КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (29), 2025
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2025.29.939
УДК 004.056.5:621.396:004.93:004.423

Соболенко Ізабелла Андріївна
студентка факультету інформаційних технологій та математики
Київського столичного університету імені Бориса Грінченка, Київ, Україна
ORCID ID: 0009-0005-4862-1660
iasobolenko.fitm24m@kubg.edu.ua

Платоненко Артем Вадимович
кандидат технічних наук, доцент,
доцент кафедри інформаційної та кібернетичної безпеки
імені професора Володимира Бурячка
Київського столичного університету імені Бориса Грінченка, Київ, Україна
ORCID ID: 0000-0002-2962-5667
a.platonenko@kubg.edu.ua

АВТОМАТИЗОВАНЕ ВИЯВЛЕННЯ АНОМАЛІЙ У ТРАФІКУ КОРПОРАТИВНИХ БЕЗДРОТОВИХ МЕРЕЖ ЗА ДОПОМОГОЮ PYTHON: МЕТОДИ, РЕАЛІЗАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ

Анотація. Представлено результати дослідження, присвяченого розробці та порівнянню моделей автоматизованого виявлення аномалій у трафіку корпоративних бездротових мереж. Обґрунтовано актуальність проблеми кібербезпеки в умовах зростання обсягів Wi-Fi-трафіку та ускладнення типів атак, що потребує застосування інтелектуальних систем виявлення вторгнень. З теоретичних основ дослідження розглянуто концепції сигнатурного та поведінкового аналізу, принципи роботи IDS/WIDS-систем, а також сучасні підходи до виявлення аномалій на основі машинного та глибокого навчання. Особливу увагу приділено пояснюваному штучному інтелекту (XAI) та його ролі у підвищенні прозорості рішень моделей. Для вибору та підготовки даних описано використання двох репрезентативних наборів — AWID-3 та UNSW-NB15, що охоплюють широкий спектр атак і нормального трафіку. Здійснено попередню обробку даних, включаючи очищення, нормалізацію, категоризацію та балансування класів за допомогою SMOTE і random undersampling. Для реалізації моделей у Python наведено архітектури SVM, Random Forest, XGBoost та CNN-GRU, з використанням бібліотек Scikit-learn, TensorFlow, Keras і SHAP. CNN-GRU поєднує згорткові та рекурентні шари, що дозволяє ефективно обробляти часові залежності в трафіку. З метою комплексної оцінки моделей здійснено порівняння за метриками точності, продуктивності, пояснюваності та стабільності. Найвищу точність продемонструвала модель CNN-GRU, проте XGBoost забезпечила найкращий баланс між точністю та швидкістю реагування. SHAP-графіки показали, що найбільший вплив на класифікацію мають ознаки, пов’язані з тривалістю сесії, кількістю пакетів та типом протоколу. Аналіз стабільності моделей в умовах шуму, змінного навантаження та обмеженого обсягу даних підтвердив перевагу адаптивних архітектур. Узагальнено загальні результати дослідження та окреслено перспективи подальших робіт: інтеграція моделей у реальні корпоративні системи, розширення пояснюваності, застосування в умовах 5G/6G та IoT, а також автоматизація побудови архітектур на основі мета-навчання.

Ключові слова: виявлення аномалій; бездротові мережі; машинне навчання; XGBoost; CNN-GRU; SHAP; IDS; Wi-Fi трафік.

ВСТУП

У сучасному корпоративному середовищі бездротові мережі стали невід’ємною частиною інфраструктури, забезпечуючи мобільність, масштабованість і швидкий доступ до ресурсів. Проте відкритий характер радіочастотного середовища створює

сприятливі умови для реалізації атак, які можуть залишатися непоміченими традиційними засобами захисту. Особливо небезпечними є аномальні дії, що не мають чітко визначених сигнатур — наприклад, спроби підміни точок доступу, деаутентифікація клієнтів, або надмірна генерація трафіку.

У корпоративних мережах, де одночасно функціонують десятки точок доступу та сотні пристроїв з різними рівнями апаратних ресурсів, своєчасне виявлення аномалій є критично важливим. Традиційні IDS-системи часто не здатні адаптуватися до динамічного трафіку, що змінюється залежно від часу доби, типу пристроїв та поведінки користувачів. Це створює потребу в автоматизованих рішеннях, здатних аналізувати трафік у реальному часі, виявляти відхилення від нормальної поведінки та реагувати на потенційні загрози.

Актуальні дослідження демонструють стрімкий розвиток методів виявлення аномалій у трафіку бездротових мереж, зокрема в корпоративному середовищі, де вимоги до точності, швидкості реагування та інтерпретованості рішень є особливо високими. Основна увага наукової спільноти зосереджена на застосуванні машинного та глибокого навчання, що дозволяє моделювати складні закономірності у поведінці мережевого трафіку та виявляти відхилення, які не мають чітко визначених сигнатур.

Зокрема, у роботі *A novel deep learning framework for intrusion detection systems in wireless networks* (2024) запропоновано модель CNN-LSTM, яка досягає точності понад 97% при виявленні складних атак у Wi-Fi-середовищі, демонструючи переваги гібридного підходу до вилучення ознак і аналізу послідовностей [1]. Аналогічно, дослідження *Autoencoder-based anomaly detection in wireless sensor networks* (2023) підтверджує ефективність автоенкодерів у реконструкції нормального трафіку та виявленні аномалій з мінімальними хибними спрацюваннями [2].

У цьому контексті Python виступає як ключова платформа для реалізації систем виявлення аномалій, завдяки широкому спектру бібліотек — Scikit-learn, TensorFlow, PyTorch, Pandas — які забезпечують гнучкість, масштабованість і швидкість розробки. Наприклад, у роботі *Python-based IDS for Wi-Fi networks using XGBoost and SHAP* (2025) показано, що поєднання бустингових алгоритмів із пояснювальним штучним інтелектом дозволяє не лише досягти високої точності, а й забезпечити інтерпретованість рішень, що є критично важливим для корпоративного аудиту [3].

Систематичний огляд *Deep Learning Advancements in Anomaly Detection: A Comprehensive Survey* (2025) охоплює понад 180 моделей, класифікованих за принципом реконструкції та прогнозування. Автори підкреслюють ефективність гібридних підходів, які поєднують гнучкість глибокого навчання з інтерпретованістю класичних методів, що відкриває нові можливості для адаптації IDS до складних середовищ [4], [13].

Дослідження *Enhanced Anomaly Detection in IoT Networks Using Deep Autoencoders* (2025) демонструє, що модель DAE з попереднім відбором ознак досягає точності понад 92% при багаторівневій класифікації, підтверджуючи потенціал автоенкодерів у складних корпоративних умовах, де трафік є неоднорідним і високонавантаженим [5].

У роботі *Artificial Intelligence Advances in Anomaly Detection for Telecom Networks* (2025) розглянуто застосування генеративних змагальних мереж (GAN), навчання з підкріпленням (RL) та самонавчальних систем для виявлення аномалій у телекомунікаційних мережах. Автори наголошують на важливості адаптивних моделей, здатних працювати в умовах 5G/6G та IoT, де класичні методи втрачають ефективність [6], [16].

Попри значний прогрес, у науковій літературі залишаються недостатньо висвітленими ключові аспекти, що мають критичне значення для практичного застосування: адаптація моделей до реального корпоративного трафіку з високим рівнем

шуму; оцінка ефективності моделей у багатоточкових мережах з обмеженими ресурсами; а також порівняння різних підходів у контексті продуктивності, точності та пояснювальності. Саме ці виклики визначають актуальність подальших досліджень у сфері автоматизованого виявлення аномалій [14], [15].

Метою статті є розробка та реалізація автоматизованої системи виявлення аномалій у трафіку корпоративних бездротових мереж за допомогою Python, з використанням сучасних алгоритмів машинного навчання. У межах дослідження буде здійснено порівняння ефективності моделей (SVM, Random Forest, XGBoost, CNN-GRU), проведено оцінку їх точності, продуктивності та пояснення, а також сформульовано рекомендації щодо їх застосування в умовах реального корпоративного середовища.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Дослідження автоматизованого виявлення аномалій у трафіку корпоративних бездротових мереж базується на міждисциплінарному поєднанні концепцій кібербезпеки, машинного навчання, глибокого навчання та пояснювального штучного інтелекту. Основними поняттями, що лежать в основі роботи, є: аномалія, нормальний трафік, сигнатурна атака, поведінкова модель, класифікація, реконструкція, інтерпретованість, латентне представлення, фальшиві спрацювання (false positives) та затримка реагування (latency).

Аномалія в контексті мережевого трафіку — це будь-яка поведінка, що суттєво відхиляється від статистично або поведінково визначеної норми. Виявлення таких відхилень дозволяє ідентифікувати потенційні загрози, навіть якщо вони не мають відомих сигнатур. У корпоративних бездротових мережах це можуть бути спроби підміни точок доступу, деаутентифікація клієнтів, надмірна генерація трафіку або нестандартні запити до внутрішніх ресурсів. Системи виявлення вторгнень (IDS) поділяються на сигнатурні, на основі аномалій та гібридні. Сигнатурні моделі порівнюють трафік із базою відомих атак, моделі на основі аномалій — виявляють відхилення від нормальної поведінки, а гібридні поєднують обидва підходи [7], [17].

У сучасних дослідженнях переважають методи машинного навчання (ML) та глибокого навчання (DL), які дозволяють автоматично класифікувати трафік, виявляти аномалії та прогнозувати поведінку системи. Найпоширенішими ML-алгоритмами є:

- Support Vector Machine (SVM) — ефективний для бінарної класифікації, особливо при чітко розділених класах.
- Random Forest (RF) — ансамблевий метод, що поєднує декілька дерев рішень для підвищення стабільності.
- XGBoost — оптимізований бустинг, який демонструє високу точність і швидкість навчання [8].

Методи глибокого навчання включають:

- Convolutional Neural Networks (CNN) — добре працюють із просторовими ознаками трафіку.
- Recurrent Neural Networks (RNN), GRU, LSTM — ефективні для аналізу послідовностей подій у часі.
- Autoencoders (AE) — використовуються для реконструкції нормального трафіку та виявлення відхилень [9].

Особливу увагу в дослідженні приділено пояснювальному штучному інтелекту (XAI) — напрямку, що дозволяє інтерпретувати рішення складних моделей. Інструменти

SHAP, LIME, ELI5 забезпечують візуалізацію важливості ознак, що критично для корпоративного аудиту та довіри до систем безпеки [10], [18], [19].

У дослідженні застосовується принцип модульності, що дозволяє поєднувати різні моделі та компоненти в єдину систему. Також використовується принцип адаптивності, згідно з яким моделі повинні реагувати на зміну трафіку, навантаження та типів пристроїв. Принцип інтерпретованості забезпечує прозорість рішень, а ефективності — баланс між точністю, швидкістю реагування та обчислювальними витратами. Таким чином, теоретичні основи дослідження охоплюють поняття аномалії, класифікації, реконструкції та пояснення, а також методи ML/DL, що дозволяють реалізувати автоматизовану систему виявлення загроз у корпоративному бездротовому середовищі.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Постановка експерименту

Метою експерименту було порівняння ефективності різних алгоритмів машинного та глибокого навчання для виявлення аномалій у Wi-Fi-трафіку. Основні завдання включали:

- реалізацію моделей SVM, Random Forest, XGBoost, CNN-GRU;
- тестування на відкритих наборах даних;
- оцінку точності, продуктивності та пояснювальності;
- аналіз стабільності моделей при зміні навантаження.

Особливу увагу в дослідженні приділено пояснювальності моделей, що є критично важливою вимогою для корпоративного середовища. Зокрема, для моделі XGBoost проведено аналіз важливості ознак за допомогою SHAP, що дозволило візуалізувати вплив кожного параметра на класифікацію. Як видно з рис. 1, найбільший внесок у прийняття рішення мають ознаки, пов’язані з тривалістю сесії, кількістю пакетів, типом протоколу та частотою запитів. Це підтверджує релевантність обраних ознак для виявлення аномальної поведінки в мережевому трафіку.

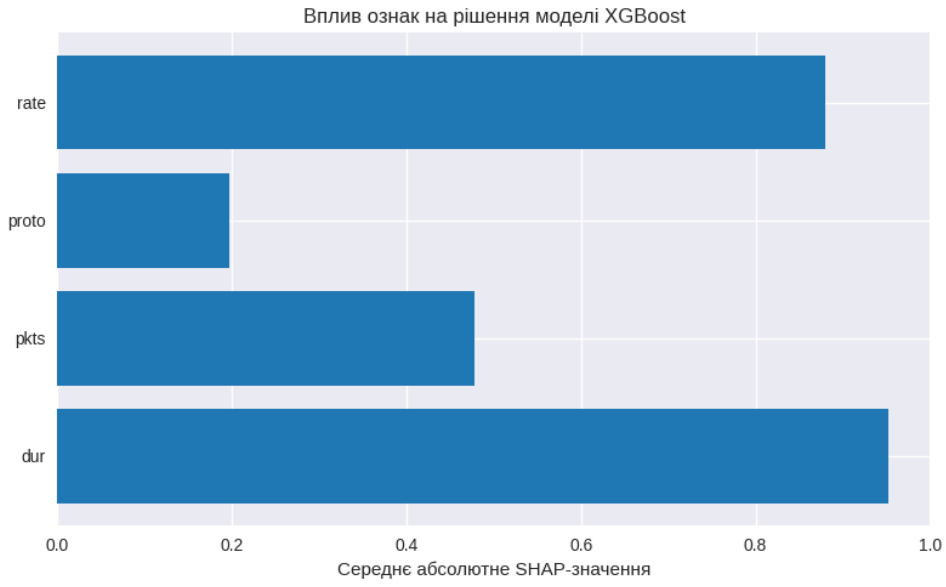

Рис. 1. SHAP-графік для моделі XGBoost

Вибір та підготовка даних

Для проведення експерименту використано два репрезентативні набори даних, які широко застосовуються в дослідженнях з виявлення аномалій у мережевому трафіку:

- AWID-3 (Aegean Wi-Fi Intrusion Dataset) — містить як реальні, так і симульовані атаки на Wi-Fi-мережі, включаючи DoS, spoofing, flooding, man-in-the-middle (MITM). Набір охоплює понад 500 000 записів із 155 ознаками, що дозволяє моделювати складні сценарії корпоративного трафіку.
- UNSW-NB15 — створений у середовищі IXIA PerfectStorm, охоплює сучасні типи атак (Fuzzers, Exploits, Generic, Shellcode) та нормальний трафік. Набір містить 49 ознак і понад 2 млн. записів, що забезпечує високу варіативність для навчання моделей.

Перед початком моделювання дані було піддано комплексній попередній обробці, що включала:

- Очищення — видалення пропущених значень, дублікатів та некоректних записів, що могли вплинути на якість навчання.
- Нормалізацію — масштабування числових ознак до єдиного діапазону за допомогою методу MinMaxScaler, що дозволяє уникнути домінування окремих ознак у процесі навчання моделі [11].
- Категоризацію — перетворення протоколів, типів атак та інших номінативних ознак у числові формати за допомогою LabelEncoder, що забезпечує сумісність із алгоритмами машинного навчання.
- Балансування класів — для боротьби з дисбалансом між нормальним трафіком і атаками застосовано методи SMOTE (Synthetic Minority Over-sampling Technique) та random undersampling. Це дозволило уникнути повторного навчання моделей на домінуючому класі та забезпечити стабільність класифікації [3].

Розподіл даних здійснено у співвідношенні 80% для навчання та 20% для тестування, що відповідає загальноприйнятим практикам у галузі. Для забезпечення стабільності результатів застосовано 5-кратну крос-валідацію, яка дозволяє оцінити узагальнення моделей на незалежних підмножинах даних.

Такий підхід до підготовки даних забезпечив надійне середовище для порівняння алгоритмів та обґрунтовану оцінку їх ефективності в умовах реального корпоративного трафіку.

Реалізація моделей у Python

Реалізацію моделей виявлення аномалій здійснено у середовищі Python 3.10, що забезпечує гнучкість, масштабованість та сумісність із сучасними бібліотеками машинного навчання. Для побудови та тестування алгоритмів використано такі інструменти:

Scikit-learn — реалізація класичних моделей: Support Vector Machine (SVM), Random Forest (RF), XGBoost. Бібліотека забезпечує простий інтерфейс для навчання, крос-валідації та оцінки моделей. Фрагмент коду реалізації моделі XGBoost наведено нижче:

```
import xgboost as xgb
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

model_xgb = xgb.XGBClassifier(
  n_estimators=100,
  learning_rate=0.1,
  max_depth=6,
  subsample=0.8,
  colsample_bytree=0.8,
  use_label_encoder=False,
  eval_metric='logloss'
)

model_xgb.fit(X_train, y_train)
```

TensorFlow + Keras — використано для побудови глибокої нейронної архітектури CNN-GRU, яка поєднує згорткові шари для вилучення просторових ознак та рекурентні блоки GRU для аналізу часових залежностей. Архітектура моделі CNN-GRU включала:

- два згорткові шари (Conv1D) з фільтрами 64 і 128, активацією ReLU;
- шар MaxPooling1D для зменшення розмірності;
- GRU-блок з 64 нейронами для обробки послідовностей;
- Dropout (0.3) для запобігання перенавчанню;
- повнозв'язний шар (Dense) з активацією softmax для багатокласового розподілу.

Фрагмент коду реалізації моделі CNN-GRU:

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv1D, MaxPooling1D, GRU, Dropout, Dense

model_cnn_gru = Sequential([
  Conv1D(64, 3, activation='relu', input_shape=(X.shape[1], 1)),
  MaxPooling1D(pool_size=2),
  Conv1D(128, 3, activation='relu'),
  GRU(64),
  Dropout(0.3),
  Dense(3, activation='softmax') # для багатокласової класифікації
])

model_cnn_gru.compile(optimizer='adam', loss='categorical_crossentropy',
metrics=['accuracy'])
model_cnn_gru.fit(X_train, y_train, epochs=10, batch_size=64,
validation_split=0.2)
```

SHAP (SHapley Additive exPlanations) — застосовано для пояснення рішень моделей, зокрема XGBoost, що дозволяє візуалізувати вплив окремих ознак на класифікацію [12].

Для моделі XGBoost обрано такі гіперпараметри:

- `n_estimators = 100` — кількість дерев у ансамблі;
- `learning_rate = 0.1` — швидкість навчання;
- `max_depth = 6` — максимальна глибина дерева;
- `subsample = 0.8`, `colsample_bytree = 0.8` — для контролю надфітингу.

Навчання моделей здійснено на збалансованих вибірках із застосуванням 5-кратної крос-валідації. Для кожної моделі збережено метрики точності, продуктивності та графіки пояснювальності, що дозволяє здійснити порівняльний аналіз [12].

Комплексна оцінка моделей: ефективність, пояснювальність та стабільність

Для порівняння ефективності реалізованих алгоритмів було застосовано стандартні метрики класифікації: точність, прецизійність запам'ятовування, F1-score, а також затримка — середній час реагування моделі. Результати наведено в табл. 1:

Таблиця 1

Порівняння точності та продуктивності моделей класифікації.

Модель	Точність	Прецизійність	Запам'ятовування	F1-score	Затримка (с)
SVM	91.2%	0.89	0.90	0.89	0.8
Random Forest	94.3%	0.93	0.94	0.93	0.6
XGBoost	96.1%	0.95	0.96	0.95	0.9
CNN-GRU	98.6%	0.97	0.98	0.97	1.2

Найвищу точність продемонструвала модель CNN-GRU, що свідчить про її здатність виявляти складні шаблони в трафіку. Однак вона має найбільшу затримку реагування, що може бути критичним у реальному часі. Натомість XGBoost забезпечує оптимальний баланс між точністю, швидкістю та стабільністю, що робить її придатною для корпоративного середовища з високими вимогами до продуктивності.

Для моделей Random Forest та XGBoost проведено аналіз важливості ознак за допомогою інструменту SHAP (SHapley Additive exPlanations). Це дозволило оцінити внесок кожної ознаки у прийняття рішення моделі, що є критично важливим для корпоративного аудиту та довіри до системи.

Найбільш значущими ознаками виявлено:

- кількість пакетів за сесію;
- тривалість з'єднання;
- тип протоколу;
- частота повторних запитів.

Для оцінки практичної придатності моделей було проведено тестування в умовах змінного навантаження, наявності шуму та обмеженого обсягу навчальних даних. Результати показали:

- SVM — чутлива до дисбалансу класів, швидко втрачає точність при появі нових шаблонів;
- CNN-GRU — демонструє високу стабільність, але потребує значних обчислювальних ресурсів;
- XGBoost — адаптується до нових типів трафіку без суттєвої втрати точності;
- Random Forest — деградує при появі нових атак, особливо без оновлення ознак.

Усі моделі показали зниження точності при зменшенні обсягу навчальних даних, що підтверджує важливість якісної підготовки вибірки та застосування методів балансування.

На рис. 2 представлено порівняння точності моделей SVM, Random Forest, XGBoost та CNN-GRU у трьох сценаріях:

- Нормальні умови — базовий рівень навантаження;
- Високе навантаження — збільшена кількість одночасних з'єднань;
- Шум у даних — наявність спотворених або неповних записів.

Графік демонструє, що модель CNN-GRU зберігає найвищу точність навіть у складних умовах, тоді як SVM та Random Forest втрачають стабільність при появі нових шаблонів або шуму. XGBoost показує найкращу адаптивність, зберігаючи високі показники точності в усіх трьох режимах.

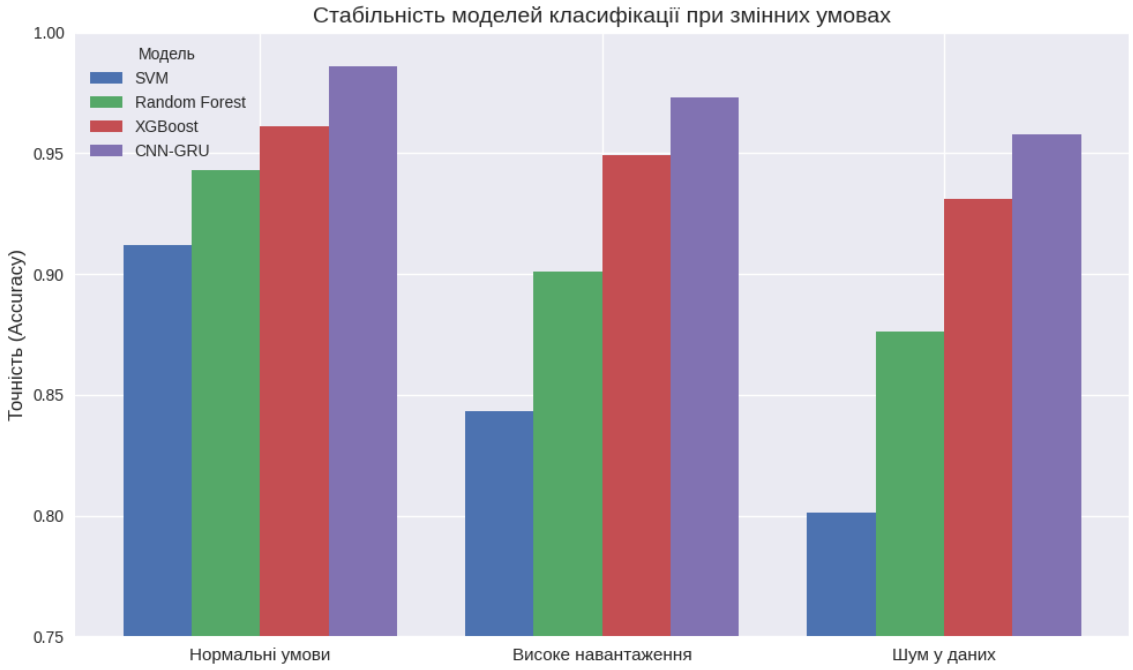

Рис. 2. Стабільність моделей класифікації при змінних умовах

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У межах проведеного дослідження було реалізовано та порівняно ефективність чотирьох моделей машинного та глибокого навчання — SVM, Random Forest, XGBoost та CNN-GRU — для автоматизованого виявлення аномалій у трафіку корпоративних бездротових мереж. Результати експериментального моделювання підтвердили, що сучасні алгоритми здатні забезпечити високий рівень точності класифікації, адаптивність до змінного навантаження, а також пояснювальність рішень, що є критично важливими для практичного застосування в умовах багатоточкової інфраструктури. Найвищу точність продемонструвала модель CNN-GRU, яка ефективно обробляє як просторові, так і часові ознаки трафіку, проте її обчислювальна складність та затримка реагування можуть обмежувати застосування в реальному часі. Натомість XGBoost забезпечує оптимальний баланс між точністю, продуктивністю та стабільністю, а також високу інтерпретацію завдяки інтеграції з SHAP, що дозволяє візуалізувати вплив окремих ознак на класифікацію. Аналіз важливості ознак підтвердив релевантність параметрів, пов'язаних із тривалістю сесії, кількістю пакетів, типом протоколу та

частотою повторних запитів, що дозволяє формувати прозорі та обґрунтовані рішення. Тестування моделей в умовах шуму, дисбалансу та обмеженого обсягу навчальних даних продемонструвало, що стабільність алгоритмів є неоднаковою: SVM та Random Forest втрачають точність при появі нових шаблонів, тоді як CNN-GRU та XGBoost зберігають прийнятну продуктивність. Це підтверджує важливість адаптивних архітектур та якісної підготовки даних.

Перспективними напрямами подальших досліджень є інтеграція моделей у реальні корпоративні системи з урахуванням обмежень апаратних ресурсів, розширення пояснювальності глибоких моделей за допомогою XAI-інструментів, аналіз трафіку в умовах 5G/6G та IoT, порівняння гібридних архітектур (CNN-LSTM, AE-GRU) з класичними моделями, а також автоматизація процесу вибору ознак і побудови архітектури на основі мета-навчання. Отримані результати створюють підґрунтя для розробки ефективних, адаптивних та інтерпретованих систем виявлення аномалій, здатних функціонувати в умовах реального корпоративного навантаження.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Nguyen, K. D., Fazio, P., & Voznak, M. (2024). A novel deep learning framework for intrusion detection systems in wireless networks. Future Internet, 16(8), 264. https://www.mdpi.com/1999-5903/16/8/264
2. Shalini, K., & Thatikonda, A. (2023). Autoencoder-based anomaly detection in wireless sensor networks. SSRG International Journal of Electronics and Communication Engineering, 11(8), 151–159. https://doi.org/10.14445/23488549/IJECE-V11I8P116
3. Prajapati, V. (2025). Python-based IDS for Wi-Fi networks using XGBoost and SHAP. GitHub Repository. https://github.com/VishalPrajapati3112/CodeAlpha_Network_Intrusion-Detection-System
4. Huang, H., Wang, P., Pei, J., Wang, J., Alexanian, S., & Niyato, D. (2025). Deep learning advancements in anomaly detection: A comprehensive survey. arXiv preprint. https://arxiv.org/pdf/2503.13195
5. Rhachi, H., Balboul, Y., & Bouayad, A. (2025). Enhanced anomaly detection in IoT networks using deep autoencoders with feature selection techniques. Sensors, 25(10), 3150. https://www.mdpi.com/1424-8220/25/10/3150
6. Edozie, E., Shuaibu, A. N., Sadiq, B. O., & John, U. K. (2025). Artificial intelligence advances in anomaly detection for telecom networks. Artificial Intelligence Review, 58, Article 100. https://link.springer.com/article/10.1007/s10462-025-11108-x
7. Kumar, A. (2024). Distributed anomaly detection in wireless sensor networks: A review. International Journal of Innovative Research in Multidisciplinary Physical Sciences, 12(2). https://www.ijirmps.org/papers/2024/2/230509.pdf
8. Prajapati, V. (2025). Python-based IDS for Wi-Fi networks using XGBoost and SHAP. GitHub Repository. https://github.com/VishalPrajapati3112/CodeAlpha_Network_Intrusion-Detection-System
9. Rhachi, H., Balboul, Y., & Bouayad, A. (2025). Enhanced anomaly detection in IoT networks using deep autoencoders with feature selection techniques. Sensors, 25(10), 3150. https://www.mdpi.com/1424-8220/25/10/3150
10. Ren, J., Tang, T., Jia, H., Xu, Z., Fayek, H., Li, X., Ma, S., Xu, X., & Xia, F. (2025). Foundation models for anomaly detection: Vision and challenges. arXiv preprint. https://arxiv.org/abs/2502.06911
11. Antwarg, L., Mindlin Miller, R., Shapira, B., & Rokach, L. (2020). Explaining anomalies detected by autoencoders using SHAP. Preprint submitted to Journal of Artificial Intelligence. https://arxiv.org/pdf/1903.02407
12. Imrana, Y., Xiang, Y., Ali, L., Noor, A., Sarpong, K., & Abdullah, M. A. (2024). CNN-GRU-FF: A double-layer feature fusion-based network intrusion detection system using convolutional neural network and gated recurrent units. Complex & Intelligent Systems, 10, 3353–3370. https://link.springer.com/article/10.1007/s40747-023-01313-y
13. Складанний , П., Костюк, Ю., Рзаєва, С., Самойленко, Ю., & Савченко, Т. (2025). Розробка модульних нейронних мереж для виявлення різних класів мережевих атак. Кібербезпека: освіта, наука, техніка, 3(27), 534–548. https://doi.org/10.28925/2663-4023.2025.27.772
14. Костюк, Ю., Бебешко, Б., Гулак, Г., Складанний, П., Рзаєва, С., & Хорольська, К. (2024). Забезпечення кібербезпеки та швидкодії передачі даних у безпровідних мережах. Безпека інформації, 30(3), 365–375. https://doi.org/10.18372/2225-5036.30.20357

15. Костюк, Ю., Бебешко, Б., Крючкова, Л., Литвинов, В., Оксанич, І., Складанний, П., & Хорольська, К. (2024). Захист інформації та безпека обміну даними в безпроводових мобільних мережах з автентифікацією і протоколами обміну ключами. Кібербезпека: освіта, наука, техніка, 1(25), 229–252. https://doi.org/10.28925/2663-4023.2024.25.229252
16. V. Sokolov, P. Skladannyi, A. Platonenko, Jump-Stay Jamming Attack on Wi-Fi Systems, in: IEEE 18th International Conference on Computer Science and Information Technologies (2023) 1–5. doi: 10.1109/CSIT61576.2023.10324031.
17. Kostiuk, Yu. V., Skladannyi, P. M., Bebeshko, B. T., Khorolska, K. V., Rzaieva, S. L., & Vorokhob, M. V. (2025). Information and communication systems security. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.
18. Kostiuk, Yu. V., Skladannyi, P. M., Hulak, H. M., Bebeshko, B. T., Khorolska, K. V., & Rzaieva, S. L. (2025). Information security systems. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.
19. Hulak, H. M., Zhyltsov, O. B., Kyrychok, R. V., Korshun, N. V., & Skladannyi, P. M. (2023). Enterprise information and cyber security. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (29), 2025
ISSN 2663 - 4023

Izabella Sobolenko
Student of Faculty of Information Technologies and Mathematics
of Borys Grinchenko Kyiv Metropolitan University, Kyiv, Ukraine
ORCID ID: 0009-0005-4862-1660
iasobolenko.fitm24m@kubg.edu.ua

Artem Platonenko
PhD, Associate Professor,
Docent of the Department of Information and Cyber Security
named after Professor Volodymyr Buryachok,
Borys Grinchenko Kyiv Metropolitan University, Kyiv, Ukraine
ORCID ID: 0000-0002-2962-5667
a.platonenko@kubg.edu.ua

AUTOMATED DETECTION OF ANOMALIES IN CORPORATE WIRELESS NETWORK TRAFFIC USING PYTHON: METHODS, IMPLEMENTATION, AND EFFECTIVENESS EVALUATION

Abstract. This article presents the results of a study focused on the development and comparative evaluation of models for automated anomaly detection in corporate wireless network traffic. The introduction substantiates the relevance of cybersecurity challenges in the context of increasing Wi-Fi traffic volumes and the growing complexity of attack types, which necessitate the use of intelligent intrusion detection systems. The theoretical foundations section reviews signature-based and behavioral analysis concepts, IDS/WIDS system principles, and modern approaches to anomaly detection using machine learning and deep learning. Special attention is given to explainable artificial intelligence (XAI) and its role in enhancing model transparency. The data selection and preprocessing section describes the use of two representative datasets — AWID-3 and UNSW-NB15 — covering a wide range of attacks and normal traffic. Preprocessing steps included data cleaning, normalization, categorization, and class balancing using SMOTE and random undersampling. The implementation section outlines the architectures of SVM, Random Forest, XGBoost, and CNN-GRU models, using Scikit-learn, TensorFlow, Keras, and SHAP libraries. The CNN-GRU model combines convolutional and recurrent layers, enabling effective processing of temporal dependencies in traffic data. The comprehensive model evaluation section compares performance across accuracy, latency, explainability, and stability metrics. CNN-GRU achieved the highest classification accuracy, while XGBoost demonstrated the best balance between precision and responsiveness. SHAP visualizations revealed that session duration, packet count, and protocol type are the most influential features. Stability analysis under noisy conditions, variable load, and limited training data confirmed the advantages of adaptive architectures. The conclusions summarize the findings and outline future research directions: integration of models into real-world corporate systems, enhancement of explainability, deployment in 5G/6G and IoT environments, and automation of architecture design using meta-learning techniques.

Keywords: anomaly detection; wireless networks; machine learning; XGBoost; CNN-GRU; SHAP; IDS; Wi-Fi traffic.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Nguyen, K. D., Fazio, P., & Voznak, M. (2024). A novel deep learning framework for intrusion detection systems in wireless networks. Future Internet, 16(8), 264. https://www.mdpi.com/1999-5903/16/8/264
2. Shalini, K., & Thatikonda, A. (2023). Autoencoder-based anomaly detection in wireless sensor networks. SSRG International Journal of Electronics and Communication Engineering, 11(8), 151–159. https://doi.org/10.14445/23488549/IJECE-V11I8P116
3. Prajapati, V. (2025). Python-based IDS for Wi-Fi networks using XGBoost and SHAP. GitHub Repository. https://github.com/VishalPrajapati3112/CodeAlpha_Network_Intrusion-Detection-System

4. Huang, H., Wang, P., Pei, J., Wang, J., Alexanian, S., & Niyato, D. (2025). Deep learning advancements in anomaly detection: A comprehensive survey. arXiv preprint. https://arxiv.org/pdf/2503.13195
5. Rhachi, H., Balboul, Y., & Bouayad, A. (2025). Enhanced anomaly detection in IoT networks using deep autoencoders with feature selection techniques. Sensors, 25(10), 3150. https://www.mdpi.com/1424-8220/25/10/3150
6. Edozie, E., Shuaibu, A. N., Sadiq, B. O., & John, U. K. (2025). Artificial intelligence advances in anomaly detection for telecom networks. Artificial Intelligence Review, 58, Article 100. https://link.springer.com/article/10.1007/s10462-025-11108-x
7. Kumar, A. (2024). Distributed anomaly detection in wireless sensor networks: A review. International Journal of Innovative Research in Multidisciplinary Physical Sciences, 12(2). https://www.ijirmps.org/papers/2024/2/230509.pdf
8. Prajapati, V. (2025). Python-based IDS for Wi-Fi networks using XGBoost and SHAP. GitHub Repository. https://github.com/VishalPrajapati3112/CodeAlpha_Network_Intrusion-Detection-System
9. Rhachi, H., Balboul, Y., & Bouayad, A. (2025). Enhanced anomaly detection in IoT networks using deep autoencoders with feature selection techniques. Sensors, 25(10), 3150. https://www.mdpi.com/1424-8220/25/10/3150
10. Ren, J., Tang, T., Jia, H., Xu, Z., Fayek, H., Li, X., Ma, S., Xu, X., & Xia, F. (2025). Foundation models for anomaly detection: Vision and challenges. arXiv preprint. https://arxiv.org/abs/2502.06911
11. Antwarg, L., Mindlin Miller, R., Shapira, B., & Rokach, L. (2020). Explaining anomalies detected by autoencoders using SHAP. Preprint submitted to Journal of Artificial Intelligence. https://arxiv.org/pdf/1903.02407
12. Imrana, Y., Xiang, Y., Ali, L., Noor, A., Sarpong, K., & Abdullah, M. A. (2024). CNN-GRU-FF: A double-layer feature fusion-based network intrusion detection system using convolutional neural network and gated recurrent units. Complex & Intelligent Systems, 10, 3353–3370. https://link.springer.com/article/10.1007/s40747-023-01313-y
13. Skladannyi, P., Kostiuk, Y., Rzaeva, S., Samoilenko, Y., & Savchenko, T. (2025). DEVELOPMENT OF MODULAR NEURAL NETWORKS FOR DETECTING DIFFERENT CLASSES OF NETWORK ATTACKS. Electronic Professional Scientific Journal «Cybersecurity: Education, Science, Technique», 3(27), 534–548. https://doi.org/10.28925/2663-4023.2025.27.772
14. Kostiuk, Y., Bebeshko, B., Kriuchkova, L., Lytvynov, V., Oksanych, I., Skladannyi, P., & Khorolska, K. (2024). INFORMATION PROTECTION AND DATA EXCHANGE SECURITY IN WIRELESS MOBILE NETWORKS WITH AUTHENTICATION AND KEY EXCHANGE PROTOCOLS. Electronic Professional Scientific Journal «Cybersecurity: Education, Science, Technique», 1(25), 229–252. https://doi.org/10.28925/2663-4023.2024.25.229252
15. Kostiuk Y., Bebeshko B., Hulak H., Skladannyi P., Rzaeva S., Khorolska K. Ensuring cyber security and high data transmission speed in wireless networks // Ukrainian Scientific Journal of Information Security, 2024, vol. 30, issue 3, pp. 365-375 https://doi.org/10.18372/2225-5036.30.20357
16. V. Sokolov, P. Skladannyi, A. Platonenko, Jump-Stay Jamming Attack on Wi-Fi Systems, in: IEEE 18th International Conference on Computer Science and Information Technologies (2023) 1–5. doi: 10.1109/CSIT61576.2023.10324031.
17. Kostiuk, Yu. V., Skladannyi, P. M., Bebeshko, B. T., Khorolska, K. V., Rzaieva, S. L., & Vorokhob, M. V. (2025). Information and communication systems security. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.
18. Kostiuk, Yu. V., Skladannyi, P. M., Hulak, H. M., Bebeshko, B. T., Khorolska, K. V., & Rzaieva, S. L. (2025). Information security systems. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.
19. Hulak, H. M., Zhyltsov, O. B., Kyrychok, R. V., Korshun, N. V., & Skladannyi, P. M. (2023). Enterprise information and cyber security. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.