КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (34), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.34.1253
УДК 004.056:004.032.26:004.7:355

Загоруйко Любов Василівна
к.т.н., доцент, доцент кафедри прикладної математики та кібербезпеки
Донецький національний університет імені Василя Стуса, м.Вінниця, Україна
ORCID:0000-0002-6958-8696
l.zahoruiko@donnu.edu.ua

Луценко Алла Володимирівна
доктор філософії, старший викладач, завідувач кафедри прикладної математики та кібербезпеки
Донецький національний університет імені Василя Стуса, м.Вінниця, Україна
ORCID:0000-0002-3459-2679
lucenko.alla32@gmail.com

Комаров Василь Федорович
к.т.н., старший викладач кафедри прикладної математики та кібербезпеки
Донецький національний університет імені Василя Стуса, м.Вінниця, Україна
ORCID:0000-0002-8797-1023
v.komarov@donnu.edu.ua

Юкальчук Андрій Іванович
магістр кафедри прикладної математики та кібербезпеки
Донецький національний університет імені Василя Стуса, м.Вінниця, Україна
ORCID:0009-0005-2811-7918
ayukalchuk2224@gmail.com

АРХІТЕКТУРИ НЕЙРОННИХ МЕРЕЖ ДЛЯ ЗАПОБІГАННЯ ТА ВИЯВЛЕННЯ РІЗНИХ ТИПІВ КІБЕРАТАК НА МЕРЕЖЕВІ РЕСУРСИ В ПЕРІОД ВІЙСЬКОВОГО СТАНУ

Анотація. У роботі обґрунтовано комплексний підхід до розробки інтелектуальних систем виявлення та запобігання вторгненням (IDS/IPS) на базі глибоких нейронних мереж для захисту об'єктів критичної інфраструктури в умовах військового стану. Доведено, що фізична деградація телекомунікацій та нестабільність зв'язку під час бойових дій генерують аномальний фоновий мережевий трафік, який нівелює ефективність традиційних сигнатурних методів. Особливу увагу приділено вирішенню фундаментальної проблеми екстремального дисбалансу даних: масивні атаки маскують міноритарні, але критично небезпечні загрози (ескалацію привілеїв U2R, віддалений доступ R2L), характерні для цілеспрямованих багатокрокових операцій урядових хакерських угруповань (APT).

Для вибору оптимальних математичних моделей проведено експериментальну перевірку некерованих алгоритмів (карти Кохонена SOM), базових мереж (MLP, CNN) та ансамблевих гібридних архітектур (CNN+MLP, Hopfield+Transformer). Оцінка ефективності здійснювалась на трьох наборах даних різного покоління та складності: класичних KDD99, NSL-KDD та сучасному дата-сеті UNSW-NB15, що виступав екстремальним стрес-тестом. Під час моделювання успішно імплементовано механізми зваженого балансування класів та процедури адаптивного масштабування обчислювальної ємності нейромереж.

Експериментальні результати підтверджують, що некеровані та базові моделі схильні ігнорувати міноритарні класи, демонструючи для них критично низьку повноту виявлення (Recall). Натомість гібридні рішення забезпечують високу здатність до просторово-часового узагальнення. Архітектура CNN+MLP показала найкращий баланс виявлення всіх класів атак, а модель Hopfield+Transformer підтвердила стійкість у розпізнаванні обфускованих патернів. Обґрунтовано перспективність впровадження мультимодальних ансамблів для надійного детектування еволюціонуючих кіберзагроз у високозашумленому мережевому середовищі.

Ключові слова: системи виявлення вторгнень; архітектура нейронної мережі; гібридні моделі; машинне навчання; дисбаланс класів; кібератака, військовий стан.

ВСТУП

Сучасні комп'ютерні мережі є критичною складовою функціонування державних установ, підприємств та інфраструктури, особливо в умовах військового стану. Під час повномасштабної збройної агресії кіберпростір трансформується у високоінтенсивний театр бойових дій, де цифрові операції розгортаються паралельно з кінетичними ударами [1, 2]. Ця кінетико-кібернетична синергія має на меті максимізацію деструктивного ефекту: кібератаки часто передують або супроводжують ракетні обстріли для придушення систем зв'язку, дезорганізації управління та створення паніки серед населення [2]. Наслідки таких комбінованих ударів є катастрофічними, що підтверджується зупинками роботи ключових державних реєстрів, паралічем митного оформлення вантажів та логістичним колапсом [3]. Крім того, фізичне руйнування телекомунікаційних вузлів та постійні блекаути генерують масиви аномального фонового трафіку (обриви сесій, тайм-аути), що призводить до появи величезної кількості хибних спрацьовувань у традиційних системах захисту [4].

Масштаб цифрового протистояння демонструє сталу ескалацію: у 2025 році було зафіксовано 5927 кіберінцидентів, що на 37,4% перевищує статистичні дані за 2024 рік [5]. Сучасний ландшафт загроз формується високоресурсними урядовими хакерськими угрупованнями (APT), які використовують вразливості «нульового дня» (zero-day), техніки обфускації та тактики «Living-off-the-Land» (LotL). Це робить традиційні сигнатурні системи IDS неефективними проти цілеспрямованих атак, створюючи вимогу переходу до адаптивних нейромережевих архітектур, здатних до вилучення глибоких просторово-часових ознак [6].

Водночас кібервійна характеризується екстремальним дисбалансом мережевих даних: масивні терабайтні DDoS-атаки відбуваються на тлі тихих, поодиноких спроб горизонтального переміщення (R2L) або ескалації привілеїв (U2R), які майже не створюють мережевого шуму [7, 8]. Звичайні моделі машинного навчання природно схильні ігнорувати такі статистичні меншості, що робить їх розгортання на об'єктах критичної інфраструктури вкрай небезпечним [9]. У зв'язку з цим зростає потреба не лише в розробці складних архітектур, здатних до узагальнення, але й у їхньому адекватному тестуванні. Практична корисність таких алгоритмів може бути доведена лише шляхом комплексного оцінювання на наборах даних різного покоління (від класичних еталонів до сучасних складних масивів), що дозволяє відстежити еволюцію загроз і підтвердити здатність систем виявляти як масові аномалії, так і вкрай рідкісні цілеспрямовані атаки в реальному часі.

Постановка проблеми. В умовах кібервійни комп'ютерні мережі функціонують у середовищі екстремальної нестабільності, де деструктивні кібератаки поєднуються з фізичними руйнуваннями інфраструктури. Це неминуче призводить до появи аномального фонового трафіку та критичного зростання кількості хибних спрацьовувань систем захисту (False Positives). Основна науково-практична проблема полягає у тому, що традиційні сигнатурні засоби та класичні методи машинного навчання виявляються неефективними проти сучасних складних загроз (зокрема APT) через їхню вразливість до екстремального дисбалансу даних, коли міноритарні атаки «губляться» на тлі масових. Відповідно, існує гостра необхідність у розробці та валідації новітніх адаптивних нейромережевих архітектур, стійких до зашумленого середовища.

Аналіз останніх досліджень і публікацій. У науковій літературі представлено низку робіт, присвячених виявленню вторгнень на основі алгоритмів глибокого навчання. Останніми роками значна частина досліджень у сфері кібербезпеки зосереджена на використанні штучного інтелекту (ШІ). Хоча у цьому напрямі досягнуто вагомих результатів, тестування моделей часто обмежується ізольованим використанням конкретних наборів даних, таких як KDD99, NSL-KDD, CICIDS2017 та UNSW-NB15, що ускладнює оцінку їхньої ефективності в умовах еволюції загроз.

Hleha та співавтори [10] досліджували оптимізацію методів пояснюваного штучного інтелекту (Explainable Artificial Intelligence, XAI) для створення нейромережевих систем виявлення вторгнень із мінімальною затримкою. У своїй роботі автори провели системний огляд і порівняльний аналіз різних архітектур глибокого навчання (CNN, LSTM, GRU, автоенкодери, гібридні моделі CNN-LSTM) та методів XAI (SHAP, LIME, Integrated Gradients, DeepLIFT, Grad-CAM, Anchors). Експериментальна перевірка на наборах даних CICIDS2017, NSL-KDD та UNSW-NB15 показала, що гібридні моделі CNN-LSTM забезпечили точність виявлення понад 98% із часом інференсу (логічного висновку) менше 10 мс, а також покращили повноту виявлення атак до 91,6%.

Chao та Xie [11] представили DeepNetGuard – інноваційний алгоритм глибокого навчання для ефективного виявлення потенційних загроз безпеки у великомасштабному мережевому трафіку. Модель використовує багатовекторну стратегію вилучення ознак, що поєднує базові, статистичні та поведінкові характеристики, а також інтегрує технології автоенкодерів і генеративно-змагальних мереж (GAN) для виявлення як відомих, так і невідомих загроз. DeepNetGuard продемонстрував вищу продуктивність

порівняно з традиційними системами виявлення та іншими DL-моделями, досягнувши значних показників точності.

Mohale та Obagbuwa [12] вдосконалили системи виявлення вторгнень шляхом інтеграції методів XAI для вирішення проблеми непрозорості алгоритмів («чорної скриньки») та підвищення їхньої надійності за умови збереження високої продуктивності. Використовуючи набір даних UNSW-NB15, автори розробили та оцінили декілька моделей машинного навчання, серед яких дерева прийняття рішень, багатошарові персептрони (MLP), XGBoost, випадковий ліс (Random Forest), CatBoost, логістична регресія та наївний баєсівський класифікатор. Моделі XGBoost та CatBoost досягли найвищої точності на рівні 87% із часткою хибнопозитивних (FPR) та хибнонегативних (FNR) спрацьовувань 0,07 і 0,12 відповідно.

У роботі [13] досліджено застосування методів глибокого навчання, зокрема згорткових (CNN) та рекурентних (RNN) нейронних мереж, для виявлення кіберзагроз і реагування на них у режимі реального часу. Гібридні архітектури CNN-RNN продемонстрували значну перевагу над традиційними підходами машинного навчання завдяки здатності одночасно вилучати просторові та часові ознаки, що дозволяє детектувати складні кіберзагрози з високою точністю та повнотою (recall) на різних наборах даних.

Jyothi та співавтори [14] запропонували оптимізовану модель нейронної мережі для виявлення кібератак із використанням вдосконаленого алгоритму оптимізації EWOA. Вона розроблена для протидії атакам із підстановкою облікових даних (credential stuffing), виявлення атак типу «відмова в обслуговуванні» (DoS) та прогнозування загроз. Експериментальні результати підтвердили високу ефективність моделі в умовах реального часу.

Abbas та ін. [15] провели оцінку методів глибокого навчання для виявлення кібератак в IoT-мережах, використовуючи дата-сет CICIoT2023, який містить 47 ознак та 33 типи атак. Автори порівняли ефективність різних архітектур (DNN, CNN та RNN) і підтвердили результативність запропонованого підходу для прогнозування векторів атак.

Hwaitat та ін. [16] представили концепцію адаптивних нейронних мереж на основі еволюційного підходу для покращення виявлення та класифікації вторгнень. Автори запропонували новий алгоритм навчання багатошарового персептрона (MLP), що використовує методи еволюційних обчислень для динамічної оптимізації ваг і зсувів. Модель була протестована на п'яти стандартизованих дата-сетах (NSL-KDD, CICIDS2017, UNSW-NB15, Bot-IoT та CSE-CIC-IDS2018), продемонструвавши перевагу над 10 сучасними алгоритмами оптимізації. Окремий напрям досліджень стосується захисту самих нейромережевих моделей та їхнього функціонування у специфічних, агресивних середовищах.

Barr [17] розробив архітектуру CNN, стійку до змагальних (adversarial) атак, для застосування на ресурсно-обмежених IoT-пристроях. У дослідженні використано інноваційну техніку APE-GAN для донавчання CNN, що значно покращило її стійкість до таких методів атак, як Deepfool та L-BFGS (перевірено на дата-сеті MNIST).

Alzaidy та Binsalleeh [18] дослідили методи змагальних атак та механізми захисту для CNN і RNN у задачах класифікації шкідливого програмного забезпечення (ПЗ). Автори реалізували атаки типу «білий ящик» (white-box) – JSMA та C&W – на виконувані файли Windows PE і оцінили два механізми захисту (дистиляцію та змагальне навчання), зафіксувавши рівень хибної класифікації у 73,5%.

Dalal та ін. [19] запропонували мережу Extremely Boosted Neural Network для більш точного прогнозування багатоетапних кібератак у хмарному середовищі. На наборі даних Multi-Step Cyber-Attack Dataset (MSCAD) модель досягла точності 99,72%, значно перевершивши показники Quest Model (94,09%), баєсівських мереж (97,29%) та стандартних нейромереж (99,09%).

Wang та ін. [20] розробили DDoS-MSCT – метод виявлення DDoS-атак на основі багатомасштабної згортки та архітектури Transformer. Запропонована модель містить модулі локального (LFEM) та глобального (GFEM) вилучення ознак, досягаючи точності 99% на дата-сеті CIC-DDoS2019.

Neto та ін. [9] представили CICIoT2023 – масштабний дата-сет для моделювання атак в IoT-середовищі у реальному часі. Набір даних створено з використанням розгалуженої топології зі 105 реальних IoT-пристроїв; він охоплює 33 типи атак, розподілених на 7 класів (DDoS, DoS, Recon, Web-based, Brute Force, Spoofing та Mirai), що забезпечує високореалістичне середовище для тестування алгоритмів машинного та глибокого навчання.

Saini та співавтори [21] здійснили всебічний огляд концепції змагального навчання (adversarial learning) у контексті виявлення вторгнень, проаналізувавши вектори атак і контрзаходи. Дослідження виявило, що графові нейронні мережі (GNN) є особливо перспективними для ідентифікації мережевих аномалій завдяки їхній здатності обробляти складні топологічні залежності та взаємозв'язки. Автори наголосили на критичній важливості розробки стійких механізмів захисту для протидії порушенням мережевої безпеки та конфіденційності.

Chhetri та Namin [22] дослідили застосування трансформерних моделей для прогнозування наслідків кібератак. Їхня робота демонструє ефективність архітектури BERT та ієрархічних мереж уваги (HAN) у задачах аналізу текстових описів інцидентів і передбачення їхнього потенційного впливу, що є важливим кроком до автоматизації систем оцінювання загроз.

Sharma [23] запропонував модель, що поєднує аналіз мережевого трафіку (NTA) з даними апаратних лічильників продуктивності (HPC). Дослідження зосереджене на створенні автоенкодера, який агрегує апаратні та мережеві характеристики для ефективної класифікації загроз. Тестування проводилось на стандартних наборах даних CICIDS2017, NSL-KDD та D.A.V.I.D.E HPC.

Rasikha та співавтори [24] розробили ансамблевий метод глибокого навчання для виявлення вторгнень. Автори підкреслили перевагу нейромережевих підходів, які не потребують ручного вилучення ознак (feature engineering) або попередніх припущень щодо статистичного розподілу даних. Запропонований підхід продемонстрував високу ефективність у розпізнаванні різноманітних типів атак в IoT-середовищах.

Ahmim та ін. [25] розробили ієрархічну IDS, що базується на комбінації дерев прийняття рішень та експертних правил. На першому етапі модель поєднує алгоритм дерева прийняття рішень зі зменшенням помилок (REPTree) та алгоритм JRip для первинної класифікації трафіку. На другому етапі класифікатор Forest PA використовує результати першого рівня разом із початковими ознаками для формування остаточного висновку. Порівняльний аналіз із 11 відомими класифікаторами на дата-сеті CICIDS2017 підтвердив високу здатність моделі: вона показала найкращі результати для семи класів атак та найнижчий рівень хибних спрацьовувань (FPR) для легітимного трафіку.

Tang та ін. [26] запропонували модель DNN для виявлення аномалій на рівні мережевих потоків. Архітектура включає один вхідний, три приховані та один вихідний шари. Експерименти на наборі даних NSL-KDD довели, що навіть така базова модель DNN здатна ефективно виявляти атаки, перевершуючи за показниками низку традиційних методів машинного навчання.

Забезпечення безперебійної та безпечної роботи мережевої інфраструктури вимагає автоматизації процесів моніторингу. Peng та ін. [27] розробили метод виявлення вторгнень, де глибока нейронна мережа слугує екстрактором ознак із даних мережевого моніторингу, а нейромережа зі зворотним поширенням помилки (BP-мережі) виконує фінальну класифікацію загроз. Оцінка моделі на дата-сеті KDDCup 99 показала точність 95,45%, що є значним покращенням порівняно з класичними алгоритмами.

Kolosnjaji та ін. [28] розробили ієрархічну нейромережеву архітектуру, що поєднує згорткові (CNN) та рекурентні (RNN) шари для вилучення складних ознак у системах виявлення шкідливого ПЗ. Цей підхід агрегує переваги локальної згортки просторових даних та послідовного моделювання часових рядів. Згодом автори розширили цей метод [29], інтегрувавши до моделі ознаки, вилучені із заголовків виконуваних файлів. Таке злиття даних (data fusion) дозволило досягти високих показників точності та повноти (recall) класифікації мережевого трафіку.

Мета статті. Розробка та комплексне оцінювання адаптивних моделей виявлення кіберзагроз на основі нейромережевих архітектур, здатних ефективно ідентифікувати як масові аномалії, так і рідкісні цілеспрямовані атаки в умовах динамічного та зашумленого мережевого середовища, шляхом аналізу просторово-часових характеристик мережевого трафіку та тестування на наборах даних різного покоління.

МЕТОДИКА ДОСЛІДЖЕННЯ

Характеристика наборів даних для валідації моделей. Для забезпечення об’єктивності та всебічної оцінки розроблених нейромережевих архітектур [30] експериментальні дослідження проводилися на трьох різних за складністю та поколінням наборах даних (дата-сетах).

KDD99: Класичний еталонний набір даних, що містить 41 ознаку та 5 базових класів (Normal, DoS, Probe, R2L, U2R). Його перевагою є широка представленість у науковій літературі, що дозволяє легко порівнювати результати з іншими роботами, а головним недоліком є наявність великої кількості дубльованих записів (до 78%), що часто призводить до перенавчання моделей та штучного завищення їхньої точності.

NSL-KDD: Очищена та збалансована версія KDD99, з якої повністю видалено надлишкові дублікати. Цей дата-сет зберігає оригінальну структуру з 41 ознаки, проте значно ускладнює задачу для алгоритмів машинного навчання, виступаючи об'єктивним бенчмарком для перевірки здатності моделей до узагальнення, а не простого «запам'ятовування» патернів.

UNSW-NB15: Сучасний, високоструктурований дата-сет, що відображає реалії актуальних кіберзагроз (включаючи гібридні атаки, використання експлойтів, хробаків та фаззерів). Для можливості співставлення з попередніми наборами, 10 оригінальних категорій атак UNSW-NB15 можна

концептуально звести до 5 базових класів або обмежитись оцінками ефективності моделей в межах оригінальних категорій. Цей дата-сет виступає в ролі екстремального стрес-тесту: він містить значно більший рівень фонового шуму та складніші мікро-патерни атак, що робить його більш наближеним до умов військового стану.

Конфігурація моделей та особливості гіперпараметрів. У ході експерименту досліджувались як класичні некеровані алгоритми (SOM), так і глибокі керовані архітектури (MLP, CNN) та складні ансамблеві моделі (CNN+MLP, Hopfield+Transformer. Також окремо був розглянутий комплексний мультимодальний гібрид (CNN+Transformer+Hopfield), який потребує суттєво більших обчислювальних ресурсів. Для забезпечення коректної та ефективної роботи моделей з різнорідними даними було імплементовано конвеєр предобробки та механізм динамічного масштабування гіперпараметрів:

Кодування та масштабування: Для обробки категоріальних ознак застосовано метод One-Hot Encoding. У випадку сучасного дата-сету UNSW-NB15 це призвело до розширення простору ознак з 41 (характерних для KDD) до ~190, що дозволило моделям точніше розрізняти протоколи та стани мережі. Вхідні дані стандартизовано за допомогою StandardScaler для багатошарових мереж та перетворено через MinMaxScaler для топологічних карт Кохонена (SOM), що працюють з евклідовими відстанями.

Балансування класів: Для вирішення фундаментальної проблеми ігнорування міноритарних класів (U2R, R2L, розвідка чи експлойти) у багатошарових мережах імплементовано зважену функцію втрат (Weighted CrossEntropyLoss). Вага кожного класу розраховувалась обернено пропорційно кількості його зразків у навчальній вибірці, що змусило алгоритми приділяти підвищену увагу виявленню рідкісних, але критично небезпечних векторів атак.

Адаптивне масштабування архітектурної ємності: З огляду на суттєву різницю в розмірності векторів ознак та складності трафіку, конфігурації нейромереж були чітко розділені на два структурні профілі:

- Профіль для класичних дата-сетів (KDD99, NSL-KDD): Застосовувались базові конфігурації для уникнення ефекту перенавчання (overfitting) на простих ознаках. Розмірність сітки SOM становила 20x20 вузлів; приховані шари MLP складалися з [128, 64, 32] нейронів; у трансформерних моделях використовувалося 2 блоки уваги з розмірністю ембедингів d_model=64.
- Профіль для високоскладного дата-сету (UNSW-NB15): Обчислювальна "ємність" мереж була радикально розширена для ідентифікації прихованих нелінійних залежностей сучасного трафіку. Топологічну сітку SOM було збільшено до 30x30 (900 вузлів) для уникнення «злипання» 10 підкласів атак; шари персептрона (MLP) поглиблено до матриці [256, 128, 64, 32]; згорткові мережі (CNN) отримали розширені набори фільтрів [32, 64, 128]; а для Transformer кількість шарів збільшено до 3 із розширенням внутрішньої розмірності до d_model=128.

Оптимізація та регуляризація навчання: Навчання глибоких архітектур проводилося пакетами (batch size) по 256 зразків із максимальною тривалістю до 500 епох. Для стабілізації градієнтного спуску використано оптимізатор AdamW зі швидкістю навчання (learning rate) 10^{-3}. Для ефективної протидії перенавчанню імплементовано L2-регуляризацію (weight_decay = 10^{-4}) та шари відсівання Dropout (0.2-0.3). Крім того, було застосовано механізм ранньої зупинки (early stopping) із розширеним «терпінням» (patience) у 20 епох, що дозволило складним моделям уникати передчасного переривання навчання під час обробки високозашумлених даних (зокрема для UNSW-NB15) та впевнено досягати глобального оптимуму функції втрат.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Зведені результати моделювання. Результати загальної правильності (Overall Accuracy) для всіх архітектур на трьох дата-сетах наведено в Таблиці 1 та відповідному рисунку 1.

Таблиця 1

Результати тестування нейромережевих архітектур

Архітектура моделі	Загальна правильність (Overall Accuracy), %		
	KDD99	NSL-KDD	UNSW-NB15
SOM (Self-Organizing Map)	99.51	97.37	~78.33
MLP (багатошаровий персептрон)	99.73	98.63	~80.71
CNN (згорткова мережа)	99.78	99.78	~80.91
Hopfield (автоасоціативна пам'ять)	74.08	80.10	~38.19
CNN + MLP (гібрид)	99.77	99.21	~84.16

Архітектура моделі	Загальна правильність (Overall Accuracy), %		
	KDD99	NSL-KDD	UNSW-NB15
Hopfield + Transformer (гібрид)	99.33	97.68	~59.51

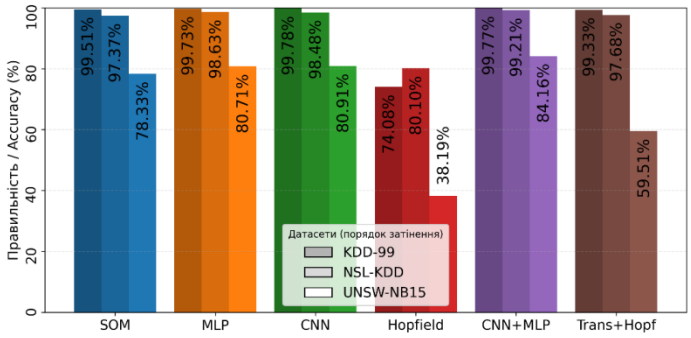

Рис. 1. Зведена діаграма загальної частки правильних відповідей (Overall Accuracy) досліджуваних моделей.

Аналіз ефективності класифікації за типами атак. Аналіз метрик повноти (Recall) демонструє суттєві відмінності у поведінці моделей, особливо при переході від масових до рідкісних класів атак.

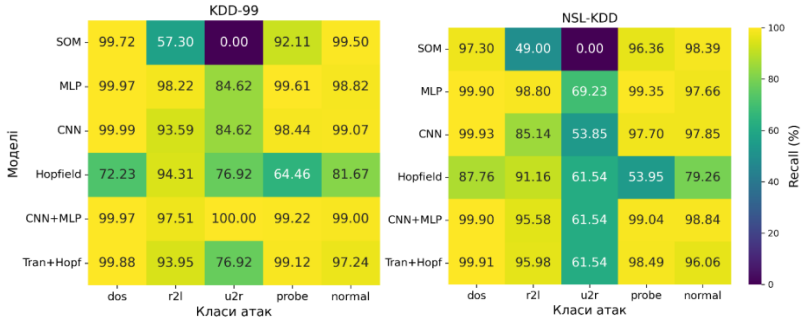

Рис 2. Теплові карти повноти виявлення різних типів кібератакдосліджуваними нейромережами на датасетах KDD-99 та NSL-KDD.

Масові класи (DoS, Probe, Normal): Майже всі оптимізовані моделі демонструють винятково високі показники повноти (близько 99% на NSL-KDD) для масових атак та нормального трафіку (рис. 3). Згорткові шари (CNN) бездоганно вилучають просторові сигнатури DoS-атак, а топологічна карта SOM формує чіткі візуалізаційні кордони (U-matrix) для відокремлення сканування мережі (Probe).

Рідкісні цілеспрямовані атаки (R2L, U2R): Ці класи є найскладнішими для ідентифікації через їхню мізерну представленість у загальному потоці. Базові архітектури без вагових коефіцієнтів (такі як чистий MLP без налаштувань балансу або некерована SOM) демонструють нульовий або критично низький Recall (0.00% – 57.64%) для цих атак (рис. 4). Натомість, гібридна архітектура CNN+MLP та комбінація Hopfield+Transformer показали безпрецедентну чутливість, досягнувши показників від 93.95% до 97.51% для R2L на базі NSL-KDD. Гібрид CNN+MLP виявився найкраще збалансованою моделлю: згортковий екстрактор відфільтровує фоновий шум, а глибокий персептрон класифікує неочевидні мікро-патерни ескалації привілеїв.

Своєрідний профіль продемонстрував гібрид Hopfield+Transformer. Незважаючи на надзвичайно високі показники виявлення всіх типів атак, на зашумлених даних він схильний класифікувати частину нормального трафіку як загрозу (зниження Recall для класу Normal). Це пояснюється роботою асоціативної пам'яті Хопфілда, яка робить межу прийняття рішень більш агресивною.

Окремі висновки щодо сучасного трафіку (UNSW-NB15): Як видно з Таблиці 1 та рисунку 1, на складному дата-сеті UNSW-NB15 точність усіх без винятку моделей закономірно знижується до діапазону 81-84%. Це підтверджує факт еволюції кіберзагроз: сучасні гібридні атаки (хробаки, шелл-коди, розвідка) успішно маскуються під легітимні з'єднання. Отримані результати доводять, що для ефективного захисту інфраструктури від новітніх загроз недостатньо класичного збільшення кількості

нейронів; концепція вимагає розробки ще більш складних (багаторівневих) ансамблевих гібридів, пропозиції щодо яких сформульовано у висновках до роботи.

У підсумку, для розгортання в умовах реального часу оптимальним архітектурним вибором є гібрид CNN+MLP (найкращий баланс точності та швидкодії). Якщо ж система функціонує в закритому військовому контурі, де пріоритетом є виявлення загрози за будь-яку ціну (навіть ціною хибних тривог), доцільним є використання асоціативних архітектур типу Hopfield+Transformer.

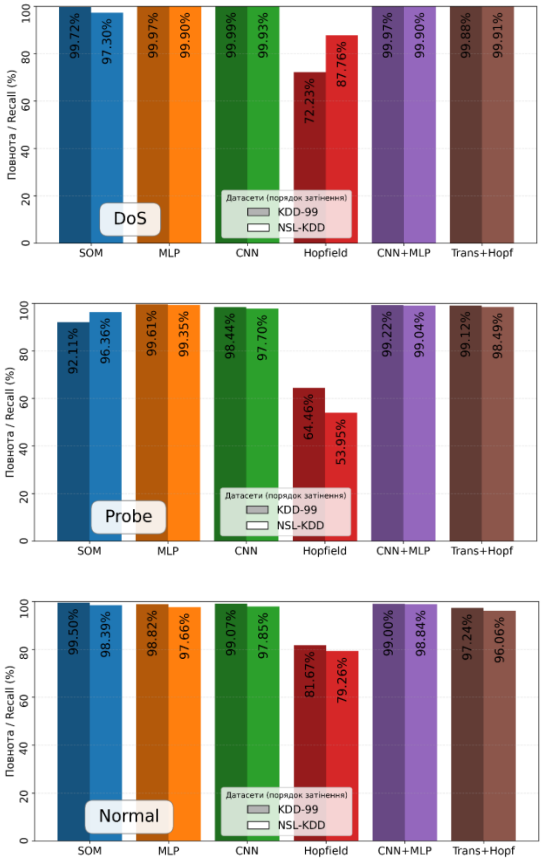

Рис. 3. Порівняльний графік повноти (Recall) виявлення масових класів атак (DoS, Probe) та звичайного трафіку (Normal).

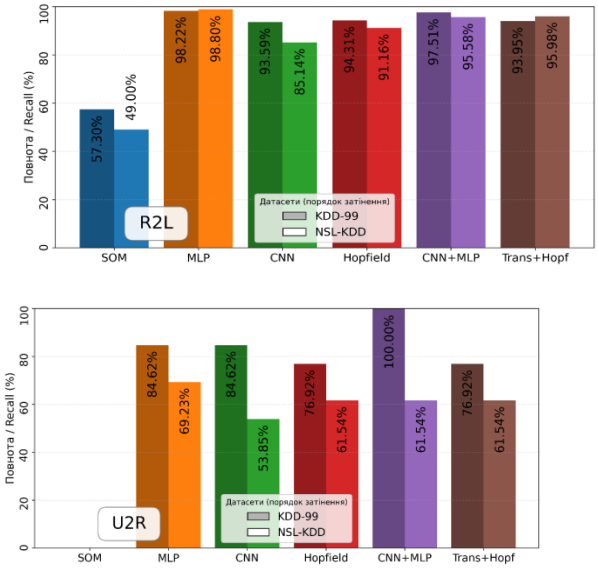

Рис. 4. Порівняльний графік повноти (Recall) виявлення міноритарних атак класу R2L та U2R

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Застосування та порівняння різних архітектур нейронних мереж на класичному наборі даних KDD дає чіткі практичні висновки: гібридні підходи (зокрема CNN+MLP) забезпечують найвищу загальну ефективність і збалансованість між виявленням масових і рідкісних атак, тоді як чисті архітектури (MLP, CNN) добре працюють для домінуючих класів, але суттєво втрачають чутливість до мало представлених категорій (R2L, U2R). Моделі із асоціативною складовою (Hopfield, Transformer+Hopfield) демонструють високу здатність виявляти окремі типи атак, проте іноді це досягається ціною зростання хибних спрацьовувань на нормальний трафік. Самоорганізовані карти (SOM) виявилися корисними як інструмент кластеризації та візуалізації – вони ефективно структурують великі кластери (dos, normal), але мало ефективні на рідкісних шаблонах.

Результати підкреслюють важливість двох практичних напрямів: по-перше, використання гібридних і ансамблевих архітектур, які поєднують локальне витягування ознак і глобальні/асоціативні механізми; по-друге, інтеграція заходів для боротьби з дисбалансом (вагові втрати, зважена семплінг-стратегія, синтетичні дані) – саме ці механізми значно підвищують чутливість до рідкісних атак. Оцінювання слід проводити багаторівнево (покласова повнота, точність, F1-міра, матриці помилок та загальна правильність), оскільки загальна метрика може приховувати критично слабкі місця моделей при виявленні атак.

Результати експериментальних досліджень підтверджують, що для практичних IDS найефективнішими є гібридні архітектури з вбудованими механізмами обробки дисбалансу і можливістю інтерпретації рішень; подальша робота повинна бути спрямована на валідацію таких рішень на реальних даних, оптимізацію для реального часу та розробку процедур адаптації моделей під змінні умови загроз.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Healey, J. (2024). Cyber effects in warfare: Categorizing the where, what, and why. *Texas National Security Review*. https://doi.org/10.26153/TSW/56029
2. Digital Front Lines. (2023, May 25). *The evolution of cyber operations in armed conflict*. https://digitalfrontlines.io/2023/05/25/the-evolution-of-cyber-operations-in-armed-conflict/
3. Sytnik, V. (2026, February 23). Ukraine: Facing the intensification of Russian cyber attacks. *Regard sur l'Est*. https://regard-est.com/ukraine-facing-the-intensification-of-russian-cyber-attacks
4. Popov, O., & Drahuntsov, R. (2025). Key challenges for cybersecurity operations centers in the context of full-scale war [In Ukrainian]. *Information Technology and Society, 3*(18), 136-144. https://doi.org/10.32689/maup.it.2025.3.19
5. State Service of Special Communications and Information Protection of Ukraine. (2026). *CERT-UA processed almost 6,000 cyber incidents in 2025: The number of hostile attacks increased by 37%* [In Ukrainian]. https://cip.gov.ua/ua/news/cert-ua-u-2025-roci-opracyuvala-maizhe-6000-kiberincidentiv-kilkist-vorozhikh-atak-zrosla-na-37
6. Google Cloud. (2026). *Cybersecurity forecast 2026*. https://services.google.com/fh/files/misc/cybersecurity-forecast-2026-en.pdf
7. ESET. (2026). *ESET Research APT report: Russian cyberattacks in Ukraine intensify, Sandworm unleashes new destructive wiper*. https://www.eset.com/us/about/newsroom/research/eset-research-apt-report-russian-cyberattacks-in-ukraine-intensify-sandworm-unleashes-new-destructive-wiper/
8. Cybersecurity and Infrastructure Security Agency. (2022). *Update: Destructive malware targeting organizations in Ukraine*. https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-057a
9. Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., & Ghorbani, A. A. (2023). CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors, 23*(13), 5941. https://doi.org/10.3390/s23135941
10. Hleha, K., & Hol, V. (2025). XAI optimization for low-latency neural-based intrusion detection systems in network environments. *Bulletin of V. N. Karazin Kharkiv National University, Series "Mathematical Modeling. Information Technology. Automated Control Systems", 66*, 19–36. https://doi.org/10.26565/2304-6201-2025-66-02
11. Chao, J., & Xie, T. (2024). Deep learning-based network security threat detection and defense. *International Journal of Advanced Computer Science and Applications, 15*(11).
12. Mohale, V. Z., & Obagbuwa, I. C. (2025). Evaluating machine learning-based intrusion detection systems with explainable AI: Enhancing transparency and interpretability. *Frontiers in Computer Science, 7*, 1520741.
13. Okafor, M. O. (2024). Deep learning in cybersecurity: Enhancing threat detection and response. *World Journal of Advanced Research and Reviews, 24*(3), 1116-1132.

14. Jyothi, K. K., et al. (2024). A novel optimized neural network model for cyber attack detection using enhanced whale optimization algorithm. *Scientific Reports, 14*, 5590. https://doi.org/10.1038/s41598-024-55098-2
15. Abbas, S., et al. (2024). Evaluating deep learning variants for cyber-attacks detection and multi-class classification in IoT networks. *PeerJ Computer Science, 10*, e1793.
16. Al Hwaitat, A. K., & Fakhouri, H. N. (2024). Adaptive cybersecurity neural networks: An evolutionary approach for enhanced attack detection and classification. *Applied Sciences, 14*, 9142. https://doi.org/10.3390/app14199142
17. Barr, M. (2025). A robust neural network against adversarial attacks. *Engineering, Technology & Applied Science Research, 15*(2), 20609-20615.
18. Alzaidy, S., & Binsalleeh, H. (2024). Adversarial attacks with defense mechanisms on convolutional neural networks and recurrent neural networks for malware classification. *Applied Sciences, 14*, 1673. https://doi.org/10.3390/app14041673
19. Dalal, S., et al. (2023). Extremely boosted neural network for more accurate multi-stage cyber attack prediction in cloud computing environment. *Journal of Cloud Computing, 12*, 14. https://doi.org/10.1186/s13677-022-00356-9
20. Wang, B., et al. (2024). DDoS-MSCT: A DDoS attack detection method based on multiscale convolution and transformer. *IET Information Security, 2024*, 1056705.
21. Saini, S., Chennamaneni, A., & Sawyerr, B. (2024). *A review of the duality of adversarial learning in network intrusion: Attacks and countermeasures* [Preprint]. arXiv. https://arxiv.org/abs/2412.13880
22. Chhetri, B., & Namin, A. S. (2025). The application of transformer-based models for predicting consequences of cyber attacks. In *2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE.
23. Sharma, A., et al. (2024). *Neural network based zero-day-attack detection: A machine learning approach to enhancing cybersecurity*. SSRN. https://ssrn.com/abstract=4848658
24. Rasikha, V., & Marikkannu, P. (2024). An ensemble deep learning-based cyber attack detection system using optimization strategy. *Knowledge-Based Systems, 301*, 112211.
25. Ahmim, A., et al. (2019). A novel hierarchical intrusion detection system based on decision tree and rules-based models. In *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)* (pp. 228-233). IEEE.
26. Tang, T. A., et al. (2016). Deep learning approach for network intrusion detection in software defined networking. In *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)* (pp. 258-263). IEEE.
27. Peng, W., et al. (2019). Network intrusion detection based on deep learning. In *2019 International Conference on Communications, Information System and Computer Engineering (CISCE)* (pp. 431-435). IEEE.
28. Kolosnjaji, B., Zarras, A., Webster, G., & Eckert, C. (2016). Deep learning for classification of malware system call sequences. In *Australasian Joint Conference on Artificial Intelligence* (pp. 137-149). Springer.
29. Kolosnjaji, B., Eraisha, G., Webster, G., Zarras, A., & Eckert, C. (2017). Empowering convolutional networks for malware classification and analysis. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 3838-3845). IEEE.
30. Zahoruiko, L., et al. (2024). Mathematical model and structure of a neural network for detection of cyber attacks on information and communication systems. *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska, 14*(3), 49-55. https://doi.org/10.35784/iapgos.6155

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (34), 2026
ISSN 2663 – 4023

Liubov Zahoruiko
Ph.D. in Engineering, Associate Professor,
Associate Professor of the Department of Applied Mathematics and Cybersecurity
Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine
ORCID:0000-0002-6958-8696
l.zahoruiko@donnu.edu.ua

Alla Lutsenko
Ph.D. in Mathematics, Senior Lecturer, Head of the Department of Applied Mathematics and Cybersecurity
Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine
ORCID:0000-0002-3459-2679
a.lutsenko@donnu.edu.ua

Vasyl Komarov
Ph.D. in Engineering, Senior Lecturer of the Department of Applied Mathematics and Cybersecurity
Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine
ORCID:0000-0002-8797-1023
v.komarov@donnu.edu.ua

Andrii Yukalchuk
Master's degree holder
Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine
ORCID:0009-0005-2811-7918
ayukalchuk2224@gmail.com

NEURAL NETWORK ARCHITECTURES FOR PREVENTING AND DETECTING VARIOUS TYPES OF CYBER ATTACKS ON NETWORK RESOURCES DURING MARTIAL LAW

Abstract. This paper substantiates a comprehensive approach to developing intelligent intrusion detection and prevention systems (IDS/IPS) based on deep neural networks to protect critical infrastructure facilities under martial law conditions. It is proven that the physical degradation of telecommunications and communication instability during combat operations generate anomalous background network traffic, which negates the effectiveness of traditional signature-based methods. Particular attention is paid to solving the fundamental problem of extreme data imbalance: massive attacks mask minority but critically dangerous threats (such as U2R privilege escalation and R2L remote access) characteristic of targeted multi-stage operations by Advanced Persistent Threats (APTs).

To select the optimal mathematical models, an experimental verification of unsupervised algorithms (Kohonen Self-Organizing Maps, SOM), base networks (MLP, CNN), and ensemble hybrid architectures (CNN+MLP, Hopfield+Transformer) was conducted. The performance evaluation was carried out on three datasets of different generations and complexity: the classic KDD99, NSL-KDD, and the modern UNSW-NB15 dataset, which served as an extreme stress test. During the simulation, weighted class balancing mechanisms and procedures for adaptive scaling of the computational capacity of neural networks were successfully implemented.

Experimental results confirm that unsupervised and base models tend to ignore minority classes, demonstrating a critically low detection completeness (Recall) for them. In contrast, hybrid solutions provide a high capability for spatiotemporal generalization. The CNN+MLP architecture showed the best balance in detecting all attack classes, while the Hopfield+Transformer model proved its robustness in recognizing obfuscated patterns. The prospects of implementing multimodal ensembles for the reliable detection of evolving cyber threats in a highly noisy network environment are substantiated.

Keywords: intrusion detection systems; neural network architecture; hybrid models; machine learning; class imbalance; cyber attack; martial law.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Healey, J. (2024). Cyber effects in warfare: Categorizing the where, what, and why. *Texas National Security Review.* https://doi.org/10.26153/TSW/56029

2. Digital Front Lines. (2023, May 25). *The evolution of cyber operations in armed conflict*. https://digitalfrontlines.io/2023/05/25/the-evolution-of-cyber-operations-in-armed-conflict/
3. Sytnik, V. (2026, February 23). Ukraine: Facing the intensification of Russian cyber attacks. *Regard sur l'Est*. https://regard-est.com/ukraine-facing-the-intensification-of-russian-cyber-attacks
4. Popov, O., & Drahuntsov, R. (2025). Key challenges for cybersecurity operations centers in the context of full-scale war [In Ukrainian]. *Information Technology and Society, 3*(18), 136-144. https://doi.org/10.32689/maup.it.2025.3.19
5. State Service of Special Communications and Information Protection of Ukraine. (2026). *CERT-UA processed almost 6,000 cyber incidents in 2025: The number of hostile attacks increased by 37%* [In Ukrainian]. https://cip.gov.ua/ua/news/cert-ua-u-2025-roci-opracyuvala-maizhe-6000-kiberincidentiv-kilkist-vorozhikh-atak-zrosla-na-37
6. Google Cloud. (2026). *Cybersecurity forecast 2026*. https://services.google.com/fh/files/misc/cybersecurity-forecast-2026-en.pdf
7. ESET. (2026). *ESET Research APT report: Russian cyberattacks in Ukraine intensify, Sandworm unleashes new destructive wiper*. https://www.eset.com/us/about/newsroom/research/eset-research-apt-report-russian-cyberattacks-in-ukraine-intensify-sandworm-unleashes-new-destructive-wiper/
8. Cybersecurity and Infrastructure Security Agency. (2022). *Update: Destructive malware targeting organizations in Ukraine*. https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-057a
9. Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., & Ghorbani, A. A. (2023). CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors, 23*(13), 5941. https://doi.org/10.3390/s23135941
10. Hleha, K., & Hol, V. (2025). XAI optimization for low-latency neural-based intrusion detection systems in network environments. *Bulletin of V. N. Karazin Kharkiv National University, Series "Mathematical Modeling. Information Technology. Automated Control Systems", 66*, 19–36. https://doi.org/10.26565/2304-6201-2025-66-02
11. Chao, J., & Xie, T. (2024). Deep learning-based network security threat detection and defense. *International Journal of Advanced Computer Science and Applications, 15*(11).
12. Mohale, V. Z., & Obagbuwa, I. C. (2025). Evaluating machine learning-based intrusion detection systems with explainable AI: Enhancing transparency and interpretability. *Frontiers in Computer Science, 7*, 1520741.
13. Okafor, M. O. (2024). Deep learning in cybersecurity: Enhancing threat detection and response. *World Journal of Advanced Research and Reviews, 24*(3), 1116-1132.
14. Jyothi, K. K., et al. (2024). A novel optimized neural network model for cyber attack detection using enhanced whale optimization algorithm. *Scientific Reports, 14*, 5590. https://doi.org/10.1038/s41598-024-55098-2
15. Abbas, S., et al. (2024). Evaluating deep learning variants for cyber-attacks detection and multi-class classification in IoT networks. *PeerJ Computer Science, 10*, e1793.
16. Al Hwaitat, A. K., & Fakhouri, H. N. (2024). Adaptive cybersecurity neural networks: An evolutionary approach for enhanced attack detection and classification. *Applied Sciences, 14*, 9142. https://doi.org/10.3390/app14199142
17. Barr, M. (2025). A robust neural network against adversarial attacks. *Engineering, Technology & Applied Science Research, 15*(2), 20609-20615.
18. Alzaidy, S., & Binsalleeh, H. (2024). Adversarial attacks with defense mechanisms on convolutional neural networks and recurrent neural networks for malware classification. *Applied Sciences, 14*, 1673. https://doi.org/10.3390/app14041673
19. Dalal, S., et al. (2023). Extremely boosted neural network for more accurate multi-stage cyber attack prediction in cloud computing environment. *Journal of Cloud Computing, 12*, 14. https://doi.org/10.1186/s13677-022-00356-9
20. Wang, B., et al. (2024). DDoS-MSCT: A DDoS attack detection method based on multiscale convolution and transformer. *IET Information Security, 2024*, 1056705.
21. Saini, S., Chennamaneni, A., & Sawyerr, B. (2024). *A review of the duality of adversarial learning in network intrusion: Attacks and countermeasures* [Preprint]. arXiv. https://arxiv.org/abs/2412.13880
22. Chhetri, B., & Namin, A. S. (2025). The application of transformer-based models for predicting consequences of cyber attacks. In *2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE.
23. Sharma, A., et al. (2024). *Neural network based zero-day-attack detection: A machine learning approach to enhancing cybersecurity*. SSRN. https://ssrn.com/abstract=4848658

24. Rasikha, V., & Marikkannu, P. (2024). An ensemble deep learning-based cyber attack detection system using optimization strategy. *Knowledge-Based Systems, 301*, 112211.
25. Ahmim, A., et al. (2019). A novel hierarchical intrusion detection system based on decision tree and rules-based models. In *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)* (pp. 228-233). IEEE.
26. Tang, T. A., et al. (2016). Deep learning approach for network intrusion detection in software defined networking. In *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)* (pp. 258-263). IEEE.
27. Peng, W., et al. (2019). Network intrusion detection based on deep learning. In *2019 International Conference on Communications, Information System and Computer Engineering (CISCE)* (pp. 431-435). IEEE.
28. Kolosnjaji, B., Zarras, A., Webster, G., & Eckert, C. (2016). Deep learning for classification of malware system call sequences. In *Australasian Joint Conference on Artificial Intelligence* (pp. 137-149). Springer.
29. Kolosnjaji, B., Eraisha, G., Webster, G., Zarras, A., & Eckert, C. (2017). Empowering convolutional networks for malware classification and analysis. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 3838-3845). IEEE.
30. Zahoruiko, L., et al. (2024). Mathematical model and structure of a neural network for detection of cyber attacks on information and communication systems. *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska, 14*(3), 49-55. https://doi.org/10.35784/iapgos.6155

Отримано редакцією журналу / Received: 02.02.26
Прорецензовано / Revised: 16.02.26
Схвалено до друку / Accepted: 24.09.26