УДК 004. 056.55:004.312.2 **doi: 10.26906/SUNZ.2025.3.193-198**

Є. Ю. Глоба, В. Р. Смірнов, М. С. Нараєвський, В. М. Федорченко

Харківський національний університет радіоелектроніки, Харків, Україна

МЕТОД ВИЯВЛЕННЯ АНОМАЛІЙ В КОРПОРАТИВНІЙ МЕРЕЖІ

Анотація. Актуальність дослідження полягає в тому, що існуючі методи виявлення аномалій нерідко демонструють недостатню точність та оперативність у реальних умовах експлуатації. **Об'єкт дослідження**: мережевий трафік корпоративної інформаційної системи, який аналізується з метою виявлення аномальної поведінки, що потенційно свідчить про кібератаки, порушення політик безпеки або внутрішні загрози. **Мета статті**: розробка, реалізація та експериментальне обґрунтування методу виявлення аномалій у корпоративній мережі на основі моделі автоенкодера, який дозволяє ефективно ідентифікувати відхилення від нормального мережевого трафіку без потреби в попередньо мічених даних. **Результати дослідження**. У процесі моделювання було згенеровано дані, які імітують типову та аномальну активність у мережі. Гістограми та boxplot вказують на те, що нормальні зразки характеризуються низьким і стабільним значенням середньоквадратичної похибки, тоді як аномальні – мають виражені відхилення. ROC-крива з AUC = 1.00 підтверджує, що модель здатна безпомилково розрізняти обидві категорії. Матриця плутанини показала практично ідеальне співвідношення між передбаченими та фактичними мітками, що підтверджує високу точність і чутливість моделі. Тренувальна крива свідчить про стабільне навчання без перенавчання, а обраний динамічний поріг дозволив знизити рівень хибнопозитивних спрацювань до мінімуму. Запропонований підхід може бути інтегрований у системи моніторингу безпеки для виявлення нетипових активностей у режимі реального часу. **Висновки.** Проведені експерименти продемонстрували високу точність класифікації, чітке розділення нормальних і аномальних зразків за похибкою реконструкції, а також стабільність навчання моделі. Отримані результати підтверджують доцільність використання глибокого навчання для автоматизованого моніторингу мережевої безпеки, а запропонований метод може бути успішно застосований у реальних умовах корпоративної IT-інфраструктури для виявлення інцидентів безпеки в режимі реального часу.

Ключові слова: корпоративна мережа. виявлення аномалій, корпоративна мережа, мережевий трафік, автоенкодер, машинне навчання, кібербезпека, реконструктивна похибка, навчання без вчителя, нейронна мережа, профілювання поведінки, інформаційна безпека.

Вступ

Постановка проблеми. Сучасні корпоративні мережі є невід'ємною складовою цифрової інфраструктури будь-якої компанії, незалежно від її розміру та сфери діяльності. Вони забезпечують ключові процеси, такі як передача інформації, підтримка операційних та управлінських систем, комунікація між підрозділами, а також зовнішню взаємодію з клієнтами й партнерами. В умовах постійного зростання обсягу переданих і оброблюваних даних, а також ускладнення топології мереж, зростають ризики виникнення небезпечних ситуацій, зокрема аномальних подій, що можуть становити загрозу безпеці та стабільності інформаційних систем підприємств. Аномалії у корпоративних мережах проявляються як нестандартні або нетипові шаблони поведінки мережевих вузлів, користувачів чи програмного забезпечення. Вони можуть бути пов’язані як з потенційними кіберзагрозами (атаками на систему, вторгненням, зловмисним програмним забезпеченням, несанкціонованим доступом), так і з різного роду технічними проблемами (відмовами обладнання, помилками програмного забезпечення, неефективністю налаштувань чи конфігурацій). Особливість таких аномалій полягає у тому, що їх важко своєчасно розпізнати за допомогою стандартних інструментів моніторингу й аналізу трафіку, оскільки вони часто не вписуються у попередньо задані моделі та правила, і можуть залишатися прихованими у великих обсягах даних. Загрози, що виникають у результаті аномальних подій, мають серйозні наслідки, серед яких особливо небезпечними є витоки конфіденційних даних, порушення працездатності бізнес-процесів, зниження продуктивності співробітників та суттєві репутаційні втрати. Тому актуальним завданням є своєчасне і точне виявлення подібних подій з метою оперативного реагування та зменшення негативних наслідків. Сьогодні на ринку представлено велику кількість рішень для моніторингу мережевого трафіку, проте існуючі методи виявлення аномалій демонструють ряд суттєвих недоліків. До них належать висока кількість хибних спрацювань (помилкових позитивних або негативних сигналів), труднощі адаптації до змінних умов експлуатації, а також недостатня ефективність при роботі з великими обсягами даних у реальному часі. Ці недоліки стимулюють пошук нових підходів і методів, що дозволять ефективно виявляти та класифікувати аномалії в корпоративних мережах.

Аналіз останніх досліджень і публікацій. У сучасному науково-практичному середовищі проблема виявлення аномалій у корпоративних мережах стала однією з найбільш обговорюваних у сфері кібербезпеки та інтелектуального аналізу даних. Розвиток мережевої інфраструктури, зростання обсягів трафіку, поширення хмарних сервісів і віддаленої роботи значно ускладнили завдання виявлення загроз у реальному часі. Саме тому останні роки стали свідками стрімкого зростання інтересу до методів виявлення аномалій на основі штучного інтелекту, машинного навчання, гібридних та контекстно-орієнтованих підходів. Базові методи, наприклад, сигнатурний аналіз, добре зарекомендували себе у виявленні відомих загроз, однак вони нездатні розпізнавати нові або обфусковані атаки. У відповідь, з'явилися численні публікації, в яких обґрунтовується перехід до статистичних методів та автоматичної побудови профілю нормальної поведінки.

Стаття [1] присвячена огляду сучасних методів навчання без вчителя для виявлення аномалій у ча-

сових рядах. У ній розглядаються досягнення в галузі збору та аналізу даних, які призвели до накопичення великої кількості послідовних спостережень. Стаття підкреслює важливість видобування знань із таких даних, зокрема виявлення викидів або аномалій, що можуть сигналізувати про помилки, порушення чи значущі події. В роботі [2] автори проводять порівняння класичних алгоритмів машинного навчання і роблять висновок, що гібридні методи (машинне навчання + фільтрація ознак) перевищують по точності базові евристики. Особливу увагу привертають методи глибокого навчання, зокрема автоенкодери, LSTM-моделі та рекурентні нейронні мережі. Вони дозволяють моделювати складну часову структуру мережевого трафіку та виявляти аномалії без необхідності міток. У дослідженні [3] описано Kitsune – систему, що використовує ансамбль автоенкодерів для виявлення аномалій на основі flow-даних. Вона працює в реальному часі з низькою обчислювальною складністю, що робить її придатною для корпоративних мереж. Стаття [4] присвячена огляду сучасних підходів до виявлення аномалій у кіберфізичних системах. Оскільки такі системи глибоко інтегровані в інфраструктуру та автоматизацію сучасного світу, вони особливо вразливі до різноманітних загроз, включно з поломками обладнання, помилками сенсорів, збоями зв'язку та кібератаками. Аномалії в кіберфізичних системах можуть свідчити про несподівані порушення нормального функціонування, тому їх своєчасне виявлення є критично важливим для запобігання шкоді, зупинкам сервісів або навіть небезпечним аваріям. У статті систематизовано та порівняно підходи, які використовуються для виявлення таких аномалій. Зокрема, розглядаються методи машинного навчання, глибокого навчання, математичного моделювання. Також залишається проблема оцінки моделей виявлення аномалій в роботі [5] автори критикують популярні набори даних за низьку репрезентативність і пропонують створення нових сценаріїв на основі реального корпоративного трафіку.

Метою роботи є розробка удосконаленого методу виявлення аномалій у корпоративних мережах, який базується на сучасних підходах у сфері машинного навчання та штучного інтелекту. Запропонований метод має забезпечити високу точність, швидкодію, адаптивність до динамічно змінних умов і низький рівень хибних спрацювань, що є критично важливим для оперативного прийняття рішень та забезпечення надійності функціонування інформаційних систем підприємств.

Основний матеріал

Проблема виявлення аномалій у корпоративних мережах уже давно є предметом активного дослідження в галузі інформаційної безпеки. У різних підходах до її вирішення застосовуються як класичні методи аналізу даних, так і сучасні алгоритми машинного навчання та штучного інтелекту [6]. Основна мета таких методів полягає у тому, щоб вчасно виявити нетипову поведінку в мережевому трафіку, яка потенційно може свідчити про загрозу, або про критичні відхилення в роботі систем. У цьому контексті важливою є здатність системи відокремлювати нормальну активність від аномальної, навіть за відсутності явної сигнатури загрози.

Традиційні методи, зокрема сигнатурний аналіз [7], орієнтовані на виявлення вже відомих шаблонів атак, що зберігаються у базі даних сигнатур. Вони ефективні у виявленні відомих загроз, однак повністю безсилі перед новими, раніше невідомими формами атак, особливо так званими атаками нульового дня. Крім того, сигнатурні системи вимагають постійного оновлення баз даних і часто створюють велике навантаження на систему безпеки через значну кількість перевірок.

У противагу сигнатурним методам, аномаліє-орієнтовані методи ґрунтуються на побудові моделей «нормальної» поведінки мережі, з якими порівнюється поточна активність. Порушення встановленого шаблону вважається підозрілим і може бути маркером аномалії. Такі методи вимагають етапу навчання або збору статистики, проте вони здатні виявляти нові або раніше не зафіксовані загрози. Найчастіше вони реалізуються у вигляді статистичного аналізу або класифікаційних алгоритмів.

Серед статистичних методів найбільш поширеними є контроль середніх значень, дисперсії та інтервальних оцінок. У цьому випадку вся вхідна мережна інформація аналізується щодо відхилення від визначених статистичних параметрів. Висока точність можлива лише за умови стабільного трафіку та низької варіативності, що у динамічних корпоративних мережах спостерігається рідко.

Останніми роками зростає популярність методів машинного навчання, які дозволяють моделювати складні залежності між параметрами мережевого трафіку. Серед них виділяють класифікаційні підходи, алгоритми кластеризації та нейронні мережі. Такі моделі демонструють хорошу адаптивність до змін середовища, здатність до самооновлення та виявлення прихованих шаблонів, однак їх застосування потребує ретельної підготовки даних, високих обчислювальних ресурсів та кваліфікованого налаштування. У контексті корпоративної мережі важливо не лише виявити факт аномалії, а й локалізувати джерело загрози та оцінити її потенційну критичність. Для цього часто застосовують гібридні підходи, які поєднують сигнатурні методи з алгоритмами машинного навчання. Такі системи дозволяють ефективно працювати як з уже відомими шаблонами атак, так і з новими, невідомими загрозами, а також знижують ризик хибних спрацювань завдяки багаторівневій перевірці підозрілих дій.

Незважаючи на значний прогрес у цій галузі, всі існуючі методи мають певні обмеження, як у точності класифікації, так і в масштабованості та споживанні ресурсів. Це зумовлює необхідність розробки нових методів, які враховують специфіку корпоративного середовища, динаміку топології мережі, різнорідність трафіку та потребу в реальному часі виявляти аномальні дії без критичного впливу на продуктивність системи. У відповідь на виявлені обмеження традиційних підходів до виявлення аномалій постає потреба у створенні нового, більш гнучкого та адаптивного методу, здатного ефективно працювати в умовах

складного та динамічного корпоративного середовища. Пропонований у даній роботі метод поєднує переваги сучасних моделей глибокого навчання з принципами динамічного моделювання трафіку, забезпечуючи високу чутливість до нетипової активності при зниженні рівня хибнопозитивних спрацювань.

Основна ідея розробленого підходу полягає у використанні автоенкодера – симетричної нейромережевої архітектури, що навчається відтворювати нормальні шаблони поведінки мережі. Після етапу навчання мережа здатна виявляти відхилення на основі величини похибки реконструкції: чим більша похибка між вхідним і відтвореним трафіком, тим вищою є ймовірність того, що спостерігається аномалія. Завдяки цьому автоенкодер не потребує наявності явного маркування даних і може функціонувати в умовах слабкого або неповного навчального набору.

Однак, оскільки корпоративний трафік є багатошаровим, контекстно-залежним і змінним у часі, було запропоновано доповнити архітектуру рекурентними нейронними мережами типу LSTM, які дозволяють моделі зберігати інформацію про попередні стани та краще моделювати часову динаміку трафіку. У поєднанні з автоенкодером LSTM-механізм надає глибшу контекстуалізацію мережевої активності та дозволяє розрізняти короткочасні флуктуації від справжніх аномалій, пов'язаних з порушенням політик безпеки або зовнішнім впливом.

Ще однією важливою складовою методу є етап динамічного профілювання. На відміну від статичних моделей, що створюються лише один раз під час початкового навчання, у запропонованому підході профіль мережевої активності постійно оновлюється з урахуванням змін у структурі, користувацькій поведінці та політиках доступу. Це досягається через віконну адаптацію вагових коефіцієнтів моделі на основі нещодавніх «безпечних» даних, що дозволяє зменшити ймовірність хибних спрацювань у разі легітимних змін у системі (наприклад, нові користувачі, додаткові сервіси або змінений графік роботи).

З метою зменшення навантаження на обчислювальні ресурси та забезпечення масштабованості метод підтримує попереднє агрегаційне перетворення вхідного трафіку – зокрема, перехід від сирих пакетних даних до flow-структур, що узагальнюють інформацію про з’єднання (час початку, тривалість, IP-адреси, протоколи, обсяги переданої інформації тощо). Такий підхід суттєво знижує розмірність вхідного простору без втрати критично важливої інформації. Загальна структура методу включає п’ять етапів:

1) збір і попередня обробка мережевого трафіку (очищення, нормалізація, перетворення до flow-формату);

2) формування початкового профілю нормальної поведінки на основі історичних даних;

3) навчання автоенкодера з рекурентними компонентами;

4) безперервне моніторинг та обчислення похибки реконструкції з виявленням аномалій на основі адаптивного порогу;

5) оновлення профілю поведінки за допомогою віконного механізму, базується на довірених даних.

Запропонований метод поєднує глибину аналізу нейронних мереж із гнучкістю адаптивного моделювання, зберігаючи високу точність в умовах динамічних змін мережевої поведінки. У наступних розділах буде розглянуто особливості збору даних, реалізацію прототипу та результати експериментальної перевірки ефективності запропонованого методу.

На першому етапі система отримує первинні дані із корпоративної мережі у вигляді сирого мережевого трафіку. Це можуть бути: пакети даних; NetFlow або IPFIX-дані; системні логи; журнали з міжмережевих екранів, проксі-серверів, SIEM-платформ тощо. Дані піддаються первинній обробці: очищенню, нормалізації та агрегації. На другому етапі з нормалізованих історичних даних створюється модель "нормальної" поведінки мережі: визначаються типові характеристики з'єднань (тривалість, обсяг, протоколи, частота); обчислюються статистичні параметри (середнє, дисперсія); формується база "нормальних" шаблонів трафіку. На третьому етапі на основі початкового профілю проводиться тренування автоенкодера або гібридної моделі: модель навчається відтворювати нормальні шаблони без аномалій; мінімізується похибка між входом і виходом; фіксується рівень допустимої похибки реконструкції. Нейромережа працює без потреби в мічених даних, що зручно для реального застосування. Цей профіль буде використовуватись як еталон для порівняння в майбутньому.

Четвертий етап передбачає подання мережевих даних на вхід навченої моделі у реальному часі: обчислюється похибка реконструкції; якщо похибка перевищує динамічний поріг – трафік визнається аномальним; система фіксує та передає сигнал про інцидент до модуля реагування.

Рішення може бути додатково перевірено з урахуванням контексту (час доби, тип користувача, критичність ресурсу тощо).

На останньому етапі у разі, якщо спостереження не класифікується як аномальне, або після ручної перевірки інциденту фахівцем, дані можуть бути використані для: оновлення моделі "нормальної" поведінки; поступового донавчання моделі; адаптації до змін у політиках, користувачах чи сервісах.

Це забезпечує живу адаптацію системи до нових умов, не втрачаючи контроль за безпекою. Цикл повторюється постійно, що дає змогу системі не лише виявляти потенційно шкідливу активність, а й навчатися на основі нових безпечних сценаріїв поведінки. Це суттєво знижує кількість хибних спрацювань і підвищує ефективність аналізу у складних умовах великомасштабної мережі.

Практична реалізація запропонованого методу виявлення аномалій у корпоративній мережі потребує поетапного конструювання програмної системи, яка поєднує механізми збору, обробки й аналізу даних, а також нейромережеву модель для визначення аномалій. Далі опишемо архітектуру програмного прототипу, обрані засоби розробки, ключові компоненти системи та особливості її інтеграції з корпоративною інфраструктурою. Розроблений прототип складається з п’яти логічних модулів, що працюють узгоджено в єдиному інформаційному потоці:

- модуль збору даних відповідає за перехоплення та збереження мережевого трафіку або його агрегованих представлень;

- модуль попередньої обробки здійснює очистку, нормалізацію та агрегацію трафіку в набори ознак, що можуть бути подані на вхід моделі;

- нейромережева підсистема реалізує автоенкодер або автоенкодер з LSTM-компонентами для виявлення відхилень на основі похибки реконструкції.

- модуль моніторингу та логування постійно аналізує результати роботи нейромережі, фіксує потенційні аномалії та створює звітні файли;

- модуль адаптації забезпечує періодичне оновлення поведінкового профілю мережі та за потреби ініціює перенавчання моделі на нових даних.

Для реалізації програмного прототипу було обрано мову Python, як найбільш придатну для швидкого прототипування систем машинного навчання, з використанням таких бібліотек і фреймворків:

- Scapy для перехоплення та парсингу мережевого трафіку;

- Pandas, NumPy для роботи з табличними даними та масивами ознак;

- Scikit-learn для попереднього аналізу даних та побудови базових моделей класифікації або кластеризації;

- TensorFlow для реалізації нейронної мережі автоенкодера з можливістю додавання LSTM-шарів;

- Loguru для ведення журналів подій та візуалізації результатів.

Для підтримки масштабованості реалізація передбачає обробку даних у потоковому режимі, використовуючи черги повідомлень.

Нейромережевий автоенкодер побудовано за класичною симетричною структурою: вхідний шар: приймає набір нормалізованих ознак трафіку (кількість пакетів, середня затримка, протокол, тривалість сесії тощо); блок кодування: складається з кількох повнозв'язних шарів зі зменшенням розмірності; декодувальний блок: симетрично відтворює вихідний вектор;

- LSTM-шари: додаються перед/після кодування для врахування часової залежності подій.

Функцією втрат є середньоквадратична похибка між вхідними й відтвореними ознаками. Під час розгортання модель виводить поточне значення похибки реконструкції, що порівнюється з адаптивним порогом для виявлення аномалій.

Реалізований прототип демонструє можливість повноцінного виявлення аномалій у корпоративному середовищі в реальному часі, з урахуванням часової динаміки поведінки та можливістю гнучкої адаптації до нових умов роботи.

Для перевірки роботи прототипу було обрано реалістичний псевдо-набір даних, стилізований під NSL-KDD (з 10+ ознаками: тривалість, кількість байтів, кількість з'єднань тощо), і на цьому наборі вчилася модель для локального тестування алгоритму. Згенерований набір даних, який імітує поведінку мережевого трафіку в корпоративному середовищі, містить 1000 нормальних зразків (позначені як label = 0); 50 аномальних зразків (label = 1), які мають значні відхилення у ключових параметрах. Для реалізації моделі автоенкодеру використовувалося середовище Google Colab.

Аналіз графіків на рис. 1, побудованих за результатами роботи моделі виявлення аномалій у корпоративній мережі, дає змогу зробити низку важливих спостережень, що свідчать про ефективність використаного підходу. На першому графіку зображено ROC-криву, яка відображає співвідношення між рівнем хибнопозитивних та істинно позитивних спрацювань. Крива має чітко виражену форму, близьку до ідеальної, з різким підйомом у лівому верхньому куті. Площа під кривою становить 1.00, що свідчить про винятково високу точність моделі: вона безпомилково розрізняє нормальні та аномальні зразки. Такий результат підтверджує, що автоенкодер навчився точно відтворювати лише характерні шаблони нормальної поведінки, а будь-які відхилення визначаються як потенційно підозрілі з високим рівнем впевненості.

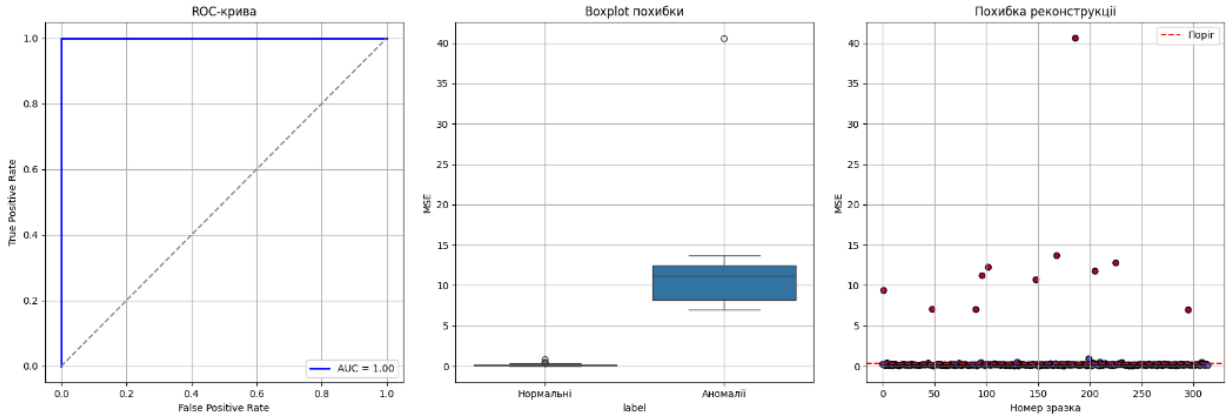

Рис. 1. Результати роботи моделі

Другий графік (boxplot похибок реконструкції) наочно демонструє контраст між нормальними та аномальними зразками. Для нормальних спостерігається надзвичайно щільне скупчення похибок у межах мінімальних значень, тоді як для аномальних характерна значна варіативність і розкид. Медіана похибки для аномалій істотно перевищує навіть найвищі значення серед нормальних, що свідчить про суттєву відмінність між цими двома класами. Це вказує на те, що модель здатна формувати стійкий профіль "нормального" трафіку, і будь-які відхилення, навіть незначні, від нього миттєво ідентифікуються

як аномальні. Таке розділення між класами на рівні статистики похибки означає, що обраний метод має хорошу здатність до генералізації. Третій графік, що показує розподіл похибок реконструкції по кожному зразку, підтверджує попередні спостереження. Більшість точок, що відповідають нормальним зразкам, зосереджені в нижній частині графіка і розміщені нижче встановленого порогу, позначеного червоною пунктирною лінією. Водночас аномальні зразки мають виражені викиди – їхні похибки значно перевищують порогове значення. Це свідчить про чітке розмежування між звичайною та підозрілою поведінкою. Видимість таких різких відхилень ще раз підкреслює здатність автоенкодера не лише точно реконструювати "правильні" дані, а й виявляти нетипові шаблони, які виходять за межі навчального простору.

У сукупності результати, візуалізовано на графіках, підтверджують ефективність розробленого підходу до виявлення аномалій. Модель демонструє здатність до точного виявлення навіть одиничних відхилень без потреби уявно маркованих даних або надмірного втручання користувача. Це відкриває широкі перспективи її практичного застосування в умовах корпоративних мереж, де висока швидкість реагування на нестандартні ситуації та мінімізація хибнопозитивних спрацювань є критично важливими для підтримання кібербезпеки та цілісності інформаційної інфраструктури.

По графіках на рис. 2 можна зробити висновок, що матриця плутанини демонструє практично ідеальну класифікацію: усі нормальні зразки визначено правильно, а більшість аномальних також виявлено.

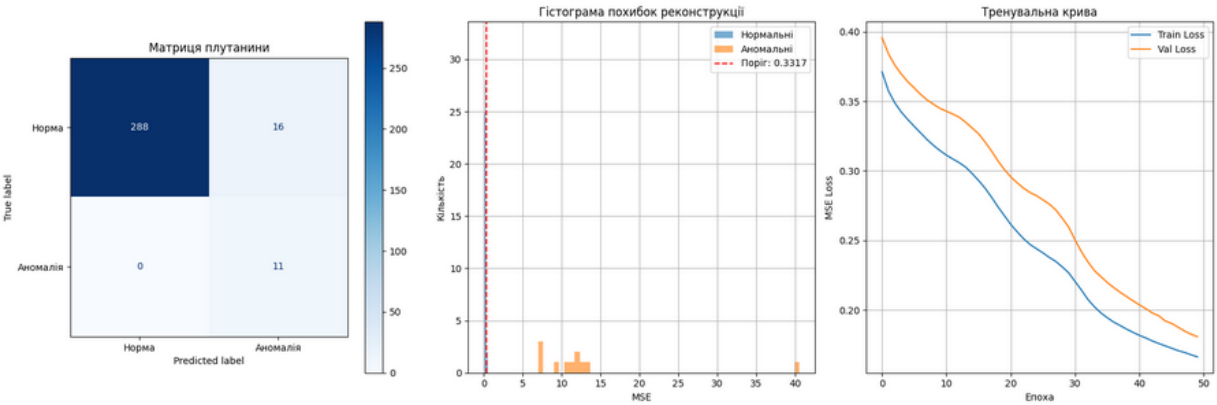

Рис. 2. Результати роботи

Це підтверджує надійність моделі. Гістограма похибок реконструкції показує чітке розділення: нормальні зразки мають низьку похибку, а аномальні – значно вищу. Обраний поріг (≈ 0.0237) добре відокремлює ці класи. Тренувальна крива ілюструє поступове зниження похибки як на train-, так і на validation-наборах без перенавчання. Модель стабільно сходиться, що свідчить про ефективне навчання.

Таким чином, розроблений автоенкодер стабільно та точно виявляє аномалії у даних, з низьким рівнем помилок і високою відокремлюваністю класів.

Висновки

Обґрунтовано, реалізовано та експериментально підтверджено ефективність методу виявлення аномалій у корпоративній мережі, що базується на використанні нейронної мережі автоенкодерного типу. Такий підхід передбачає створення моделі, здатної самостійно навчатися характеристикам нормального мережевого трафіку без залучення попередньо маркованих даних. Це особливо важливо для корпоративних інформаційних систем, де швидкість реагування та адаптивність до нових загроз мають критичне значення, а обмеження в доступі до мічених прикладів ускладнюють використання традиційних методів контролю.

В основі методу лежить принцип, згідно з яким автоенкодер добре реконструює лише ті зразки, які відповідають звичним, часто повторюваним шаблонам поведінки мережі. Усі нетипові чи невідомі варіанти трафіку, що не були представлені під час навчання, викликають суттєве зростання похибки реконструкції, що і дозволяє трактувати їх як потенційно небезпечні чи сумнівні. Аналіз результатів моделювання, що включав побудову кількох графіків, показав високий рівень точності, стабільності та узгодженості роботи моделі. Модель чітко розмежовувала нормальні та аномальні зразки за допомогою порогу, визначеного на основі реконструктивної похибки. Водночас відсутність ознак перенавчання, що видно на графіку тренувальної динаміки, свідчить про здатність мережі узагальнювати поведінку без втрати точності.

Отримані результати доводять, що обраний підхід дозволяє досягти високої ефективності у завданні автоматичного виявлення відхилень у корпоративному трафіку. Перевагами методу є не лише його точність, а й гнучкість до масштабування, мінімальні вимоги до мічених даних та можливість інтеграції у вже існуючі системи інформаційної безпеки (наприклад, SIEM або SOC-платформи). Це дає підстави стверджувати, що подальше вдосконалення автоенкодерних моделей, зокрема шляхом врахування часової динаміки або впровадження контекстуальних ознак, може значно розширити спектр задач, які успішно вирішуються в рамках архітектур корпоративної безпеки. Таким чином, розроблена модель автоенкодера підтвердила свою здатність ефективно виконувати виявлення аномалій у середовищі з високим рівнем динаміки та ін-

формаційного навантаження. Вона може бути використана як основа для побудови інтелектуальних систем раннього виявлення загроз у реальних корпоративних мережах, де важливо своєчасно реагувати на будь-які нетипові дії або порушення у звичній схемі функціонування інформаційної інфраструктури.

СПИСОК ЛІТЕРАТУРИ

1. Blazquez-Gartfa, A., Conde A., Mori U., Lozano J. A review on outlier/anomaly detection in time series data, ACM Comput. Surv. Vol. 54. No. 3. 2021. DOI: http://dx.doi.org/10.1145/3444690.
2. Vaishali Bhatia; Shabnam Choudhary; K.R Ramkumar. A Comparative Study on Various Intrusion Detection Techniques Using Machine Learning and Neural Network. 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020. https://doi.org/10.1109/ICRITO48877.2020.9198008
3. Y. Mirsky, T. Doitshman, Y.Elovici, A. Shabtai. Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection. Cornell University. Computer Science,2018. 15 p. https://doi.org/10.48550/arXiv.1802.09089
4. D. Abshari, M. Sridhar. A Survey of Anomaly Detection in Cyber-Physical Systems, 2025. https://arxiv.org/html/2502.13256v1
5. M. Ring, S. Wunderlich, D. Scheuring, D. Landes, A. Hotho. A survey of network-based intrusion detection data sets. Elsevier. ScinceDirect. Computers & Security, vol. 86, 2019. P. 147-167. https://doi.org/10.1016/j.cose.2019.06.005 .
6. Flach P. A. Machine Learning: The Art and Science of Algoritms that Makes Sense of Data. Cambridge: Cambridge University Press, 2012. 291 p. https://doi.org/10.1017/CBO9780511973000 .
7. R.Abu-Zaid, A.Hammad. Streamlining Data Processing Efficiency in Large-Scale Applications: Proven Strategies for Optimizing Performance, Scalability, and Resource Utilization in Distributed Architectures. International Journal of Machine Intelligence. International Journal of Machine Intelligence for Smart Applications, 14(8), 2024. P. 31-49. https://dljournals.com/index.php/IJMISA/article/view/27 .

Received (Надійшла) 02.06.2025
Accepted for publication (Прийнята до друку) 13.08.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Глоба Єлизавета Юріївна– студентка кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Yelyzaveta Hloba – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: yelyzaveta.hloba@nure.ua; ORCID Author ID: http://orcid.org/0009-0002-2447-000X.

Смірнов Владислав Русланович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Vladyslav Smirnov – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: vladyslav.smirnov@nure.ua; ORCID Author ID: https://orcid.org/0009-0008-8997-0324.

Нараєвський Микола Сергійович – бакалавр кафедри інформаційних управляючих систем, Харківський національний університет радіоелектроніки, Харків, Україна;
Mykola Naraievskyi – bachelor, Department of Information Control System, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: mykolanaraievskyi@gmail.com; ORCID Author ID: https://orcid.org/0009-0002-4599-6975.

Федорченко Володимир Миколайович кандидат технічних наук, електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Volodymyr Fedorchenko – Candidate of Technical Sciences, Associate Professor, Associate Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: volodymyr.fedorchenko@nure.ua ; ORCID Author ID: https://orcid.org/0000-0001-7359-1460 .

Method for detecting anomalies in corporate networks

Yelyzaveta Hloba, Vladyslav Smirnov, Mykola Naraievskyi, Volodymyr Fedorchenko

Abstract. **The relevance of this study** lies in the fact that existing anomaly detection methods often demonstrate insufficient accuracy and responsiveness in real-world operational environments. **The object of research**: the network traffic of a corporate information system, which is analyzed to detect abnormal behavior potentially indicating cyberattacks, security policy violations, or internal threats. **Purpose of the article**: the development, implementation, and experimental validation of an anomaly detection method for corporate networks based on an autoencoder model, which makes it possible to effectively identify deviations from normal network traffic without the need for pre-labeled data. **Research results**. In the modeling process, synthetic data simulating both typical and anomalous network activity was generated. Histograms and boxplots indicate that normal samples are characterized by low and stable mean squared error values, while anomalous samples demonstrate significant deviations. The ROC curve with AUC = 1.00 confirms that the model can reliably distinguish between the two categories. The confusion matrix showed a nearly perfect match between predicted and actual labels, indicating high accuracy and sensitivity of the model. The training curve demonstrates stable learning without overfitting, and the selected dynamic threshold minimized the false positive rate. The proposed approach can be integrated into security monitoring systems to detect atypical activities in real time. **Conclusions.** The conducted experiments demonstrated high classification accuracy, clear separation between normal and anomalous samples based on reconstruction error, and stable model training. The obtained results confirm the feasibility of using deep learning for automated network security monitoring, and the proposed method can be successfully applied in real corporate IT infrastructures to detect security incidents in real time.

Keywords: corporate network, anomaly detection, network traffic, autoencoder, machine learning, cybersecurity, reconstruction error, unsupervised learning, neural network, behavior profiling, information security.