УДК 004.056.5:004.8 DOI https://doi.org/10.35546/kntu2078-4481.2025.4.3.31

О. І. ФЕДЮШИН
кандидат технічних наук, доцент,
доцент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-3600-405X

П. В. ШУЛІК
кандидат технічних наук,
старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0004-6200-2172

О. М. ТОВМА
аспірант кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0009-7225-3297

В. П. КОЦЮБА
викладач кафедри радіоелектронних систем пунктів управління Повітряних Сил
Харківський національний університет Повітряних Сил
імені Івана Кожедуба
ORCID: 0000-0001-6336-8193

ВИЯВЛЕННЯ АНОМАЛІЙ У МЕРЕЖЕВОМУ ТРАФІКУ ВЕБ-ДОДАТКІВ З ВИКОРИСТАННЯМ АЛГОРИТМУ RANDOM FOREST

Забезпечення високої точності та ефективності систем виявлення вторгнень є критичним для уникнення помилкових спрацьовувань та пропусків веб-атак. Мережевий трафік є основним каналом комунікації клієнт-серверної моделі у веб-додатках, тому його аналіз є критично важливим для виявлення шкідливої активності. Саме через нього передаються дані запитів користувача, відповіді сервера, API-виклики, а також взаємодія з зовнішніми сервісами. Використання машинного навчання для прогнозування атак на веб-додатки на основі аналізу мережевого трафіку може покращити ці характеристики. В роботі проводиться вибір технік машинного навчання, що можуть допомагати враховувати не лише структуру, а й семантику трафіку, також розглянуто актуальну проблему захисту веб-додатків від сучасних кіберзагроз, зокрема атак типу SQL Injection, XSS та Brute Force.

Проаналізовано недоліки традиційних сигнатурних методів захисту (WAF, IDS), які демонструють низьку ефективність проти атак «нульового дня» та модифікованих шкідливих запитів. Обґрунтовано доцільність використання методів машинного навчання (Machine Learning) для побудови адаптивних систем виявлення вторгнень. У якості інструментарію дослідження обрано алгоритм Random Forest (випадковий ліс) через його стійкість до перенавчання та здатність ефективно працювати з багатовимірними даними. Експериментальну частину виконано на базі датасету CICIDS2017, який містить актуальні профілі нормального та аномального трафіку. Описано етапи попередньої обробки даних: очищення від шумів, кодування категоріальних ознак, балансування класів за допомогою бібліотеки Imbalanced-learn.

В результаті проведеного експерименту розроблена модель показала загальну точність класифікації (Accuracy) на рівні 98%. Проаналізовано матрицю невідповідностей, яка підтвердила високу здатність моделі розрізняти легітимний трафік та атаки сканування, при цьому виявлено певні обмеження при класифікації складних веб-атак зі схожими сигнатурами. Результати дослідження можуть бути використані для удосконалення існуючих систем моніторингу безпеки та інтеграції в SIEM-системи.

Ключові слова: *кібербезпека, веб-додатки, Random Forest, виявлення аномалій, CICIDS2017, машинне навчання, класифікація трафіку.*

O. I. FEDIUSHYN
Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-3600-405X

P. V. SHULIK
Candidate of Technical Sciences,
Senior Lecturer at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0004-6200-2172

O. M. TOVMA
Postgraduate Student at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0009-7225-3297

V. P. KOTSIUBA
Lecturer at the Department of Radio-Electronic Systems
of Air Force Control Points
Ivan Kozhedub Kharkiv National University of the Air Force
ORCID: 0000-0001-6336-8193

ANOMALY DETECTION IN WEB APPLICATION NETWORK TRAFFIC USING THE RANDOM FOREST ALGORITHM

Ensuring high accuracy and efficiency of intrusion detection systems is critical for avoiding false positives and missed web attacks. Network traffic is the primary communication channel of the client-server model in web applications; therefore, its analysis is critically important for detecting malicious activity. It is through this channel that user request data, server responses, API calls, and interactions with external services are transmitted. The use of machine learning to predict attacks on web applications based on network traffic analysis can improve these characteristics. The paper performs a selection of machine learning techniques that can help consider not only the structure but also the semantics of traffic. It also examines the current problem of protecting web applications from modern cyber threats, particularly SQL Injection, XSS, and Brute Force attacks.

The shortcomings of traditional signature-based protection methods (WAF, IDS), which demonstrate low effectiveness against zero-day attacks and modified malicious requests, are analyzed. The feasibility of using Machine Learning methods to build adaptive intrusion detection systems is substantiated. The Random Forest algorithm was chosen as the research tool due to its resistance to overfitting and ability to work effectively with high-dimensional data. The experimental part was performed based on the CICIDS2017 dataset, which contains current profiles of normal and anomalous traffic. Data preprocessing stages are described in detail: noise cleaning, encoding categorical features, and class balancing using the Imbalanced-learn library. As a result of the experiment, the developed model showed an overall classification Accuracy of 98%. The confusion matrix was analyzed, confirming the model's high ability to distinguish legitimate traffic and scanning attacks, while identifying certain limitations in classifying complex web attacks with similar signatures. The research results can be used to improve existing security monitoring systems and for integration into SIEM systems.

Key words: *cybersecurity, web applications, Random Forest, anomaly detection, CICIDS2017, machine learning, traffic classification.*

Постановка проблеми

Пасивні та активні методи збору даних становлять основу сучасного аналізу трафіку. Пасивні методи передбачають моніторинг мережевих пакетів без втручання у потік даних, що забезпечує високу ефективність при виявленні аномалій. До таких методів належать снифінг трафіку за допомогою інструментів типу Wireshark або Tcpdump. Активні методи, такі як сканування портів або тестування пропускної здатності, дозволяють отримувати додаткову інформацію про стан мережі, але можуть впливати на її продуктивність.

Роль трафіку у забезпеченні безпеки полягає в тому, що саме в ньому можна виявити багато типів атак. До того ж аналіз трафіку – це не лише засіб виявлення атак після їх здійснення, а й інструмент для проактивного захисту. Регулярне логування, моніторинг та застосування систем виявлення вторгнень, які аналізують трафік у реальному часі, дозволяють зупинити атаку ще до того, як вона призведе до наслідків. Таким чином, трафік виступає джерелом для побудови моделей поведінки користувачів, виявлення аномалій та автоматизованого реагування.

Мережевий трафік веб-додатку має специфічну структуру, яка визначається протоколами та технологіями, що використовуються для взаємодії між клієнтом і сервером. У протоколах HTTP/HTTPS кожна сесія складається

з запиту клієнта та відповіді сервера. Запит зазвичай включає метод запиту, URL-шлях та параметри, заголовки зі службовою інформацією (куки, токени тощо), тіло запиту, що може містити дані форм або JSON-об'єкти. Відповідь сервера має код статусу, який сигналізує про результат обробки, заголовки з інформацією про тип вмісту і політики кешування, тіло відповіді. Важливо враховувати специфіку протоколів, які використовуються у сучасних односторінкових або мобільних додатках, де активніше застосовуються API-запити та WebSocket-з'єднання. У таких випадках структура трафіку відрізняється від класичних веб-сторінок, і аналіз має фокусуватися на JSON-повідомленнях, структурі API-викликів і підтримці з'єднання.

Розуміння структури та динаміки трафіку дає можливість не лише виявити шкідливу активність, а й забезпечити адаптивний захист. Традиційні методи захисту досі залишаються одним із найпоширеніших та актуальних варіантів. Веб-брандмауери (WAF) аналізують запити в реальному часі, блокуючи відомі шаблони атак за допомогою сигнатурних баз. Брандмауери веб-додатків аналізують HTTP/HTTPS-запити й блокують потенційно шкідливі, базуючись на наборах сигнатур. Ці рішення ефективні проти стандартних загроз, таких як SQL-ін'єкції або XSS, проте мають обмеження щодо нових, невідомих векторів атак. Порівняно з традиційними сигнатурними методами, підходи з боку машинного навчання є більш гнучкими й здатними до самонавчання – це означає, що моделі можна оновлювати у відповідь на зміну поведінки атакувальників, що також зменшує навантаження на аналітиків безпеки. Сюди входить виявлення аномалій, побудоване на визначенні профілю нормальної поведінки веб-додатку, а також використання систем управління подіями безпеки (SIEM), які централізують логування, моніторинг і кореляцію подій у реальному часі.

Разом з перевагами сучасні підходи до виявлення атак часто не враховують динамічність та послідовність мережевих подій, що призводить до зниження точності моделей, більшість моделей машинного навчання виявляють проблеми з дисбалансом класів та наявністю шуму в даних, що впливає на їхню продуктивність, і ці питання вимагають вчасного вирішення.

Завданням даного дослідження є розробка моделі машинного навчання для прогнозування атак на веб-додатки на основі аналізу мережевого трафіку, яка буде враховувати структуру пакетів окремих видів трафіку, а також специфіку їх обробки за числовими та категоріальними характеристиками.

Аналіз останніх досліджень і публікацій

За останні роки застосування штучного інтелекту у сфері кібербезпеки стало предметом численних наукових досліджень, опублікованих у провідних виданнях, таких як IEEE, ACM, а також на популярній платформі arXiv. Найбільш цитовані роботи присвячені як теоретичним основам машинного навчання, так і практичним застосуванням для виявлення різних типів атак на мережеві та веб-додатки. Аналіз наукових публікацій за 2022-2025 роки дозволяє виділити три основні напрями, за якими розвиваються системи виявлення атак на основі використання штучного інтелекту: методи глибокого навчання (Deep Learning), класичні алгоритми машинного навчання (Machine Learning) та гібридні підходи.

Одним із ключових напрямків досліджень є використання глибокого навчання для аналізу мережевого трафіку. Зокрема, нейронні мережі, такі як рекурентні (RNN) [1, 2] та трансформери [3,4], показали високу ефективність у виявленні аномалій, DDoS-атак і навіть складних APT-кампаній. Наприклад, моделі на основі Long Short-Term Memory (LSTM) здатні аналізувати часові послідовності в трафіку, виявляючи незвичайні патерни, які можуть свідчити про атаку. Останні роботи також демонструють потенціал графових нейронних мереж (GNN) для аналізу взаємозв'язків між вузлами в мережі [5, 6], що особливо актуально для виявлення скоординованих атак.

Дослідники в працях [7,8] демонструють високу ефективність згорткових нейронних мереж (CNN) та мереж довгої короткострокової пам'яті (LSTM) для класифікації трафіку датасету CICIDS2017. Наприклад, у роботі [7] досягнуто точності 99.1% завдяки здатності нейромереж автоматично виділяти приховані ознаки.

Головним недоліком описаних вище методів є висока обчислювальна складність ("black-box problem") та потреба у потужному апаратному забезпеченні (GPU), що ускладнює їх інтеграцію в системи реального часу (Real-time WAF) з обмеженими ресурсами.

Окремий кластер досліджень присвячено аналізу текстових логів. Зокрема, Al-Omari та ін. (2024) у роботі *«Anomaly Detection in HTTP Logs»* [9] пропонують фреймворк для виявлення аномалій безпосередньо в журналах HTTP-серверів. Хоча цей метод ефективний для виявлення атак прикладного рівня (L7), він ігнорує мережеві параметри нижніх рівнів (TCP/IP flags, packet flow duration), що робить систему вразливою до розподілених атак на відмову в обслуговуванні (DDoS), які не завжди коректно відображаються в логах веб-сервера.

Потужний напрям – це класичні та ансамблеві методи машинного навчання, до яких належить і дане дослідження. У роботах [10-11] проводиться порівняльний аналіз алгоритмів SVM, k-NN та Random Forest. Перевага SVM, k-NN це простота реалізації, зрозуміла математична модель, недоліками є: низька швидкість на великих вибірках (k-NN), чутливість до шуму. Автори дослідження [11] зазначають, що Random Forest демонструє найкращий баланс між швидкістю навчання та точністю класифікації на незбалансованих даних.

Попри велику кількість робіт, більшість досліджень [7, 8] розглядають датасет CICIDS2017 як основний для навчання та тестування моделей. При цьому завдання ставляться дуже узагальнено, намагаючись

класифікувати всі 14 типів атак одночасно, що часто призводить до зниження точності виявлення саме веб-загроз (Web Attacks) на фоні масових DDoS-атак. Тому існує потреба у створенні спеціалізованої моделі, оптимізованої саме для захисту веб-додатків, яка поєднувала б глибину аналізу пакетного трафіку (на відміну від аналізу логів [9]) з високою швидкодією ансамблевих методів (на відміну від важких DL-моделей). Саме цю задачу покликана вирішити дана робота, пропонуючи оптимізовану модель Random Forest для специфічного профілю загроз веб-застосунків.

Формулювання мети дослідження

Метою роботи є підвищення рівня захищеності веб-додатків шляхом розробки моделі машинного навчання на основі алгоритму Random Forest, здатної з високою точністю класифікувати мережевий трафік та виявляти аномалії, характерні для сучасних кібератак.

Для досягнення мети потрібно дослідити існуючі методи аналізу мережевого трафіку та алгоритми машинного навчання, що застосовуються в задачах безпеки веб-додатків; провести вибір та обґрунтування оптимальних типів моделей машинного навчання з урахуванням особливостей мережевого трафіку та задачі прогнозування атак; розробити та реалізувати модель машинного навчання для прогнозування атак на основі зібраних даних мережевого трафіку; провести тестування та оцінку якості розробленої моделі за допомогою відповідних метрик (Precision, Recall, F1-score).

Викладення основного матеріалу дослідження

Найбільш поширеним напрямком у застосуванні машинного навчання до кібербезпеки є контрольоване навчання (Supervised Learning), яке потребує попередньо розмічених даних (наприклад, «атака» чи «нормальний трафік»). Тут використовуються моделі:

– Random Forest – дозволяє будувати ансамбль дерев рішень і добре працює з ознаками трафіку;

– Support Vector Machine – ефективна для задач класифікації з чіткою границею між класами, особливо при високій розмірності простору ознак;

– XGBoost – градієнтний бустинг, який демонструє високу точність при виявленні атак завдяки оптимізації втрат і регуляризації;

– MLP (багатошарові перцептрони) – використовуються для більш складного представлення нелінійних залежностей у трафіку;

– Convolutional Neural Networks (CNN) – хоча більше відомі в комп'ютерному зорі, можуть бути адаптовані до задач безпеки, наприклад, при аналізі трафіку у вигляді часових рядів або образів.

Контрольоване навчання найбільш ефективне при наявності якісних і репрезентативних наборів даних, де класи атак добре розрізняються. Водночас основним викликом залишається отримання та оновлення таких даних, особливо для нових типів атак.

Для виявлення невідомих або атак нульового дня, коли немає маркованих прикладів, застосовують неконтрольоване навчання (Unsupervised Learning). Ці моделі не потребують розмітки та фокусуються на виявленні аномальної поведінки, відмінної від «нормального» трафіку:

– Isolation Forest – ізолює аномалії шляхом випадкового поділу ознак; добре працює з великими наборами даних;

– DBSCAN – кластеризує запити за відстанню та щільністю; незвичні запити автоматично класифікуються як шум (аномалії).

– Autoencoders – нейронні мережі, що навчаються реконструювати «нормальні» зразки; великі відхилення в реконструкції сигналізують про потенційну атаку.

Ці методи ідеально підходять для середовищ з постійно змінюваним трафіком, де оновлення сигнатур не встигає за новими загрозами.

Створення ефективної системи виявлення кібератак вимагає ретельної побудови експериментального пайплайну, який охоплює всі етапи – від завантаження даних до отримання прогнозів. Цей процес починається з підготовки даних, де особливу увагу приділяємо очищенню, нормалізації та балансуванню набору. Кожен етап має вирішальне значення для кінцевого результату, оскільки навіть найдосконаліші алгоритми не зможуть показати хороші результати на неякісних даних. Важливою частиною пайплайну є створення модульної структури, яка дозволяє легко змінювати окремі компоненти для експериментів з різними алгоритмами та підходами.

Для дослідження з побудови моделі було обрано виключно алгоритм Random Forest, що обумовлено низкою критичних переваг цього методу для задач аналізу мережевого трафіку [7, 8]. Random Forest є ансамблевим алгоритмом на основі дерев рішень, який особливо добре проявляє себе при роботі з даними, що мають численні ознаки та складні залежності між ними. Ця характеристика ідеально відповідає природі мережевого трафіку, де кожне з'єднання може описуватися десятками параметрів.

Головною перевагою Random Forest є його здатність ефективно працювати з даними, що мають високий рівень шуму та часткові відсутні значення – типові проблеми реальних наборів даних про мережевий трафік. Алгоритм

демонструє виняткову стійкість до перенавчання завдяки механізму випадкового вибору підмножин ознак для кожного дерева в ансамблі. Ця властивість є особливо важливою для нашої задачі, оскільки набір CICIDS2017 містить значну кількість корельованих ознак.

Для проведення експерименту було обрано набір даних CICIDS2017 (Canadian Institute for Cybersecurity Intrusion Detection System). Цей датасет є одним з найбільш репрезентативних для задач кібербезпеки, оскільки він містить сучасні сценарії атак (DoS, DDoS, Brute Force, Web Attacks, Infiltration) та повні пейлоади пакетів, на відміну від застарілих наборів на кшталт KDD99.

У рамках дослідження було виконано наступні етапи попередньої обробки даних (Data Preprocessing):

1) очищення даних (Data Cleaning): видалення дублікатів та записів із пропущеними значеннями (NaN), а також обробка значень *Infinity*, які виникають при розрахунку статистик потоку (наприклад, ділення на нуль при розрахунку швидкості пакетів);

2) кодування ознак (Encoding): категоріальні дані (наприклад, текстові мітки класів атак) було перетворено у числовий формат;

3) нормалізація (Scaling): оскільки ознаки трафіку мають різні діапазони значень (наприклад, тривалість з'єднання та кількість байтів), було застосовано масштабування для покращення збіжності алгоритму;

4) балансування класів: використано підходи бібліотеки *Imbalanced-learn* для усунення дисбалансу між класом нормального трафіку (Benign), якого значно більше, та класами атак.

Після завершення всіх етапів очищення датасет було збережено у новий файл CICIDS2017_cleaned.csv. Фінальна кількість записів після усіх операцій очищення склала 2830628 рядків. Порівняно з початковим обсягом 2830744 рядки даних це становить 99,996% від оригінального набору, що вказує на відносно невелику частку аномальних записів. Отриманий очищений датасет містить узгоджені та коректні дані, готові до подальшого аналізу та використання в моделях машинного навчання.

Наступним етапом був експлораторний аналіз даних. Він є критично важливим етапом у процесі аналізу даних, який дозволяє досліднику ознайомитися з основними характеристиками датасету, виявити закономірності, аномалії та взаємозв'язки між змінними. У контексті дослідження мережевого трафіку, аналіз допомагає зрозуміти структуру даних, розподіл різних типів трафіку, а також виявити потенційні проблеми, які можуть вплинути на подальше моделювання.

Для кращого розуміння було створено не тільки текстовий варіант, а й візуальний. Візуалізація розподілу типів трафіку допомагає виявити дисбаланс класів, що може вплинути на вибір метрик оцінки моделі та стратегії навчання. Для визначення найбільш інформативних ознак використано ANOVA F-test, який оцінює вплив кожної ознаки на цільову змінну. Було відібрано топ-20 ознак із найвищими значеннями F-критерію (рис. 1).

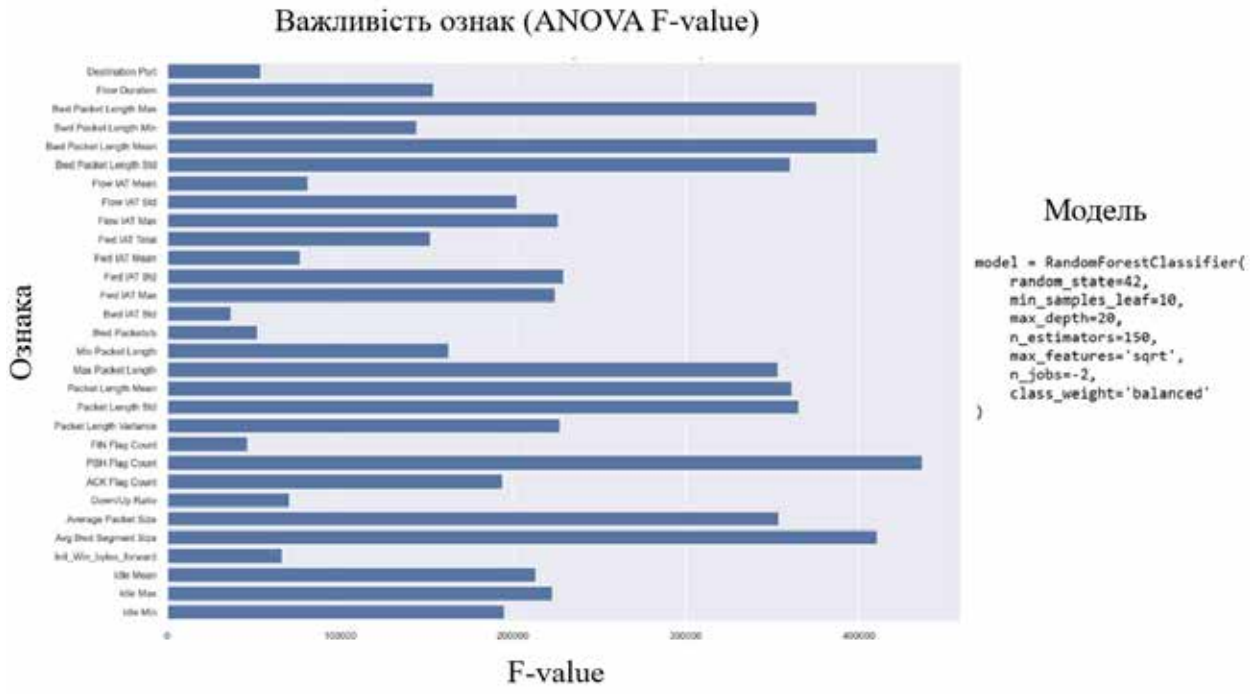

Рис. 1. Інформативні ознаки трафіку та модель класифікації

Практична реалізація моделі виконувалася мовою програмування Python 3.13 з використанням бібліотек Scikit-learn для побудови моделі та Pandas для маніпуляцій з даними. Після навчання моделі на тренувальній вибірці (70% даних) було проведено валідацію на тестовій вибірці (30%). Результати оцінки ефективності класифікатора наведено на рис. 2. Як видно з отриманих даних, модель досягла високої загальної точності у 98%.

Для детальнішого аналізу помилок було побудовано матрицю невідповідностей (Confusion Matrix). Модель Random Forest продемонструвала високу ефективність у класифікації мережевого трафіку на три категорії: 0 – атаки, 1– звичайний трафік (BENIGN) та 2 – розвідувальні дії (Recon). Матриця невідповідностей показала, що модель правильно ідентифікувала 117091 випадків атак, 669953 випадків звичайного трафіку та 48006 випадків розвідувальних дій. При цьому спостерігається невелика кількість помилок класифікації: 1924 випадки звичайного трафіку були помилково класифіковані як атаки, а 11433 випадки атак були помилково віднесені до звичайного трафіку. Для класу Recon кількість помилок була мінімальною – лише 509 випадків.

Метрики ефективності моделі

Random Forest Performance:

	precision	recall	f1-score	support
Attack	0.91	0.98	0.95	119025
BENIGN	1.00	0.98	0.99	681895
Recon	0.99	0.99	0.99	48269
accuracy			0.98	849189
macro avg	0.97	0.99	0.98	849189
weighted avg	0.98	0.98	0.98	849189

Рис. 2. Матриця невідповідностей та метрики ефективності моделі Random Forest

Аналіз матриці показує, що:

1) модель майже безпомилково класифікує нормальний трафік (Benign). Це критично важливо для уникнення блокування легітимних користувачів (мінімізація False Positives);

2) висока точність спостерігається для атак типу Recon (Port Scan) та DoS, які мають виражені статистичні аномалії (частота пакетів, тривалість сесії);

3) певні труднощі виникають при розрізненні підтипів веб-атак (наприклад, XSS та SQL Injection), що пояснюється схожістю їхніх ознак на рівні мережевих пакетів у зашифрованому трафіку (HTTPS).

Висновки

Завдяки здатності моделей машинного навчання аналізувати великі обсяги трафіку, виявляти приховані залежності та нетипові шаблони, такі підходи можуть розпізнати атаки, які залишаються непоміченими для класичних сигнатурних систем. Наприклад, під час автоматизованих атак з ботнетів, трафік з кожного окремого вузла може виглядати звичайним, однак моделі кластеризації або виявлення аномалій можуть виявити підозрілу синхронність або повторюваність дій. У випадку підвищення привілеїв ML-моделі можуть помітити нетипові шаблони доступу до адміністративних функцій з облікових записів звичайних користувачів.

У роботі було розроблено та досліджено модель виявлення аномалій у мережевому трафіку на базі алгоритму Random Forest. Використання набору даних CICIDS2017 дозволило наблизити умови експерименту до реальних. Отримані результати (точність 98%) підтверджують гіпотезу про те, що машинне навчання є ефективним інструментом для посилення безпеки веб-додатків. Модель демонструє високу здатність до узагальнення та може бути використана як другий ешелон захисту в комплексі з традиційними WAF. Подальші дослідження будуть спрямовані на оптимізацію гіперпараметрів моделі для зменшення часу обробки пакетів та інтеграцію підходів глибокого навчання для кращого розпізнавання складних веб-атак.

Список використаної літератури

1. Adefemi, K. O., Mutanga, M. B., & Alimi, O. A. (2025). A Hybrid CNN–GRU Deep Learning Model for IoT Network Intrusion Detection. Journal of Sensor and Actuator Networks, 14(5), 96. https://doi.org/10.3390/jsan14050096.

2. Alahmadi, A. A., Aljabri, M., Alhaidari, F., Alharthi, D. J., Rayani, G. E., Marghalani, L. A., Alotaibi, O. B., & Bajandouh, S. A. (2023). DDoS Attack Detection in IoT-Based Networks Using Machine Learning Models: A Survey and Research Directions. Electronics, 12(14), 3103. https://doi.org/10.3390/ electronics 12143103.

3. Long, Z., Yan, H., Shen, G. et al. A Transformer-based network intrusion detection approach for cloud security. J Cloud Comp 13, 5 (2024). https://doi.org/10.1186/s13677-023-00574-9.

4. Wu, Zihan & Zhang, Hong & Wang, Penghai & Sun, Zhibo. (2022). RTIDS: A Robust Transformer-Based Approach for Intrusion Detection System. IEEE Access. 10. 1-1. 10.1109/ACCESS.2022.3182333.

5. Bilot, T., Madhoun, N.E., Agha, K.A., & Zouaoui, A. (2023). Graph Neural Networks for Intrusion Detection: A Survey. IEEE Access, 11, 49114-49139.

6. Sun, Z., Teixeira, A. M. H., & Toor, S. (2024). GNN-IDS: Graph neural network based intrusion detection system. In Proceedings of the 19th International Conference on Availability, Reliability and Security (ARES 2024) (Article 14, 12 pages). https://doi.org/10.1145/3664476.3664515.

7. Kartiwi, M. CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System. IEEE Access. https://doi.org/10.1109/ACCESS. 2022.3206425.

8. Psychogyios K, Papadakis A, Bourou S, Nikolaou N, Maniatis A, Zahariadis T. (2024). Deep Learning for Intrusion Detection Systems (IDSs) in Time Series Data. Future Internet. 16(3):73. https://doi.org/10.3390/fi16030073.

9. Ahmad, Waqas & Amjad, Muhammad. (2024). Anomaly Detection in HTTP Logs: Leveraging Machine Learning for Uncovering Anomalous Traffic Patterns with SIEM Integration. 622-629. 10.1109/IBCAST61650.2024.10877256.

10. Rahman, M. M., Al Shakil, S. & Mustakim, M. R. (2025) A survey on intrusion detection system in iot networks. Cyber Secur. Appl. 3, 100082. https://doi.org/10.1016/j.csa.2024.100082.

11. Pradhan, Reshamlal. (2022). Decision Tree Based Classifications on CICIDS 2017 Dataset for the Identification of DDoS, Botnet, and Web Attack. NeuroQuantology. 20. 4468-4475. 10.48047/NQ.2022.20.12.NQ77771.

References

1. Adefemi, K. O., Mutanga, M. B., & Alimi, O. A. (2025). A Hybrid CNN–GRU Deep Learning Model for IoT Network Intrusion Detection. Journal of Sensor and Actuator Networks, 14(5), 96. https://doi.org/10.3390/jsan14050096.

2. Alahmadi, A. A., Aljabri, M., Alhaidari, F., Alharthi, D. J., Rayani, G. E., Marghalani, L. A., Alotaibi, O. B., & Bajandouh, S. A. (2023). DDoS Attack Detection in IoT-Based Networks Using Machine Learning Models: A Survey and Research Directions. Electronics, 12(14), 3103. https://doi.org/10.3390/ electronics 12143103.

3. Long, Z., Yan, H., Shen, G. et al. A Transformer-based network intrusion detection approach for cloud security. J Cloud Comp 13, 5 (2024). https://doi.org/10.1186/s13677-023-00574-9.

4. Wu, Zihan & Zhang, Hong & Wang, Penghai & Sun, Zhibo. (2022). RTIDS: A Robust Transformer-Based Approach for Intrusion Detection System. IEEE Access. 10. 1-1. 10.1109/ACCESS.2022.3182333.

5. Bilot, T., Madhoun, N.E., Agha, K.A., & Zouaoui, A. (2023). Graph Neural Networks for Intrusion Detection: A Survey. IEEE Access, 11, 49114-49139.

6. Sun, Z., Teixeira, A. M. H., & Toor, S. (2024). GNN-IDS: Graph neural network based intrusion detection system. In Proceedings of the 19th International Conference on Availability, Reliability and Security (ARES 2024) (Article 14, 12 pages). https://doi.org/10.1145/3664476.3664515.

7. Kartiwi, M. CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System. IEEE Access. https://doi.org/10.1109/ACCESS. 2022.3206425.

8. Psychogyios K, Papadakis A, Bourou S, Nikolaou N, Maniatis A, Zahariadis T. (2024). Deep Learning for Intrusion Detection Systems (IDSs) in Time Series Data. Future Internet. 16(3):73. https://doi.org/10.3390/fi16030073.

9. Ahmad, Waqas & Amjad, Muhammad. (2024). Anomaly Detection in HTTP Logs: Leveraging Machine Learning for Uncovering Anomalous Traffic Patterns with SIEM Integration. 622-629. 10.1109/IBCAST61650.2024.10877256.

10. Rahman, M. M., Al Shakil, S. & Mustakim, M. R. (2025) A survey on intrusion detection system in iot networks. Cyber Secur. Appl. 3, 100082. https://doi.org/10.1016/j.csa.2024.100082.

11. Pradhan, Reshamlal. (2022). Decision Tree Based Classifications on CICIDS 2017 Dataset for the Identification of DDoS, Botnet, and Web Attack. NeuroQuantology. 20. 4468-4475. 10.48047/NQ.2022.20.12.NQ77771.

Дата першого надходження рукопису до видання: 22.11.2025
Дата прийнятого до друку рукопису після рецензування: 17.12.2025
Дата публікації: 31.12.2025