнтовані на перехоплення та аналіз пакетів безпосередньо у визначеному сегменті мережі [32]. Такі системи функціонують переважно у пасивному режимі, здійснюючи моніторинг трафіку, що спрямований до багатьох вузлів одночасно. Головною перевагою NIDS є можливість централізованого охоплення великої інфраструктури без створення додаткового навантаження на обчислювальні ресурси кінцевих хостів. Крім того, завдяки відсутності активного втручання в потік даних, ці сенсори залишаються майже непомітними для зовнішнього зловмисника. Проте існують і суттєві технологічні бар'єри: NIDS важко справляються з аналізом пакетів у надшвидкісних комп’ютерних мережах, мають складнощі в роботі з комутованою архітектурою та принципово не здатні аналізувати вміст зашифрованого трафіку без наявності ключів розшифрування. Додатковим деструктивним фактором є висока чутливість NIDS до специфічних методів обходу захисту, таких як фрагментація пакетів або штучне зашумлення трафіку, які зловмисники використовують для десинхронізації баз сигнатур. Це призводить до стрімкого зростання рівня хибнопозитивних спрацьовувань та створює надмірне навантаження на аналітичні системи моніторингу [33]. На рисунку 2.1 зображено мережеву систему.

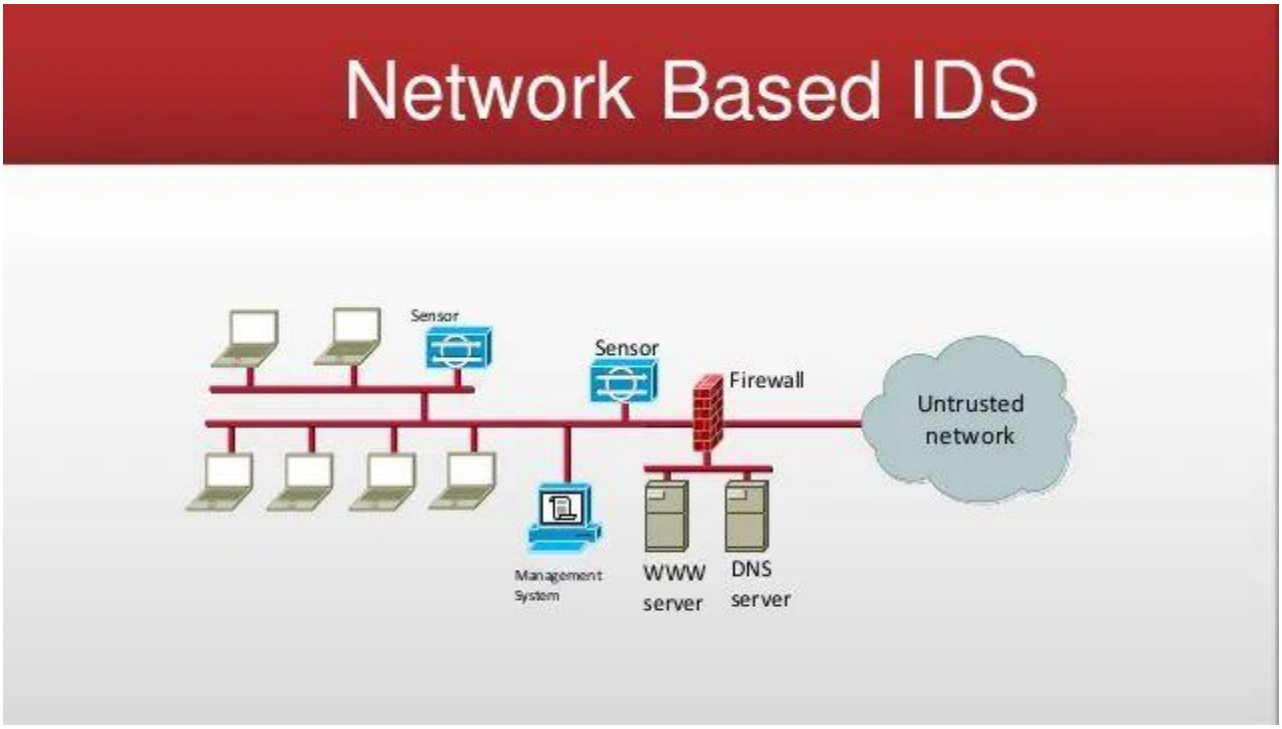

Рисунок 2.1 – Мережева система

На противагу мережевим рішенням, вузлові системи (Host-based IDS, HIDS) розгортаються безпосередньо на конкретних обчислювальних одиницях. Їхня робота базується на вивченні внутрішніх системних подій, аналізі журналів аудиту та контролі цілісності критичних файлів операційної системи. Ключова цінність HIDS полягає у здатності фіксувати безпосередні наслідки атаки, такі як несанкціонована зміна системних конфігурацій, а також у можливості аналізувати дані вже після їхнього дешифрування на рівні хоста. Водночас розгортання HIDS вимагає значних адміністративних витрат, оскільки кожен вузол потребує індивідуальної інсталяції та підтримки. Крім того, компрометація самого хоста часто призводить до виведення з ладу локального модуля IDS, а мережеві атаки на етапі сканування зазвичай залишаються поза увагою таких систем [34]. На рисунку 2.2 зображено вузлову систему.

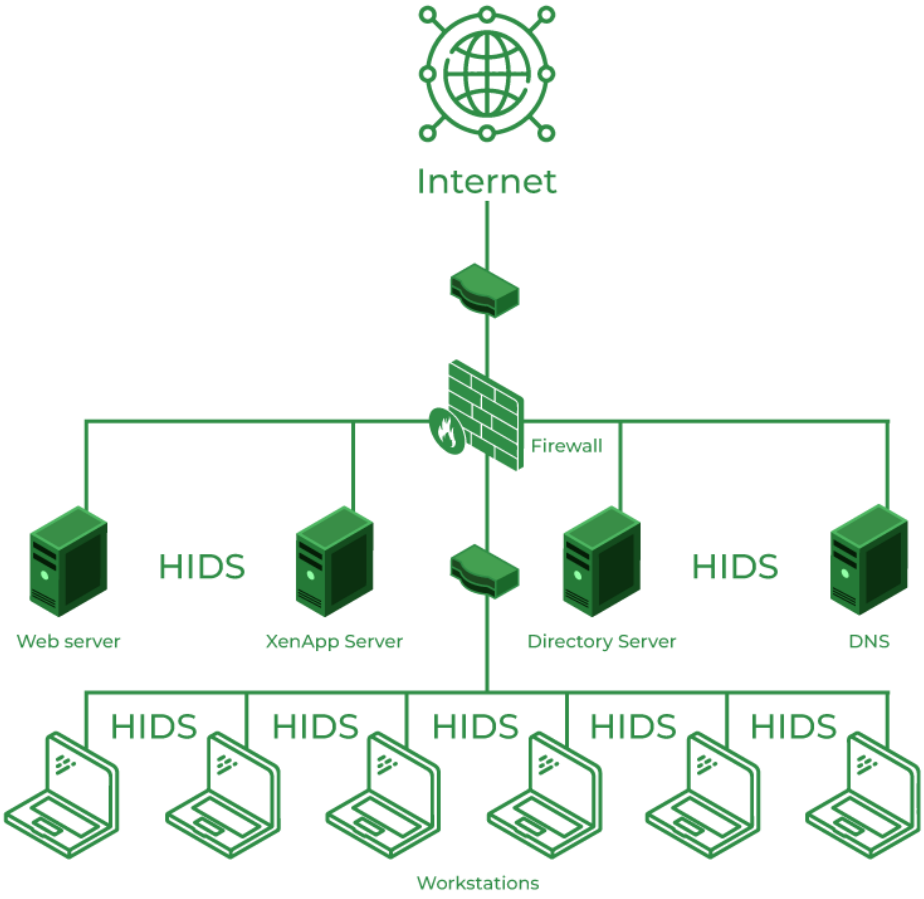

Рисунок 2.2 – Вузлова система

Прикладні системи (Application-based IDS, AIDS) виступають вузькоспеціалізованим відгалуженням вузлових рішень, фокусуючись виключно на подіях всередині конкретних програмних додатків, таких як веб-сервери або системи управління базами даних [35]. Глибоке розуміння контексту та семантики роботи конкретного ПЗ дозволяє AIDS ідентифікувати зловмисні дії навіть з боку легітимних користувачів, що є складним завданням для систем інших типів. Проте через свою спеціалізацію вони виявляються неефективними проти загальносистемних загроз і залишаються вразливими до технік маніпуляції лог-файлами, якщо зловмисник отримує прямий доступ до механізмів журналювання програми.

Найбільш ефективною стратегією побудови сучасної архітектури безпеки є впровадження гібридних систем, які інтегрують можливості мережевого та вузлового моніторингу в єдине інформаційне середовище. У такій структурі дані з різних джерел передаються до централізованого модуля управління, де

здійснюється їх кореляція та комплексний аналіз. Використання гібридного підходу дозволяє нівелювати індивідуальні недоліки кожного типу систем: мережеві сенсори забезпечують загальну видимість периметра, тоді як вузлові агенти гарантують детальний контроль за станом конкретних хостів. Це створює багаторівневий ешелон захисту, здатний ефективно протидіяти як масовим мережевим атакам, так і цілеспрямованим вторгненням, що використовують методи прихованого переміщення всередині інфраструктури.

Логічним розвитком систем виявлення є перехід до систем запобігання вторгненням, які, на відміну від пасивних засобів моніторингу, здатні втручатися в процес передачі даних для негайної нейтралізації загрози. У мережевому виконанні такі системи зазвичай розгортаються «в розрив» трафіку, що дозволяє їм аналізувати кожен пакет до того, як він потрапить до кінцевого адресата. Це дає можливість автоматично скидати підозрілі пакети, розривати встановлені TCP-сесії або динамічно переналаштовувати правила мережевого екрана для блокування IP-адреси зловмисника. Такий підхід мінімізує час реакції на інцидент, оскільки виключає затримку, пов'язану з людським фактором, що є критично важливим при протидії швидкісним атакам, таким як черв'яки або експлуатація вразливостей прикладного рівня.

Snort є відкритою системою виявлення та запобігання вторгненням, що базується на методі сигнатурного аналізу. Функціонування системи полягає у перехопленні мережевих пакетів та їх порівнянні з базою заздалегідь визначених правил, які описують патерни відомих кіберзагроз.

Процес обробки трафіку в Snort включає три основні етапи:

1. Нормалізація мережевих пакетів та підготовка їх до аналізу.
2. Порівняння структури та вмісту пакетів із сигнатурами в реальному часі за допомогою аналітичного рушія.
3. Генерація сповіщень або активне блокування трафіку при виявленні збігів.

Головною перевагою Snort є висока ефективність проти відомих атак та гнучкість налаштування правил. Проте ключовим недоліком залишається неспроможність системи ідентифікувати загрози нульового дня, що зумовлює

актуальність впровадження інтелектуальних методів детекції на основі машинного навчання. На рисунку 2.3 зображено середовище Snort.

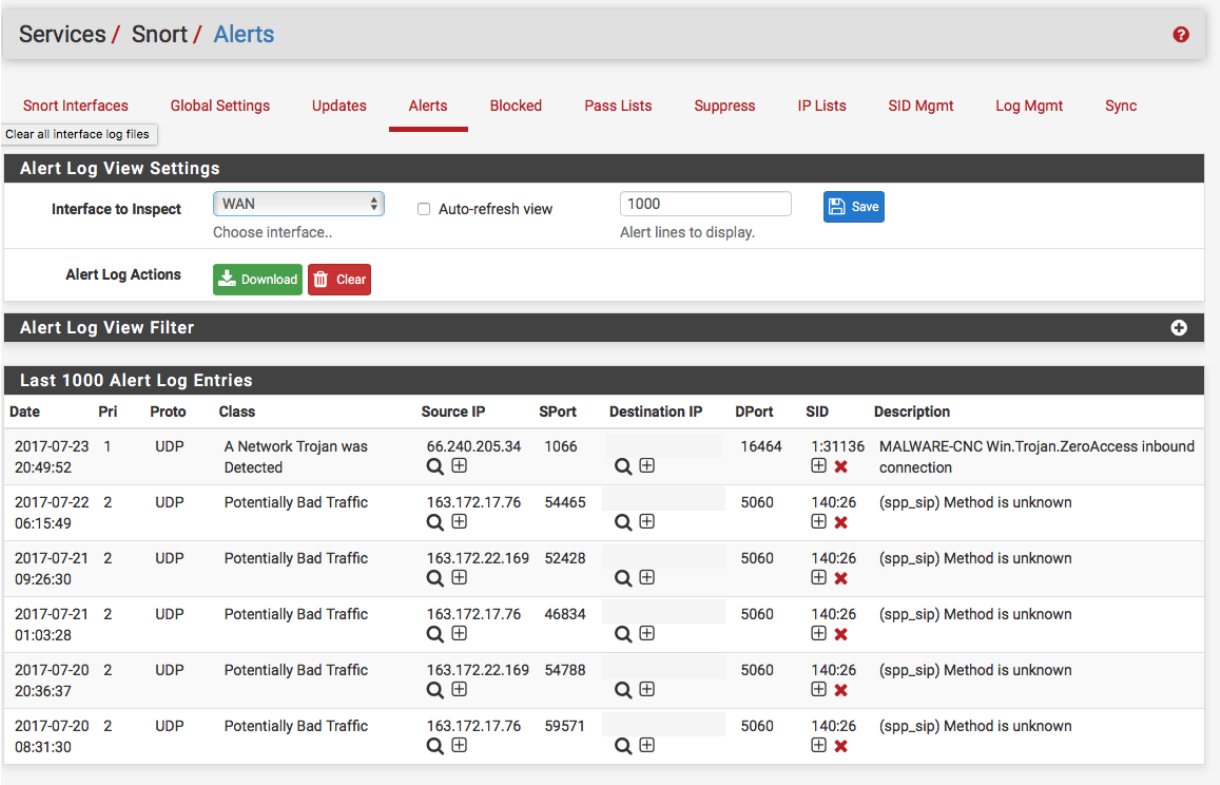

Рисунок 2.3 – IDS Snort

Suricata – це сучасна система виявлення та запобігання вторгненням (IDS/IPS) із відкритим вихідним кодом, яка була розроблена фондом OISF як наступник традиційних сигнатурних систем. Головною архітектурною особливістю Suricata є підтримка багатопотоковості, що дозволяє ефективно розподіляти навантаження між усіма ядрами процесора та забезпечувати стабільну роботу в мережах із пропускною здатністю 10 Гбіт/с і вище.

На відміну від класичних рішень, Suricata має ряд розширених можливостей:

- вбудовані парсери для протоколів прикладного рівня (HTTP, DNS, TLS, SMB);
- система підтримує роботу не лише із сигнатурами, а й зі складними логічними правилами;

- Suricata здатна самостійно ідентифікувати протокол у потоці даних, незалежно від порту.

Використання Suricata в інфраструктурі дозволяє поєднувати традиційну сигнатурну детекцію з елементами поведінкового аналізу. Проте, як і інші IDS, вона потребує регулярного оновлення баз правил і демонструє обмежену ефективність проти нових, раніше не задокументованих типів атак, що підкреслює необхідність впровадження адаптивних моделей на базі машинного навчання. На рисунку 2.4 зображено інтерфейс Suricata.

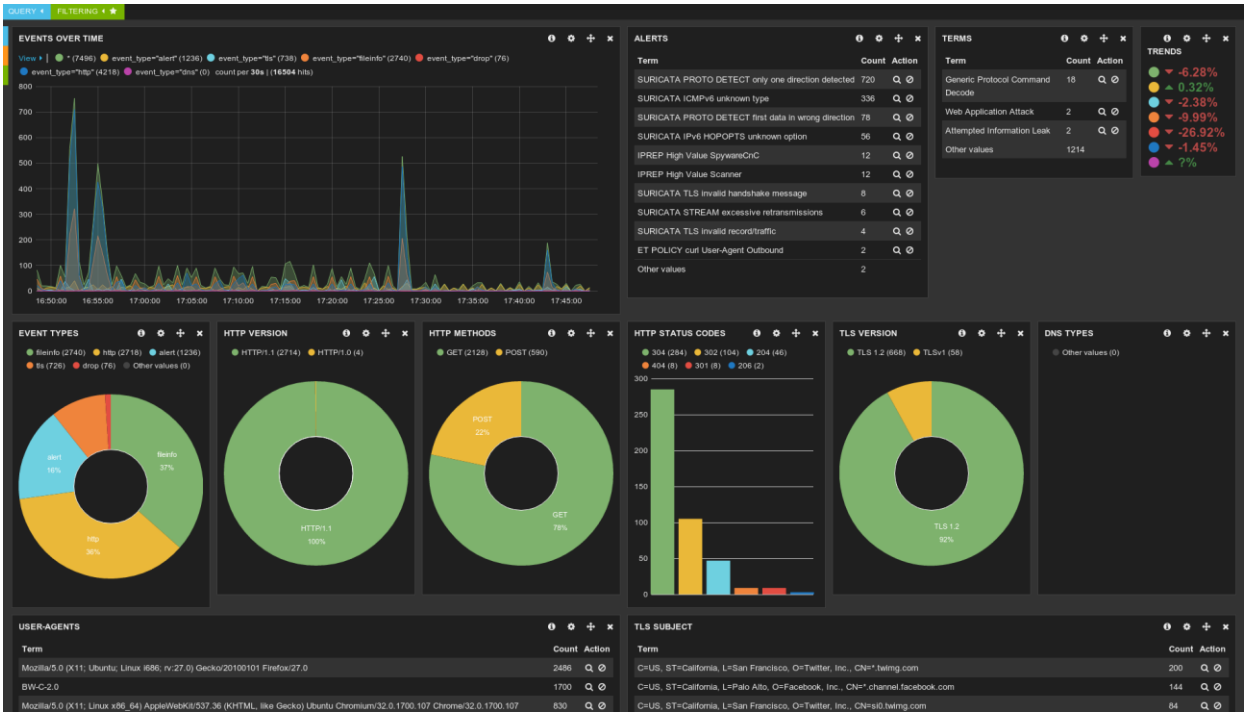

Рисунок 2.4 – Інтерфейс Suricata

Окреме місце в ієрархії засобів мережевого моніторингу посідає система Zeek, яка за своєю архітектурою відрізняється від класичних сигнатурних IDS. У той час як Snort або Suricata фокусуються на негайному виявленні атак через зіставлення пакетів із відомими шаблонами, Zeek функціонує як потужна платформа для глибокого аналізу мережевого контексту [36]. Вона перетворює сирий потік пакетів у структуровані журнали транзакцій, що описують активність на прикладному рівні. Це дозволяє адміністратору безпеки бачити не просто факт з’єднання, а деталізовані дані: назви запитуваних доменів у DNS,

заголовки HTTP-запитів, метадані сертифікатів шифрування та навіть структуру файлів, що передаються мережею.

Функціональна перевага Zeek полягає у використанні власної скриптової мови, орієнтованої на події, що дозволяє реалізовувати складну логіку виявлення аномалій, недоступну для звичайних сигнатурних правил. Система дозволяє відстежувати стан з'єднань протягом тривалого часу, що є критично важливим для ідентифікації прихованих методів передачі даних або повільного сканування мережі. Крім того, Zeek виконує роль джерела високоякісних метаданих для систем класу SIEM, оскільки надає розширений контекст інциденту, необхідний для проведення цифрової форензики та ретроспективного розслідування складних кіберзагроз. На рисунку 2.5 зображено інтерфейс IDS Zeek.

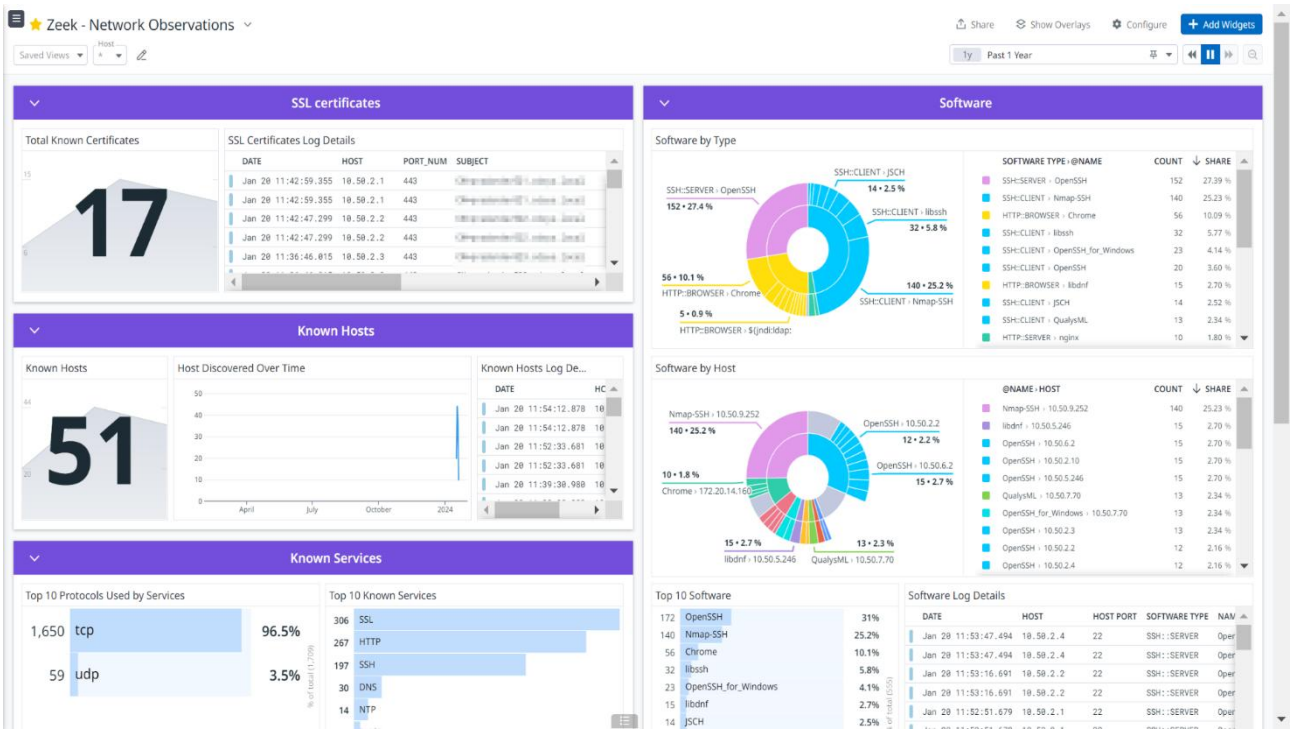

Рисунок 2.5 – IDS Zeek

Проведено порівняльний аналіз технічних характеристик розглянутих систем. Результати зіставлення наведено у таблиці 2.1.

Таблиця 2.1 – Порівняння характеристик IDS

Характеристика	**Snort**	**Suricata**	**Zeek (Bro)**
Основна роль	Сигнатурна IDS/IPS	Багатопотокова IDS/IPS	Аналізатор мережевих подій
Метод аналізу	Пошук за шаблоном	Шаблони + Глибокий аналіз	Скриптова логіка + Метадані
Блокування	Так	Так	Ні
Деталізація логів	Базова	Середня	Надзвичайно висока
Фокус використання	Периметровий захист	Високонавантажені канали	Розслідування та форензика

Результати порівняльного аналізу свідчать, що кожне з розглянутих рішень має свою специфічну сферу застосування, проте жодне з них самостійно не забезпечує повного циклу захисту. Якщо Snort та Suricata орієнтовані на оперативне реагування за відомими шаблонами атак, то Zeek виступає як інструмент глибокої ретроспективної аналітики та збору метаданих.

Системи класу SIEM призначені для забезпечення цілісного моніторингу стану безпеки інформаційної системи шляхом агрегації та кореляції подій з різнорідних джерел. На відміну від локальних засобів захисту, які оперують даними лише в межах одного сегмента або вузла, SIEM збирає логи з маршрутизаторів, міжмережевих екранів, антивірусних комплексів, серверів та систем виявлення вторгнень. На рисунку 2.6 зображено принцип роботи SIEM.

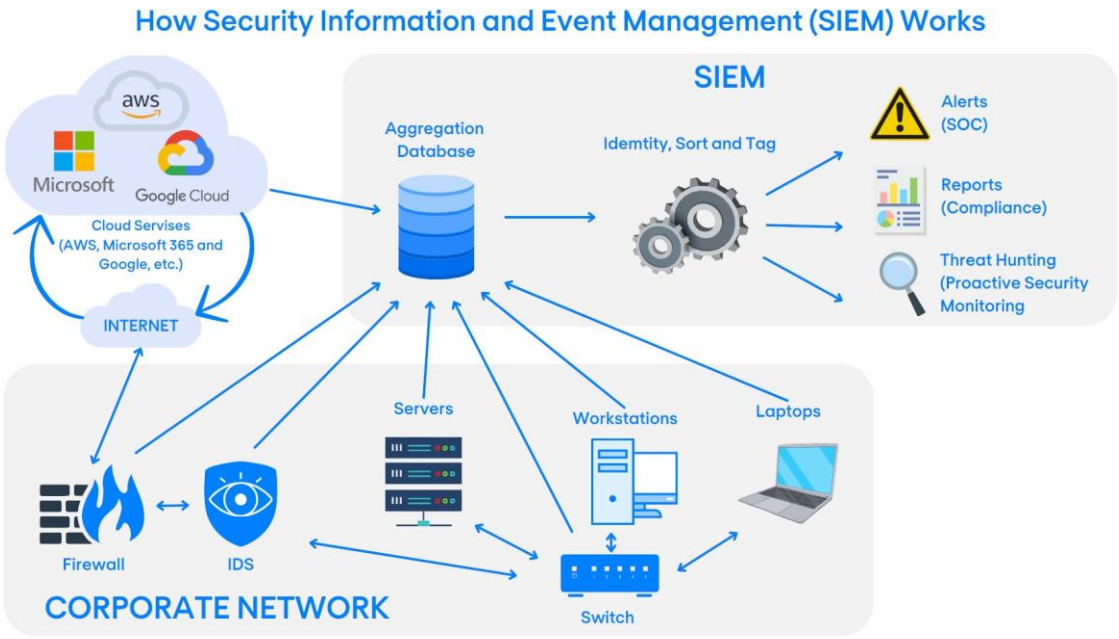

Рисунок 2.6 – Принцип роботи SIEM

Основною цінністю таких систем є здатність виявляти складні ланцюжки атак, що рознесені у часі та просторі. Наприклад, серія невдалих спроб входу на одному сервері та запуск нетипового процесу на іншому можуть розцінюватися як окремі незначні події, проте SIEM автоматично пов'язує їх у єдиний інцидент, що свідчить про спробу несанкціонованого просування зловмисника всередині мережі.

Функціональний цикл роботи SIEM-системи включає кілька ключових етапів: збір даних через спеціальні агенти або мережеві протоколи , нормалізацію отриманої інформації до єдиного формату та її довгострокове зберігання для ретроспективного аналізу. Проте найважливішим компонентом є двигун кореляції, який у реальному часі перевіряє потік подій на відповідність заданим правилам безпеки. Сучасні системи також активно впроваджують модулі поведінкової аналітики користувачів, що дозволяє ідентифікувати аномалії не лише за жорсткими правилами, а й через виявлення відхилень від типового профілю роботи кожного окремого облікового запису. Це забезпечує можливість раннього виявлення компрометації облікових даних або дій інсайдерів, що є однією з найскладніших задач у сучасній кібербезпеці.

Splunk Enterprise Security представляє собою платформу класу SIEM, яка призначена для агрегації, кореляції та аналізу великих масивів даних, що надходять з різних джерел IT-інфраструктури. На відміну від локальних систем виявлення вторгнень, Splunk забезпечує комплексний моніторинг стану безпеки на рівні всього підприємства, дозволяючи ідентифікувати складні цілеспрямовані атаки.

Основними функціональними можливостями платформи є:

- збирання даних з мережевих пристроїв, серверів, антивірусних систем та IDS в єдине сховище;
- можливість створення логічних зв'язків між розрізненими подіями в мережі;
- використання статистичних методів для виявлення відхилень від типової поведінки вузлів або користувачів;
- автоматизація процесів реагування на інциденти та візуалізація життєвого циклу атаки.

На рисунку 2.7 зображено SIEM Splunk Enterprise Security.

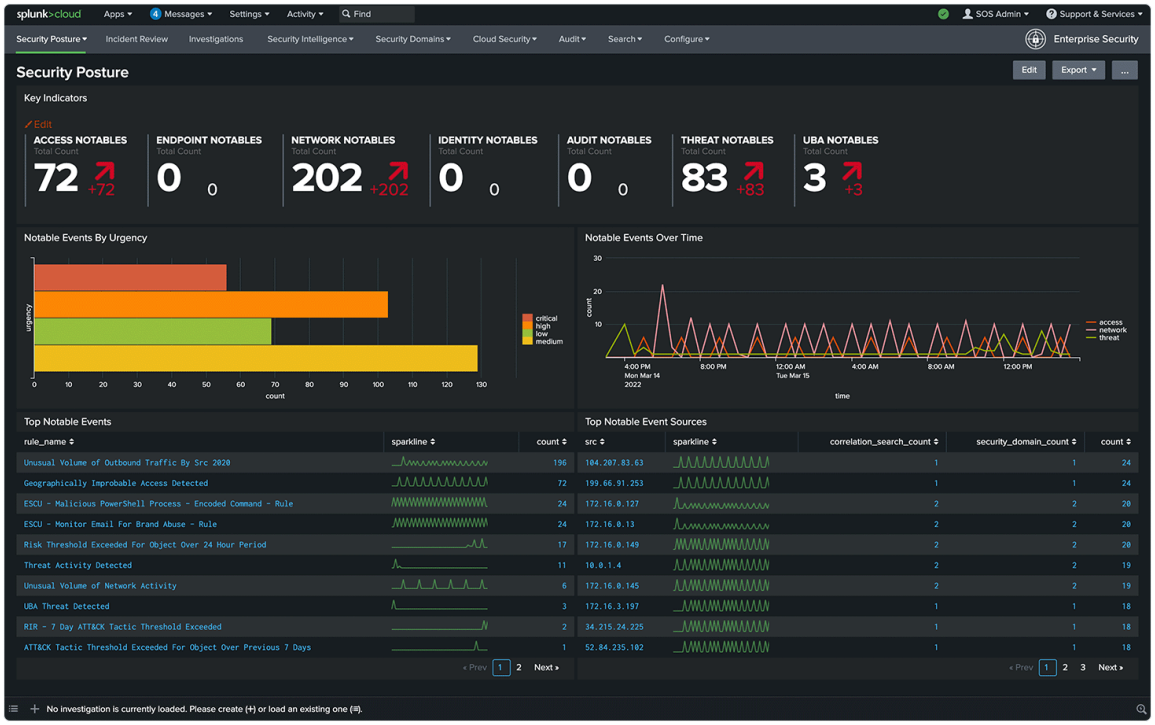

Рисунок 2.7 – Splunk Enterprise Security

Wazuh є платформою з відкритим вихідним кодом, призначеною для моніторингу безпеки, контролю цілісності та виявлення вторгнень на рівні хостів і хмарних інфраструктур. Вона базується на архітектурі OSSEC, проте суттєво розширює її можливості за рахунок інтеграції з Elastic Stack для візуалізації та аналізу великих масивів подій безпеки.

Основними функціональними напрямками Wazuh у контексті забезпечення кібербезпеки є:

- постійне сканування критичних системних файлів для виявлення несанкціонованих змін у їхньому вмісті, правах доступу або атрибутах;
- збір та автоматичний аналіз лог-файлів операційних систем і додатків у реальному часі;
- автоматичне виявлення застарілого ПЗ та відомих вразливостей на кінцевих точках, що дозволяє превентивно усувати вектори атак;
- можливість автоматичного виконання сценаріїв протидії загрозам.

На рисунку 2.8 зображено панель Wazuh.

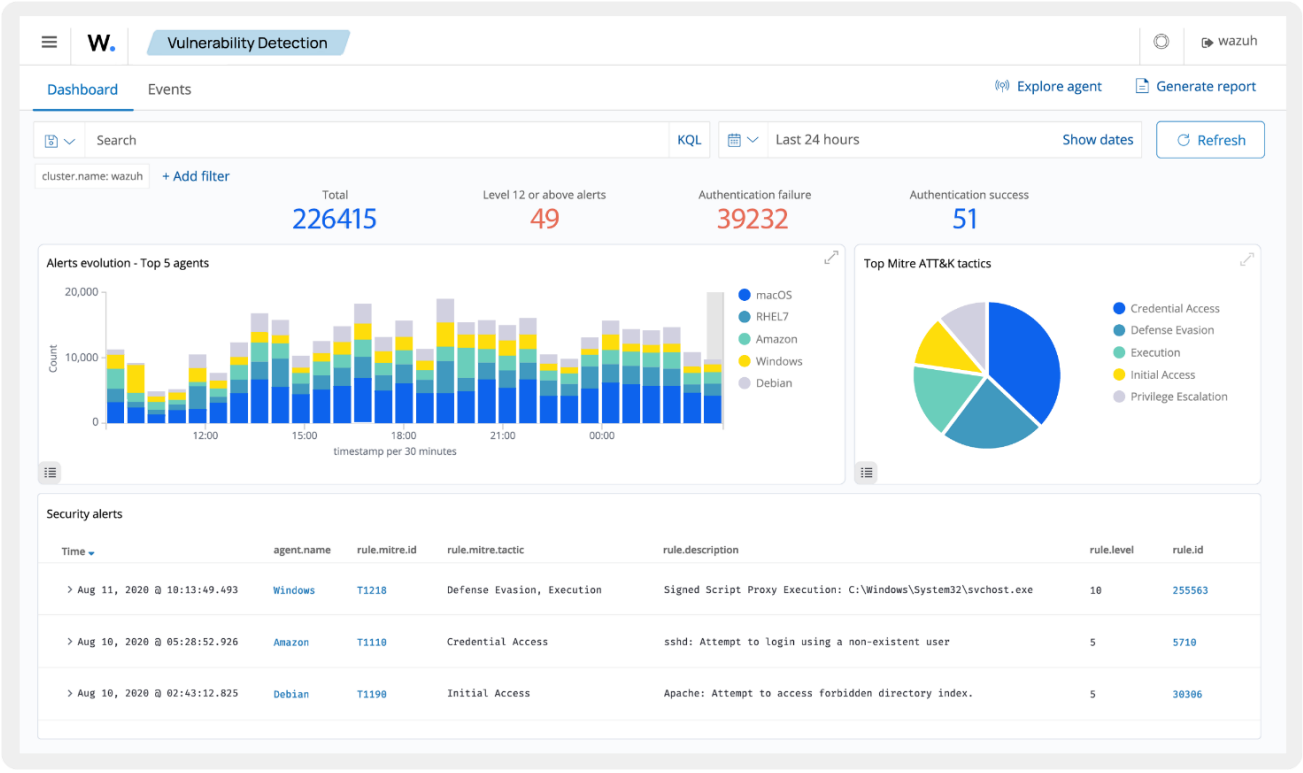

Рисунок 2.8 – Панель Wazuh

Особливістю Wazuh є здатність виступати як сполучна ланка між вузловим моніторингом та централізованою аналітикою.

Результати порівняльного аналізу найбільш поширених SIEM-рішень, що відрізняються підходом до обробки даних та вартістю впровадження наведено у таблиці 2.2.

Таблиця 2.2 – Порівняння характеристик SIEM

Характеристика	**Splunk Enterprise Security**	**Wazuh**
Призначення	Професійна аналітика	Моніторинг хостів та SIEM
Метод аналізу	Глибокий пошук та кореляція	Аналіз логів та цілісності файлів
Джерела даних	Будь-які (універсальні конектори)	Агенти на хостах + Syslog
Візуалізація	Складні аналітичні дашборди	Спеціалізований інтерфейс безпеки
Складність	Висока	Середня
Вартість	Висока	Безкоштовно

Аналіз продемонстрував, що Splunk є найбільш потужним інструментом, проте його висока вартість та закритість коду роблять його недоступним для багатьох дослідницьких проєктів. Wazuh пропонує вже готові набори правил для виявлення атак, вбудовані модулі контролю цілісності файлів та аналізу вразливостей.

Традиційним рівнем захисту хостів є антивірусне програмне забезпечення, основна задача якого полягає у виявленні та блокуванні шкідливих файлів до моменту їх виконання. Сучасні антивірусні комплекси еволюціонували від простого порівняння сигнатур до використання евристичних методів та хмарної репутаційної аналітики [37]. Це дозволяє ідентифікувати не лише відомі віруси, а й нові модифікації програм-вимагачів або шпигунського ПЗ. Проте класичні антивіруси часто виявляються неефективними проти складних атак, що не

використовують файли або експлуатують легітимні системні інструменти, такі як PowerShell, для виконання шкідливого коду безпосередньо в оперативній пам’яті.

Для протидії таким просунутим загрозам було розроблено клас систем виявлення та реагування на кінцевих точках (EDR – Endpoint Detection and Response). На відміну від антивірусів, EDR фокусується не на статичних файлах, а на безперервному моніторингу поведінки всіх процесів у системі [38]. Такі інструменти фіксують створення мережевих з’єднань, зміну ключів реєстру, ін’єкції в пам’ять та маніпуляції з правами доступу. Головною перевагою EDR є можливість візуалізації повного ланцюжка атаки, що дозволяє адміністратору безпеки побачити, як саме зловмисник потрапив у систему та які дії він встиг виконати. Крім того, ці системи надають інструменти активного реагування, включаючи дистанційну ізоляцію скомпрометованого хоста від мережі або відкат шкідливих змін у файловій системі.

Одним із найбільш актуальних рішень для реалізації функцій EDR у проєктах з відкритим кодом є інтеграція агента Wazuh із модулем Sysmon. Таке поєднання дозволяє збирати деталізовану інформацію про активність процесів у середовищі Windows та корелювати її з правилами виявлення атак. Для систем на базі Linux подібний функціонал реалізується через підсистему Auditd або eBPF, що забезпечує глибоку видимість подій на рівні ядра операційної системи. Використання технологій EDR у складі розроблюваної системи захисту дозволяє значно підвищити стійкість інфраструктури до сучасних цілеспрямованих атак, які здатні обходити традиційні периметрові засоби контролю. Порівняння можливостей засобів захисту хостів наведено в таблиці 2.3.

Таблиця 2.3 – Порівняння характеристик засобів захисту хостів

Характеристика	**Антивірус**	**Система EDR**
Об’єкт моніторингу	Файли та шкідливі коди	Поведінка процесів та події ОС
Метод виявлення	Сигнатури та евристика	Поведінковий аналіз та кореляція
Реакція на загрозу	Видалення або карантин файлу	Ізоляція хоста, блокування процесів
Фокус захисту	Запобігання зараженню	Виявлення та розслідування інцидентів
Складність	Низька	Висока

Отже, порівняльний аналіз свідчить, що традиційні антивіруси спрямовані на автоматизоване запобігання відомим загрозам на рівні файлів, тоді як системи EDR забезпечують глибокий поведінковий моніторинг та можливості для розслідування складних інцидентів у реальному часі.

2.2 Характеристика ансамблевих алгоритмів машинного навчання

Для розв'язання задачі виявлення аномалій у мережевому трафіку, яка в межах даного дослідження розглядається як задача бінарної або багатокласової класифікації, обрано алгоритм XGBoost. Вибір зумовлений високою обчислювальною ефективністю та здатністю моделі працювати з великими масивами табличних даних, характерних для сучасних систем виявлення вторгнень. XGBoost є вдосконаленою реалізацією алгоритму градієнтного бустингу на деревах рішень [39]. Основна ідея методу полягає в ітеративному побудуванні ансамблю дерев рішень, де кожне наступне дерево навчається на помилках попередньої сукупності моделей. На рисунку 2.9 зображено принцип роботи алгоритму XGBoost.

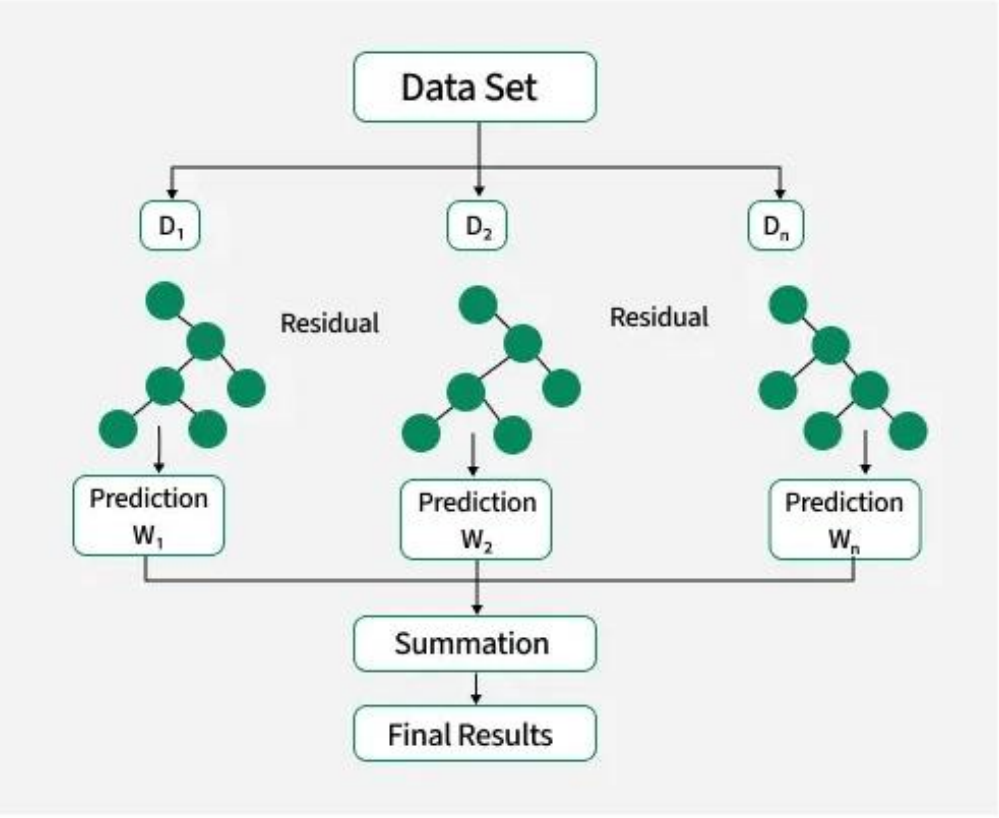

Рисунок 2.9 – Алгоритм XGBoost

Вибір алгоритму XGBoost зумовлений низкою його технічних переваг, серед яких ключове місце посідає використання механізмів регуляризації. На відміну від стандартного градієнтного бустингу, цей підхід включає вбудовані методи контролю складності моделі, що є критично важливим аспектом при аналізі мережевого трафіку, який зазвичай характеризується значним рівнем специфічного «шуму». Окрім цього, алгоритм демонструє високу ефективність під час роботи з незбалансованими даними, оскільки мережеві атаки у реальному потоці складають лише незначну частку від загального обсягу даних. Можливість гнучкого налаштування параметрів дозволяє штучно підвищити чутливість моделі до рідкісних аномальних подій, успішно вирішуючи проблему дисбалансу класів.

Важливою функціональною особливістю XGBoost є також вбудований інструментарій для оцінки важливості ознак. Завдяки цьому дослідник може кількісно визначити вплив кожного окремого параметра трафіку, як-от тривалість сесії чи кількість специфічних TCP-прапорців, на фінальне рішення моделі. Такий аналіз дає змогу оптимізувати архітектуру системи виявлення

вторгнень шляхом відкидання малоінформативних ознак, що суттєво прискорює процес обробки даних у реальному часі.

Алгоритм Random Forest представляє собою один із найбільш стійких методів ансамблевого навчання, що базується на поєднанні великої кількості незалежних дерев рішень для формування єдиного прогнозного результату. У системах виявлення вторгнень цей алгоритм виступає надійним інструментом класифікації трафіку завдяки своїй здатності ефективно протидіяти проблемі перенавчання, яка є критичною при аналізі великих обсягів мережевих даних.

Фундаментальний механізм роботи Random Forest базується на принципі агрегування результатів багатьох слабких класифікаторів за допомогою методу беггінгу. Під час фази навчання кожне дерево будується на випадковій вибірці даних, що дозволяє моделі охопити різні варіації мережевої активності та забезпечити високу узагальнюючу здатність. Важливою особливістю алгоритму є використання методу випадкових ознак, згідно з яким при кожному розщепленні вузла дерева обирається найкращий параметр лише з обмеженої випадкової підмножини всіх доступних характеристик трафіку. Це гарантує низьку кореляцію між окремими деревами в лісі, що безпосередньо підвищує точність ідентифікації аномальних пакетів у загальному потоці [40]. На рисунку 2.10 зображено принцип алгоритму Random Forest.

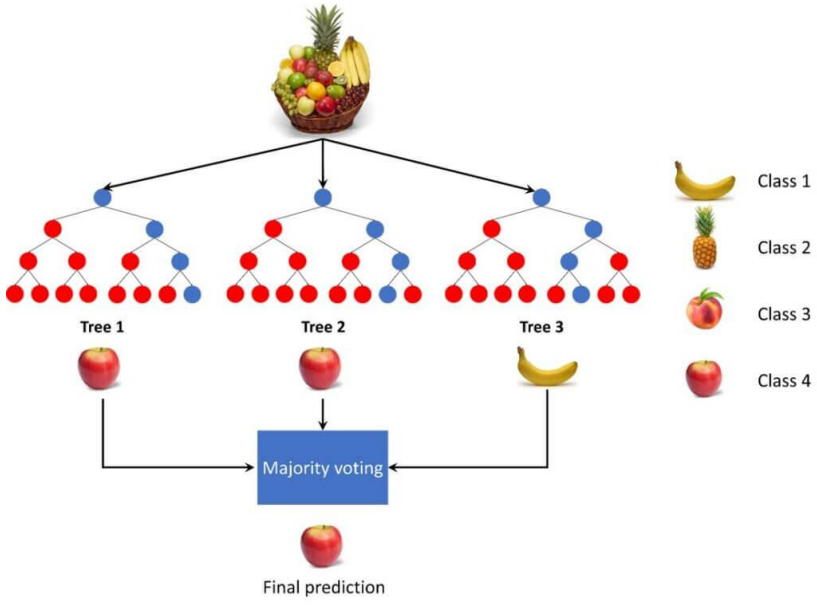

Рисунок 2.10 – Принцип Random Forest

Математично процес прийняття рішення в Random Forest описується як процедура мажоритарного голосування, де кожне дерево видає свій клас для заданого вектора ознак, а фінальний висновок системи відповідає класу, який набрав найбільшу кількість голосів. Такий підхід дозволяє нівелювати помилки окремих дерев та забезпечує стабільність моделі навіть за наявності значного обсягу шуму або аномальних викидів у вхідних даних, що характерно для інтенсивних мережевих атак.

Окрім високої точності класифікації, Random Forest надає можливість кількісної оцінки важливості ознак, що дозволяє аналітику визначити найбільш інформативні параметри мережевих протоколів, які свідчать про наявність шкідливою активності. Використання даного алгоритму забезпечує ефективне розрізнення нормальної поведінки користувачів від складних цілеспрямованих атак, гарантуючи при цьому прозорість та інтерпретованість отриманих результатів, що є важливою перевагою при проєктуванні інтелектуальних систем захисту інформації.

2.3 Порівняльна характеристика наборів даних

Якість роботи будь-якої системи виявлення вторгнень, що базується на машинному навчанні, фундаментально залежить від репрезентативності та актуальності навчальної вибірки. Процес формування еталонних наборів даних для IDS пройшов довгий шлях розвитку, відображаючи трансформацію методів мережевих атак та архітектур побудови інформаційних систем. Історично першим значущим кроком у створенні загальнодоступного середовища для тестування IDS став набір даних DARPA 98, розроблений лабораторією MIT Lincoln. На його базі згодом було сформовано KDD Cup 99, який протягом майже двох десятиліть залишався найбільш цитованим датасетом у науковій спільноті. Проте, попри свою історичну важливість, KDD Cup 99 має низку критичних недоліків, які сьогодні роблять його непридатним для реальних досліджень [41]. До них відносяться величезна кількість дублікатів та застарілий характер

трафіку, який не містить сучасних протоколів, таких як HTTPS, або сучасних типів атак, як-от фішинг чи ботнети [42].

Спробою усунути статистичні дефекти попередника став набір даних NSL-KDD. Його розробники провели ретельну роботу з видалення надлишкових записів та балансування вибірки, що дозволило моделям краще розпізнавати рідкісні атаки. Проте NSL-KDD успадкував головну проблему – самі дані все ще базувалися на застарілій структурі мереж кінця 90-х років. Це зумовило появу другого покоління наборів, таких як ISCX 2012 та UNSW-NB15. Вони запропонували більш сучасні профілі активності користувачів, проте часто мали проблеми з коректністю розмітки трафіку або недостатньою різноманітністю сценаріїв вторгнень [43].

Набори даних CIC-IDS2017 та особливо CSE-CIC-IDS2018 стали новим стандартом де-факто для досліджень у сфері мережевої безпеки. Вибір CSE-CIC-IDS2018 для даної роботи обґрунтований його унікальною методологією збору. Дані були отримані шляхом моделювання роботи великої корпоративної мережі в інфраструктурі AWS, що включала сотні вузлів на базі різних операційних систем. Методологія збору базувалася на створенні профілів поведінки користувачів за допомогою протоколу M-Profile, що дозволило згенерувати максимально реалістичний фоновий трафік, на який згодом накладалися сценарії атак.

Склад залучених мережевих протоколів у CSE-CIC-IDS2018 охоплює не лише базові TCP/UDP, а й прикладні протоколи HTTP, HTTPS, SSH, FTP, що є критичним для виявлення сучасних веб-атак та несанкціонованого віддаленого доступу. Особливу увагу в наборі приділено реалістичності відтворених сценаріїв атак. Зокрема, сценарій Botnet включає активність заражених вузлів, що намагаються встановити зв'язок із командними серверами, що дозволяє моделям навчатися на складних часових паттернах зв'язку. Атаки класу DDoS представлені різними векторами (LOIC, HOIC), що демонструють аномальне зростання інтенсивності трафіку. Сценарій Infiltration імітує реальне проникнення в мережу через експлуатацію вразливостей програмного забезпечення з подальшим скануванням портів та розповсюдженням по

інфраструктурі, що є викликом для алгоритмів через подібність такої активності до дій системних адміністраторів.

На відміну від застарілих пакетно-орієнтованих наборів, CSE-CIC-IDS2018 надає дані у форматі потоків, кожен з яких описується 80 ознаками, витягнутими за допомогою інструменту CICFlowMeter. Ці ознаки включають тривалість потоку, кількість пакетів, середній розмір заголовків, а також статистичні параметри інтервалів прибуття пакетів (IAT), що дозволяє проводити глибокий аналіз динаміки трафіку. Така деталізація та сучасність сценаріїв дозволяють використовувати цей набір для побудови моделей, що будуть стійкими до реальних кіберзагроз, а не лише до лабораторних прикладів. У таблиці 2.4 наведено порівняльну характеристику еталонних наборів даних мережевого трафіку.

Таблиця 2.4 – Порівняльна характеристика еталонних наборів даних мережевого трафіку

Критерій порівняння	**KDD Cup 99**	**NSL-KDD**	**UNSW-NB15**	**CSE-CIC-IDS2018**
Рік створення	1999	2009	2015	2018
Тип трафіку	Штучно згенерований	Модифікований KDD	Реальний та симульований	Реалістичний
Кількість ознак	41	41	49	80
Кількість записів	~4.9 млн	~150 тис.	~2.5 млн	~16 млн
Формат даних	Пакетно-орієнтований	Пакетно-орієнтований	Потоковий	Потоковий
Наявність дублікатів	Висока	Відсутні	Низька	Мінімальна
Сучасні атаки	Відсутні	Відсутні	Частково	Повний спектр
Актуальність	Застарілий	Обмежена	Середня	Висока

Аналіз порівняльних характеристик показує, що набір даних CSE-CIC-IDS2018 є найбільш актуальним та репрезентативним для дослідження, оскільки він містить найбільшу кількість ознак і записів, базується на реалістичному потоковому трафіку та охоплює повний спектр сучасних кібератак. На противагу йому, старіші набори, як-от KDD Cup 99 та NSL-KDD, є застарілими через відсутність актуальних загроз та обмеженість форматів даних.

2.4 Математична формалізація процесів обробки, нормалізації та балансування даних

Процес інтелектуального аналізу мережевого трафіку починається з етапу первинної обробки, метою якого є усунення технічних артефактів та аномалій, що виникають під час збору та генерації статистичних ознак. Набір даних CSE-CIC-IDS2018, попри свою високу якість, містить значну кількість записів із некоректними значеннями, що є результатом помилок при розрахунку динамічних параметрів потоку інструментом CICFlowMeter. Формалізація цього процесу передбачає виявлення та обробку двох основних типів дефектів: відсутніх значень та математичних нескінченностей.

Виникнення значень типу нескінченності зазвичай пов’язане з розрахунком швидкості передачі даних. Математично це пояснюється діленням об’єму переданої інформації на тривалість потоку, яка у випадку надшвидких з’єднань або помилок таймінгу може дорівнювати нулю. Оскільки більшість алгоритмів машинного навчання, включаючи нейронні мережі та методи на основі градієнтного спуску, оперують дійсними числами у скінченних межах, наявність нескінченності призводить до розбіжності ваг моделі та виходу системи з ладу [44].

Математичне обґрунтування процесу фільтрації також включає аналіз впливу очищення на обсяг вибірки. Видалення некоректних записів у датасеті CSE-CIC-IDS2018 може призвести до втрати певної кількості даних, що вимагає перевірки збереження балансу між класами [45]. Якщо більшість видалених записів належали до малопредставленого класу атак, це може погіршити здатність моделі до навчання. Таким чином, первинна фільтрація – це не просто технічне видалення помилок, а стратегічне балансування між повнотою даних та математичною стабільністю алгоритмів класифікації.

Етап очищення завершується дедуплікацією даних. Наявність ідентичних записів у навчальній та тестовій вибірках призводить до ефекту «запам'ятовування» замість узагальнення, що штучно завищує метрики точності. Тільки після успішного проходження всіх етапів очищення – видалення

некоректних значень та дедуплікації – масив даних вважається математично придатним для наступного кроку: нормалізації та масштабування ознак.

Після завершення етапу очищення даних від технічних аномалій постає задача узгодження масштабів 80 ознак, що складають вектор стану мережевого потоку. У наборі даних CSE-CIC-IDS2018 спостерігається значний розрив у діапазонах значень: наприклад, ознака «Flow Duration» може набувати значень у межах десятків мільйонів мікросекунд, тоді як «Total Fwd Packets» може вимірюватися одиницями. Для більшості алгоритмів машинного навчання, зокрема тих, що використовують евклідову відстань або градієнтні методи оптимізації, така невідповідність є критичною. Ознаки з більшими абсолютними значеннями будуть домінувати при розрахунку функції втрат, фактично нівелюючи вплив параметрів з малими значеннями, які можуть бути навіть більш інформативними для детекції атаки.

Теоретичне обґрунтування необхідності масштабування полягає в аналізі поверхні функції втрат. У випадку ненормованих ознак лінії рівня функції втрат мають вигляд сильно витягнутих еліпсів. Це призводить до того, що градієнтний спуск починає коливатися навколо оптимального шляху, що суттєво сповільнює швидкість збіжності або взагалі призводить до розбіжності алгоритму [46]. Після проведення нормалізації поверхня стає більш сферичною, що дозволяє кроку градієнта бути більш спрямованим до глобального мінімуму, забезпечуючи стабільність навчання нейронних мереж.

Метод Min-Max Scaling цей підхід лінійно трансформує дані у фіксований діапазон, зазвичай [0,1]. Метод ефективний для алгоритмів, що не роблять припущень про розподіл даних. Проте його головним недоліком є висока чутливість до викидів: якщо в трафіку є поодинокий аномальний сплеск, усі інші значення будуть «стиснуті» в дуже вузький діапазон, що призведе до втрати інформативності.

Z-оцінка або стандартизація передбачає центрування даних навколо нуля з одиничним стандартним відхиленням. На відміну від Min-Max, цей метод є більш стійким до помірних викидів і забезпечує кращу роботу лінійних моделей та алгоритмів на основі ядер. Він базується на припущенні про нормальний

розподіл ознак, що часто зустрічається в мережевих параметрах завдяки центральній граничній теоремі.

Robust Scaling враховуючи специфіку мережевих атак, які часто проявляються як екстремальні викиди (наприклад, раптовий сплеск кількості пакетів при DoS), метод Robust Scaler є найбільш доречним. Він використовує медіану та інтерквартильний розмах. Оскільки медіана та квартилі не залежать від екстремальних значень так сильно, як середнє та дисперсія, цей метод дозволяє зберегти структуру аномалій, одночасно привівши основну масу даних до єдиного масштабу.

Вибір конкретного методу для системи виявлення вторгнень залежить від обраної архітектури класифікатора [47]. Для глибоких нейронних мереж часто використовують комбінацію: спочатку Robust Scaling для нівелювання впливу екстремальних шумів, а потім стандартизацію для прискорення навчання.

Однією з найбільш складних проблем при розробці інтелектуальних систем виявлення вторгнень є значний числовий дисбаланс між класами легітимного та шкідливого трафіку. У реальних корпоративних мережах частка аномальної активності зазвичай складає лише малу частку відсотка від загального об’єму переданих даних. Датасет CSE-CIC-IDS2018 відображає цю закономірність: у той час як кількість записів класу «Benign» обчислюється мільйонами, окремі типи атак, такі як «Infiltration» або «Heartbleed», представлені лише кількома сотнями або тисячами прикладів. Для алгоритмів машинного навчання такий розподіл створює ефект «упередженості мажоритарного класу». Модель, намагаючись мінімізувати глобальну функцію втрат, схиляється до стратегії ігнорування міноритарного класу [48]. У результаті система може демонструвати загальну точність понад 80%, проте при цьому не ідентифікувати жодної реальної атаки, оскільки помилка на рідкісних прикладах майже не впливає на підсумковий статистичний показник.

Теоретичне обґрунтування необхідності балансування вибірки полягає в коригуванні меж прийняття рішень. Коли даних одного класу занадто багато, класифікатор «розширює» простір цього класу, поглинаючи області, де можуть

знаходитися рідкісні атаки. Для вирішення цієї проблеми застосовуються два основні підходи: методи зменшення вибірки та методи збільшення вибірки.

Випадкове зменшення мажоритарного класу полягає у вилученні частини записів легітимного трафіку для досягнення паритету з класом атак. Перевагою цього методу є суттєве скорочення часу навчання моделі за рахунок зменшення загального обсягу даних. Однак головним ризиком є втрата потенційно важливої інформації: видаляючи записи «Benign», ми можемо випадково прибрати приклади нормальної поведінки, які є критично схожими на атаки, що призведе до зростання кількості хибних тривог.

Алгоритм синтетичної генерації даних SMOTE (Synthetic Minority Over-sampling Technique) пропонує більш інтелектуальний підхід до збільшення кількості прикладів міноритарного класу. На відміну від простого копіювання існуючих записів, яке призводить до перенавчання, SMOTE створює нові, синтетичні екземпляри атак [49]. Математично це реалізується шляхом пошуку k найближчих сусідів для кожного прикладу рідкісного класу в просторі ознак. Новий об’єкт створюється як випадкова точка на відрізку, що з’єднує обраний приклад з одним із його сусідів. Це дозволяє моделі «заповнити» прогалини в просторі ознак атак, роблячи межу між нормою та аномалією більш чіткою та стійкою до варіацій.

Застосування методів балансування дозволяє змістити фокус навчання моделі з простої оптимізації точності на максимізацію здатності виявляти загрози. Це критично важливо для таких сценаріїв, як «Infiltration», де атака розвивається повільно і вкрай схожа на дії легітимного користувача. Без попереднього балансування за допомогою SMOTE або комбінованих методів (наприклад, SMOTE+Tomek links), інтелектуальна система залишатиметься «сліпою» до найбільш небезпечних, але малочисельних загроз, що робить цей етап підготовки ознак обов’язковим компонентом проектування надійних IDS.

2.5 Стратегії вибору ознак, сегментації аномалій та вибір метрик ефективності

У сучасних системах виявлення вторгнень, що працюють із датасетами високої розмірності, виникає явище, відоме в машинному навчанні як «прокляття розмірності». Теоретична суть цієї проблеми полягає в тому, що зі зростанням кількості ознак об’єм простору зростає експоненціально, внаслідок чого наявні дані стають розрідженими. Для алгоритмів класифікації це означає, що статистична значущість спостережень падає, а ризик перенавчання моделі суттєво зростає. Модель починає знаходити помилкові закономірності в шумі, замість того, щоб виявляти реальні ознаки атак [50]. Крім того, надлишковість даних створює величезне обчислювальне навантаження, що є неприпустимим для систем, які мають працювати в режимі реального часу.

Першочерговим етапом оптимізації простору ознак є вилучення неінформативних метаданих. До них належать ідентифікатори сесій, мітки часу та мережеві реквізити (IP-адреси джерела та призначення, номери портів). Хоча ці дані є критичними для розслідування інцидентів, для навчання узагальненої моделі вони є шкідливими. Якщо залишити IP-адреси у вибірці, модель може просто «запам'ятати», що атака йшла з конкретної адреси, замість того, щоб вивчити поведінкові патерни самої атаки. Таке «навчання» робить систему нездатною розпізнати ту саму атаку, якщо вона буде здійснена з іншого вузла. Видалення метаданих дозволяє перемістити фокус класифікатора на статистичну динаміку трафіку – тривалість потоків, інтенсивність пакетів та варіативність часових інтервалів. Для підвищення прозорості прийняття рішень та обґрунтування результатів класифікації мережевих загроз доцільно спиратися на архітектурні підходи інформаційних систем прогнозування та інтерпретації, які дозволяють математично оцінити вагомість окремих ознак трафіку [51].

Для автоматизованого відбору найбільш інформативних ознак застосовуються три основні стратегії.

Методи фільтрації на основі статистичних показників, найбільш поширеним підходом є аналіз коефіцієнта кореляції Пірсона. Математично він

дозволяє виявити мультиколінеарність – ситуацію, коли дві ознаки сильно корелюють між собою. Якщо r близький до 1, одна з ознак є надлишковою і може бути видалена без втрати інформації. Крім того, застосовується аналіз ентропії за методом Шеннона, який дозволяє оцінити «кількість інформації», що міститься в ознаці. Ознаки з нульовою або низькою варіативністю (де значення майже не змінюються) мають низьку ентропію і є практично марними для розрізнення класів «норма» та «атака».

Методи «обгорток» одним із найпотужніших алгоритмів цієї групи є рекурсивне виключення ознак. Принцип його роботи полягає в ітеративному навчанні базової моделі (наприклад, випадкового лісу) та ранжуванні ознак за їхньою значущістю [52]. На кожному кроці найменш важлива ознака видаляється, і процес повторюється до досягнення оптимальної кількості параметрів. Це дозволяє врахувати взаємозв’язки між ознаками, які методи фільтрації могли б пропустити.

Вбудовані методи, ці методи реалізують відбір ознак безпосередньо під час процесу навчання. Прикладом є використання регуляризації L1, яка «штрафує» ваги неважливих ознак, прирівнюючи їх до нуля. Також популярним є аналіз «важливості ознак» в ансамблевих методах, де модель самостійно розраховує, наскільки сильно кожна ознака впливає на зменшення неоднорідності при розбитті вузлів дерева.

Впровадження стратегій зменшення розмірності дозволяє скоротити простір ознак CSE-CIC-IDS2018 з 80 до 20–25 найбільш релевантних параметрів. Це не лише підвищує швидкість роботи системи в кілька разів, а й робить її більш стійкою до нових варіацій атак, оскільки модель фокусується на найбільш стабільних характеристиках аномальної активності. Таким чином, інтелектуальний аналіз ознак є необхідною ланкою, що трансформує теоретичну модель у прикладний інструмент кіберзахисту.

У процесі інтелектуального аналізу мережевих ознак особливе місце посідає виявлення викидів – спостережень, які настільки суттєво відхиляються від загальної сукупності даних, що виникає припущення про їхню іншу природу. У контексті кібербезпеки такі викиди часто є прямим індикатором

цілеспрямованої атаки або серйозного технічного збою [53]. На відміну від класичної класифікації, методи виявлення викидів не потребують розмічених даних, що дозволяє використовувати їх як інструмент попередньої сегментації трафіку для виявлення невідомих раніше загроз.

Одним із найбільш ефективних алгоритмів для роботи з високовимірними мережевими даними є Isolation Forest. Математична концепція цього методу базується на припущенні, що аномалії є «нечисленними та відмінними». Алгоритм будує ансамбль випадкових дерев рішень, де для кожного спостереження розраховується кількість розбиттів, необхідних для відокремлення даного об’єкта від інших. Оскільки аномальні записи (наприклад, поодинокі спроби сканування портів або інфільтрації) мають специфічні значення ознак, вони «ізолюються» значно швидше – ближче до кореня дерева. Чим коротшим є шлях від кореня до листа, тим вищою є ймовірність того, що даний запис є викидом.

Іншим важливим методом є Local Outlier Factor (LOF), який базується на аналізі локальної щільності даних. Основна ідея LOF полягає в порівнянні локальної щільності об’єкта зі щільністю його k найближчих сусідів. Якщо щільність об’єкта суттєво нижча за щільність його оточення, це свідчить про те, що він знаходиться в розрідженій області простору ознак і є викидом. Це особливо корисно для виявлення складних, контекстуальних аномалій, де значення ознак самі по собі можуть бути в межах норми, але їхня комбінація в конкретному локальному просторі є нетиповою.

Попереднє виявлення викидів на етапі аналізу ознак виконує дві стратегічні функції. По-перше, воно дозволяє сегментувати трафік і виділити найбільш підозрілі сесії для подальшого детального аналізу або донавчання моделі [54]. По-друге, видалення статистичного «шуму» (викидів, що не є атаками, а технічними дефектами) дозволяє підготувати «чисту» навчальну вибірку, що суттєво зменшує кількість хибних спрацювань системи в майбутньому. Математична формалізація меж аномальності дозволяє гнучко налаштовувати чутливість IDS залежно від критичності сегмента мережі, що забезпечує баланс між безпекою та безперервністю бізнес-процесів.

Оцінювання ефективності систем виявлення вторгнень вимагає застосування комплексного математичного апарату, оскільки проста перевірка відсотка вгадувань не здатна відобразити реальну надійність системи в умовах кіберзагроз. Традиційна метрика точності Accuracy, яка розраховується як відношення кількості правильних передбачень до загальної кількості записів, у задачах IDS часто виявляється оманливою [55]. Це зумовлено раніше розглянутою проблемою дисбалансу класів: якщо в тестовій вибірці 99% трафіку є легітимним, то модель, яка просто завжди класифікує будь-який пакет як «Benign», отримає показник Accuracy 99%, попри повну неспроможність виявити хоча б одну реальну атаку. Таким чином, для об’єктивного аналізу необхідно використовувати метрики, що базуються на матриці помилок.

Матриця помилок формалізує результати класифікації за чотирма напрямками:

1. True Positive. Атака, яку система успішно розпізнала.
2. True Negative. Легітимний трафік, який система коректно пропустила.
3. False Positive. «Хибна тривога», коли нормальний трафік помилково прийнято за атаку.
4. False Negative. «Пропуск атаки», коли шкідлива активність була класифікована як норма.

У контексті інформаційної безпеки пріоритетним завданням є мінімізація False Negative, оскільки кожен пропущений інцидент може призвести до компрометації всієї інфраструктури. Для вимірювання здатності моделі виявляти загрози використовується метрика повноти Recall.

Високий рівень Recall гарантує, що більшість атак буде зафіксовано, проте це часто призводить до зростання кількості хибних тривог. Велика кількість FP виснажує ресурси аналітиків безпеки та може призвести до блокування легітимних бізнес-процесів. Тому важливо одночасно контролювати метрику Precision, яка показує частку реальних атак серед усіх спрацювань системи.

Для досягнення системного балансу між цими двома показниками використовується F1-Score – гармонійне середнє між Precision та Recall. На відміну від середнього арифметичного, F1-Score значно знижується, якщо хоча

б одна з метрик має низьке значення, що робить його еталонним показником якості при роботі з незбалансованими датасетами, такими як CSE-CIC-IDS2018.

Крім того, для глибокого аналізу ефективності ознак доцільно використовувати криві ROC-AUC та Precision-Recall Curve. Площа під ROC-кривою (AUC) демонструє здатність моделі розділяти класи при різних порогах чутливості [56]. Високий показник AUC свідчить про те, що обрані ознаки мають високу роздільну здатність і дозволяють системі ефективно диференціювати аномальну активність від фонового шуму. Такий системний підхід до оцінювання дозволяє не лише констатувати факт виявлення вторгнень, а й оптимізувати параметри попередньої обробки ознак для досягнення максимальної стійкості системи захисту.

2.6 Висновок до другого розділу

В другому розділі кваліфікаційної роботи досліджено архітектуру сучасних систем виявлення вторгнень (IDS/IPS) та засобів захисту кінцевих точок, а також проведено математичну формалізацію процесів підготовки даних, що заклало методичну основу для вибору ансамблевих алгоритмів машинного навчання як найбільш ефективних інструментів ідентифікації аномальної активності.

3 СТВОРЕННЯ МОДЕЛІ ТА ОЦІНКА ЕФЕКТИВНСТІ

3.1 Підготовка та попередня обробка набору даних

Процес підготовки розпочинається з вирішення проблеми критичної незбалансованості класів, де категорія нормального трафіку (Benign) становить понад 80% від загального обсягу в 13 мільйонів записів. На рисунку 3.1 зображено розподіл записів набору даних.

```
Label distribution:
Label
Benign                      11193663
DDOS attack-HOIC              686012
DDoS attacks-LOIC-HTTP        576191
DoS attacks-Hulk              461912
DoS attacks-SlowHTTPTest      139890
DoS attacks-GoldenEye          41508
DoS attacks-Slowloris          10990
DDOS attack-LOIC-UDP            1730
Brute Force -Web                 611
Brute Force -XSS                 230
SQL Injection                     87
Name: count, dtype: int64
```

Рисунок 3.1 – Розподіл значень набору даних

Для запобігання перенавчанню моделі на фоновому трафіку та забезпечення стабільної роботи алгоритму XGBoost застосовано стратегію керованого зменшення вибірки. З кожного файлу датасету відбирається обмежена кількість записів, що дозволяє вирівняти пропорції між легітимними транзакціями та різними типами атак, такими як DDoS, DoS та Brute Force. На рисунку 3.2 зображено результати після змінення вибірки деяких значень.

```
Final balanced distribution:
Label
DDOS attack-HOIC            100000
DoS attacks-Hulk            100000
DDoS attacks-LOIC-HTTP      100000
Normal                      100000
DoS attacks-SlowHTTPTest    100000
DoS attacks-GoldenEye        41508
DoS attacks-Slowloris        10990
Brute Force -XSS             10000
Brute Force -Web             10000
DDOS attack-LOIC-UDP         10000
SQL Injection                10000
Name: count, dtype: int64
```

Рисунок 3.2 – Зміна вибірки значень

Очищення даних включає видалення технічних ідентифікаторів (Timestamp, Source/Destination IP), які можуть спричинити помилкову кореляцію, а також обробку некоректних значень (Infinity та NaN), що виникають при розрахунку швидкості потоків. Усі числові ознаки проходять етап робастного масштабування за допомогою StandardScaler для приведення їх до єдиного діапазону значень.

3.2 Відбір ознак

Ключовим викликом при розробці системи виявлення аномалій є фундаментальна відмінність між структурами даних, що використовуються для навчання, та даними, які отримуються під час живого перехоплення трафіку. Використаний у роботі набір даних CIC-IDS2018 базується на ознаках потоків, що являють собою агреговану статистику за певну кількість пакетів, тоді як безпосереднє захоплення трафіку надає доступ лише до ознак окремих пакетів. Це створює ситуацію, коли модель навчається на високорівневих статистичних патернах, але у реальному часі стикається з розрізненими атрибутами одиничних одиниць інформації.

Для подолання цієї невідповідності було розроблено підхід, що базується на семантичному наближенні ознак. Було відібрано 17 ключових параметрів із набору CIC-IDS2018, які мають найбільш близьке логічне значення до характеристик, що містяться безпосередньо в мережевих пакетах. До переліку тренувальних ознак увійшли номери портів, ідентифікатори протоколів, довжини пакетів, набір TCP-прапорців (SYN, ACK, FIN, RST, PSH, URG), а також похідні показники, такі як інтенсивність пакетів та довжини заголовків, розмір корисного навантаження, час життя пакета (TTL), розмір вікна та порядкові номери пакетів.

Слід враховувати, що такий метод вибору ознак є апроксимацією, оскільки ідеального збігу між статистичним даними та окремим пакетом досягти неможливо. Модель у такому випадку фокусується на вивченні аномальних шаблонів у межах окремих транзакцій, а не довготривалих з’єднань. Для забезпечення коректної роботи системи на етапі розгортання впроваджується спеціальний шар відображення ознак, який трансформує вхідні дані з мережевого інтерфейсу у формат, сумісний із навченою моделлю XGBoost.

3.3 Модель XGBoost

Процес побудови інтелектуальної моделі виявлення аномалій базується на використанні алгоритму екстремального градієнтного бустингу (XGBoost), який реалізує ансамблевий підхід на основі дерев рішень. Через значний обсяг вихідного датасету CSE-CIC-IDS2018, що унеможливлює його повне завантаження в оперативну пам'ять, у роботі застосовано стратегію інкрементального навчання. Дані обробляються послідовними блоками, де з кожного файлу виокремлюється збалансована частина записів, що дозволяє моделі поступово адаптуватися до різних типів атак (DDoS, Brute Force, Web Attacks тощо) без втрати загальної прогностичної здатності. Для прискорення математичних обчислень та побудови гістограм розподілу ознак задіюється апаратне прискорення на базі графічних процесорів (архітектура CUDA), що критично важливо для обробки мільйонів транзакцій за прийнятний час.

Архітектура моделі налаштована для вирішення задачі багатокласової класифікації , що дозволяє отримувати ймовірнісний розподіл для кожної з одинадцяти категорій трафіку. Ключовим етапом оптимізації є підбір гіперпараметрів, які забезпечують баланс між точністю та здатністю до узагальнення. Максимальна глибина дерев встановлена на рівні 10 рівнів), що дозволяє моделі фіксувати складні нелінійні залежності між портами, довжиною пакетів та часовими інтервалами.

Для об'єктивного контролю якості безпосередньо під час навчання використовується окремий валідаційний набір даних, який складається з 10% записів від кожної ітераційної порції. Це дозволяє ідентифікувати ознаки перенавчання. Використання вбудованої регуляризації автоматично обмежує вплив окремих шумів у трафіку, гарантуючи, що модель виділяє лише стійкі патерни, характерні для реальних втручань. Отримана в результаті навчання модель зберігається у вигляді серіалізованого об'єкта, що містить оптимальні параметри ансамблю, готові до розгортання в модулі аналізу трафіку в реальному часі. На рисунку 3.3 зображено метрики ефективності моделі.

	precision	recall	f1-score	support
Brute Force -Web	0.84	0.75	0.80	2000
Brute Force -XSS	0.92	0.77	0.84	2000
DDOS attack-HOIC	0.98	1.00	0.99	20000
DDOS attack-LOIC-UDP	0.99	1.00	1.00	2000
DDoS attacks-LOIC-HTTP	1.00	1.00	1.00	20000
DoS attacks-GoldenEye	1.00	1.00	1.00	8302
DoS attacks-Hulk	1.00	1.00	1.00	20000
DoS attacks-SlowHTTPTest	1.00	1.00	1.00	20000
DoS attacks-Slowloris	1.00	1.00	1.00	2198
Normal	0.98	0.99	0.98	20000
SQL Injection	0.87	0.94	0.90	2000
accuracy			0.99	118500
macro avg	0.96	0.95	0.95	118500
weighted avg	0.99	0.99	0.99	118500

✓ Overall Accuracy: 0.9873

Рисунок 3.3 – Метрики ефективності моделі XGBoost

Аналіз результатів тестування розробленої моделі дозволяє об’єктивно оцінити її здатність до виявлення аномалій у складних мережевих середовищах. Загальний показник точності системи склав 98,73%, що є високим результатом

для класифікації на одинадцять категорій трафіку. Така ефективність досягається завдяки здатності алгоритму XGBoost чітко розрізняти статистичні відбитки різних типів атак навіть за умови використання обмеженого набору з сімнадцяти ознак. Отримані метрики прецизійності та повноти для більшості класів наближаються до одиниці, що підтверджує правильність обраної стратегії попередньої обробки та балансування даних.

Найвищу ефективність модель продемонструвала при ідентифікації масових загроз, таких як DDoS-атаки та різні варіації DoS-активності, де показники повноти досягли максимального значення. Це свідчить про те, що інтенсивні потоки пакетів мають настільки специфічні параметри тривалості та об’єму, що система виявляє їх без жодних пропусків. Аналогічно високі результати зафіксовані для ботнет-активності, що дозволяє стверджувати про надійність моделі у розпізнаванні скоординованих дій заражених вузлів. Легітимний трафік також класифікується з високою точністю, що є критично важливим для уникнення хибних спрацювань, які могли б порушити роботу реальної мережі.

Дещо складнішою виявилася ідентифікація веб-атак, зокрема спроб перебору паролів через HTTP та XSS-ін'єкцій. Для цих категорій спостерігається певне зниження повноти до рівня сімдесяти п'яти відсотків, що пояснюється високою схожістю таких аномалій із нормальною поведінкою користувачів при взаємодії з веб-додатками. Проте показник прецизійності для цих атак залишається високим, а отже, система не створює зайвого навантаження на адміністратора через помилкові тривоги. Атаки типу SQL-ін’єкцій розпізнаються впевненіше, що дозволяє забезпечити надійний захист баз даних інформаційної системи.

Отримані дані підтверджують, що градієнтний бустинг є оптимальним вибором для систем виявлення вторгнень нового покоління. Високе значення зваженої F1-міри, яка становить дев'яносто дев'ять відсотків, доводить збалансованість моделі та її стійкість до незбалансованості вхідних даних. Таким чином, результати експерименту повністю підтверджують робочу гіпотезу про можливість побудови швидкої та точної системи захисту, яка здатна працювати

з живим трафіком, зберігаючи при цьому високу якість детектування широкого спектра сучасних кіберзагроз.

На рисунку 3.4 зображено матрицю невідповідностей для XGBoost.

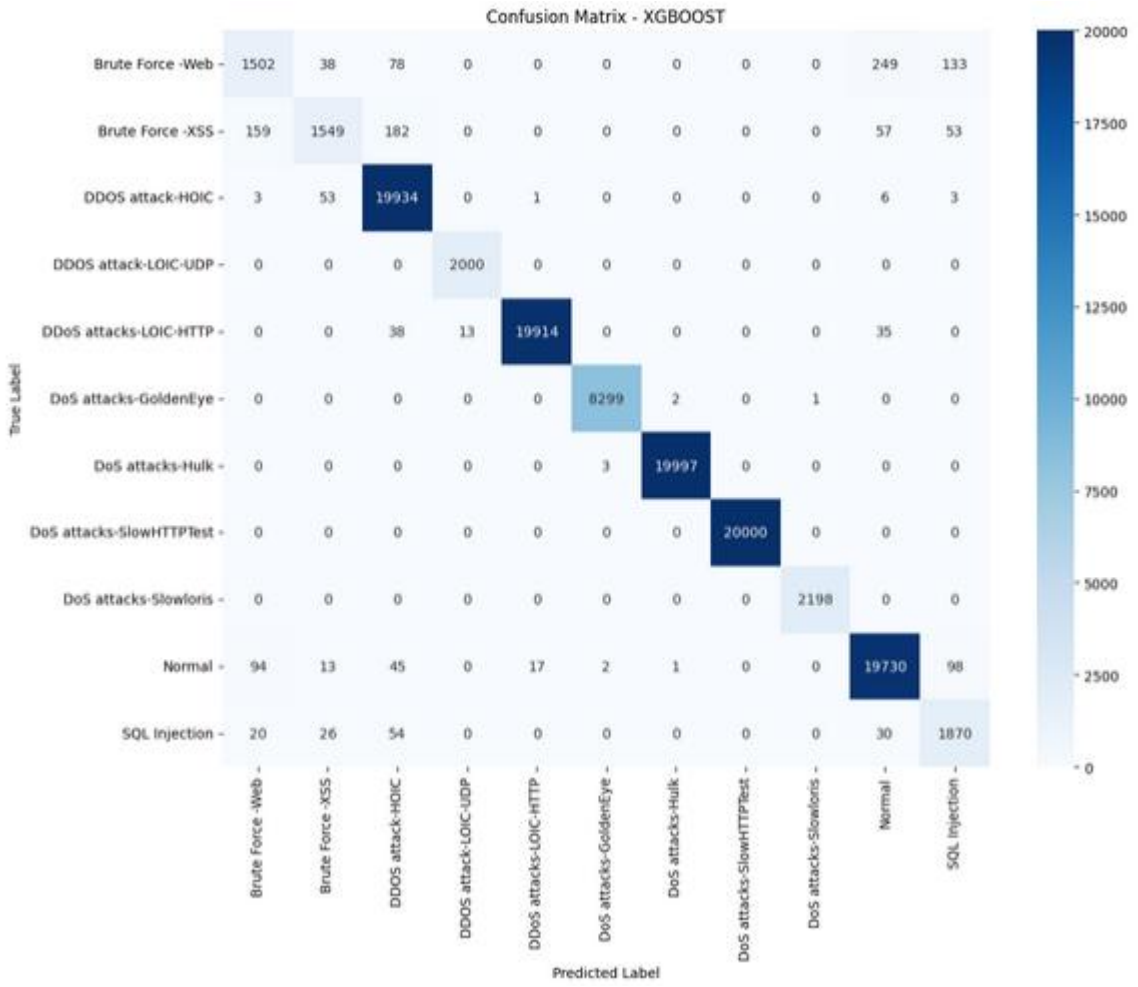

Рисунок 3.4 – Матриця невідповідностей XGBoost

Матриця невідповідностей XGBoost свідчить про високу точність алгоритму в детекції DoS та DDoS атак, де для категорій SlowHTTPTest та Slowloris зафіксовано нульовий рівень помилок. Модель демонструє надійність у розпізнаванні масового трафіку Hulk та HOIC, а також коректно класифікує 98,6% легітимних з’єднань.

3.4 Модель Random Forest

Процес навчання моделі на базі алгоритму Random Forest у межах даного дослідження базується на використанні збалансованого набору даних,

отриманого з декількох днів активності датасету CSE-CIC-IDS2018. Для забезпечення відтворюваності результатів та стабільності навчання використовується фіксований стан генератора випадкових чисел за допомогою, що дозволяє проводити коректне порівняння різних ітерацій моделі. Основним етапом підготовки до навчання є розподіл вибірки на тренувальну та тестову у співвідношенні 80% на 20% відповідно, причому застосування методу стратифікації гарантує збереження пропорцій одинадцяти цільових класів атак у обох підмножинах.

На етапі конфігурації архітектури «лісу» встановлюється кількість дерев рішень 100, що забезпечує достатню статистичну потужність ансамблю без надмірного навантаження на обчислювальні ресурси. Для запобігання перенавчанню моделі вводиться обмеження максимальної глибини кожного дерева значенням 20, що дозволяє алгоритму узагальнювати патерни аномалій, не фокусуючись на одиничних викидах у мережевому трафіку. Оптимізація швидкості навчання досягається шляхом паралелізації обчислень на всіх доступних ядрах процесора.

Особлива увага приділяється відповідності ознак під час навчання можливостям реального перехоплення трафіку, тому модель тренується на 17 семантично відібраних ознаках, таких як порти призначення, протоколи та статистичні показники довжини пакетів. Це забезпечує сумісність навченої моделі з компонентами попередньої обробки даних у режимі реального часу. Результатом навчання є робастна модель, яка проходить фінальну валідацію на тестових даних, де оцінюється її здатність класифікувати як нормальний трафік, так і складні вектори атак, включаючи Botnet, DoS та Brute Force. Збереження навченої моделі разом із параметрами масштабування StandardScaler у форматі серіалізованих об'єктів дозволяє використовувати результати оптимізації для подальшого розгортання в інформаційних системах. На рисунку 3.4 зображено метрики ефективності моделі Random Forest.

```
Classification Report:
============================================================
                          precision    recall  f1-score   support

       Brute Force -Web       0.90      0.82      0.86      2000
       Brute Force -XSS       0.95      0.90      0.92      2000
       DDOS attack-HOIC       0.99      1.00      0.99     20000
    DDOS attack-LOIC-UDP       0.99      1.00      1.00      2000
  DDoS attacks-LOIC-HTTP       1.00      1.00      1.00     20000
   DoS attacks-GoldenEye       1.00      1.00      1.00      8302
        DoS attacks-Hulk       1.00      1.00      1.00     20000
DoS attacks-SlowHTTPTest       1.00      1.00      1.00     20000
   DoS attacks-Slowloris       1.00      1.00      1.00      2198
                  Normal       0.99      0.99      0.99     20000
           SQL Injection       0.92      1.00      0.96      2000

                accuracy                           0.99    118500
               macro avg       0.98      0.97      0.97    118500
            weighted avg       0.99      0.99      0.99    118500


✓ Overall Accuracy: 0.9919
```

Рисунок 3.4 – Метрики ефективності моделі Random Forest

Об'єктивний аналіз результатів тестування навченої моделі дозволяє оцінити її реальний потенціал у виявленні аномалій у мережевих структурах. Загальний показник точності системи склав 99,19%, що є винятково високим результатом для багатокласової класифікації одинадцяти категорій трафіку. Така ефективність підтверджує, що алгоритм Random Forest здатен успішно виділяти ключові ознаки аномальної активності навіть при використанні оптимізованого набору з сімнадцяти параметрів пакетного рівня. Метрики прецизійності, повноти та F1-міри для більшості досліджуваних класів практично досягають одиниці, що свідчить про високу якість попередньої підготовки даних та ефективність застосованої гібридної стратегії балансування вибірки.

Найвищу результативність модель продемонструвала при детектуванні критичних масових загроз, зокрема DDoS-атак (HOIC, LOIC-HTTP) та різних варіацій DoS-активності (Hulk, SlowHTTPTest, Slowloris, GoldenEye). Для цих категорій показники повноти досягли максимального значення 1,00, що означає здатність системи ідентифікувати 100% подібних інцидентів без жодних пропусків. Це доводить, що статистичні патерни інтенсивних атак є достатньо специфічними для впевненого розпізнавання ансамблем дерев рішень. Важливо підкреслити, що легітимний трафік класифікується з точністю 99%, що мінімізує ймовірність хибнопозитивних спрацювань та гарантує стабільність роботи інформаційної системи для кінцевих користувачів.

Певні труднощі виникли лише при ідентифікації веб-атак, пов'язаних із підбором паролів та XSS-ін'єкціями. Для цих класів спостерігається зниження показника повноти до рівня 82–90%, що зумовлено високою подібністю таких аномалій до стандартної активності користувачів у веб-інтерфейсах. Проте показник прецизійності для цих категорій залишається стабільно високим, що запобігає генерації зайвих помилкових тривог. При цьому атаки типу SQL-ін'єкцій розпізнаються з повною ефективністю, забезпечуючи надійний рівень захисту баз даних від несанкціонованого втручання.

Отримані в ході експерименту дані доводять, що використання випадкового лісу є виправданим для розв'язання задач мережевої безпеки. Високе значення зваженої F1-міри на рівні 99% підтверджує стійкість моделі до незбалансованості класів та її здатність до узагальнення складних залежностей у даних. Таким чином, результати тестування повністю підтверджують можливість реалізації надійної системи виявлення вторгнень, яка поєднує в собі високу швидкість обробки інформації та точність ідентифікації широкого спектра сучасних кіберзагроз у режимі реального часу. На рисунку 3.5 зображено матрицю невідповідностей.

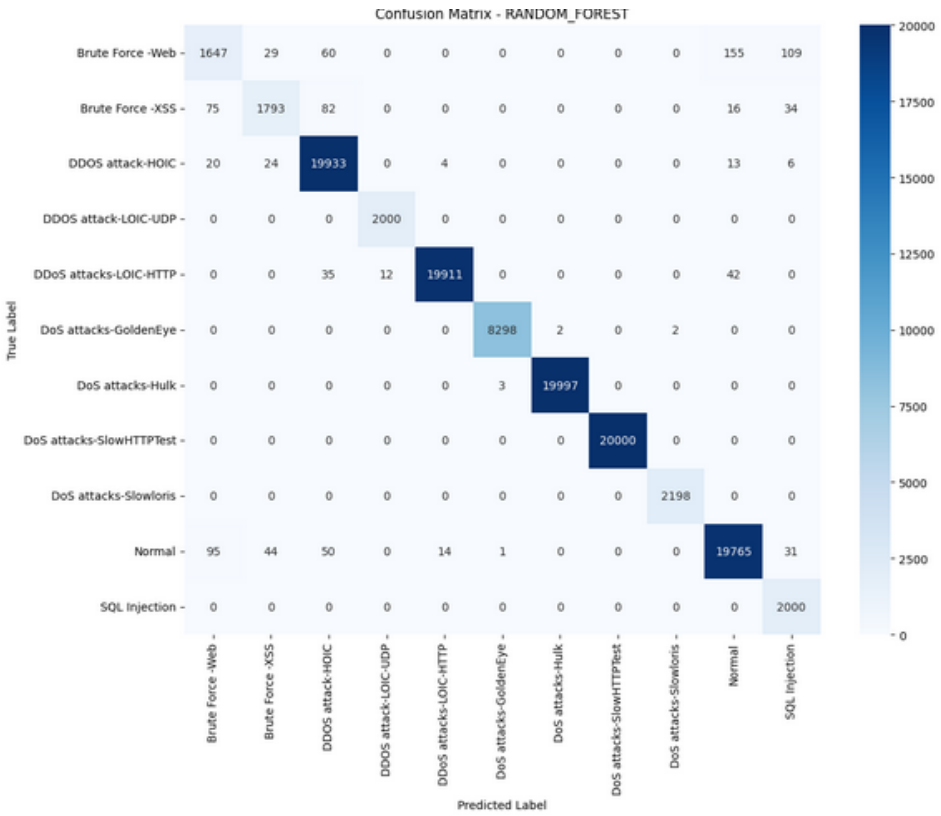

Рисунок 3.5 – Матриця невідповідностей для Random Forest

Матриця невідповідностей підтверджує високу вибірковість алгоритму Random Forest, який продемонстрував безпомилкову ідентифікацію атак типу SlowHTTPTest, LOIC-UDP та SQL Injection. Модель успішно розрізняє масові DoS-загрози, допускаючи лише одиничні похибки при класифікації тисяч зразків трафіку Hulk та HOIC. Висока точність детекції легітимної активності на рівні 99% мінімізує ймовірність хибних спрацювань, що є критично важливим для стабільної роботи мережі. Основні зони плутанини локалізовані в межах веб-атак, де невелика частина спроб підбору паролів (Brute Force Web) помилково приймається за нормальний трафік або XSS через схожість їхніх статистичних ознак. Проте загальна стійкість діагоналі матриці доводить надійність системи у розпізнаванні найбільш небезпечних векторів вторгнень.

3.5 Висновок до третього розділу

В третьому розділі кваліфікаційної роботи проаналізовано результати проведеного обчислювального експерименту з використанням моделей XGBoost та Random Forest, у ході якого було виконано відбір найбільш значущих ознак трафіку та підтверджено високу точність обраних моделей у виявленні різноманітних сценаріїв мережевих атак.

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Організація безпечного робочого місця користувача ПК

Організація безпечного робочого місця користувача електронно-обчислювальних машин у сучасних умовах цифровізації є критично важливим завданням, що потребує інтегрованого підходу до ергономіки, інженерної безпеки та гігієни праці. Даний процес суворо регламентується державною нормативною базою, де фундаментальне значення мають мінімальні вимоги щодо безпеки та захисту здоров’я працівників під час роботи з екранними пристроями, затверджені наказом Міністерства соціальної політики України № 207 від 14.02.2018 року [57]. Відповідно до положень цього документа, створення належних умов праці є прямим обов’язком роботодавця, що передбачає регулярну оцінку ризиків для зору та фізичного стану персоналу. Особлива увага приділяється просторовому плануванню, оскільки кожне робоче місце повинно мати достатню площу (не менше 6 квадратних метрів) та об’єм (не менше 20 кубічних метрів), що забезпечує необхідну циркуляцію повітря та мінімізує психологічне навантаження від перебування в обмеженому просторі.

Ергономіка робочого місця безпосередньо впливає на працездатність та довгострокове здоров’я фахівця, тому конструкція робочих столів та крісел повинна відповідати антропометричним характеристикам конкретного користувача. Робочий стіл має бути оснащений регульованою за висотою поверхнею з матовим покриттям, що виключає появу світлових відблисків, які викликають швидку втомлюваність очей. Робоче крісло обов’язково повинно мати п’ятипроменеву стійку опору з колесами, анатомічну спинку з підтримкою попереку та механізми регулювання кута нахилу та висоти сидіння, які легко доступні в сидячому положенні. Важливо забезпечити таку посадку, при якій стопи ніг повністю торкаються підлоги, а кут у колінних та ліктьових суглобах становить приблизно дев'яносто градусів, що запобігає застійним явищам у судинах та знижує статичне навантаження на хребет.

Світлове середовище в приміщеннях з комп’ютерною технікою підлягає суворому контролю згідно з Державними будівельними нормами ДБН В.2.5-28:2018 «Природне і штучне освітлення» [58]. Дані норми встановлюють, що рівень освітленості на робочій поверхні має бути в межах від 300 до 500 люксів, при цьому світловий потік має бути рівномірним, без різких тіней. Монітори слід розташовувати перпендикулярно до віконних прорізів, щоб уникнути як прямої засвітки очей сонячним світлом, так і дзеркальних відблисків на екрані. Відстань від очей до монітора має становити шістсот-сімсот міліметрів, а верхній край екрана повинен бути на рівні очей або дещо нижче, що дозволяє тримати м’язи шиї в розслабленому стані.

Мікрокліматичні умови в робочих зонах регламентуються Санітарними нормами мікроклімату виробничих приміщень ДСН 3.3.6.042-99 [59]. Враховуючи високу щільність електронного обладнання, яке постійно генерує тепло, необхідно забезпечувати ефективну примусову вентиляцію та кондиціювання повітря. Оптимальна температура в робочій зоні має підтримуватися в діапазоні від 22 до 25 °С при відносній вологості від 40 до 60 %. Порушення цих параметрів призводить до пересихання слизових оболонок, зниження концентрації уваги та підвищення ймовірності помилок під час виконання критично важливих завдань із мережевої безпеки.

Технічна безпека експлуатації електроустановок споживачів вимагає наявності надійного контуру захисного заземлення для всього парку комп’ютерної техніки. Відповідно до галузевих правил, усі струмопровідні частини корпусів повинні мати електричний зв’язок із землею, що унеможливлює накопичення небезпечного статичного заряду та захищає персонал у разі пробою ізоляції. Прокладання кабельних ліній живлення та передачі даних має виконуватися в захищених каналах, що виключає їх механічне пошкодження. Регулярний контроль опору ізоляції та перевірка справності штепсельних з’єднань є обов’язковими заходами для запобігання електротравматизму.

У системі протипожежного захисту приміщень з ЕОМ ключове значення має дотримання Правил пожежної безпеки в Україні [60]. Оскільки комп’ютерна

техніка містить велику кількість полімерних матеріалів та працює під високою напругою, приміщення мають бути обладнані автоматичними датчиками диму та температури. У разі виникнення загоряння першочерговою дією персоналу є повне знеструмлення обладнання та сповіщення служби порятунку. Для локалізації вогню на ранніх стадіях слід використовувати вуглекислотні вогнегасники, які, на відміну від пінних чи порошкових, не проводять струм і не забруднюють внутрішні мікросхеми серверів та робочих станцій залишками вогнегасної речовини.

Підтримання працездатності систем також вимагає суворого дотримання Правил технічної експлуатації електроустановок споживачів [61]. Це стосується використання джерел безперебійного живлення, які не лише захищають дані від раптової втрати, а й запобігають перегріву блоків живлення при стрибках напруги, що часто стає причиною задимлення. У разі виникнення нещасних випадків на робочому місці, зокрема ураження електричним струмом, персонал повинен володіти навичками домедичної допомоги згідно з Порядком, затвердженим МОЗ України [62]. Такий системний підхід, що охоплює ергономіку, електробезпеку та готовність до НС, є необхідною умовою для безпечної та продуктивної професійної діяльності в сучасному ІТ-середовищі.

4.2 Захист від статичної електрики в серверних приміщеннях

Забезпечення надійної роботи сучасних інформаційних систем вимагає суворого контролю фізичних небезпечних факторів усередині апаратних комплексів. Одним із найбільш підступних та поширених явищ, що здатні дестабілізувати роботу мережевого обладнання, є накопичення та розряд статичної електрики. Електростатичний розряд виникає внаслідок трибоелектричного ефекту під час тертя двох різних матеріалів і може досягати потенціалу в кілька тисяч вольт. Для людини такий розряд зазвичай є безпечним, проте для високочутливих напівпровідникових компонентів, мікросхем, процесорів та модулів оперативної пам'яті він є руйнівним. Попадання статичного заряду на елементи системних плат викликає миттєвий тепловий

пробій переходів, деградацію напівпровідників або приховані дефекти, які проявляються з часом у вигляді спорадичних збоїв, зависань серверів та втрати пакетів даних під час аналізу трафіку.

Основним джерелом генерації статичних зарядів у серверних приміщеннях є постійний рух повітряних потоків, що створюються потужними прецизійними системами кондиціювання та вентиляції. Сухе повітря, проходячи на високій швидкості через металеві повітроводи, захисні решітки та корпуси обладнання, інтенсивно електризує навколишні поверхні. Ситуація суттєво погіршується, якщо вологість повітря в приміщенні падає нижче критичної межі, що суворо контролюється загальними санітарними нормами мікроклімату виробничих приміщень. Додатковими факторами ризику є переміщення персоналу по кімнаті, тертя одягу з синтетичних тканин, а також використання невідповідних матеріалів для покриття підлоги чи оздоблення стін. Звичайне покриття здатне накопичувати критичний потенціал лише за умови звичайного ходіння працівника, що створює постійну загрозу для апаратної інфраструктури.

Захист від статики регламентується галузевими правилами безпечної експлуатації електроустановок споживачів та державними стандартами щодо захисту електронних пристроїв від електростатичних явищ [63]. Комплексна система захисту базується на трьох основних принципах: заземлення всіх провідних поверхонь, підтримання оптимального мікроклімату та використання спеціальних антистатичних матеріалів. Першочерговим інженерним заходом є облаштування технологічного заземлення, яке має бути інтегроване із загальним контуром захисного заземлення будівлі відповідно до правил технічної експлуатації електроустановок. Усі серверні шафи, стійки, металеві лотки для прокладання кабелів, а також корпуси комутаторів та маршрутизаторів повинні мати надійне гальванічне з’єднання із заземлювальною шиною, причому опір розтікання струму такого заземлювального пристрою не повинен перевищувати встановлених нормами 4 ом.

Важливу роль у запобіганні трибоелектричному ефекту відіграє суворе дотримання параметрів виробничого середовища, що забезпечується інтеграцією систем клімат-контролю з автоматичними зволожувачами повітря. Відносна

вологість повітря в комп'ютерних залах має постійно підтримуватися на рівні, який дозволяє утворювати на поверхнях предметів мікроскопічну плівку конденсату. Ця плівка є природним провідником і дозволяє статичному заряду самостійно та безпечно стікати в землю, не накопичуючись до небезпечних значень. Для фінішного покриття підлоги в серверних залах дозволяється використовувати виключно спеціальний антистатичний лінолеум або фальшпідлогу із заземленою мідною сіткою, що відповідає сучасним вимогам безпеки до обладнання інформаційних технологій.

Будь-який персонал, який виконує технічне обслуговування, модернізацію чи заміну апаратних модулів безпосередньо всередині серверних стійок, зобов'язаний використовувати індивідуальні засоби захисту від електростатичного розряду [64]. До них належать антистатичні браслети, які одягаються на зап’ястя та підключаються за допомогою крученого шнура із вбудованим резистором номіналом в один мегаом до спеціальних технологічних гнізд заземлення на серверній шафі. Такий резистор є обов'язковим елементом безпеки, оскільки він захищає самого працівника від ураження струмом у разі випадкового контакту з фазною напругою обладнання. Лише такий системний інженерний підхід до фізичної безпеки середовища дозволяє мінімізувати ризик раптової відмови «заліза», захистити внутрішні мікросхеми від пробою та забезпечити стабільний моніторинг інформаційних потоків.

4.3 Дії персоналу при виникненні пожежі в комп’ютерному залі

Ефективність ліквідації загоряння в приміщеннях з високою концентрацією електронно-обчислювальної техніки безпосередньо залежить від оперативності та злагодженості дій персоналу, що базуються на чіткому знанні протоколів безпеки. Відповідно до чинних законодавчих актів, кожен працівник зобов’язаний знати правила пожежної безпеки та сприяти проведенню рятувальних заходів у разі виникнення надзвичайної ситуації [65]. Першочерговою дією при виявленні ознак горіння, задимлення або характерного запаху горілої ізоляції є негайне повідомлення оперативно-рятувальної служби

за телефоном 101. Під час виклику необхідно чітко вказати адресу об’єкта, місце виникнення пожежі, наявність людей у небезпечній зоні та власне прізвище. Після цього працівник має сповістити про подію керівництво та задіяти систему тривожної сигналізації для оповіщення всього персоналу будівлі.

Критично важливим етапом, що передує безпосередньому гасінню вогню, є повне знеструмлення апаратної частини та систем кондиціювання повітря, оскільки робота електроустановок у режимі короткого замикання призводить до стрімкого поширення вогню. Вимкнення електроенергії здійснюється шляхом активації пристроїв аварійної зупинки або вимкнення автоматичних вимикачів у розподільчих щитах. Після знеструмлення приміщення персонал повинен негайно розпочати евакуацію, рухаючись до найближчих виходів згідно із затвердженими планами евакуації, що розробляються для кожного об’єкта [66]. Оскільки комп'ютерний зал часто має складну інженерну інфраструктуру, вогонь може поширюватися приховано через кабельні канали, що вимагає особливої обережності під час виходу з небезпечної зони.

Для гасіння пожежі власними силами до прибуття підрозділів рятувальної служби дозволяється використовувати лише ті первинні засоби, що призначені для відповідного класу пожеж. Категорично заборонено застосовувати воду або пінні вогнегасники, оскільки вони є провідниками електрики та спричиняють остаточне псування дорогого електронного обладнання. Найбільш ефективними в умовах комп'ютерного залу є вуглекислотні вогнегасники, які забезпечують швидке зниження температури в осередку горіння та витісняють кисень, не залишаючи при цьому слідів на серверних стійках. Під час використання таких засобів у замкненому просторі необхідно забезпечити захист органів дихання через ризик різкого зростання концентрації вуглекислого газу в повітрі.

У разі, якщо під час пожежі персонал отримав травми або опіки, необхідно забезпечити надання допомоги відповідно до встановлених медичних протоколів. Важливим аспектом є також проведення регулярних інструктажів та практичних тренувань, що дозволяє виробити у працівників автоматизм дій у критичних умовах [67]. Такий комплексний підхід до підготовки кадрів та технічного оснащення приміщень системами виявлення вогню дозволяє

мінімізувати матеріальні збитки та гарантувати безпеку людей під час надзвичайних ситуацій на об’єктах інформаційної інфраструктури.

4.4 Висновок до четвертого розділу

В четвертому розділі кваліфікаційної роботи описано заходи з охорони праці та безпеки в надзвичайних ситуаціях, що спрямовані на створення безпечного виробничого середовища під час експлуатації комп’ютерних систем. Подано детальний аналіз вимог до організації робочого місця користувача ПК з урахуванням ергономічних, світлотехнічних та мікрокліматичних нормативів, а також визначено чіткий алгоритм дій персоналу при виникненні пожежі в комп’ютерному залі для мінімізації ризиків і збереження здоров'я працівників. Окрему увагу приділено захисту від виникнення електростатичного розряду в зонах із високою щільністю обчислювальної інфраструктури.

ВИСНОВКИ

У кваліфікаційній роботі освітнього рівня «Магістр» проведено комплексне дослідження методів машинного навчання для виявлення аномалій у мережевому трафіку. За результатами роботи зроблено такі висновки:

В першому розділі кваліфікаційної роботи:

- подано таксономію мережевих аномалій та їх класифікацію за характером впливу на інформаційну систему;
- розглянуто сучасний стан кіберзагроз та роль інтелектуального аналізу даних у сучасних системах захисту;
- висвітлено характеристики та ключові атрибути мережевого трафіку як джерела даних для виявлення вторгнень;
- проаналізовано наявні підходи до моніторингу безпеки та виявлено обмеження традиційних сигнатурних методів;
- досліджено концепцію цілеспрямованих стійких загроз та механізми їх прихованого перебування в мережі;
- обґрунтовано доцільність впровадження методів машинного навчання для автоматизації детекції невідомих раніше аномалій;
- сформовано теоретичний базис для подальшого моделювання процесів розпізнавання шкідливого трафіку.

В другому розділі кваліфікаційної роботи:

- описано принципи роботи та архітектурні відмінності систем класу IDS, SIEM та EDR;
- досліджено математичний апарат ансамблевих методів класифікації, зокрема алгоритмів Random Forest та XGBoost;
- подано порівняльний опис еталонних наборів даних, за результатами якого для проведення експериментів обрано датасет CSE-CIC-IDS2018 як найбільш актуальний та репрезентативний.

В третьому розділі кваліфікаційної роботи:

- розроблено методику попередньої обробки даних, що включає нормалізацію, очищення та балансування вибірок для підвищення якості навчання моделей;
- запропоновано алгоритм відбору найбільш значущих ознак трафіку, що дозволило оптимізувати обчислювальні витрати без втрати точності;
- спроектовано програмні моделі на основі ансамблевих алгоритмів для автоматизованої класифікації мережевих подій;
- протестовано розроблені рішення на основі метрик Accuracy, Precision та Recall, що підтвердило високу ефективність обраного підходу у виявленні сучасних типів атак.

У розділі «Охорона праці та безпека в надзвичайних ситуаціях» проаналізовано нормативно-правові вимоги та санітарно-гігієнічні параметри організації робочого місця користувача ПК, що забезпечують мінімізацію професійних ризиків і підтримання високої працездатності персоналу. Обґрунтовано комплекс інженерно-технічних заходів захисту від статичної електрики в серверних приміщеннях, що включає облаштування технологічного заземлення, контроль відносної вологості повітря та використання спеціальних покриттів для підлоги з метою запобігання апаратним збоям мережевого обладнання. Описано чіткий алгоритм дій фахівців при виникненні пожежі в комп’ютерному залі, включаючи порядок оповіщення, правила знеструмлення обладнання та застосування первинних засобів пожежогасіння для гарантування безпеки людей і збереження технічної інфраструктури.

ПЕРЕЛІК ДЖЕРЕЛ

1. ENISA Threat Landscape 2024. Enisa. 2024. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024.

2. Війна Росії проти України: хронологія кібератак. European Parlament. URL: https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733549/EPRS_BRI(2022)733549_XL.pdf.

3. Reddy Kothamali P., Banik S. Limitations of Signature-Based Threat Detection. Limitations of Signature-Based Threat Detection. 2022. URL: https://www.researchgate.net/publication/388494583_Limitations_of_Signature-Based_Threat_Detection.

4. Jeon J. Possibility for Proactive Anomaly Detection. arxiv. 15.04.2025. URL: https://arxiv.org/abs/2504.11623.

5. Ikechukwu Okoli U. Machine learning in cybersecurity: A review of threat detection and defense mechanisms. ResearchGate. 09.02.2024. URL: https://www.researchgate.net/publication/378208150_Machine_learning_in_cybersecurity_A_review_of_threat_detection_and_defense_mechanisms.

6. Haas S. Zeek-Osquery: Host-Network Correlation for Advanced Monitoring and Intrusion Detection. HAL open science. 04.02.2021. URL: https://inria.hal.science/hal-03440828/preview/497034_1_En_17_Chapter.pdf.

7. Moruff Oyelakin A. Overview and Exploratory Analyses of CICIDS 2017 Intrusion Detection Dataset. ResearchGate. 16.09.2023. URL: https://www.researchgate.net/publication/373994633_Overview_and_Exploratory_Analyses_of_CICIDS_2017_Intrusion_Detection_Dataset.

8. Arcos-Argudo M. A Deterministic Comparison of Classical Machine Learning and Hybrid Deep Representation Models for Intrusion Detection on NSL-KDD and CICIDS2017. mdpi. 08.05.2025. URL: https://www.mdpi.com/1999-4893/18/12/749.

9. de Carvalho Bertoli G. Bridging the gap to real-world for network intrusion detection systems with data-centric approach. arxiv. 13.05.2021. URL: https://arxiv.org/abs/2110.13655.

10. zakon rada. 08.05.2025. URL: https://zakon.rada.gov.ua/laws/show/2163-19#Text.

11. ISO/IEC 27001:2022. ISO. URL: https://www.iso.org/cms/live/live/en/sites/isoorg/contents/data/standard/08/28/82875.html.

12. ISO/IEC 27005:2022. ISO. URL: https://www.iso.org/standard/80585.html.

13. S. Lupenko, O. Orobchuk, I. Kateryniuk, R. Kozak, H. Lypak. Secure information system for Chinese Image medicine knowledge consolidation. Proceedings of the 4th International Workshop on Information Technologies: Theoretical and Applied Problems 2024 (ITTAP 2024), Ternopil, Ukraine and Opole, Poland, October 23-25, 2024, CEUR Workshop Proceedings, Vol-3896. pp. 509-519. ISSN 1613-0073. URL: https://ceur-ws.org/Vol-3896/paper23.pdf.

14. Rahman M., Atiquzzaman M. A Survey on Network Security Traffic Analysis and Anomaly Detection Techniques. ResearchGate. 27.05.2024. URL:

https://www.researchgate.net/publication/380903277_A_Survey_on_Network_Security_Traffic_Analysis_and_Anomaly_Detection_Techniques.

15. What is Anomaly Detection? VMware by Broadcom. URL: https://www.vmware.com/topics/anomaly-detection.

16. Gujral E. Anomaly Detection Methods: A Survey. MADLab. 2023. URL: https://www.cs.ucr.edu/~egujr001/ucr/madlab/publication/EG_2023_Anomaly_Detection_Methods.pdf.

17. Ahmed M., Mahmood A. N., Hu J. A comprehensive survey on network anomaly detection. ResearchGate. 02.07.2016. URL: https://www.researchgate.net/publication/326136935_A_comprehensive_survey_on_network_anomaly_detection.

18. Anomaly Detection: A Guide to Algorithms & Use Cases. Encord. 15.02.2024. URL: https://encord.com/blog/anomaly-detection/.

19. Суботін С. О. Побудова інтелектуальних систем моніторингу та діагностики: навчальний посібник. КПІ ім. Ігоря Сікорського. 2023. URL: https://ela.kpi.ua/server/api/core/bitstreams/ecdf61dd-487e-46e2-817a-c80f2198eff7/content.

20. Point Anomaly Detection: Identifying Single Unusual Data Points. UX for AI. 24.08.2023. URL: https://uxforai.com/p/point-anomaly-detection.

21. Gupta P., Goudar V. S., Singh P. Network intrusion detection using machine learning and deep learning: a review. PMC. 16.12.2024. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC11723367/.

22. What Is an Advanced Persistent Threat (APT)? Palo Alto Networks. URL: https://www.paloaltonetworks.com/cyberpedia/what-is-advanced-persistent-threat-apt.

23. Network Traffic Analysis (NTA) Fundamentals. Rapid7. URL: https://www.rapid7.com/fundamentals/network-traffic-analysis.

24. Early S., Stephens C., Doyle R. Anomaly Detection in Network Traffic. UCSD CSE. 2001. URL: https://cseweb.ucsd.edu/classes/fa01/cse221/projects/group10.pdf.

25. Duda, O., Pasichnyk V., Lypak H., Matsiuk O., Mudrokha V. Formation of integrated repositories of social and communication data by consolidating the resources of museums, libraries and archives in smart cities projects. CEUR Workshop Proceedings, 2021, 2870, pp. 1420–1430. URL: http://ceur-ws.org/Vol-2870/paper104.pdf.

26. Supervised Learning. ScienceDirect. URL: https://www.sciencedirect.com/topics/computer-science/supervised-learning.

27. What is unsupervised learning? Google Cloud. URL: https://cloud.google.com/discover/what-is-unsupervised-learning.

28. What is Reinforcement Learning? GeeksforGeeks. 20.01.2025. URL: https://www.geeksforgeeks.org/machine-learning/what-is-reinforcement-learning.

29. Ясінський О. Методи машинного навчання для виявлення аномалій у мережевому трафіку сучасних інформаційних систем. Природничі та гуманітарні

науки. Актуальні питання: збірник тез доповідей IX Міжнародної студентської науково-технічної конференції. Тернопіль: ТНТУ, 2026. С. 264-265.

30. Липак Г., Липак Т., Кунанець Н. Проєктування інформаційної системи на основі машинного навчання для збереження та класифікації артефактів документальної спадщини. Вісник Хмельницького національного університету: технічні науки. Т. 334 № 4 (2024). С. 176-182. https://doi.org/10.31891/2307-5732-2024-339-4-29.

31. Липак Г. І., Коломийчук Д. А. Вплив бекдор атак на можливості навчання нейронних мереж // Матеріали МНТК „Фундаментальні та прикладні проблеми сучасних технологій“, Тернопіль, 28-29 травня 2025 року. 2025. С. 207.

32. What is an example of NIDS? Stamus Networks. 27.09.2023. URL: https://www.stamus-networks.com/blog/what-is-an-example-of-nids.

33. Network-Based Intrusion Detection System (NIDS). CyberHoot. URL: https://cyberhoot.com/cybrary/network-based-intrusion-detection-system-nids.

34. Difference between HIDS and NIDS. GeeksforGeeks. 29.05.2024. URL: https://www.geeksforgeeks.org/computer-networks/difference-between-hids-and-nids.

35. Deepanshu, Saini V. Network Intrusion Detection System using Machine Learning. Serials Journals. 2021. URL: https://serialsjournals.com/abstract/51172_44-deepanshu.pdf.

36. What is the Difference Between Snort and Zeek? Stamus Networks. 11.10.2023. URL: https://www.stamus-networks.com/blog/what-is-the-difference-between-snort-and-zeek.

37. How Does Antivirus Software Work? Security.org. URL: https://www.security.org/antivirus/how-does-antivirus-work.

38. What Is Endpoint Detection and Response (EDR)? CrowdStrike. URL: https://www.crowdstrike.com/en-us/cybersecurity-101/endpoint-security/endpoint-detection-and-response-edr.

39. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. arXiv. 10.06.2016. URL: https://arxiv.org/abs/1603.02754.

40. Introduction to Random Forest. DataHacker. 25.04.2020. URL: https://datahacker.rs/012-machine-learning-introduction-to-random-forest.

41. Sharafaldin I. Toward a Reliable Dataset for Intrusion Detection. ResearchGate. 01.02.2018. URL: https://www.researchgate.net/publication/322923753_Toward_a_Reliable_Dataset_for_Intrusion_Detection.

42. Ясінський О. О., Липак Г. І. Проблема репрезентативності наборів даних у теоретичному моделюванні сучасних кіберзагроз. Актуальні задачі сучасних технологій: збірник тез доповідей XIV Міжнародної науково-технічної конференції молодих учених та студентів. Тернопіль: ТНТУ, 2025. С. 381-382.

43. Intrusion Detection Datasets: A Review from KDD Cup 99 to CIC-IDS-2017. ScienceDirect. 15.05.2023. URL: https://www.sciencedirect.com/science/article/abs/pii/S187705092101568X.

44. Data Cleaning: Definition, Process, and Benefit. Tableau. 12.01.2024. URL: https://www.tableau.com/learn/articles/what-is-data-cleaning.

45. Brownlee J. How to Handle Missing Data with Python. Machine Learning Mastery. 02.08.2023. URL: https://machinelearningmastery.com/handle-missing-data-python/.

46. All About Feature Scaling. Towards Data Science. 14.07.2023. URL: https://towardsdatascience.com/all-about-feature-scaling-bcc0ad7d2d13.

47. Robust Scaler: Outlier Robust Feature Scaling. GeeksforGeeks. 22.11.2024. URL: https://www.geeksforgeeks.org/standardscaler-minmaxscaler-and-robustscaler-techniques-ml/.

48. SMOTE for Imbalanced Classification with Python. Machine Learning Mastery. 14.01.2024. URL: https://machinelearningmastery.com/smote-oversampling-for-imbalanced-classification/.

49. Handling Class Imbalance with Synthetic Data Generation. KDnuggets. 28.08.2023. URL: https://www.kdnuggets.com/2023/08/handling-class-imbalance-synthetic-data-generation.html.

50. Feature Selection Techniques in Machine Learning. Analytics Vidhya. 15.06.2024. URL: https://www.analyticsvidhya.com/blog/2020/10/feature-selection-techniques-in-machine-learning/.

51. Кліщ, М., Липак, Г., Кунанець, Н., Пасічник, С., & Липак, Т. Структура інформаційної системи передбачення та інтерпретації зміни стану користувача сервісу. Вісник Національного Університету “Львівська Політехніка”. Інформаційні системи та мережі, 17 (2025). С. 226 - 238. https://doi.org/10.23939/sisn2025.17.226.

52. Understanding the Curse of Dimensionality. Towards Data Science. 05.09.2023. URL: https://towardsdatascience.com/the-curse-of-dimensionality-50dc6e49aa12.

53. Outlier Detection: Methods, Models and Applications. DataCamp. 10.12.2024. URL: https://www.datacamp.com/blog/outlier-detection-methods.

54. Isolation Forest Algorithm for Anomaly Detection. GeeksforGeeks. 25.10.2024. URL: https://www.geeksforgeeks.org/isolation-forest-algorithm-anomly-detection/.

55. Precision vs. Recall: Which Metric Should You Use?. HubSpot. 04.11.2024. URL: https://blog.hubspot.com/service/precision-vs-recall.

56. Understanding Confusion Matrix and Its Metrics. Analytics Vidhya. 12.02.2025. URL: https://www.analyticsvidhya.com/blog/2020/06/confusion-matrix-machine-learning/.

57. Мінімальні вимоги щодо безпеки та захисту здоров’я працівників під час роботи з екранними пристроями. Наказ Міністерства соціальної політики України від 14.02.2018 № 207. URL: https://zakon.rada.gov.ua/laws/show/z0308-18.

58. ДБН В.2.5-28:2018. Природне і штучне освітлення. К.: Мінрегіонбуд України, 2018. URL: https://online.budstandart.com/ua/catalog/doc-page.html?id_doc=79885.

59. ДСН 3.3.6.042-99. Санітарні норми мікроклімату виробничих приміщень. МОЗ України. URL: https://zakon.rada.gov.ua/rada/show/va042282-99.

60. Правила пожежної безпеки в Україні. Наказ Міністерства внутрішніх справ України від 30.12.2014 № 1417. URL: https://zakon.rada.gov.ua/laws/show/z0252-15.

61. Правила технічної експлуатації електроустановок споживачів. Наказ Міністерства енергетики та вугільної промисловості України від 13.02.2012 № 91. URL: https://zakon.rada.gov.ua/laws/show/z0350-12.

62. Порядок надання домедичної допомоги особам при підозрі на ураження електричним струмом. Наказ Міністерства охорони здоров'я України від 09.03.2022 № 441. URL: https://zakon.rada.gov.ua/laws/show/z0343-22.

63. ДСТУ EN 61340-5-1:2018. Електростатика. Частина 5-1. Захист електронних пристроїв від електростатичних явищ. Загальні вимоги. К.: ДП «УкрНДНЦ», 2018.

64. ДСТУ EN 60950-1:2015. Обладнання інформаційних технологій. Безпека. Частина 1. Загальні вимоги. К.: ДП «УкрНДНЦ», 2015.

65. Кодекс цивільного захисту України. Верховна Рада України. URL: https://zakon.rada.gov.ua/laws/show/5403-17.

66. Порядок здійснення навчання населення діям у надзвичайних ситуаціях. Постанова Кабінету Міністрів України від 26.06.2013 № 444. URL: https://zakon.rada.gov.ua/laws/show/444-2013-п.

67. ДСТУ ISO 23601:2019. Ідентифікація безпеки. Знаки безпеки на планах евакуації. К.: ДП «УкрНДНЦ», 2019.

ДОДАТКИ

Додаток А

Тези конференцій

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Тернопільський національний технічний університет імені Івана Пулюя
Маріборский університет (Словенія)
Технічний університет у Кошице (Словаччина)
Вільнюський технічний університет ім. Гедиміноса (Литва)
Краківський економічний університет (Польща)
Вроцлавський економічний університет (Польща)
Університет «Опольська Політехніка» (Польща)
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
Вінницький національний аграрний університет
Львівський національний університет ім. І. Франка
Головне управління Пенсійного фонду в Тернопільській області
Наукове товариство ім. Шевченка
Тернопільський обласний комунальний інститут післядипломної педагогічної освіти
Сумський державний педагогічний університет
Запорізький національний університет

АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ

Збірник

тез доповідей

XIV Міжнародної науково-технічної конференції молодих учених та студентів

11-12 грудня 2025 року

УКРАЇНА
ТЕРНОПІЛЬ – 2025

УДК 004.8
О.О. Ясінський; Г.І. Липак, к. соц. ком.
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ПРОБЛЕМА РЕПРЕЗЕНТАТИВНОСТІ НАБОРІВ ДАНИХ У ТЕОРЕТИЧНОМУ МОДЕЛЮВАННІ СУЧАСНИХ КІБЕРЗАГРОЗ

O.O. Yasinskyi; H.I. Lypak Ph.D

THE PROBLEM OF DATA SET REPRESENTATIVENESS IN THEORETICAL MODELING OF MODERN CYBERTHREATS

У 2025 році темпи еволюції кіберзагроз значно перевищують швидкість оновлення більшості дослідницьких наборів даних, що використовуються для тренування моделей систем виявлення вторгнень. За оцінками фахівців, понад 40% нових типів атак, зафіксованих протягом останніх двох років, не мають ані загальнодоступних еталонних наборів даних, ані формалізованих поведінкових профілів. Це створює критичний розрив між теоретичними результатами, отриманими в умовах лабораторного тестування, та реальними можливостями ML-моделей протистояти модернізованим, адаптивним і часто зашифрованим атакам у корпоративних мережах. Проблема репрезентативності даних стає ключовим обмеженням для розвитку IDS, оскільки сучасні атаки є динамічними, контекстними й дедалі краще оптимізованими для обходу алгоритмічних механізмів захисту.

Ефективність систем виявлення вторгень (англ. Intrusion Detection System, IDS) на базі машинного навчання (англ. *Machine learning*, ML) критично залежить від якості даних, на яких вони навчаються. Фундаментальна проблема полягає у значному розриві між високою лабораторною точністю ML-моделей та їхньою практичною неспроможністю виявляти новітні кіберзагрози. Це відбувається тому, що більшість популярних наборів даних є застарілими або неповно відображають характеристики атак.

NSL-KDD, що є покращеною версією набору даних KDD99, досі залишається найбільш поширеним для академічних порівнянь. Однак його трафік частково синтетичний, а загрози абсолютно неактуальні у 2025 році, якщо порівнювати з діями APT (англ. Advanced Persistent Threat, розвинена стала загроза) чи атаками де є шифрування. Моделі, навчені на ньому, фактично вчаться розпізнавати "артефакти" самого набору даних.

Набір даних CICIDS2017 був значним кроком уперед, оскільки він містив дані 2017 року з реалістичними профілями атак (DDoS, XSS, Botnet). Проте за понад 8 років він також застарів. Крім того, його атаки були "профільованими" – виконаними за чітким сценарієм – що не відображає реальні "тихі" та повільні (low-and-slow) атаки типу APT, які маскуються під легітимний трафік.

Проблема в тому, що дослідники продовжують публікувати роботи, де їхні нові нейронні мережі показують високий відсоток точності на NSL-KDD. Це створює хибне відчуття вирішеної проблеми. Насправді така модель просто ідеально "вивчила" застарілий, нерелевантний набір даних.

При перенесенні у реальну корпоративну мережу, модель стикається з трафіком, профілі якого ніколи не були представлені під час тренування.

В таблиці 1 подано аналіз загроз, які відсутні в сучасних популярних наборах даних.

Таблиця 1. Загрози, що відсутні в наборах даних

Тип загрози	Опис	Причина застарілості
Adversarial AI Attacks	Атаки, що цілеспрямовано "обманюють" ML-моделі, вносячи ледь помітні зміни в трафік.	Набори даних не містять таких прикладів.
Fileless Malware	Атаки "без файлів", що живуть у оперативній пам'яті системи.	Більшість наборів даних аналізують мережевий трафік, а не поведінку хоста.
Атаки на IoT	Специфічні атаки на "розумні" пристрої та промислові системи.	Використовують специфічні протоколи (MQTT, Modbus), яких немає в CICIDS чи NSL-KDD.
Зашифрований C&C трафік	Командні центри, що спілкуються через зашифровані канали	Аналіз метаданих зашифрованого трафіку – задача, не відображена в старих наборах даних.

Висновок. Використання застарілих та нерепрезентативних наборів даних або наборів даних, що швидко втрачають актуальність, є серйозною перешкодою для розвитку ефективних IDS. Теоретичне моделювання показує завищені результати, що не транслюються у реальну безпеку. Вирішення проблеми не релевантних наборів даних вимагає нових підходів, а саме використання генеративних моделей для імітації нових атак, навчання на реальних даних та фокус на метаданих замість зашифрованого вмісту.

Література

1. Sharafaldin I., Lashkari A. H., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSP. 2018. Vol. 1. P. 108–116.URL: https://www.scitepress.org/papers/2018/66398/66398.pdf

2. Tavallaee M., Bagheri E., Lu W., Ghorbani A. A. A detailed analysis of the KDD CUP 99 data set. 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. 2009. URL: https://www.researchgate.net/publication/48446353_A_detailed_analysis_of_the_KDD_CUP_99_data_set

3. Ring M., Wunderlich S., Grüdl D., Landes D., Hotho A. A survey of network-based intrusion detection data sets. Computers & Security. 2019. URL: https://arxiv.org/abs/1903.02460 .

4. Engelen T., van Ede T., Doerr C. Troubleshooting an Intrusion Detection Dataset: the CICIDS2017 Case Study. Proceedings of the 2021 Workshop on Theoretic and Experimental Benchmarking of BGP Attacks and Defenses (WTMC 2021). 2021.URL: https://www.researchgate.net/publication/353107141_Troubleshooting_an_Intrusion_Detection_Dataset_the_CICIDS2017_Case_Study

Міністерство освіти і науки України
Тернопільський національний технічний університет
імені Івана Пулюя
Маріборський університет (Словенія)
Технічний університет в Кошице (Словаччина)
Каунаський технологічний університет (Литва)
Львівський національний університет
імені Івана Франка
Гірничо-металургійна академія ім. Станіслава Сташиця (Польща)
Луцький національний технічний університет
Чернівецький національний університет
імені Юрія Федьковича
Вроцлавський економічний університет (Польща)
Університет технологій та економіки
імені Хелени Ходковської (Польща)
Донбаська державна машинобудівна академія

Студентське наукове товариство

IX МІЖНАРОДНА

студентська науково - технічна конференція

"ПРИРОДНИЧІ ТА ГУМАНІТАРНІ НАУКИ. АКТУАЛЬНІ ПИТАННЯ"

24-25 квітня 2026 р.

(збірник тез конференції)

Тернопіль 2026

IX Міжнародна студентська науково - технічна конференція
"ПРИРОДНИЧІ ТА ГУМАНІТАРНІ НАУКИ. АКТУАЛЬНІ ПИТАННЯ"

УДК 004.8
Ясінський О. - ст. гр. СНнм-61
Тернопільський національний технічний університет імені Івана Пулюя

МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ У МЕРЕЖЕВОМУ ТРАФІКУ СУЧАСНИХ ІНФОРМАЦІЙНИХ СИСТЕМ

Науковий керівник: к. соц. ком Липак Г.І.

Yasinskyi O.
Ternopil Ivan Puluj National Technical University

MACHINE LEARNING METHODS FOR DETECTING ANOMALIES IN NETWORK TRAFFIC IN MODERN INFORMATION SYSTEMS

Supervisor: Ph.D Lypak H.I.

Ключові слова: машинне навчання, аномалія, мережевий трафік
Keywords: machine learning, anomaly, network traffic

Виявлення аномалій у мережевому трафіку сучасних інформаційних систем є критично важливим завданням, оскільки традиційні методи на основі сигнатур дедалі частіше пасують перед новими типами кібератак та шифрованим трафіком. Основна ідея застосування машинного навчання тут полягає у створенні профілю «нормальної» поведінки мережі, на фоні якого будь-які суттєві відхилення ідентифікуються як потенційні загрози. Це дозволяє системам захисту реагувати на невідомі раніше вразливості, хоча й створює певний ризик хибнопозитивних спрацювань через природну динамічність мережевих процесів [1]. При цьому сучасні рішення мають бути адаптивними, оскільки характер мережевої активності постійно змінюється, що вимагає від алгоритмів здатності до самонавчання у реальному часі без постійного переналаштовування параметрів людиною.

При порівнянні різних підходів варто виділити методи навчання з учителем, такі як дерева рішень або випадкові ліси, які демонструють високу точність у класифікації вже відомих атак, проте вони майже безпорадні, якщо в навчальній вибірці відсутні приклади конкретної аномалії [2]. Натомість методи навчання без учителя, зокрема алгоритми кластеризації типу k-means або DBSCAN, набагато краще підходять для реальних умов, де мітки даних часто відсутні, оскільки вони шукають ізольовані точки або групи даних, що стоять осторонь від основної маси трафіку. Окреме місце посідають методи на основі статистичних моделей та опорних векторів (One-Class SVM), які фокусуються на описі межі нормальності, що робить їх ефективними для виявлення тонких відхилень у структурі пакетів. Особливу складність сьогодні становить аналіз зашифрованих пакетів, де неможливо перевірити вміст, тому методи машинного навчання все частіше орієнтуються на статистичні характеристики потоків, як-от розмір вікна чи інтервали між пакетами, для виявлення прихованих загроз [3].

Сучасні нейронні мережі, особливо автоенкодери, пропонують ще глибший аналіз, оскільки вони здатні самостійно стискати дані та відновлювати їх, при цьому аномальний трафік відновлюється з високою похибкою, що і слугує маркером загрози. Порівняно з простими метричними методами, глибоке навчання потребує значних обчислювальних ресурсів, проте воно значно краще справляється з величезними

обсягами даних у великих корпоративних мережах [4]. Окрім потужності, важливою стає і стійкість самих моделей до маніпуляцій, адже зловмисники можуть намагатися "обманути" алгоритм, поступово підмішуючи шкідливий трафік у навчальну вибірку.

В таблиці 1 подано порівняльний аналіз методів навчання.

Таблиця 1. Порівняння методів навчання

Метод навчання	Тип вхідних даних	Виявлення нових атак	Обчислювальна складність	Рівень хибних спрацювань
Навчання з учителем (Random Forest, SVM)	Розмічені дані (мітки атак)	Низький	Середня	Низький
Навчання без учителя (k-means, Isolation Forest)	Нерозмічені дані	Високий	Низька / Середня	Високий
Однокласові методи (One-Class SVM)	Тільки нормальний трафік	Середній	Середня	Середній
Глибоке навчання (Автоенкодери, RNN)	Сирі дані або Потоки	Дуже високий	Висока	Середній

Вибір конкретного методу залежить від специфіки системи: для жорстко контрольованих промислових мереж краще підходять статистичні та однокласові моделі, тоді як для динамічного інтернет-трафіку найбільш перспективним є поєднання методів кластеризації та глибокого навчання. Оптимальним рішенням для сучасних систем стають гібридні моделі, які поєднують швидкість класичних алгоритмів із гнучкістю нейромереж.

Література:

1. Шевченко А. С., Застело Г. І., Шпачинський Є. О. Аналіз застосування методів машинного навчання на основі штучних нейронних мереж для виявлення кіберзагроз. ela.kpi.ua. 2019. URL: https://ela.kpi.ua/items/1da6e657-b91a-4842-85f7-ea72ae928ad2.

2. Shone N., Tran Nguyen N. A Deep Learning Approach to Network Intrusion Detection. ResearchGate. 2018. URL: https://www.researchgate.net/publication/322866638_A_Deep_Learning_Approach_to_Network_Intrusion_Detection.

3. Kwon D., Kim J. A survey of deep learning-based network anomaly detection. ResearchGate. 2019. URL: https://www.researchgate.net/publication/320066760_A_survey_of_deep_learning-based_network_anomaly_detection.

4. Chalapathy R. Deep Learning for Anomaly Detection: A Survey. arxiv. 2019. URL: https://arxiv.org/abs/1901.03407