УДК 004.056.53 DOI https://doi.org/10.35546/kntu2078-4481.2026.1.44

А. В. СОКОЛОВ
доктор технічних наук, професор,
завідувач кафедри кібербезпеки
Національний університет «Одеська юридична академія»
ORCID: 0000-0003-0283-7229

В. М. СЛАТВІНСЬКА
доктор філософії, асистент кафедри кібербезпеки
Національний університет «Одеська юридична академія»
ORCID: 0000-0002-6082-981X

В. І. БЕВЗА
магістр І курсу
Національний університет «Одеська юридична академія»
ORCID: 0009-0007-2695-969X

ВИЯВЛЕННЯ АНОМАЛІЙ У ТРАФІКУ МОБІЛЬНИХ ПРИСТРОЇВ ЗА ДОПОМОГОЮ МАШИННОГО НАВЧАННЯ

Кількість мобільних пристроїв у мережі швидко зростає. Разом із цим зростають обсяги передаваних даних. Через це кіберзахист мережевого трафіку ускладнюється. Сигнатурні методи дедалі частіше не спрацьовують. Нові атаки не мають стабільних ознак. Вони маскуються у легітимному трафіку. Тому для виявлення аномалій застосовують методи машинного навчання. Йдеться про автоматизовану обробку мережевого трафіку мобільних пристроїв.

Метою роботи є розробка та дослідження методів машинного навчання для виявлення аномальної поведінки у трафіку мобільних пристроїв, забезпечення раннього виявлення невідомих загроз та підвищення ефективності кіберзахисту.

Наукова новизна. Основа запропонованого підходу – алгоритми неконтрольованого навчання. Використано LSTM Autoencoder та Isolation Forest. Вони аналізують багатовимірні характеристики мобільного трафіку. Далі задається адаптивний поріг. Рівень хибних спрацювань зменшується. Надійність детекції зростає. Порівняння – з класичними статистичними методами.

Результати. Розроблено програмну систему виявлення аномалій на основі Python. У реалізації використано бібліотеки TensorFlow та scikit-learn. Алгоритми перевірено експериментально на реальних наборах мобільного трафіку. Найвищу точність виявлення аномалій показав LSTM Autoencoder – 96,1 %. Isolation Forest зберігає стабільну точність і високу швидкість обробки. Тому він придатний для систем з обмеженими ресурсами.

Висновки. Методи машинного навчання ефективні для інтелектуального аналізу мережевого трафіку мобільних пристроїв. LSTM Autoencoder забезпечує високоточну детекцію аномалій. Isolation Forest має меншу обчислювальну складність. Тому він придатний для реального часу. Результати підтверджують перспективність ML-підходів. Вони підтримують раннє виявлення невідомих кіберзагроз.

Ключові слова: *машинне навчання, виявлення аномалій, мобільний трафік, нейронні мережі, python, кібербезпека, мережевий моніторинг.*

A. V. SOKOLOV
Doctor of Technical Sciences, Professor,
Head of the Department of Cybersecurity
National University “Odesa Law Academy”
ORCID: 0000-0003-0283-7229

V. M. SLATVINSKA
Doctor of Philosophy,
Assistant Professor at the Department of Cybersecurity
National University “Odesa Law Academy”
ORCID: 0000-0002-6082-981X

V. I. BEVZA
1st year Master's Degree Student
National University "Odesa Law Academy"
ORCID: 0009-0007-2695-969X

DETECTION OF ANOMALIES IN MOBILE TRAFFIC USING MACHINE LEARNING

The number of mobile devices and data growth is giving a challenge to cybersecurity network traffic. Traditional signature-based approaches are losing their effectiveness as new types of attack have no consistent pattern and are also hidden in legitimate traffic. In this context, the application of machine learning techniques for the automated detection of anomalies in the mobile network traffic is of particular relevance.

The purpose of the paper is to develop and research machine learning methods for detecting anomalous behavior in mobile device traffic, enabling early identification of previously unknown threats, and enhancing the effectiveness of cyber defense.

Scientific novelty. An approach that relies on unsupervised learning algorithms LSTM Autoencoder and Isolation Forest is proposed for analyzing the multidimensional characteristics of mobile traffic. Using the adaptive threshold allows to decrease the false-positive rates and to increase detection reliability (compared to traditional statistical methods).

Results. A Python-based anomaly detection system was developed using the TensorFlow and scikit-learn libraries. The algorithms were experimentally evaluated on real mobile traffic datasets. The LSTM Autoencoder achieved the highest anomaly detection accuracy (96.1 %), while Isolation Forest offers a balance of stable accuracy and high processing speed, making it suitable for resource-constrained systems.

Conclusions. Machine learning methods are found to be fruitful in intelligent analysis of mobile network traffic. The use of LSTM Autoencoder allows finding very accurately, while algorithms like Isolation Forest have a lower computational complexity and can be applied in real-time. The results provide proof of concept of the potential of machine learning approaches for the early identification of previously unknown cyber threats.

Key words: *machine learning, anomaly detection, mobile traffic, neural networks, Python, cybersecurity, network monitoring.*

Постановка проблеми

Кількість мобільних пристроїв у мережі зростає. Це наслідок цифровізації сучасного суспільства. Зростання має експоненційний характер. Обсяги мережевого трафіку різко збільшуються. Поверхня атаки розширюється. Для зловмисників це додає можливостей. Традиційні засоби захисту не завжди адаптуються оперативно. Мережеве середовище змінюється динамічно. У мобільних мережах з'являються нові типи загроз. Серед них – розподілені атаки типу DDoS. Також фіксується діяльність ботнетів. Окремо йдеться про приховане шкідливе програмне забезпечення. На цьому тлі виявлення аномалій у мережевому трафіку стає ключовим завданням кібербезпеки. Відхилення від нормальної поведінки трафіку може означати компрометацію системи. Воно також може вказувати на цілеспрямовану атаку.

Актуальність дослідження – це практика сучасного кіберзахисту. Потрібне оперативне виявлення загроз у мережевому трафіку. Затримки критичні для функціонування інформаційних систем. Журнали подій вручну не масштабуєш. Особливо на високошвидкісні мережеві потоки. Фіксовані сигнатури і шаблонні правила швидко «старіють». Структура та характеристики трафіку постійно змінюються. Алгоритми машинного навчання беруть великі обсяги мережевих даних. Далі вони знаходять приховані патерни й нелінійні залежності. У явні правила це не згортається. Методи машинного навчання (ML) дають надійність і адаптивність систем кіберзахисту. Умови – динамічне та різноманітне мережеве середовище.

Аналіз останніх досліджень та публікацій

Виявлення аномалій у мережевому трафіку лишається актуальною задачею. Мережеві середовища ускладнюються. З'являються нові й раніше невідомі типи загроз. Тому дослідження тривають. Багато робіт спираються на неконтрольоване машинне навчання. Розмічені дані для цього не потрібні.

Зокрема, у роботі Neri M., Baldoni S. [1] запропоновано використання автоенкодерів із представленням мережевого трафіку у вигляді зображень, що дозволяє ефективно виявляти аномальні патерни без попереднього знання про типи атак. Питання підвищення довіри до результатів машинного навчання та їх інтерпретації розглядаються у дослідженні Singh K., Kashyap A., Cherukuri A. K. [2], де для пояснення рішень моделей у зашифрованому трафіку використовується метод SHAP.

Напрям адаптивного аналізу мережевих потоків представлено в роботі Miguel-Diez A. та співавт. [3], у якій досліджується застосування неконтрольованого онлайн-навчання для динамічних мережевих середовищ. Для врахування часових залежностей у трафіку Noonari N. та співавт. [4] пропонують багатомасштабні згорткові LSTM-мережі з використанням трансферного навчання, орієнтовані на стільникові мережі.

Окрему групу робіт присвячено питанням ефективності та практичної реалізації систем виявлення аномалій. Так, Chen Z. та співавт. [5] аналізують можливості детектування аномалій в інфраструктурі машинного навчання на основі апаратної телеметрії, тоді як Benmachiche A., Rais K., Slimi H. [6] зосереджуються на методах машинного навчання реального часу для вбудованих систем. У контексті мереж наступних поколінь (6G) Rzym G., Masny A., Chołda P. [7] досліджують використання динамічної телеметрії та глибоких нейронних мереж у програмно-керованих мережах.

Комплексні підходи до проєктування та оцінки систем виявлення мережевих аномалій на основі машинного навчання наведено у роботі Schummer P. та співавт. [8], а методи раннього виявлення загроз розглянуто у дослідженні Thwaini M. H. [9]. Узагальнюючий огляд сучасних методів машинного навчання для виявлення мережевих аномалій представлено у роботі Mahmood N. та співавт. [10].

Є прогрес у застосуванні методів машинного навчання. Вимоги практичних систем при цьому не знімаються. Перша вимога – висока точність детектування. Друга – адаптивність до динамічних характеристик мобільного трафіку. Третій обмежувач – прийнятна обчислювальна складність. У системах реального часу ці параметри мають виконуватися одночасно.

Викладення основного матеріалу дослідження

Основним завданням даного дослідження є розробка моделі, здатної ідентифікувати аномальні патерни у потоках мережевого трафіку мобільних пристроїв в умовах відсутності повної інформації про типи можливих атак. З урахуванням гетерогенності мобільного трафіку та високої динаміки його характеристик у роботі застосовано підходи неконтрольованого машинного навчання, які не потребують розмічених наборів даних і є придатними для виявлення раніше невідомих загроз, що узгоджується з результатами, наведеними у [1, с. 2].

Для забезпечення практичної придатності запропонованої моделі окрему увагу приділено інтерпретованості результатів детектування, що є важливим чинником довіри до систем машинного навчання в задачах кіберзахисту, як зазначено в роботі [2, с. 3]. Аналіз трафіку здійснюється на рівні мережевих потоків, що дозволяє узагальнювати поведінкові характеристики пристроїв і зменшувати чутливість до шуму, характерного для пакетного рівня, відповідно до підходів, описаних у [3, с. 4].

Для експериментальної перевірки ефективності запропонованого підходу було розроблено програмний модуль мовою Python. У роботі використано рекурентні нейронні мережі типу LSTM (Long Short-Term Memory), які демонструють високу ефективність при аналізі часових рядів мережевого трафіку та дозволяють враховувати часові залежності між послідовностями переданих даних, що підтверджується результатами досліджень [4, с. 2].

Крім того, під час реалізації моделей було враховано апаратні обмеження середовищ виконання, зокрема вимоги до швидкодії та обчислювальних ресурсів, що є критичним для систем реального часу. Відповідно до рекомендацій, наведених у [5, с. 3], програмну реалізацію оптимізовано з метою забезпечення прийнятної продуктивності без суттєвої втрати точності детектування.

У процесі дослідження було розроблено алгоритмічну схему автоматизованого виявлення аномалій у мережевому трафіку мобільних пристроїв на основі методу Isolation Forest для виявлення аномалій. Цей метод був обраний як базовий завдяки його швидкодії, що важливо для систем реального часу, як зазначають Benmachiche A., Rais K., Slimi H. [6, с. 2]. Запропонований підхід реалізує послідовний конвеєр обробки даних, що включає попередню підготовку трафіку, формування векторів ознак, навчання моделі та прийняття рішень щодо наявності аномальної поведінки. Загальна логіка функціонування системи представлена у вигляді блок-схеми (рис. 1).

Як видно з рис. 1, на першому етапі виконується збір та попередня обробка мережевого трафіку, що передбачає очищення даних і нормалізацію ознак з метою усунення масштабних відмінностей між параметрами. Далі формується багатовимірний вектор ознак, який використовується для навчання моделі неконтрольованого навчання Isolation Forest.

Модель навчено – далі рахують оцінку аномальності. Розрахунок виконують для кожного мережевого з'єднання. Потім застосовують порогове значення. За ним трафік класифікують як нормальний або аномальний. Фінал – аналіз результатів. Ефективність алгоритму оцінюють за precision, recall та F1-мірою.

У контексті мереж майбутнього покоління (6G), як зазначають Rzym G., Masny A., Chołda P. [7, с. 384], важливим є використання динамічної телеметрії, що забезпечує безперервний моніторинг стану мережі в режимі реального часу. Зазначений підхід може бути інтегрований у запропоновану у даній роботі систему виявлення аномалій на етапі збору та попередньої обробки мережевих даних.

Schummer P., del Rio A., Serrano J., Jimenez D., Sánchez G., Llorente Á. [8, с. 2975] підкреслюють визначальну роль етапу оцінювання моделей машинного навчання, зокрема з використанням кількісних метрик ефективності. У межах даного дослідження оцінювання здійснюється на основі стандартних показників якості, що дозволяє об'єктивно порівняти результати роботи досліджуваних алгоритмів.

У табл. 1 наведено порівняльні характеристики ключових параметрів моделей машинного навчання. Таблиця узагальнює результати експериментального дослідження ефективності алгоритмів для задачі виявлення аномалій у мобільному мережевому трафіку.

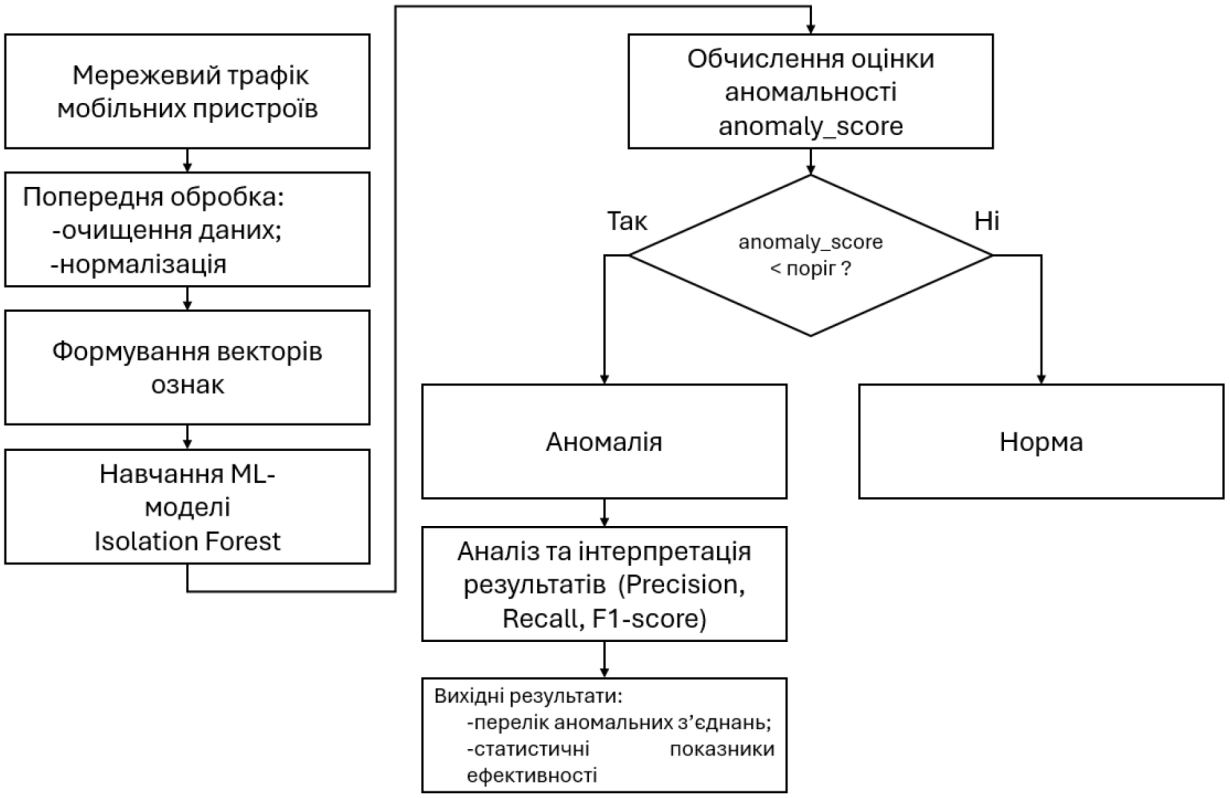

Рис. 1. Блок-схема процесу автоматизованого виявлення аномалій у мережевому трафіку мобільних пристроїв на основі алгоритму Isolation Forest

Джерело: авторська розробка

Таблиця 1

Порівняльні характеристики алгоритмів машинного навчання для виявлення аномалій у мобільному трафіку

Алгоритм	Точність (%)	Час навчання (с)	F1-міра	Складність
Isolation Forest	94.2	12.4	0.91	Середня
One-Class SVM	91.8	45.3	0.88	Висока
LSTM Autoencoder	96.1	120.7	0.94	Висока
Random Forest	92.5	18.2	0.89	Середня
K-Means	87.3	8.1	0.82	Низька

Джерело: авторська розробка

Як видно з табл. 1, алгоритм LSTM Autoencoder демонструє найвищу точність виявлення аномалій (96,1 %), однак характеризується підвищеними обчислювальними витратами та більшим часом навчання. Водночас метод Isolation Forest забезпечує кращий компроміс між точністю та швидкодією, що робить його придатним для застосування в умовах обмежених обчислювальних ресурсів.

З огляду на необхідність раннього виявлення загроз, на що вказує Thwaini M. H. [9, с. 72], у межах дослідження було проаналізовано кілька алгоритмів машинного навчання з різними характеристиками. Узагальнюючий огляд, наведений у роботі Mahmood N., Hussein D. H., Askar S., Ibrahim M. A. [10, с. 4], підтверджує відсутність універсального підходу до задачі виявлення мережевих аномалій, що обґрунтовує доцільність порівняльного аналізу моделей.

Висновки

Результати дослідження підтверджують доцільність застосування методів машинного навчання для виявлення аномалій у мережевому трафіку мобільних пристроїв.

У роботі виконано аналіз і порівняння підходів на основі LSTM Autoencoder та алгоритму Isolation Forest, які дають змогу виявляти раніше невідомі загрози без використання великих розмічених наборів даних. Експериментальні результати показали, що LSTM Autoencoder забезпечує найвищу точність детектування аномалій (96,1 %), однак потребує значних обчислювальних ресурсів і тривалого часу навчання, що може обмежувати застосування в системах реального часу.

Водночас алгоритм Isolation Forest забезпечує оптимальний баланс між швидкодією та точністю (94,2 %), що робить його доцільним для використання в умовах обмежених обчислювальних ресурсів.

Подальші дослідження доцільно спрямувати на розроблення гібридних підходів, які поєднуватимуть високу точність глибоких нейронних мереж із швидкодією та ресурсною ефективністю класичних алгоритмів машинного навчання.

Список використаної літератури

1. Neri M., Baldoni S. Unsupervised Network Anomaly Detection with Autoencoders and Traffic Images (Version 1). arXiv. 2025. DOI: https://doi.org/10.48550/arXiv.2505.16650 URL: https://arxiv.org/abs/2505.16650

2. Singh K., Kashyap A., Cherukuri A. K. Interpretable Anomaly Detection in Encrypted Traffic Using SHAP with Machine Learning Models (Version 1). arXiv. 2025. DOI: https://doi.org/10.48550/arXiv.2505.16261 URL: https://arxiv.org/abs/2505.16261

3. Miguel-Diez A., Campazas-Vega A., Guerrero-Higueras Á. M., Álvarez-Aparicio C., Matellán-Olivera V. Anomaly detection in network flows using unsupervised online machine learning (Version 1). arXiv. 2025. DOI: https://doi.org/10.48550/arXiv.2509.01375 URL: https://arxiv.org/abs/2509.01375

4. Noonari N., Corujo D., Aguiar R. L., Ferrao F. J. Multi-Scale Convolutional LSTM with Transfer Learning for Anomaly Detection in Cellular Networks (Version 1). arXiv. 2024. DOI: https://doi.org/10.48550/arXiv.2410.03732 URL: https://arxiv.org/abs/2410.03732

5. Chen Z., Chien S. W. D., Qian P., Zilberman N. Detecting Anomalies in Machine Learning Infrastructure via Hardware Telemetry (Version 1). arXiv. 2025. DOI: https://doi.org/10.48550/arXiv.2510.26008 URL: https://arxiv.org/abs/2510.26008

6. Benmachiche A., Rais K., Slimi H. Real-Time Machine Learning for Embedded Anomaly Detection (Version 1). arXiv. 2025. DOI: https://doi.org/10.48550/arXiv.2512.19383 URL: https://arxiv.org/abs/2512.19383

7. Rzym G., Masny A., Chołda P. Dynamic Telemetry and Deep Neural Networks for Anomaly Detection in 6G Software-Defined Networks. Electronics. 2024. Vol. 13, No. 2. DOI: https://doi.org/10.3390/electronics13020382 URL: https://www.mdpi.com/2079-9292/13/2/382

8. Schummer P., del Rio A., Serrano J., Jimenez D., Sánchez G., Llorente Á. Machine Learning-Based Network Anomaly Detection: Design, Implementation, and Evaluation. AI. 2024. Vol. 5, No. 4. P. 143. DOI: https://doi.org/10.3390/ai5040143 URL: https://www.mdpi.com/2673-2688/5/4/143

9. Thwaini M. H. Anomaly Detection in Network Traffic using Machine Learning for Early Threat Detection. Digital Medicine. 2022. DOI: 10.56294/dm202272 URL: https://is.gd/T2JCaF

10. Mahmood N., Hussein D. H., Askar S., Ibrahim M. A. Machine Learning for Network Anomaly Detection: A Review. International Journal of Computer Science. 2025. Vol. 14, No. 1. DOI: 10.33022/ijcs.v14i1.4703 URL: https://is.gd/mYVrU5

References

1. Neri, M., & Baldoni, S. (2025). Unsupervised Network Anomaly Detection with Autoencoders and Traffic Images (Version 1). arXiv. DOI: https://doi.org/10.48550/arXiv.2505.16650 Retrieved from https://arxiv.org/abs/2505.16650 [in English].

2. Singh, K., Kashyap, A., & Cherukuri, A. K. (2025). Interpretable Anomaly Detection in Encrypted Traffic Using SHAP with Machine Learning Models (Version 1). arXiv. DOI: https://doi.org/10.48550/arXiv.2505.16261 Retrieved from https://arxiv.org/abs/2505.16261 [in English].

3. Miguel-Diez, A., Campazas-Vega, A., Guerrero-Higueras, Á. M., Álvarez-Aparicio, C., & Matellán-Olivera, V. (2025). Anomaly detection in network flows using unsupervised online machine learning (Version 1). arXiv. DOI: https://doi.org/10.48550/arXiv.2509.01375 Retrieved from https://arxiv.org/abs/2509.01375 [in English].

4. Noonari, N., Corujo, D., Aguiar, R. L., & Ferrao, F. J. (2024). Multi-Scale Convolutional LSTM with Transfer Learning for Anomaly Detection in Cellular Networks (Version 1). arXiv. DOI: https://doi.org/10.48550/arXiv.2410.03732 Retrieved from https://arxiv.org/abs/2410.03732 [in English].

5. Chen, Z., Chien, S. W. D., Qian, P., & Zilberman, N. (2025). Detecting Anomalies in Machine Learning Infrastructure via Hardware Telemetry (Version 1). arXiv. DOI: https://doi.org/10.48550/arXiv.2510.26008 Retrieved from https://arxiv.org/abs/2510.26008 [in English].

6. Benmachiche, A., Rais, K., & Slimi, H. (2025). Real-Time Machine Learning for Embedded Anomaly Detection (Version 1). arXiv. DOI: https://doi.org/10.48550/arXiv.2512.19383 Retrieved from https://arxiv.org/abs/2512.19383 [in English].

7. Rzym, G., Masny, A., & Chołda, P. (2024). Dynamic Telemetry and Deep Neural Networks for Anomaly Detection in 6G Software-Defined Networks. Electronics, 13(2). DOI: https://doi.org/10.3390/electronics13020382 Retrieved from https://www.mdpi.com/2079-9292/13/2/382 [in English].

8. Schummer, P., del Rio, A., Serrano, J., Jimenez, D., Sánchez, G., & Llorente, Á. (2024). Machine Learning-Based Network Anomaly Detection: Design, Implementation, and Evaluation. AI, 5(4), 2967–2983. DOI: https://doi.org/10.3390/ai5040143 Retrieved from https://www.mdpi.com/2673-2688/5/4/143 [in English].

9. Thwaini, M. H. (2022). Anomaly Detection in Network Traffic using Machine Learning for Early Threat Detection. Digital Medicine. DOI: 10.56294/dm202272 Retrieved from https://is.gd/T2JCaF [in English].

10. Mahmood, N., Hussein, D. H., Askar, S., & Ibrahim, M. A. (2025). Machine Learning for Network Anomaly Detection: A Review. International Journal of Computer Science, 14(1). DOI: 10.33022/ijcs.v14i1.4703 Retrieved from https://is.gd/mYVrU5 [in English].

Дата першого надходження статті до видання: 10.01.2026
Дата прийняття статті до друку після рецензування: 17.02.2026
Дата публікації (оприлюднення) статті: 30.04.2026