НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»
Навчально-науковий інститут телекомунікаційних систем
Кафедра Інформаційних технологій в телекомунікаціях

«До захисту допущено»
В.О. завідувача кафедри
_____Валерій ПРАВИЛО
(підпис)
“___”_____2026 р.

Дипломна робота

на здобуття ступеня бакалавра

за освітньо-професійною програмою «Інформаційних технологій в телекомунікаціях»
спеціальності 172 Телекомунікації та радіотехніка

на тему: «Метод підвищення виявлення аномалій LPWAN на основі машинного навчання для попередження атак на кіберфізичні системи»

Виконав: студент IV курсу, групи ТІ-22
Загорулько Ілля Юрійович _____
(підпис)

Науковий керівник: старший викладач кафедри ІТТ НН ІТС, кандидат технічних наук Педан Станіслав Ігорович _____
(підпис)

Рецензент: доцент кафедри Інформаційної безпеки НН ФТІ, доктор технічних наук Прогонов Дмитро Олександрович _____
(підпис)

Засвідчую, що у цієї дипломної роботі немає запозичень з праць інших авторів без відповідних посилань.
Студент_____
(підпис)

Київ – 2026 р

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Навчально-науковий інститут телекомунікаційних систем

Кафедра Інформаційних технологій в телекомунікаціях

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 172 Телекомунікації та радіотехніка

Освітньо-професійна програма «Інформаційно-комунікаційні технології»

ЗАТВЕРДЖУЮ
В.О. завідувача кафедри
_____ Валерій Правило
(підпис)
«___»_____2026 р.

ЗАВДАННЯ
на дипломну роботу студенту
Загорулькові Іллі Юрійовичу

1. Тема роботи : «Метод підвищення виявлення аномалій LPWAN на основі машинного навчання для попередження атак на кіберфізичні системи», науковий керівник роботи Педан Станіслав Ігорович, старший викладач кафедри ІТТ ННІ ІТС, кандидат технічних наук - затверджені наказом по університету від «26» травня 2026 р. №1912-с

2. Строк подання студентом роботи 06.06.2026 р.

3. Вихідні дані до роботи : спеціальна література, ноутбук для проведення дослідження, матеріали мережі інтернет, власні роздуми, публічні датасети IoT.

4. Зміст пояснювальної записки дипломної роботи
- Архітектура та сучасні тренди розвитку сфери IoT та LPWAN-технологій
- Основні вразливості IoT/LPWAN пристроїв та кіберфізичних систем
- Огляд методів виявлення аномалій та захисту LPWAN-мереж
- Метод підвищення виявлення аномалій LPWAN на основі машинного навчання
- Практична реалізація та дослідження запропонованого методу
- Загальні висновки

5. **Перелік графічного (ілюстративного) матеріалу** (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо):

- Зміст дипломної роботи

- Мета та завдання роботи
- Архітектура LPWAN-мереж та роль IoT у розвитку розумних міст
- Основні вектори атак на LPWAN-мережі
- Порівняльний аналіз вразливостей LoRaWAN, Sigfox, NB-IoT
- Модель атакуючого для LPWAN-мереж
- Методологія виявлення аномалій на основі машинного навчання
- Структура запропонованого гібридного методу виявлення аномалій
- Практичне дослідження та результати навчання моделі Isolation Forest
- Діаграми точності, ROC-криві та порівняльні графіки
- Висновки

6. Дата видачі завдання : «12» жовтня 2026 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Строк виконання етапів роботи	Примітка
1	Архітектура та сучасні тренди розвитку сфери IoT та LPWAN-технологій	15.04-21.04	виконано
2	Аналіз критичних вразливостей технології LPWAN	22.04-28.04	виконано
3	Огляд методів виявлення аномалій та захисту LPWAN-мереж	29.04-05.05	виконано
4	Розробка методу підвищення виявлення аномалій на основі ML	06.08-12.05	виконано
5	Практична реалізація, дослідження та тестування моделі	13.05-19.05	виконано

Студент _____ Ілля ЗАГОРУЛЬКО
(підпис)

Науковий керівник роботи_____ Станіслав ПЕДАН
(підпис)

РЕФЕРАТ

Робота містить 84 сторінок, 28 рисунків та 10 таблиць. Було також використано 32 літературних джерела посилань.

У роботі розглядається метод підвищення ефективності виявлення аномалій у LPWAN-мережах на основі машинного навчання для захисту кіберфізичних систем «розумного» міста. Запропоновано гібридний підхід, який поєднує алгоритм Isolation Forest з детермінованими правилами доменних знань для виявлення атак типу False Data Injection (FDI) на рівні мережевого сервера.

Об'єкт дослідження – Процес безпечної комунікації пристроїв у бездротових мережах великого радіусу дії з низьким енергоспоживанням (LPWAN)

Предмет дослідження — Методи та моделі машинного навчання для підвищення ефективності виявлення семантичних аномалій у даних кіберфізичних систем

Мета – підвищення безпеки бездротових мереж, що використовуються в концепції розумного міста.

Наукова новизна дослідження полягає у систематизації та практичній апробації сучасних підходів до аналізу захищеності LPWAN-мереж.

В межах дипломної роботи розроблено та перевірено на реальних даних комплексний підхід до виявлення аномалій, адаптований саме для LPWAN-технологій.

Наукова новизна дослідження полягає у такому:

- Розроблено гібридну модель виявлення аномалій, яка поєднує алгоритм Isolation Forest з контекстними (просторовими та часовими) правилами доменних знань, що враховують фізичну природу процесу.
- Запропоновано вектор ознак, який включає мережеві метадані,

семантичні показники сенсорів та міжпристроєву кореляцію для виявлення як грубих, так і прихованих (stealthy) FDI-атак.

- Проведено експериментальне дослідження ефективності моделі на реальному датасеті системи розумного вуличного освітлення.

Ключові слова: LPWAN, LORAWAN, ВИЯВЛЕННЯ АНОМАЛІЙ, МАШИННЕ НАВЧАННЯ, ISOLATION FOREST, FALSE DATA INJECTION, КІБЕРФІЗИЧНІ СИСТЕМИ, РОЗУМНЕ МІСТО, ІОТ-БЕЗПЕКА, СЕМАНТИЧНИЙ АНАЛІЗ.

ABSTRACT

The work contains 84 pages, 28 figures and 10 tables. 32 literature references were also used.

The work considers a method for increasing the efficiency of anomaly detection in LPWAN networks based on machine learning to protect cyber-physical systems of a "smart" city. A hybrid approach is proposed that combines the Isolation Forest algorithm with deterministic domain knowledge rules to detect attacks such as False Data Injection (FDI) at the network server level.

The object of work is the process of secure communication of devices in wireless networks of long range with low power consumption (LPWAN)

The subject of work is methods and models of machine learning to increase the efficiency of detecting semantic anomalies in data of cyber-physical systems

The purpose of work is to increase the security of wireless networks used in the concept of a smart city

The scientific novelty of work:

The systematization and practical testing of modern approaches to the analysis of the security of LPWAN networks.

Within the framework of the thesis, a comprehensive approach to anomaly detection, adapted specifically for LPWAN technologies, was developed and tested on real data.

The scientific novelty of work is as follows:

- A hybrid anomaly detection model was developed, which combines the Isolation Forest algorithm with contextual (spatial and temporal) domain knowledge rules that take into account the physical nature of the process.
- A feature vector was proposed, which includes network metadata, semantic sensor indicators, and inter-device correlation to detect both brute and stealthy FDI attacks.
- An experimental study of the model's effectiveness was conducted on a real dataset of a smart street lighting system.

Keywords: LPWAN, LORAWAN, ANOMALY DETECTION, MACHINE LEARNING, ISOLATION FOREST, FALSE DATA INJECTION, CYBER-PHYSICAL SYSTEMS, SMART CITY, IOT SECURITY, SEMANTIC ANALYSIS.

ПЕРЕЛІК СКОРОЧЕНЬ

FDI (False Data Injection) – Атака шляхом ін’єкції фальшивих даних.

F1-score – Гармонійне середнє між точністю (precision) та повнотою (recall) результатів класифікації.

IoT (Internet of Things) – Інтернет речей.

LoRaWAN (Long Range Wide Area Network) – Протокол низькошвидкісної передачі даних на великі відстані.

LPWAN (Low Power Wide Area Network) – Енергоефективна мережа дальнього радіуса дії.

OWASP (Open Worldwide Application Security Project) – Відкритий проект із безпеки веб-додатків.

RSSI (Received Signal Strength Indicator) – Індикатор сили прийнятого сигналу.

SDR (Software Defined Radio) – Програмно-визначена радіосистема.

SNR (Signal-to-Noise Ratio) – Відношення сигнал/шум.

UML (Unified Modeling Language) – Уніфікована мова моделювання.

tau (Tau) – Часовий інтервал між послідовними пакетами даних.

ЗМІСТ

ВСТУП

Стрімкий розвиток цифрових технологій та активне впровадження концепції «розумних міст» призводять до масового розгортання мереж Інтернет речей (IoT) на базі систем інфраструктури телекомунікаційних операторів [7]. Особливої актуальності набувають енергоефективні технології LPWAN (Low-Power Wide Area Networks), які широко застосовуються для дистанційного керування системами комунального господарства, зокрема системами вуличного освітлення, моніторингу інфраструктури та інших кіберфізичних систем [1, 8].

Масштабність, розподілена архітектура та обмежені обчислювальні ресурси кінцевих пристроїв роблять LPWAN-мережі привабливою мішенню для зловмисників [11, 20]. Компрометація навіть одного IoT-пристрою може стати початком складної атаки, спрямованої на маніпуляцію даними сенсорів [15]. Особливо небезпечними є атаки типу False Data Injection (FDI), коли зловмисник ін'єктує фальшиві, але криптографічно коректні дані [18]. Такі атаки можуть призвести до неправильного керування вуличним освітленням, перевантаження енергомереж, створення аварійних ситуацій та значних економічних збитків [19, 21].

Класичні методи захисту, такі як криптографічне шифрування (AES-128) та механізми автентифікації, ефективно захищають канал зв'язку, але недостатньо протидіють семантичним атакам на рівні даних [10, 17]. У цих умовах традиційні оборонні механізми стають малоефективними, що зумовлює потребу в проактивних методах захисту — зокрема, у сучасних системах виявлення аномалій на основі машинного навчання [22].

Метою диплоної роботи є підвищення безпеки бездротових мереж, що використовуються в концепції розумного міста.

РОЗДІЛ 1

ОГЛЯД АРХІТЕКТУРИ ТА ПРИНЦИПІВ РОБОТИ LPWAN ТЕХНОЛОГІЙ

1.1 Огляд архітектури LPWAN

1.1.1 Концептуальні основи та ключові характеристики:

Архітектура LPWAN (Low-Power Wide Area Network) являє собою спеціалізовану мережеву топологію, розроблену для забезпечення енергоефективного зв'язку на великі відстані для пристроїв, що передають невеликі обсяги даних. Її фундаментальні принципи відрізняються від класичних стільникових та локальних мереж і базуються на трьох ключових характеристиках.

Низьке енергоспоживання. Метою є автономна робота пристроїв від акумулятора протягом 5–10 років. Це досягається завдяки циклічному режиму роботи «сон–активація» (sleep-wake cycle) та мінімальній складності модуляційних схем. У режимі сну типовий пристрій споживає лише 1–10 мкА, тоді як у активному режимі передачі — 30–120 мА. Завдяки тому, що активна фаза займає незначну частку загального часу роботи, батарея ємністю 2–10 А·год забезпечує автономність до 10 років без заміни [1].

Велика дальність зв'язку. Дальність між кінцевим пристроєм та базовою станцією сягає 10–15 км у сільській місцевості та 2–5 км в міських умовах завдяки використанню діапазонів суб-ГГц та стійкості до завад [2]. Кількісно ця характеристика описується через бюджет каналу зв'язку (Link Budget)(1.1):

$$P_{RX} = P_{TX} + G_{RX}G_{TX} + \ - L_{path} - L_{misc} \quad (1.1)$$

де P_{RX} — потужність прийому (дБм), P_{TX} — потужність передачі (дБм), G_{TX}, G_{RX} — коефіцієнти підсилення антен (дБі), L_{path} — втрати на трасі поширення (дБ), L_{misc} — додаткові втрати (дБ). Для типових умов розгортання LoRaWAN на частоті 868 МГц та відстані 10 км втрати на трасі становлять близько 111

дБ, що відповідає бюджету каналу понад 150 дБ — значно більше, ніж у класичних Wi-Fi або Bluetooth мережах.

Низька швидкість передачі даних. Пропускна здатність LPWAN знаходиться в діапазоні від 0,3 до 50 кбіт/с, що цілком достатньо для передачі телеметрії, показників датчиків або статусних команд. Теоретична межа пропускної здатності визначається шириною смуги та відношенням сигнал/шум відповідно до теореми Шеннона–Гартлі; однак на практиці досяжна швидкість суттєво нижча через обмеження модуляції, протокольні накладні витрати та вимоги до дальності [4].

Ці три характеристики утворюють «залізний трикутник» LPWAN: покращення одного параметра неминуче призводить до погіршення інших, що визначає архітектурні компроміси при проектуванні мереж цього класу. Наочне порівняння того, як різні технології розташовуються у цьому трикутнику, наведено на рисунку 1.1.

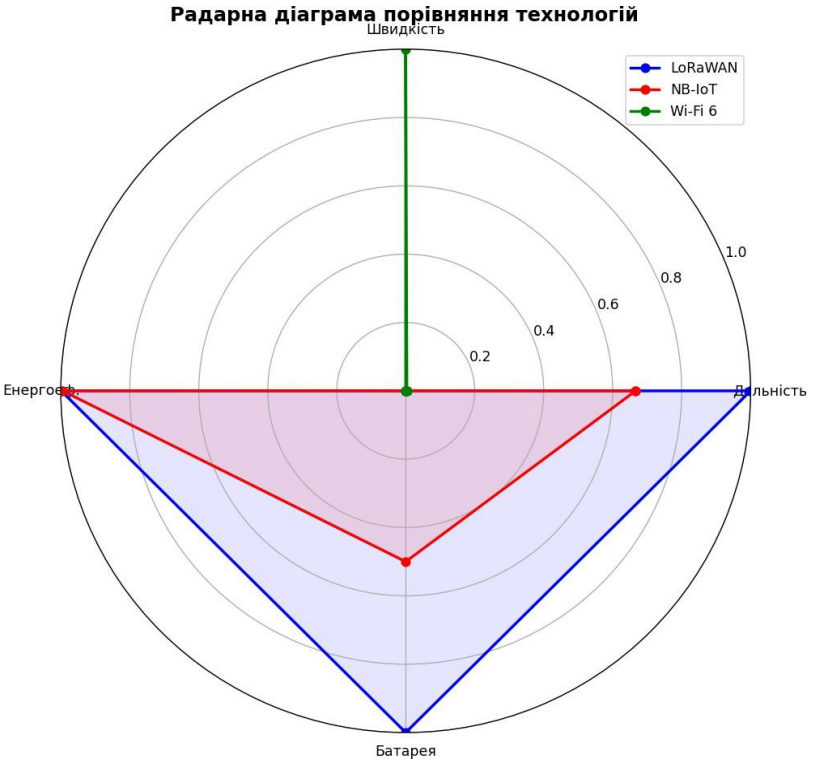

Рисунок 1.1— радарна діаграма порівняння технологій

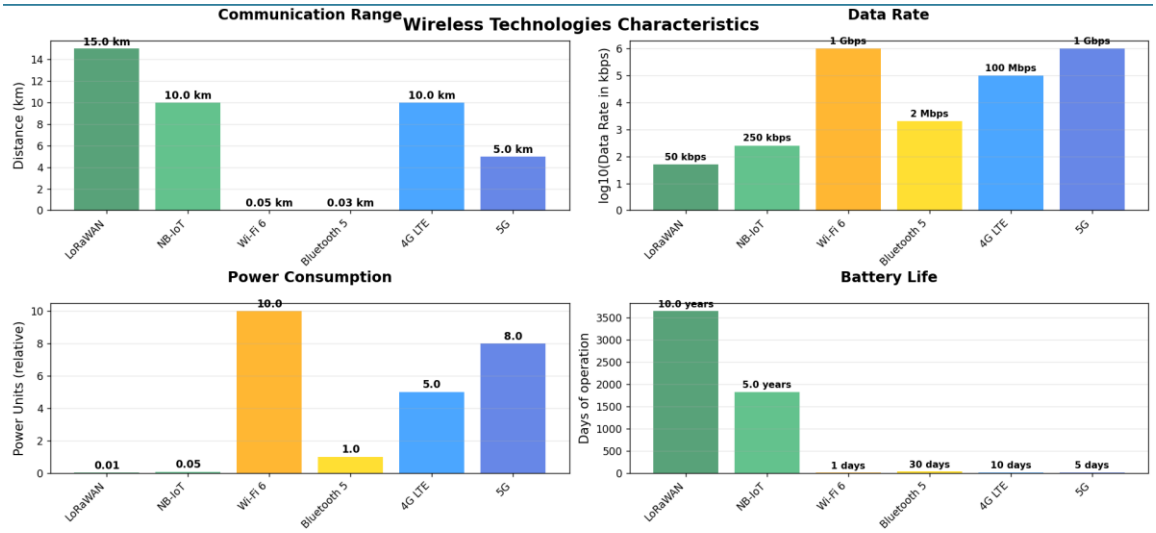

Рисунок 1.2— Характеристики бездротових технологій (Communication Range, Data Rate, Power Consumption, Battery Life)

1.1.2 Структурна архітектура мережі LPWAN

Архітектура LPWAN є багаторівневою і включає чотири ключові компоненти, кожен з яких виконує чітко визначені функції.

Кінцеві пристрої (End Devices) — це сенсори, датчики або актуатори, розгорнуті в полі. Вони відповідають за збір даних, їх первинну обробку та передачу до базових станцій. Через необхідність тривалої автономної роботи такі пристрої мають мінімальну обчислювальну складність.

Базові станції (Gateways) виконують роль концентраторів, які збирають радіофрейми від тисяч кінцевих пристроїв. Отримавши сигнал через антену, шлюз демодулює його та передає далі в мережу через IP-з’єднання (Ethernet, 4G, оптоволокно). Важливою особливістю є те, що шлюзи зазвичай є «прозорими» — вони лише ретрансують дані, не виконуючи функцій управління пристроями.

Мережевий сервер (Network Server) є центральним елементом і «мозком» усієї LPWAN-мережі. Він забезпечує маршрутизацію та дедуплікацію даних (оскільки один пакет може бути прийнятий кількома шлюзами), управляє безпекою (аутентифікація пристроїв та шифрування), регулює швидкість передачі даних за допомогою механізму Adaptive Data Rate,

а також обробляє процедури приєднання до мережі та планує downlink-повідомлення.

Сервер прикладних програм (Application Server) є заключною ланкою архітектури. Він отримує очищені корисні дані від мережевого сервера, інтерпретує їх відповідно до бізнес-логіки конкретного застосунку (наприклад, аналізує показники освітленості, температури чи рівня заповнення резервуарів) та надає інтерфейс для кінцевого користувача або систем управління.

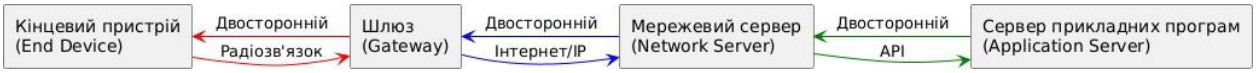

Рисунок 1.3— Структурна архітектура LPWAN-мережі [3]

1.1.3. Класифікація архітектур за типом мережі

Архітектури LPWAN поділяються на дві фундаментальні категорії залежно від використання радіочастотного спектра, що безпосередньо визначає принципи організації мережі, вартість розгортання та рівень залежності від операторів зв'язку.

Мережі з ліцензованим спектром (LoRaWAN, Sigfox) працюють у вільних ISM-діапазонах (868 МГц у Європі, 915 МГц у США) та використовують архітектуру типу «Алоха», за якої пристрої передають дані без жорсткої синхронізації зі шлюзом. Перевагами є низька вартість і повна автономність від мобільних операторів; недоліками — чутливість до радіозавад та обмеження на час передачі (Duty Cycle). Мережі з ліцензованим спектром (NB-IoT, LTE-M) базуються на стільникових стандартах 3GPP, забезпечують гарантовану якість обслуговування (QoS) та кращу захищеність, проте потребують вищих витрат на модулі та підключення. Наочне співвідношення між двома категоріями за ключовими параметрами — дальністю, швидкістю передачі та енергоспоживанням — наведено на рисунку 1.4.

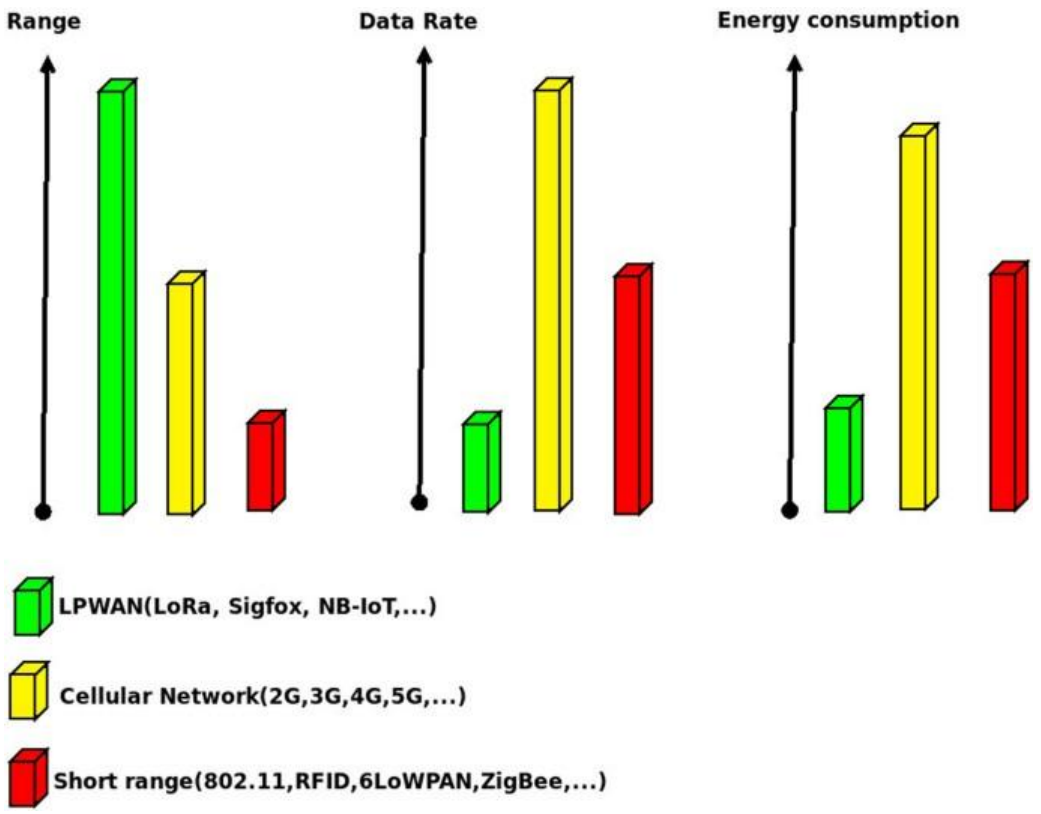

Рисунок 1.4— Порівняння структурної класифікації мереж LPWAN за типом спектру.

Зведене порівняння обох категорій за основними характеристиками наведено у таблиці 1.1.

Таблиця 1.1 — Порівняльна характеристика категорій LPWAN

Характеристика	**Неліцензований спектр (LoRaWAN, Sigfox)**	**Ліцензований спектр (NB-IoT, LTE-M)**
Частотний діапазон	ISM (868/915 МГц)	Ліцензовані смуги LTE
Модуляція (PHY)	CSS (LoRa), UNB (Sigfox)	OFDMA, SC-FDMA
Доступ до каналу (MAC)	ALOHA-подібний, без синхронізації	Планування ресурсів (Scheduling)

Якість обслуговування (QoS)	Не гарантована	Гарантована
Вартість розгортання	Низька	Висока
Залежність від оператора	Відсутня	Повна
Захищеність	Середня	Висока

Відмінності між категоріями особливо яскраво проявляються на рівні комунікаційного стеку протоколів. На відміну від класичної моделі OSI, стек LPWAN є суттєво спрощеним і обмежується кількома ключовими рівнями. На фізичному рівні (PHY) LoRaWAN застосовує технологію Chirp Spread Spectrum (CSS), тоді як NB-IoT використовує ортогональне частотне розділення (OFDMA/SC-FDMA). На канальному рівні (MAC) різниця ще помітніша: LoRaWAN реалізує простий ALOHA-подібний протокол без централізованого планування, тоді як NB-IoT використовує механізм Scheduling Request із явним виділенням ресурсів для кожного пристрою, що забезпечує кращу масштабованість при великій кількості вузлів. Детальне порівняння стеків обох технологій наведено на рисунку 1.5.

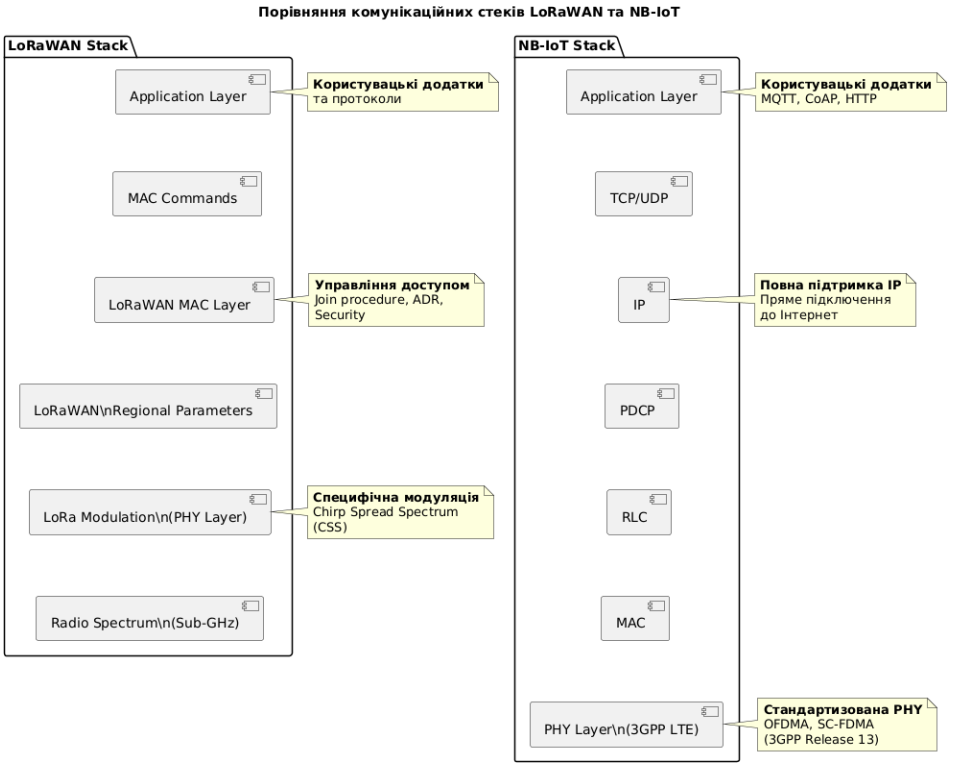

Рисунок 1.5— Порівняння комунікаційних стеків LoRaWAN та NB-IoT

Описаний поділ суттєво впливає на стратегію розгортання та, що особливо важливо в контексті даної роботи, на характер вразливостей: відкритість радіоінтерфейсу в мережах неліцензованого спектру створює специфічні вектори атак, яким присвячено наступний підрозділ.

Архітектура LPWAN є спеціалізованим і високооптимізованим рішенням для широкого класу IoT-застосунків, де пріоритетними виступають низьке енергоспоживання, велика дальність зв’язку та можливість масштабування на тисячі пристроїв. Багаторівнева структура, яка чітко розділяє функції передачі даних (шлюзи), управління мережею (мережевий сервер) та бізнес-логіку (сервер прикладних програм), забезпечує необхідну гнучкість і надійність роботи.

Важливим архітектурним рішенням є поділ технологій на дві фундаментальні категорії — мережі з неліцензованим спектром (LoRaWAN, Sigfox) та мережі з ліцензованим спектром (NB-IoT, LTE-M). Цей поділ

суттєво впливає на стратегію розгортання, вартість володіння та технічні можливості мережі, тому є одним із ключових факторів при виборі технології для конкретного проекту.

Разом з тим, особливості архітектури LPWAN, такі як відкритість радіоінтерфейсу, обмежені ресурси кінцевих пристроїв та спрощені механізми комунікації, створюють специфічні вектори загроз. Саме тому подальший аналіз присвячено вивченню відомих векторів атак та актуальних загроз інформаційній безпеці мереж LoRaWAN, Sigfox і NB-IoT.

1.2 Аналіз відомих векторів атак та актуальних загроз інформаційній безпеці мереж LoRaWAN, Sigfox, NB-IoT

1.2.1 Актуальність проблеми захищеності LPWAN-мереж в контексті їх архітектурних особливостей:

Після детального аналізу архітектурних особливостей технологій LPWAN у попередньому розділі, виникає нагальна потреба у комплексному дослідженні їх вразливостей. Унікальна архітектура LPWAN, орієнтована на енергоефективність та максимальну дальність.

Якість зв'язку, одночасно створює специфічні виклики в галузі кібербезпеки. Обмежені обчислювальні ресурси кінцевих пристроїв, відкритість радіоінтерфейсу, довгі цикли передачі даних та особливості протоколів зв'язку роблять ці мережі вразливими до спеціалізованих атак, що потребують глибокого аналізу в контексті їх архітектурної реалізації.

Для систематизації дослідження запропоновано багаторівневий підхід до аналізу загроз, що враховує архітектурні особливості кожної технології. Аналіз проводиться на наступних рівнях[7]:

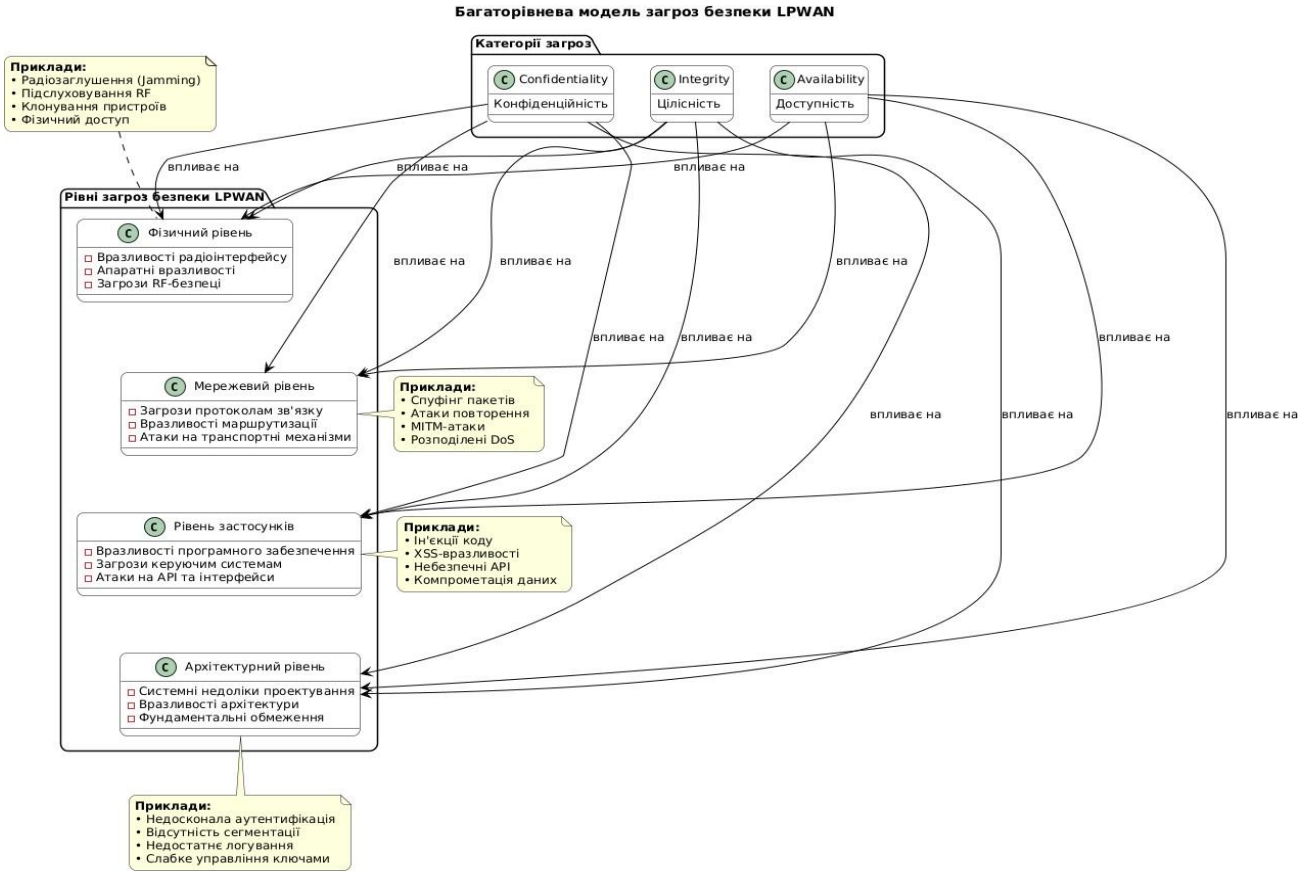

Рисунок 1.6 — Багаторівнева модель загроз безпеки LPWAN

1.2.2 Детальний аналіз вразливостей та векторів атак в технології LoRaWAN

Архітектура LoRaWAN має кілька критичних точок безпеки. Згідно з дослідженнями LoRa Alliance Security Working Group (2023), основні вразливості пов'язані з особливостями радіоінтерфейсу, фізичного та мережевого рівнів.

На рівні радіоінтерфейсу та фізичного рівня найбільш значущим є пасивне підслуховування (Eavesdropping). Відкритий характер передач в неліцензованому ISM-діапазоні дозволяє зловмисникам перехоплювати дані на відстані до кількох кілометрів — за допомогою стандартних антен трафік можна перехопити на відстані до 5 км навіть в умовах міської забудови. Механізм Over-the-Air Activation (OTAA) вразливий до спуфінгу запитів на приєднання (Join-request spoofing), що може дозволити несанкціонований доступ до мережі.

Окремої уваги заслуговує стійкість LoRaWAN до навантаження. Технологія використовує протокол Pure ALOHA, ефективність якого суттєво падає при підвищеному навантаженні на канал: математичний аналіз показує, що вже при завантаженості каналу на 50% лише близько 37% пакетів передаються успішно. Це робить мережу вразливою до організованих DoS-атак через штучне підвищення навантаження, а відсутність складних механізмів контролю доступу до середовища (CSMA/CA) робить її додатково чутливою до цілеспрямованих радіоперешкод (jamming).

На мережевому рівні серйозними загрозами є атаки повторення (Replay attacks), що стають можливими через недостатній захист часових міток. Також небезпечними є маніпуляції механізмом Adaptive Data Rate (ADR): зловмисник може навмисно спотворювати параметри передачі, що призводить до неефективного використання ресурсів мережі та експоненційного зростання колізій. При відносно невеликій кількості активних пристроїв ймовірність колізії може перевищувати 60%, що значно знижує загальну пропускну здатність мережі та створює додаткові можливості для атак на доступність.

Таким чином, архітектурні особливості LoRaWAN створюють передумови як для пасивних, так і для активних атак на всі ключові компоненти мережі. Механізми двох найбільш характерних атак — OTAA Spoofing та ACK Storm — детально проілюстровано на рисунку 1.7.

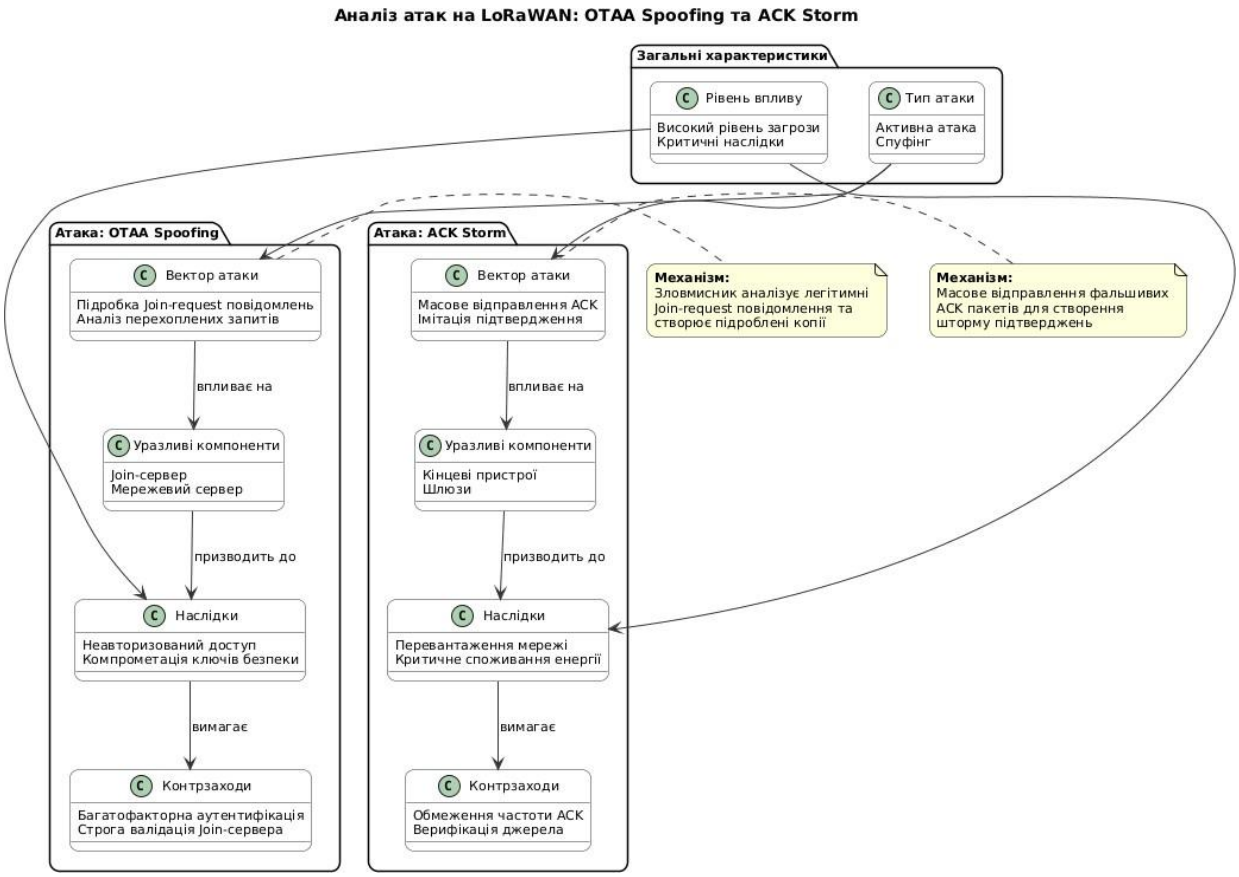

Рисунок 1.7 — Аналіз атак на LoRaWAN: OTAA Spoofing та ACK Storm

1.2.3 Комплексний аналіз вразливостей технології Sigfox

Архітектура Sigfox, що базується на ультра-вузькосмуговій технології (UNB), має фундаментальні обмеження у реалізації механізмів безпеки. Дослідження Sanchez-Iborra та Skarmeta (2022) підкреслюють, що основні проблеми зумовлені спрощеною структурою протоколу.

Sigfox характеризується відсутністю повноцінного двостороннього зв'язку в реальному часі, що суттєво обмежує можливості динамічної автентифікації та оперативного реагування на загрози. Технологія використовує статичні ключі шифрування протягом усього життєвого циклу пристрою, що створює підвищений ризик їх компрометації при тривалому використанні. Фіксована структура повідомлень також полегшує аналіз трафіку та ідентифікацію шаблонів передачі, знижуючи рівень конфіденційності.

Основними векторами атак на Sigfox [8] є цільові атаки на доступність (Targeted DoS), які реалізуються через заглушення конкретних вузьких частот (100 Гц). Також можливе клонування пристроїв та ідентифікаторів через слабкі механізми автентифікації, а обмежений набір можливих послідовностей повідомлень значно спрощує проведення атак повторення (Replay attacks).

1.2.4 Аналіз загроз безпеці NB-IoT

Архітектура NB-IoT, що базується на стільникових стандартах LTE, успадковує як сильні сторони, так і вразливості традиційних мобільних мереж. Згідно з технічним аналізом 3GPP (2023), технологія має кілька груп характерних загроз [9].

На фізичному та канальному рівнях найбільш небезпечним є цільове заглушення сигналу (Selective Jamming), коли зловмисник блокує конкретні NB-IoT-канали, знаючи їх частотне планування. Також можливі складні атаки на синхронізацію, що порушують роботу таймінгів передачі через маніпуляцію синхронізаційними сигналами.

На мережевому рівні суттєву загрозу становлять експлойти протоколу NAS (Non-Access Stratum), що дозволяють втручатися в процедури автентифікації та управління сесіями. Крім того, існують теоретичні вразливості алгоритму AKA (Authentication and Key Agreement) — криптографічного механізму, який забезпечує взаємну автентифікацію пристрою та мережі.

Специфічні загрози NB-IoT, обумовлені IoT-орієнтованістю, пов'язані переважно з обмеженими ресурсами кінцевих пристроїв. Це створює додаткові можливості для атак на енергоефективність (battery drainage) та сигнатурний аналіз (signal fingerprinting), коли зловмисник намагається ідентифікувати пристрої за характерними особливостями їхнього радіосигналу. Послідовність реалізації обох типів атак — від ініціації фальшивих запитів до ідентифікації пристроїв через аналіз RF-характеристик — наочно відображена на рисунку 1.8.

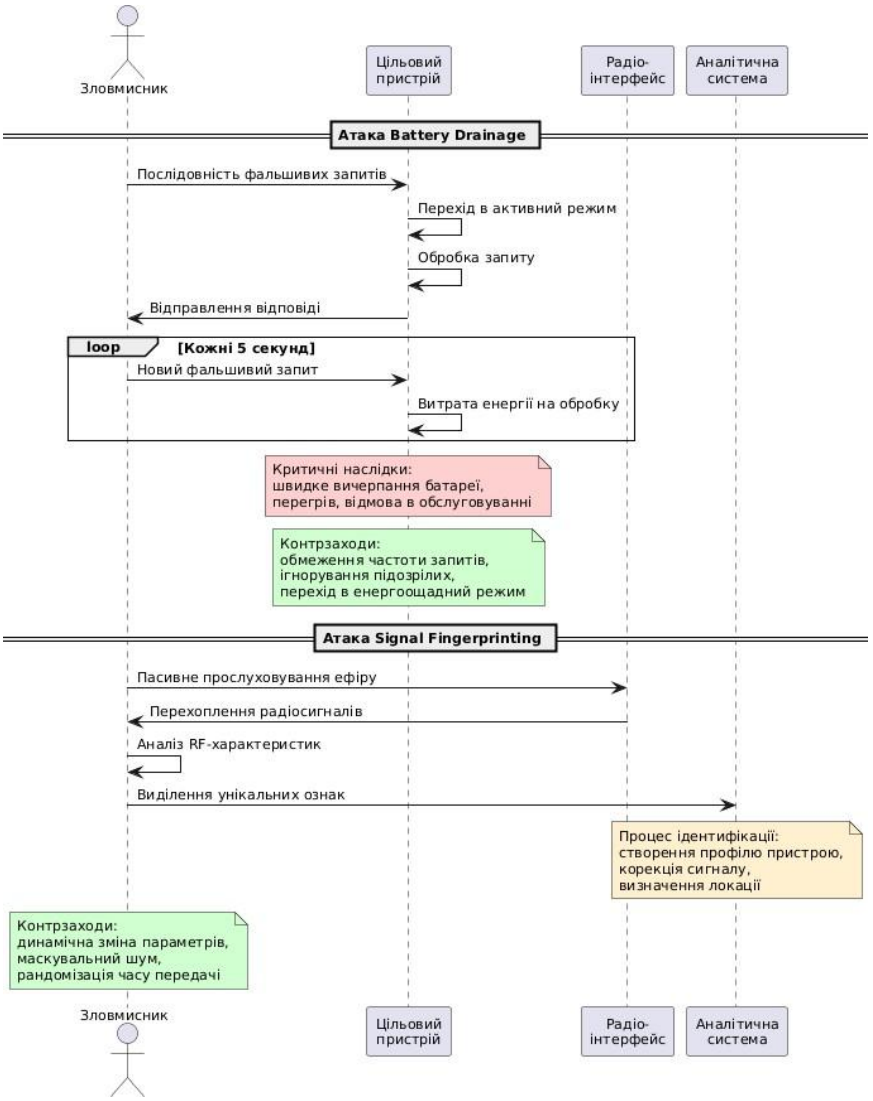

Рисунок 1.8 — Діаграма послідовності атак на енергоефективність та конфіденційність

Таким чином, хоча NB-IoT вважається найбільш захищеною технологією серед LPWAN завдяки використанню ліцензованого спектру та перевірених стільникових механізмів безпеки, вона зберігає низку вразливостей, характерних як для класичних мобільних мереж, так і для IoT-систем.

1.2.5 Порівняльний аналіз рівнів безпеки та критичності загроз

Таблиця 1.2 Порівняння вразливостей LPWAN-технологій

Загроза	LoRaWAN	Sigfox	NB-IoT
DoS-атаки / Jamming	Високий	Високий	Середній
Клонування пристроїв	Середній	Високий	Низький
Цільове заглушення	Високий	Високий	Середній
Активний спуфінг	Високий	Середній	Низький
Пасивне підслуховування	Високий	Середній	Низький
FDI-атаки (семантичні)	Високий	Високий	Середній
Архітектурні вразливості	Середній	Високий	Низький

Як видно з таблиці 1.2, серед розглянутих LPWAN-технологій NB-IoT, що показа вважається найбільш захищеною. Вона побудована на надійній інфраструктурі стільникового зв’язку та успадковує перевірені механізми безпеки 3GPP, зокрема взаємну автентифікацію на основі SIM-карти та наскрізне шифрування. Завдяки цьому критичність більшості кіберзагроз для NB-IoT є значно нижчою порівняно з іншими технологіями.

LoRaWAN забезпечує два незалежних рівні безпеки (мережевий та прикладний) з використанням алгоритму AES-128, що гарантує конфіденційність і цілісність даних. Однак через роботу в неліцензійному діапазоні технологія залишається вразливою до фізичних атак, таких як глушіння (jamming) та атаки повторного відтворення (replay attacks). Ці загрози можуть мати високу критичність, особливо щодо доступності сервісу у великомасштабних розгортаннях.

Sigfox застосовує більш простий підхід до захисту — шифрування від пристрою безпосередньо до хмари з акцентом на безпеку на рівні додатків. Як і LoRaWAN, технологія працює в неліцензійному спектрі, тому також чутлива до атак глушіння. Рівень критичності загроз для Sigfox є переважно середнім і значною мірою залежить від чутливості переданих даних та якості реалізації конкретного прикладного рішення.

Незважаючи на відмінності в рівнях захисту різних технологій, для всіх них існує спільна і особливо небезпечна категорія загроз — атаки типу False Data Injection (FDI). На відміну від класичних криптографічних атак, FDI-атаки не намагаються порушити автентичність пакета, а маніпулюють семантикою даних. Зловмисник надсилає технічно коректні пакети, які успішно проходять усі перевірки, але містять фальшиву інформацію про стан сенсорів. Саме цьому типу атак присвячено подальший детальний аналіз.

1.2.6 False Data Injection (FDI) атаки в LPWAN-мережах

Одним із найбільш небезпечних і водночас складних для виявлення типів атак на кіберфізичні системи є атаки типу False Data Injection (FDI) — ін’єкція фальшивих даних. Ця загроза займає особливе місце серед усіх векторів атак на LPWAN-мережі, оскільки діє на семантичному рівні даних, обходячи стандартні криптографічні механізми захисту [17].

На відміну від класичних криптографічних атак, FDI-атака не намагається порушити автентичність пакета. Зловмисник надсилає технічно коректний пакет, який успішно проходить усі перевірки (AES-128, MIC, Join-процедуру), але містить фальшиві значення сенсорних даних. Формально: $\Psi(P_i(t)) = 1$ (пакет автентичний), але $D_i(t) \neq \varphi(s(t))$ (дані не відповідають реальному фізичному стану). Саме ця властивість робить FDI-атаки критично небезпечними для кіберфізичних систем — система управління отримує “правильні” з криптографічної точки зору, але семантично хибні дані [18].

Реальні інциденти та задокументовані випадки FDI-атак. Реальність загрози FDI-атак підтверджується низкою задокументованих інцидентів та дослідницьких демонстрацій.

Маніпуляції з лічильниками електроенергії (США, ’Puerto Rico’, 2009–2012). Розслідування ФБР задокументувало масштабні випадки підробки показань інтелектуальних лічильників AMI (Advanced Metering Infrastructure) в Пуерто-Ріко та на півдні США. Зловмисники фізично модифікували прошивку пристроїв або використовували магніти для спотворення показань, що призводило до заниження обліку спожитої енергії. Загальні збитки оцінювалися у понад 400 мільйонів доларів щорічно лише для одного постачальника [19]. Хоча ці атаки передували масовому розгортанню LPWAN, вони наочно продемонстрували наслідки семантичних маніпуляцій у кіберфізичних системах обліку ресурсів.

Атака на LoRaWAN-інфраструктуру розумного сільського господарства (дослідження IOActive/Trend Micro, 2021–2022). Дослідники IOActive та Trend Micro опублікували технічний звіт, у якому продемонстрували практичну реалізацію FDI-атак на реальні LoRaWAN-мережі агропромислових підприємств [17]. Використовуючи SDR-обладнання вартістю менш ніж 50 доларів та відкрите програмне забезпечення GNU Radio, дослідники успішно перехоплювали пакети LoRaWAN і відтворювали модифіковані версії з підробленими показниками вологості ґрунту, температури та рівня pH. Оскільки атаки здійснювалися на компрометованих пристроях із легітимними ключами, стандартний криптографічний захист LoRaWAN (AES-128, MIC) не виявляв жодних порушень. Системи автоматичного зрошення реагували на фальшиві дані, що призводило до надмірного поливу та потенційних збитків урожаю. Дослідження підкреслило, що “найслабшою ланкою є не криптографія, а відсутність семантичної валідації даних” [17].

Компрометація систем управління будівлями (BMS) через NB-IoT (Європа, 2022–2023). Агентство Європейського Союзу з кібербезпеки (ENISA)

у своєму щорічному звіті про загрози для IoT зафіксувало серію атак на системи управління будівлями (BMS), підключені через NB-IoT [20]. У кількох випадках зловмисники отримували фізичний доступ до кінцевих пристроїв (датчиків температури та CO_2) та модифікували їхнє мікропрограмне забезпечення для передачі завідомо завищених або занижених показань. Це призводило до некоректної роботи систем вентиляції та опалення в комерційних будівлях. ENISA класифікувала ці інциденти як “семантичні атаки на рівні даних” і виділила їх в окрему категорію загроз, наголосивши на неефективності традиційних IDPS-систем для їх виявлення [20].

Атаки на системи розумного водопостачання (Флорида, США, 2021). Резонансний інцидент у м. Олдсмар, штат Флорида, у лютому 2021 року продемонстрував реальні наслідки маніпуляцій із показниками систем водопостачання [21]. Зловмисник отримав несанкціонований доступ до SCADA-системи водоочисної станції та змінив налаштування концентрації гідроксиду натрію (NaOH) зі 111 до 11 100 частин на мільйон — рівень, небезпечний для здоров’я населення. Хоча цей інцидент реалізовувався через HMI-інтерфейс, а не через LPWAN, він продемонстрував критичну вразливість: системи не мали механізмів валідації фізичної правдоподібності введених значень. Принципова аналогія з FDI-атаками на LPWAN полягає у тому, що в обох випадках маніпуляції відбуваються на рівні даних, а не мережевого протоколу [21].

Дослідницька демонстрація FDI на LoRaWAN-мережах розумного освітлення (Університет Твенте, 2023). Дослідники Університету Твенте (Нідерланди) опублікували роботу, присвячену безпеці LoRaWAN-мереж розумного вуличного освітлення [22]. У контрольованому середовищі вони продемонстрували два типи FDI-атак: одноразову грубу ін’єкцію (додавання 300% до номінального значення освітленості) та “повільну” (stealthy) атаку з поступовим кумулятивним дрейфом показника +5% на кожен пакет. Перший тип виявлявся наявними пороговими системами у 94% випадків, тоді як

другий — лише у 12%. Це підтвердило критичну потребу у методах семантичного аналізу, здатних виявляти приховані маніпуляції [22].

Обґрунтування критичності загрози. Наведені приклади демонструють спільну закономірність: традиційні методи захисту виявляються неефективними проти FDI-атак, оскільки криптографічні механізми (AES-128, MIC) успішно проходяться, а стандартні системи виявлення вторгнень (IDPS) орієнтовані переважно на мережеві аномалії, а не на фізичну семантику даних [18]. Особливо складно виявити повільні (stealthy) FDI-атаки, де зміни показників відбуваються поступово та імітують природні процеси, що підтверджується як дослідженнями [22], так і реальними інцидентами [17, 19].

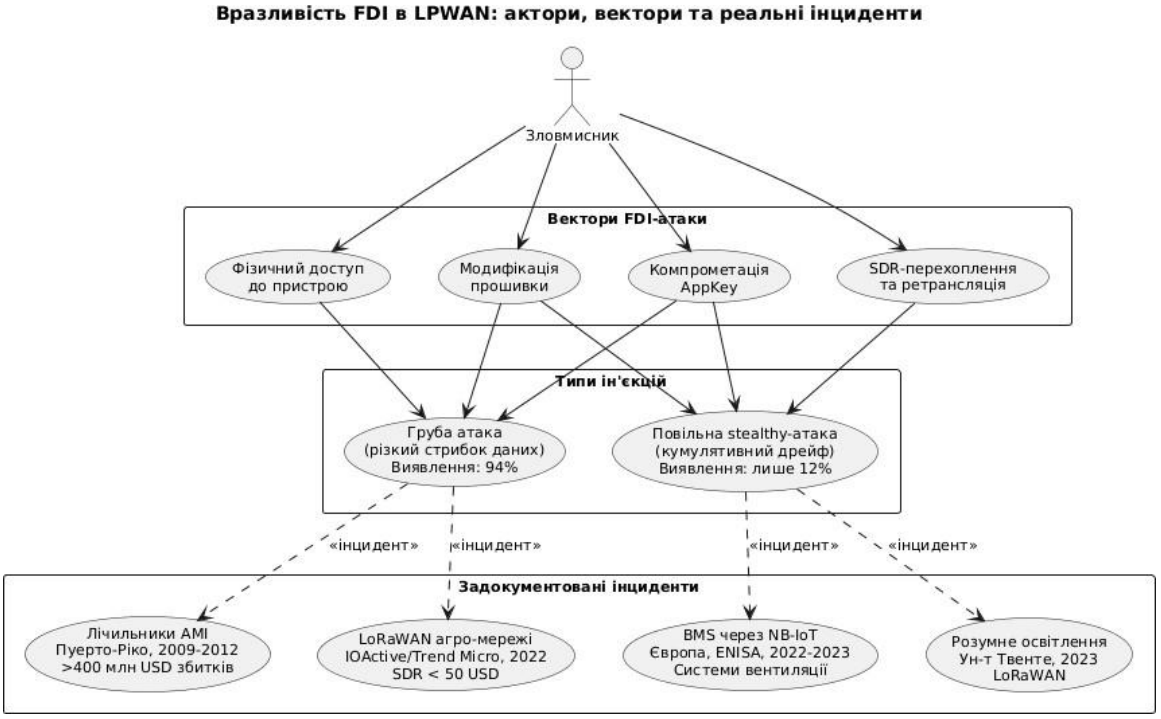

Рисунок 1.9 — Діаграма вразливості FDI. актори вектори і реальні інциденти.

Таким чином, на сьогодні існує критичний розрив між рівнем криптографічного захисту та реальною безпекою кіберфізичних систем. Задокументовані інциденти підтверджують, що FDI-атаки є реальною, а не лише теоретичною загрозою, і вже завдали відчутної шкоди інфраструктурним об’єктам у різних країнах. Саме усунення цього розриву шляхом розробки

ефективних методів виявлення семантичних аномалій є основним завданням даної роботи. Подальший аналіз присвячено актуальним тенденціям та перспективним векторам атак на LPWAN-мережі, які продовжують еволюціонувати разом із поширенням технології.

1.2.7 Актуальні тенденції та перспективні вектори атак на LPWAN-мережі

Сучасні дослідження демонструють швидку еволюцію атак на IoT-інфраструктуру у кількох напрямках. Зростає використання розподілених IoT-ботнетів для організації масованих скоординованих атак на критичну інфраструктуру; активізуються цільові атаки на ключові елементи управління мережею — шлюзи та сервери приєднання — із застосуванням експлойтів нульового дня. Окрему небезпеку становлять криптографічні атаки на ослаблені реалізації алгоритмів, зумовлені обмеженими обчислювальними можливостями кінцевих пристроїв, а також соціально-інженерні та фізичні атаки через відкриті налагоджувальні інтерфейси та компрометацію ланцюгів постачання обладнання. Спільною рисою цих тенденцій є зміщення вектора загроз від мережевого рівня до рівня даних, що додатково підкреслює актуальність семантичних методів виявлення аномалій, розроблених у даній роботі.

Висновки до розділу 1

Проведений аналіз показав суттєві відмінності в рівнях безпеки основних LPWAN-технологій, які безпосередньо випливають з їх архітектурних особливостей. NB-IoT забезпечує найвищий рівень захисту завдяки ліцензованому спектру та механізмам 3GPP. LoRaWAN пропонує гнучкі рівні безпеки з AES-128, але сильно залежить від якості реалізації. Sigfox через ультра-спрощену архітектуру має найбільші фундаментальні обмеження у захисті.

Незважаючи на ці відмінності, для всіх технологій найбільш критичними залишаються атаки на доступність сервісів та конфіденційність даних.

Ефективний захист LPWAN-інфраструктур потребує комплексного багаторівневого підходу, що поєднує криптографію, фізичну безпеку, постійний моніторинг та адаптивні системи виявлення аномалій.

Отримані теоретичні результати щодо архітектури, моделей загроз та рівня захищеності LPWAN-мереж створюють необхідну основу для подальшого практичного дослідження. У наступному розділі буде проведено аналіз захищеності та методів виявлення аномалій LPWAN-мереж, розглянуто існуючі підходи та інструменти оцінки безпеки, адаптовані до особливостей цих технологій.

РОЗДІЛ 2
АНАЛІЗ ЗАХИЩЕНОСТІ ТА МЕТОДІВ ВИЯВЛЕННЯ АНОМАЛІЙ LPWAN-МЕРЕЖ

2.1Аналіз захищеності LPWAN

Аналіз безпеки LPWAN вимагає адаптації стандартних підходів до оцінки ризиків, оскільки класичні методи, розроблені для традиційних IT-систем, часто виявляються неефективними в умовах обмеженої пропускної здатності, низького енергоспоживання та значного радіусу дії [7, 12]. Для комплексного дослідження використовується комбінований підхід, що поєднує загальновизнані стандарти OWASP з техніками, адаптованими до архітектури LPWAN-мереж [9, 10, 13].

2.1.1 Багаторівневий підхід до аналізу захищеності

В основу аналізу покладено методологію OWASP IoT Security Testing Framework [10], доповнену інтеграцією двох ключових стандартів — OWASP Web Top 10 та OWASP IoT Top 10 [9, 14]. Таке поєднання дає цілісне бачення захищеності всієї LPWAN-екосистеми: перший стандарт охоплює веб-інтерфейси управління мережею та API-взаємодію, другий — безпеку безпосередньо кінцевих пристроїв та шлюзів [13].

З OWASP Web Top 10 особливо актуальними для LPWAN є контроль доступу (A01) — для захисту веб-інтерфейсів управління мережею, криптографічні помилки (A02) — для аналізу шифрування в протоколах, та неправильні налаштування (A05) — для перевірки конфігурації мережевих компонентів [14]. З OWASP IoT Top 10 ключовими є слабкі паролі (I1) — для аналізу ключів автентифікації пристроїв, небезпечні мережеві сервіси (I2) — для тестування шлюзів та серверів, та відсутність оновлень (I4) — для перевірки процедур оновлення прошивок [9]. Структуру відповідності компонентів LPWAN обом стандартам наведено на рисунку 2.1.

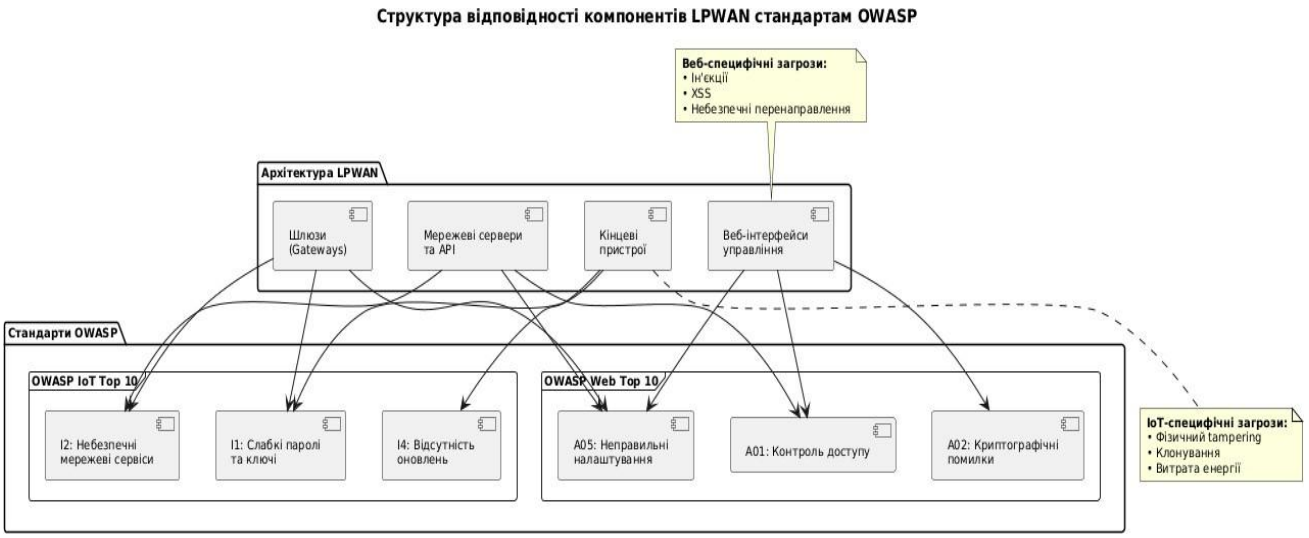

Рисунок 2.1 — Структура відповідності компонентів LPWAN стандартам OWASP

Поєднання стандартів уможливлює наскрізний (End-to-End) аналіз на всьому шляху передачі інформації — від первинного сенсора до інтерфейсу кінцевого користувача [15]. Це є критично важливим для виявлення складних комбінованих атак, які одночасно експлуатують вразливості на різних рівнях системи — наприклад, використання веб-вразливості (XSS) для компрометації облікових даних з метою подальшого несанкціонованого доступу до керування IoT-пристроями [11]. Крім того, такий підхід створює умови для реалізації семантичного аналізу даних: аналізу піддається не лише факт успішної доставки пакета, а й його відповідність поточному фізичному стану технологічного процесу, що є основою для виявлення FDI-атак [18].

Для систематизації процесу аудиту розроблено чотирифазовий регламент тестування, де результати кожної фази стають вхідними даними для наступної. Фаза 1 охоплює аналіз пристроїв за критеріями IoT Top 10 (слабкі ключі, безпека прошивок, фізичний захист); Фаза 2 — тестування мережі (безпека мережевих сервісів, конфігурація компонентів); Фаза 3 — аналіз серверів (контроль доступу до API, шифрування даних); Фаза 4 — тестування веб-інтерфейсів (ін'єкції, помилки автентифікації, недостатнє логування) [10, 13]. Така уніфікована процедура тестування забезпечує повне покриття векторів загроз, визначених у розділі 1.

Поєднання двох стандартів дає цілісне бачення безпеки всієї LPWAN-екосистеми, що є критично важливим для забезпечення сталого

функціонування кіберфізичних систем [16]. У той час як технічні стандарти регламентують механізми автентифікації і шифрування трафіку [6], організаційно-методологічні стандарти визначають вектори загроз, пов'язані з цілісністю сенсорних даних [9, 14].

Використання комбінованого підходу дозволяє розробити уніфіковані метрики оцінки безпеки, які враховують як мережеві параметри (SNR, RSSI, Duty Cycle), так і логічну структуру переданої інформації [7]. Таке розширене бачення архітектури безпеки є необхідною умовою для впровадження інтелектуальних методів аналізу, зокрема алгоритмів машинного навчання, оскільки забезпечує репрезентативність вхідних даних та чітке розмежування між нормальним функціонуванням мережі та аномальною активністю [22].

Зокрема, інтеграція підходів OWASP Web Top 10 та OWASP IoT Top 10 дозволяє нівелювати розрив між рівнем польових пристроїв та рівнем обробки даних у хмарі або на мережевому сервері [12]. Це створює умови для реалізації концепції «контекстуальної безпеки», де аналізу піддається не лише факт успішної доставки пакета, а й його семантична відповідність поточному стану технологічного процесу [18]. Наприклад, виявлення маніпуляцій з даними (FDI-атак) стає можливим через кореляцію мережевих метаданих із фізичними законами функціонування КФС [22].

Більш того, такий багаторівневий аналіз дозволяє ефективно протидіяти атакам, що маскуються під легітимний трафік [17]. Оскільки зловмисники все частіше використовують стратегії низькоінтенсивних (stealthy) втручань, які не викликають різких стрибків мережевої активності, єдиним дієвим інструментом захисту стає глибока інспекція пакетів у поєднанні з поведінковим аналізом кінцевих вузлів [22].

Це трансформує роль системи захисту з пасивного бар'єру в активний інтелектуальний компонент інфраструктури, здатний до самонавчання та адаптації до нових типів загроз. Впровадження цієї стратегії забезпечує не лише конфіденційність передачі, а й функціональну стійкість (resilience) всієї системи, що є першочерговим завданням при моніторингу об'єктів критичної інфраструктури, де ціна помилки або простою є надзвичайно високою.

2.1.2. Модель атакуючого для LPWAN-мереж.

Для комплексного аналізу інформаційної безпеки мережі LPWAN, можна віднести 4 категорії атакуючих, кожна з них має свою мотивацію, цілі та свій стек можливостей[13]. Для кількісної оцінки загроз використано модель оцінки ризику:

$$Risk = L \times I \tag{2.1}$$

де L — ймовірність реалізації загрози (0-10), а I - потенційний вплив на систему(0-10).

На основі аналізу архітектури LPWAN виділено чотири основні категорії зловмисників, їх профілі, типові вектори атак та розраховані показники ризику зведено у таблицю 2.1 [13, 15].

Таблиця 2.1 — Модель атакуючого для LPWAN-мереж

Категорія	Профіль	Типові вектори атак	L	I	Risk
Зловмисник з фізичним доступом	Колишній співробітник, вандал;	Tampering, клонування пристроїв,	7,5	6,8	51,0

	базовий рівень знань	маніпуляції з живленням			
Віддалений пасивний слухач	Радіолюбитель, дослідник; SDR-обладнання	Перехоплення RF-сигналів, аналіз трафіку, криптоаналіз	8,2	5,5	45,1
Активний мережевий атакуючий	Хакер, інсайдер; мережеві експлойти	Спуфінг пакетів, DoS/DDoS, MitM	6,3	8,9	56,1
Кримінальні та державні групи (APT)	Держструктури; необмежені ресурси	Zero-day експлойти, компрометація ланцюгів постачання	4,1	9,7	39,8

Аналіз таблиці 2.1 показує, що найвищий показник ризику має категорія активного мережевого атакуючого (Risk = 56,1), оскільки поєднує відносно високу ймовірність реалізації з критичним рівнем впливу на систему [16]. Зловмисник з фізичним доступом посідає друге місце (Risk = 51,0) — саме ця категорія найбільш релевантна до сценаріїв FDI-атак, оскільки фізичний доступ до пристрою дозволяє компрометувати ключі та модифікувати прошивку для надсилання криптографічно коректних, але семантично хибних даних [17, 19]. Ієрархію категорій та їх можливості наочно представлено на рисунку 2.2.

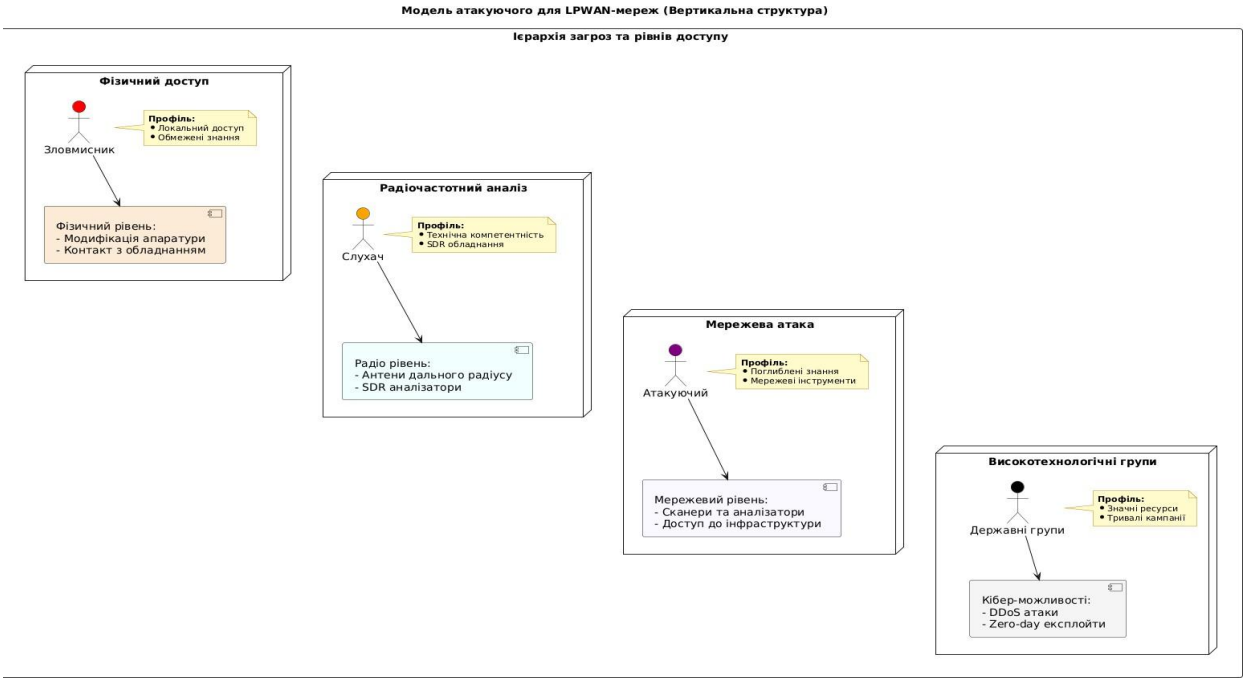

Рисунок 2.2 — Модель атакуючого для LPWAN-мереж

Висновок до розділу 2

Проведений аналіз безпеки екосистеми LPWAN та побудована модель загроз дозволяють сформулювати наступні висновки.

По-перше, багаторівнева структура мережі вимагає комплексного підходу, що здатен протидіяти широкому спектру зловмисників — від вандалів до високотехнологічних груп [13, 15]. Інтеграція стандартів OWASP Web Top 10 та IoT Top 10 забезпечує наскрізний контроль цілісності на кожному етапі передачі даних [9, 14] і створює умови для виявлення складних комбінованих атак [11, 12].

По-друге, кількісна модель ризиків підтверджує, що найбільш критичними загрозами для цілісності системи є активні мережеві втручання (Risk = 56,1) та атаки з фізичним доступом до пристроїв (Risk = 51,0) [16, 20]. Саме останні є ключовим вектором для FDI-атак, що становлять основний предмет дослідження даної роботи [17, 18].

По-третє, оскільки традиційні методи захисту на основі сигнатур виявляються неефективними проти семантичних атак [18, 20], виникає потреба у системах інтелектуального аналізу поведінки мережевих вузлів [22]. Це логічно зумовлює перехід до розділу 3, присвяченого розробці та апробації гібридного методу виявлення аномалій на основі алгоритмів машинного навчання [22].

РОЗДІЛ 3

ПРАКТИЧНА РЕАЛІЗАЦІЯ МЕТОДУ ВИЯВЛЕННЯ АНОМАЛІЙ LPWAN НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Аналіз архітектури та вразливостей LPWAN-мереж, проведений у розділах 1 та 2, показав, що стандартні механізми безпеки (AES-128 шифрування, MIC для цілісності, процедури OTAA/ABP для автентифікації) забезпечують захист на криптографічному та мережевому рівнях [1, 6], але залишають відкритою вразливість на семантичному рівні даних [18]. Зокрема, атаки типу False Data Injection (FDI) дозволяють зловмиснику ін'єктувати фальшиві показники сенсорів у технічно коректні пакети, не порушуючи криптографічного захисту [17, 18]. У контексті систем «розумного» вуличного освітлення це може призвести до критичних наслідків: помилкового увімкнення/вимкнення ліхтарів, перевантаження енергомережі, неефективного використання ресурсів або навіть створення аварійних ситуацій у міській інфраструктурі [22].

Класичні методи виявлення аномалій (порогові перевірки, статистичні фільтри типу CUSUM, прості правила на основі діапазонів) виявляються недостатньо чутливими до повільних (stealthy) FDI-атак, коли зловмисник змінює дані поступово, імітуючи природну динаміку фізичного процесу [18, 22]. Водночас чисто машинно-навчальні підходи (наприклад, ізольовані моделі без учителя) часто ігнорують фізичну семантику та просторово-часовий контекст сусідніх пристроїв, що знижує їх ефективність у реальних IoT-системах [22].

3.1 Постановка задачі практичної частини

Метою практичного етапу дослідження є апробація та кількісна оцінка запропонованого гібридного методу виявлення аномалій на рівні мережевого сервера LPWAN-мережі, що не потребує модифікації кінцевих пристроїв [22]. Запропонований метод базується на синергії алгоритму машинного навчання без учителя Isolation Forest [24], який відповідає за виявлення статистичних відхилень у багатовимірному просторі ознак, та детермінованих правил

доменних знань [22].

Задача формалізується через розгляд мережі, що складається з N кінцевих сенсорів, де кожен пристрій i у момент часу t генерує пакет$P_{\mathrm{i}}(t) = (M_{\mathrm{i}}(t), D_{\mathrm{i}}(t))$. У цій структурі $M_{\mathrm{i}}(t) = \{\tau, RSSI, SNR, ID\}$ представляє мережеві метадані LoRaWAN [1], $D_{\mathrm{i}}(t) \in \mathbb{R}^{\mathrm{m}}$— вектор безпосередніх показників сенсорів. Хоча стандартна криптографічна перевірка $\Psi(P_{\mathrm{i}}(t)) \in \{0,1\}$ гарантує автентичність пакета [6], вона залишається нечутливою до його семантичного змісту [18]. FDI-атака відбувається, коли $\Psi(P_{\mathrm{i}}(t)) = 1$, але $D_{\mathrm{i}}(t) \neq \varphi\big(s(t)\big)$, де $\mathrm{s(t)}$ — реальний фізичний стан середовища, а φ — відображення стану у показники сенсорів [18, 22].

Для вирішення цієї проблеми синтезується функція семантичного аналізу(3.1):

$$\mathcal{A}(P_{\mathrm{i}}(t), H_{\mathrm{i}}(t), C_{\mathrm{i}}(t)) \rightarrow score \in [0,1] \tag{3.1}$$

де $H_{\mathrm{i}}(t)$ — часовий контекст (останні k пакетів від пристрою i), $C_{\mathrm{i}}(t)$ — просторовий контекст (пакети від сусідніх пристроїв $\mathcal{N}(\mathrm{i})$) [22].

Пакет вважається аномальним, якщо $score > \theta$, де θ — адаптивний поріг, обчислений у ковзному режимі [24].

Критеріями успішності розв'язання задачі визначено здатність системи виявляти втручання при мінімальній кількості помилкових спрацювань [25]. Для оцінки якості класифікації використовуються метрики точності (Precision) — частка справжніх аномалій серед усіх виявлених, та повноти (Recall) — здатність системи знайти всі існуючі аномалії [25]. Узагальнюючим показником обрано F1-score, що є гармонійним середнім між ними [25], із цільовим значенням $F1 \geq 0.90$(3.2):

$$F1 = 2 \times \frac{Precision + Recal}{Precision \cdot Recall} \tag{3.2}$$

Важливою вимогою є мінімізація затримки виявлення до рівня $\Delta t<2$ пакетів для забезпечення режиму реального часу, а також здатність моделі до навчання виключно на нормальних даних без попередньої розмітки атак. Практична реалізація проводиться у середовищі Jupiner Notepad, в розробці

моделі була використана мова програмування Python, що гарантує відтворюваність експериментів та відповідність принципам тестування безпеки IoT.

3.2. Побудова датасету для експериментів

Для формування вектора ознак $f \in \mathbb{R}^6$ використано реальний публічний датасет «IoT-Enabled Smart Lighting Dataset» (Kaggle, автор Ziya, 2024) [26], який містить дані сенсорів вуличного освітлення за 2024 рік по 12 міських зонах: рівень освітленості (ambient light, lux), температура середовища, погодні умови, щільність трафіку, виявлення руху, енергоспоживання та рішення про увімкнення/вимкнення освітлення [25]. Загальна кількість записів після обробки становить 12 000 рядків × 17 колонок, кожному з яких присвоєно ідентифікатор пристрою device_id у діапазоні від 1 до 12, що відповідає 12 міським зонам.

До датасету додано синтетичні LoRaWAN-метадані — інтервал передачі τ, потужність сигналу *RSSI* та відношення сигнал/шум *SNR*, згенеровані за нормальним розподілом, характерним для міських умов розгортання ($\tau \sim N(3600, 300)$ сек, $RSSI \sim N(-100, 10)$ дБм, $SNR \sim N(10, 5)$ дБ) [1,2]. Такий підхід дозволяє відтворити реалістичні мережеві характеристики LoRaWAN без необхідності розгортання фізичної тестової інфраструктури [4].

Систему розумного вуличного освітлення на основі LoRa розглянуто, зокрема, у роботі [24], де описано практичну реалізацію моніторингу параметрів освітленості та електроенергії через LPWAN-мережу.

Імітація FDI-атак. Для формування розміченої вибірки змодельовано два типи атак ін'єкції фальшивих даних [18, 22]:

- грубі атаки — додавання до показника освітленості величини 2–4 × std зони (від ~8 000 до ~26 000 люкс) для 7% пакетів; такий масштаб гарантовано виводить значення за межі 3σ нормального розподілу, що відповідає реальному сценарію одномоментної підміни показника сенсора зловмисником [22];

- повільні атаки — поступовий кумулятивний дрейф показника з кроком 0,3 × std зони на пакет (~1 200–1 950 люкс/крок) для ~3% пакетів; після 10 кроків

загальний дрейф перевищує 3σ, що дозволяє виявити приховану маніпуляцію через правило r_3 (падіння кореляції).Кожен атакований пакет отримував мітку is_anomaly = True, що слугує ground truth для подальшої оцінки якості моделі. [18].

Кожен атакований пакет отримував мітку is_anomaly = True, що слугує ground truth для подальшої оцінки якості моделі. Результати імітації атак наведено на рисунку 3.1.

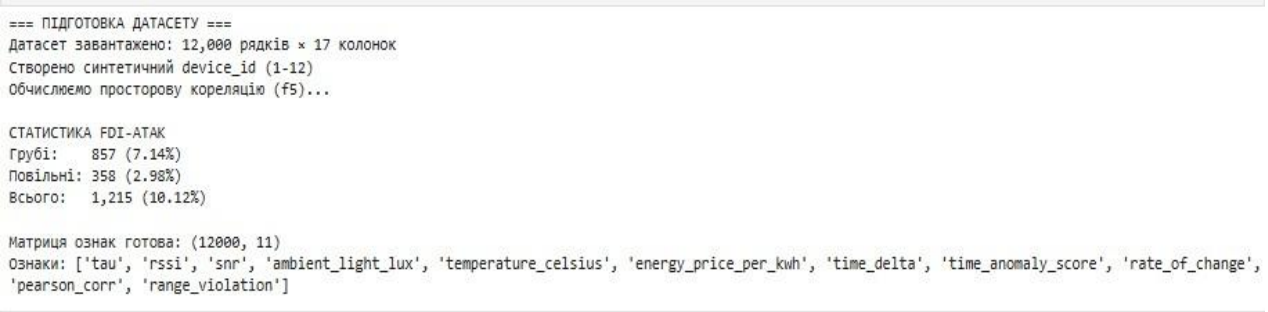

Рисунок 3.1 — Результат виконання підготовки датасету (вивід Jupyter Notebook)

Як видно з рис. 3.1, після імітації створено 1 215 аномальних записів, що становить 10,12% від загальної вибірки, з яких 857 — грубі атаки (7,14%) та 358 — повільні (2,98%). Таке співвідношення відповідає реалістичному сценарію LPWAN-мережі, де атаки є відносно рідкісними подіями на тлі нормального трафіку. Сформована матриця ознак X_raw розмірністю 12 000 × 11 передається на наступний етап — оцінку вразливості та хронологічну розбивку вибірок, що описано у підрозділах 3.2.1 та 3.4.

3.2.1 Оцінка вразливості до атак типу False Data Injection

Проведена імітація FDI-атак підтвердила наявність критичної вразливості на семантичному рівні в типових реалізаціях LPWAN-мереж «розумного» вуличного освітлення. Без додаткових механізмів семантичного контролю вразливість класифікується як критична.

Для кількісної оцінки використано CVSS v3.1 та OWASP Risk Rating Methodology.

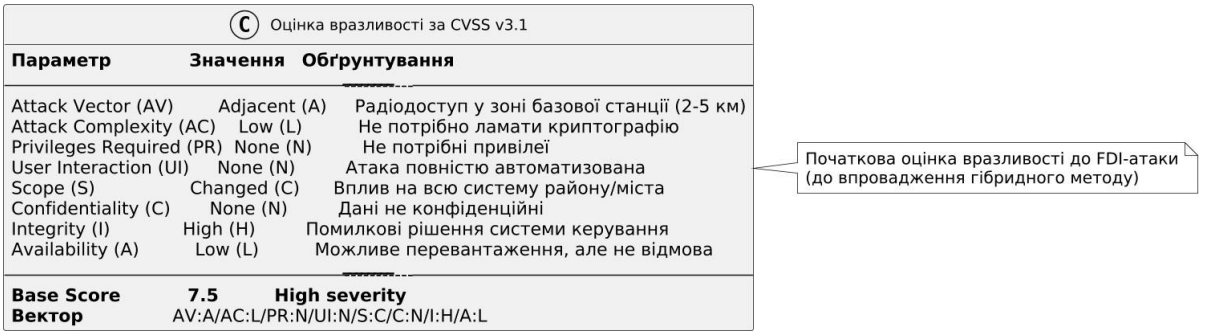

Рисунок 3.2— діаграма оцінки вразливостей за CVSS

Базовий бал 7.5 відповідає рівню High severity.

Згідно з методу до оцінки ризиків OWASP Risk Rating Methodology, було проведено комплексний аналіз ймовірності реалізації та потенційного впливу семантичних атак на систему. Ймовірність реалізації *Likelihood* класифікується як висока *High*, оскільки технічне виконання такої атаки є відносно простим і не потребує високовартісного обладнання, залишаючись доступним за умови використання SDR-технологій або модифікованих кінцевих пристроїв. Потенційний вплив *Impact* оцінюється як важкий *Severe*, що зумовлено значними економічними збитками від неефективного енергоспоживання, ризиками для безпеки дорожнього руху в нічний час та можливістю виникнення аварійних ситуацій у міській інфраструктурі.

вектор ознак включає мережеві метадані, семантичні показники сенсорів та міжпристроєву кореляцію[25] FDI-атак у сучасних LPWAN-системах «розумного» освітлення, позбавлених додаткових засобів інтелектуального контролю даних, класифікується як критична загроза, що потребує негайного впровадження превентивних механізмів захисту.

3.3 Формування вектора ознак $\mathbf{f} \in \mathbb{R}^6$

Центральним елементом запропонованого методу виявлення FDI-атак є вектор ознак $f \in \mathbb{R}^6$, який формується для кожного пакету $P_i(t)$ на основі його мережевих метаданих $M_i(t)$, показників сенсора $D_i(t)$, часового контексту $H_i(t)$

та просторового контексту $C_i(t)$. На відміну від підходів, що аналізують лише один вимір даних, запропонований вектор охоплює три групи характеристик, що є взаємодоповнюючими: мережеві, семантичні та контекстні ознаки. Використання мережевих метаданих RSSI та SNR як ознак для аналізу трафіку LPWAN підтверджено рядом досліджень [25], де ці параметри застосовувались для задач класифікації та виявлення аномалій. Така структура забезпечує виявлення як грубих одномоментних ін'єкцій, так і повільних прихованих атак, що імітують природну динаміку фізичного процесу.

Формально вектор ознак визначається як (3.3):

$$f = [f_1, f_2, f_3, f_4, f_5, f_6]^T \in \mathbb{R}^6 \quad (3.3)$$

де кожна компонента відповідає за певний аспект аномальності пакету. Зведена характеристика вектора наведена у таблиці 3.1.

Таблиця 3.1 — Структура вектора ознак $f \in \mathbb{R}^6$

Ознака	**Позначення**	**Група**	**Що вловлює**
f_1	time_delta	мережева	Порушення ритму передачі пакетів
f_2	time_anomaly_score	мережева	Нормалізоване відхилення від очікуваного τ
f_3	ambient_light_lux	семантична	Абсолютне значення показника сенсора
f_4	rate_of_change	семантична	Різкий стрибок показника між пакетами

f_5	pearson_corr	контекстна	Аномальне падіння кореляції з сусіднім пристроєм
f_6	range_violation	контекстна	Вихід значення за фізично допустимий діапазон

Оскільки компоненти вектора f мають суттєво відмінні діапазони та одиниці вимірювання — f_1 вимірюється в секундах (порядок 10^2–10^4), f_3 у люксах (10^2–10^4), тоді як f_5 є безрозмірним коефіцієнтом $\in [-1, 1]$, а f_6 — бінарним індикатором $\in \{0, 1\}$ — безпосереднє порівняння їх у єдиному просторі неможливе без попереднього масштабування. Для приведення всіх ознак до єдиного діапазону застосовується стандартизація через $z = (x - \mu) / \sigma$, що описана у підрозділі 3.4.2. Після нормалізації кожна компонента має $\mu = 0$ та $\sigma = 1$, що усуває домінування ознак з великим масштабом та забезпечує рівноважний внесок усіх шести компонент у аномальний score алгоритму Isolation Forest.

3.3.1 Мережеві ознаки (f_1, f_2)

Мережеві ознаки описують статистичну поведінку пристрою на рівні каналу передачі даних і є першим індикатором можливої аномалії[7]. При FDI-атаці зловмисник надсилає підроблені пакети, що порушує природний ритм передачі легітимного пристрою[17] — саме це відхилення фіксують ознаки f_1 та f_2.

Ознака f_1 — відхилення інтервалу передачі (time_delta). Кожен пристрій LoRaWAN надсилає пакети з певним очікуваним інтервалом τ, що відповідає його налаштуванням (в умовах датасету $\tau \sim N(3600, 300)$ сек)[1]. Ознака f_1 фіксує реальну різницю в секундах між двома послідовними пакетами одного пристрою (3.4):

$$f_1 = \tau(t) - \tau(t - 1) \tag{3.4}$$

де $\tau(t)$ — мітка часу поточного пакету, $\tau(t-1)$ — мітка часу попереднього пакету того самого пристрою. Обчислення виконується в межах групи кожного пристрою окремо через операцію groupby(device_id), щоб виключити порівняння між різними пристроями[25]. Для першого пакету кожного пристрою, де попереднє значення відсутнє, застосовується заміна на медіану по всій вибірці.

При нормальній роботі f_1 близька до очікуваного τ з невеликим відхиленням. При FDI-атаці, коли зловмисник надсилає додаткові пакети поза розкладом або пропускає передачі, f_1 набуває аномальних значень — значно менших або більших за норму[15, 18].

Ознака f_2 — нормалізоване відхилення інтервалу (time_anomaly_score). Вона нормалізує відхилення f_1 відносно очікуваного інтервалу $\bar{\tau}_i$ конкретного пристрою(3.5):

$$f_2 = |time\_delta - \bar{\tau}_i| / \bar{\tau}_i \quad (3.5)$$

де $\bar{\tau}_i$ — середній очікуваний інтервал для пристрою i, обчислений як ковзне середнє по всій його історії. Нормалізація дозволяє порівнювати відхилення між пристроями з різними інтервалами передачі на єдиній шкалі[25]. Значення f_2 близьке до 0 означає нормальний ритм, значення вище 1 — відхилення більше ніж на один інтервал, що є сильним індикатором аномалії[24].

Важливим аспектом є те, що мережеві ознаки f_1 та f_2 є відносно слабким індикатором для повільних FDI-атак — зловмисник, що використовує легітимний пристрій, зберігає нормальний ритм передачі[18, 22]. Саме тому вони доповнюються семантичними та контекстними ознаками.

3.3.2 Семантичні ознаки (f_3, f_4)

Семантичні ознаки аналізують безпосередньо показники сенсора — рівень освітленості ambient_light_lux — і є ключовими для виявлення FDI-атак на рівні даних[18]. Саме ці ознаки відповідають на питання: чи відповідає значення $D_i(t)$ реальному фізичному стану середовища $\varphi(s(t))$?

Ознака f_3 — абсолютне значення показника сенсора (ambient_light_lux), яка є значенням освітленості, що передає пристрій у пакеті $P_i(t)$. У контексті гібридної моделі f_3 використовується Isolation Forest для виявлення статистичних відхилень у багатовимірному просторі ознак[24] — пакети з аномально великими або малими значеннями освітленості отримують вищий anomaly score. При грубій FDI-атаці, коли до показника додається 2–4 × std зони (від ~8 000 до ~26 000 люкс), значення f_3 гарантовано виходить за межі 3σ і легко виявляється як Isolation Forest, так і правилом r_1(виявлення виходу за діапазон, описане у підрозділі 3.4)[22].

Ознака f_4 — швидкість зміни показника сенсора (rate_of_change), реалізує першу дискретну похідну показника сенсора по часу(3.6):

$$f_4 = (D(t) - D(t-1)) / \Delta t \quad (3.6)$$

де $D(t)$ — поточне значення освітленості, $D(t-1)$ — попереднє значення того самого пристрою, Δt — реальний інтервал між пакетами[26]. Ознака вимірює наскільки швидко змінюється показник між двома послідовними пакетами одного пристрою в одиницях люкс/секунда. Використання часових похідних (rate of change) як ознак для виявлення аномалій у IoT-даних є загальновизнаним підходом [26].

Фізична мотивація ознаки полягає в тому, що рівень вуличного освітлення є повільно змінним процесом [22] — він залежить від часу доби, хмарності та руху транспорту і змінюється плавно. Різкий стрибок f_4 в умовах нормального функціонування є фізично неможливим або вкрай малоймовірним[18]. При грубій FDI-атаці, коли до значення одномоментно додається кілька тисяч люкс, f_4 набуває аномально великого значення, що перевищує статистично встановлений поріг δ[24]. Саме на цій ознаці базується детерміноване правило r_2гібридної моделі.

При обчисленні f_4 можуть виникати нескінченні значення у випадках коли Δt наближається до нуля. Для запобігання цьому застосовується обмеження знизу через операцію clip(lower=1) та подальша заміна нескінченних значень: replace([inf, −inf], NaN) з наступним заповненням через fillna(0).

3.3.3 Контекстні ознаки (f_5, f_6)

Контекстні ознаки враховують зовнішній контекст пакету — фізичне оточення пристрою та статистичні межі нормальної поведінки [22]. Вони є найефективнішими для виявлення повільних FDI-атак, коли грубих відхилень у значеннях немає, але поведінка пристрою поступово розходиться з фізичною реальністю [18, 22].

Ознака f_5 — кореляція Пірсона з сусіднім пристроєм (pearson_corr), ґрунтується на фізичному принципі: сусідні ліхтарі на одній вулиці освітлюють одне і те саме середовище, тому їх показники освітленості мають корелювати між собою в часі [22]. При FDI-атаці один пристрій починає надсилати підроблені дані — його показники більше не відповідають реальному фізичному стану і кореляція з сусідом падає.

Формально f_5 визначається як коефіцієнт кореляції Пірсона між показниками пристрою i та його сусіда j ∈ $\mathcal{N}$(i) по ковзному вікну з W останніх спостережень(3.7) [27]:

$$f_5 = \rho Pearson(D_i(t - W{:}\,t), D\ (t - W{:}\,t)) \qquad (3.7)$$

де $\rho Pearson$ — коефіцієнт лінійної кореляції Пірсона, W = 10 спостережень, мінімальний розмір вікна для обчислення — 3 спостереження (min_periods=3). Застосування кореляції Пірсона для виявлення просторових аномалій між сусідніми IoT-пристроями обґрунтовано в роботі [27.

Технічно обчислення реалізоване через тимчасове об'єднання рядів двох пристроїв за найближчою міткою часу в межах допуску ±2 години (pd.merge_asof з direction='nearest') [25], що вирішує проблему неідеальної синхронізації пакетів у реальній LoRaWAN-мережі. При нормальній роботі f_5 близька до 1. Значення нижче порогу $\rho_{min} = \mu_{corr} - 2\sigma_{corr}$, обчисленого на нормальних тренувальних даних, є індикатором аномалії і активує детерміноване правило r_3[24].

Ознака f_6 — індикатор виходу за допустимий діапазон (range_violation) реалізує жорстку перевірку фізичних меж показника сенсора для кожного пристрою окремо(3.8):

$$f_6 = \mathbb{1}[D_{\text{i}}(t)\ D_{min}^{i} \notin [, D_{max}^{i}] \tag{3.8}$$

де $D_{min}^{i} = \mu_{\text{i}} - 3\sigma_{\text{ i}}$ та $D_{max}^{i} = \mu_{\text{i}} + 3\sigma_{\text{i}}$ — нижня та верхня межа допустимого діапазону для пристрою i, обчислені як mean ± 3σ по всій його історії. За правилом трьох сигм нормального розподілу 99,7% нормальних значень потрапляють у цей діапазон[27] — все що виходить за межі є статистичною аномалією[24].

Ознака f_6 є бінарною — приймає значення 0 (норма) або 1 (аномалія). Вона є найпростішим але найнадійнішим індикатором грубих FDI-атак: при додаванні 2–4 × std зони значення гарантовано виходить за межі 3σ, і $r_1 = 1$ у 54% аномальних пакетів тестової вибірки [22]. На цій ознаці безпосередньо базується детерміноване правило r_1гібридної моделі.

Межі D_{min} та $D_{\max}$ обчислюються окремо для кожного пристрою і приєднуються до датасету через операцію merge по device_id[25], що дозволяє враховувати індивідуальні характеристики кожного сенсора — різні базові рівні освітленості у різних зонах міста.

Висновок до підрозділу 3.3

Запропонований вектор ознак $f \in \mathbb{R}^6$ забезпечує комплексний аналіз пакету одночасно на трьох рівнях — мережевому (f_1, f_2), семантичному (f_3, f_4) та контекстному (f_5, f_6), — що дозволяє виявляти як грубі одномоментні ін'єкції, так і повільні приховані атаки, які імітують природну динаміку фізичного процесу. Сформована матриця ознак X_raw розмірністю 12 000 × 11 передається на наступний етап — підготовку навчальної та тестової вибірок без витоку даних, що описано у підрозділі 3.4.

3.4 Підготовка навчальної та тестової вибірок

Коректна підготовка навчальної та тестової вибірок є критично важливим етапом, що безпосередньо впливає на достовірність отриманих метрик якості моделі[25]. Оскільки датасет являє собою часовий ряд — послідовність пакетів від 12 пристроїв впродовж 2024 року [26] — до нього не можна застосовувати стандартні методи випадкового розбиття, характерні для незалежних спостережень. Використання випадкового розбиття на часових рядах призводить до явища витоку даних з майбутнього у минуле (data leakage), що штучно завищує метрики якості і робить оцінку нерепрезентативною для реального сценарію розгортання [24, 25].

3.4.1 Хронологічна розбивка (75/25)

Для коректної оцінки моделі застосовано хронологічну розбивку датасету: перші 75% записів за часом утворюють тренувальну вибірку, останні 25% — тестову[25]. Така схема відтворює реальний сценарій розгортання системи виявлення аномалій — модель навчається на минулих даних і перевіряється на майбутніх, яких вона не бачила під час навчання[24, 25]. Принципова різниця між хронологічним та випадковим розбиттям полягає у наступному:

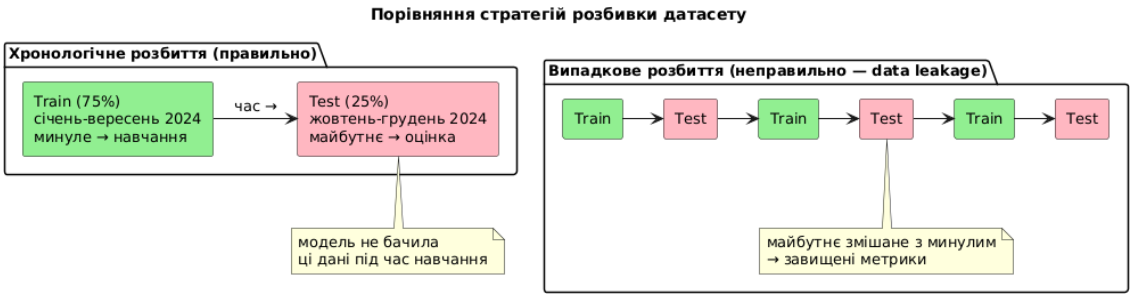

Рисунок 3.3— діаграма порівняння розбивки датасету

Для більш об'єктивної оцінки метрик Recall та F1-score для класу аномалій додатково сформовано збалансовану тестову вибірку з рівним співвідношенням нормальних та аномальних пакетів (50/50)[25]. Це є стандартною практикою при порівнянні моделей виявлення аномалій, оскільки дозволяє уникнути зміщення (bias) у бік majority-класу та отримати надійнішу оцінку здатності моделі виявляти саме атаки[25, 28]. Результат виконання розбивки наведено на рисунку 3.4.

```
=== РОЗБИВКА ДАТАСЕТУ ===
Train: 1,822 рядків | аномалій: 50.00%
Test:  608  рядків | аномалій: 50.00%

Масштабування виконано коректно:
  - scaler.fit_transform() → тільки на train
  - scaler.transform()     → тільки на test
  Тестовий сет НЕ впливає на параметри нормалізації.
```

Рисунок 3.4 — Результат виконання підготовки вибірок (вивід Jupyter Notebook)

Як видно з риунку 3.4, тренувальна вибірка містить 1 822 записи, тестова — 608 записів, у обох частка аномальних пакетів становить рівно 50,00%. Для навчання ML-компоненти Isolation Forest з тренувальної вибірки виокремлено лише нормальні записи (y_train == 0) — 911 легітимних пакетів [24]22, що повністю відповідає реальному сценарію розгортання, де приклади атак на момент навчання відсутні Нормалізацію вхідних ознак виконано за допомогою MinMaxScaler з бібліотеки scikit-learn [28], при цьому параметри нормалізатора обчислювались виключно на навчальній вибірці для уникнення витоку даних (data leakage).

3.4.2 Нормалізація без data leakage

Ознаки вектора f мають суттєво різні масштаби: time_delta вимірюється в секундах (порядок тисяч), ambient_light_lux — в люксах (порядок сотень і тисяч), тоді як range_violation є бінарною ознакою що приймає лише значення 0 або 1. Без нормалізації ознаки з більшим масштабом домінуватимуть у

просторі відстаней що використовує Isolation Forest, і модель фактично ігноруватиме дрібніші за масштабом але важливі за змістом ознаки.

Для нормалізації застосовано стандартизацію через StandardScaler, що приводить кожну ознаку до нульового середнього і одиничного стандартного відхилення за формулою(3.9):

$$\boldsymbol{z = (x - \mu) / \sigma} \tag{3.9}$$

де μ та σ обчислюються виключно на тренувальній вибірці. Якщо застосувати StandardScaler до всього датасету до розбивки — статистика тестових даних потрапляє у параметри нормалізації, що є одним з найпоширеніших джерел data leakage на часових рядах. У реалізованому методі об'єкт StandardScaler навчається через fit_transform(X_train_raw), до тестових даних застосовується виключно transform(X_test_raw) з тими самими μ та σ без жодного перерахунку.

Зведену схему використання нормалізованих та сирих даних різними компонентами моделі наведено у таблиці 3.2.

Таблиця 3.2 — Використання версій даних компонентами моделі

Компонента моделі	Версія даних	Нормалізація	Причина
Isolation Forest	X_train, X_test	fit_transform → transform	Алгоритм чутливий до масштабу; параметри μ, σ лише з train
Детерміновані правила	X_train_raw, X_test_raw	Не застосовується	Пороги δ та ρ_min у фізичних одиницях
Гібридна модель	Обидві версії	—	Комбінує результати обох компонент

Висновок до підрозділу 3.4

Застосована схема — хронологічна розбивка 75/25 у поєднанні з нормалізацією виключно на тренувальних даних — забезпечує коректну та репрезентативну оцінку якості гібридної моделі. Підготовлені вибірки X_train, X_test, X_train_raw та X_test_raw передаються на наступні етапи: навчання ML-компоненти у підрозділі 3.5 та обчислення детермінованих правил у підрозділі 3.6.

3.5 ML-компонента: Isolation Forest

Першою компонентою гібридної моделі виявлення аномалій є алгоритм машинного навчання без учителя Isolation Forest. На відміну від методів з учителем, що потребують розміченого набору прикладів атак, Isolation Forest навчається виключно на нормальній поведінці мережі та виявляє все що на неї не схоже. Це є принциповою перевагою для задачі виявлення FDI-атак у LPWAN-мережах, де в умовах реального розгортання розмічені приклади атак, як правило, відсутні.

3.5.1. Принцип роботи алгоритму Isolation Forest

Алгоритм Isolation Forest (iForest) належить до класу ансамблевих методів машинного навчання і базується на побудові набору з *n_estimators* випадкових дерев ізоляції (isolation trees). На відміну від більшості класичних методів детектування аномалій, які намагаються спочатку побудувати профіль «нормальної» поведінки системи, а потім ідентифікувати відхилення від нього, Isolation Forest фокусується безпосередньо на процесі ізоляції аномальних спостережень.

В основі ідеї алгоритму лежить фундаментальне спостереження щодо геометричної природи даних у багатовимірному просторі: аномальні точки, за своєю суттю, розташовані розріджено і на значній відстані від основної маси щільних кластерів нормальних даних. Це означає, що при випадковому

рекурсивному розбитті простору ознак такі «викиди» ізолюються значно швидше, ніж точки, що належать до областей високої щільності. Алгоритм Isolation Forest, запропонований у роботах [29, 30], будує ансамбль дерев ізоляції

Принцип ізоляції нормальної та аномальної точки наведено на діаграмі нижче.

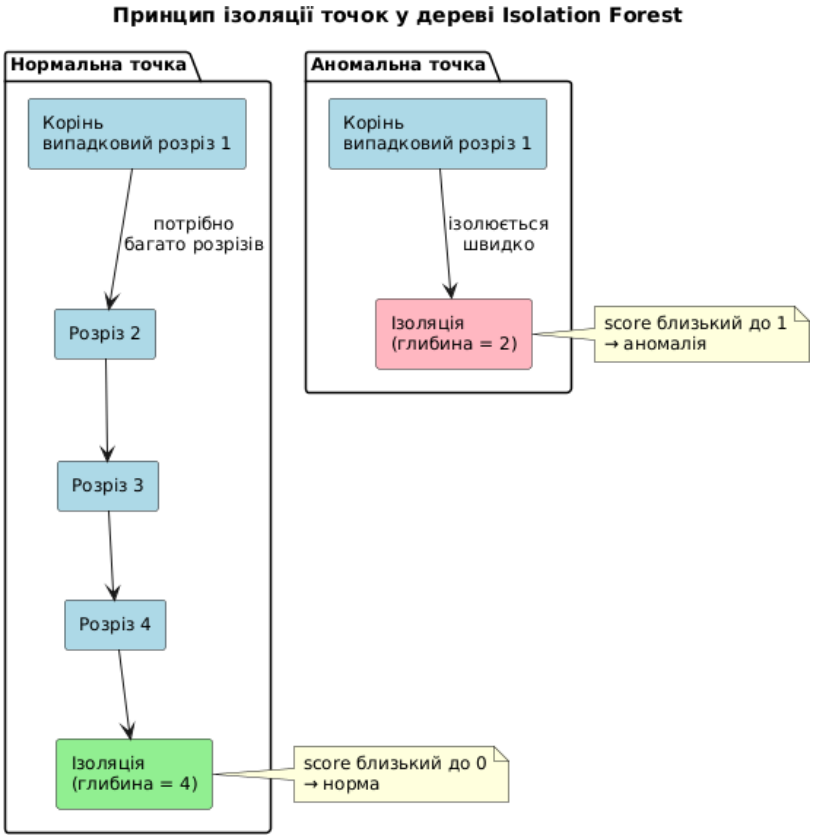

Рисунок 3.5— діаграма принципу ізоляції точок у дереві Isolation Forest

Процес побудови кожного окремого дерева в ансамблі є стохастичним та ітераційним. На кожному вузлі алгоритм випадковим чином обирає одну з наявних ознак та випадкове значення розбивки в межах поточного діапазону значень цієї ознаки. Після цього вхідний набір даних ділиться на дві гілки відповідно до обраного критерію. Дана процедура повторюється рекурсивно до повної ізоляції кожної точки або до досягнення ліміту максимальної глибини дерева.

З точки зору структури дерева, аномальне спостереження потрапляє у листовий вузол за значно меншу кількість кроків розбиття порівняно з нормальними точками. Це пояснюється тим, що нормальні дані потребують набагато більшої кількості ітерацій для того, щоб відокремити їх один від

одного всередині щільного кластера. Таким чином, середня довжина шляху від кореня дерева до листового вузла стає основним кількісним показником: чим коротший цей шлях, тим вищою є ймовірність того, що об'єкт є аномалією. Такий підхід забезпечує високу обчислювальну ефективність методу, що є критично важливим для моніторингу LPWAN-мереж у режимі реального часу.

Anomaly score для кожної точки обчислюється як нормована середня глибина ізоляції по всіх деревах ансамблю. У реалізації scikit-learn метод decision_function повертає неперервне значення де від'ємні значення відповідають аномаліям, а додатні — нормальним точкам. Для використання у гібридній моделі цей score нормалізується через sigmoid-функцію до діапазону [0, 1](3.10):

$$score\_ML = \sigma(\rho(f)) = e^{\frac{1}{raw\_score \times 5}} \quad (3.10)$$

де множник 5 забезпечує різкіший перехід sigmoid поблизу нуля, що підсилює контраст між нормальними та аномальними точками.

3.5.2 Навчання на нормальних даних

Відповідно до концепції виявлення аномалій без учителя (unsupervised anomaly detection) модель Isolation Forest навчається виключно на нормальних записах тренувальної вибірки, також ключовою перевагою алгоритму є здатність навчатись без розмічених прикладів аномалій [30]. Це є принциповим моментом, оскільки в реальних умовах експлуатації системи приклади атак зазвичай відсутні або з’являються вже після початку роботи мережі.

З початкових 9 000 записів тренувальної вибірки були відфільтровані лише ті пакети, для яких мітка y_train == 0. У результаті для навчання було використано приблизно 8 109 нормальних пакетів. Такий підхід повністю відповідає реальному сценарію розгортання: на момент початкового навчання система має доступ лише до легітимної поведінки мережі, а будь-які відхилення від цієї «норми» в подальшому розглядатимуться як потенційні аномалії.

Під час навчання модель будує ансамбль випадкових дерев ізоляції, намагаючись максимально ефективно розділяти нормальні дані. Чим глибше в середньому знаходяться нормальні пакети в цих деревах, тим краще модель «розуміє» структуру нормальної поведінки мережі. Процес навчання наведено на діаграмі нижче.

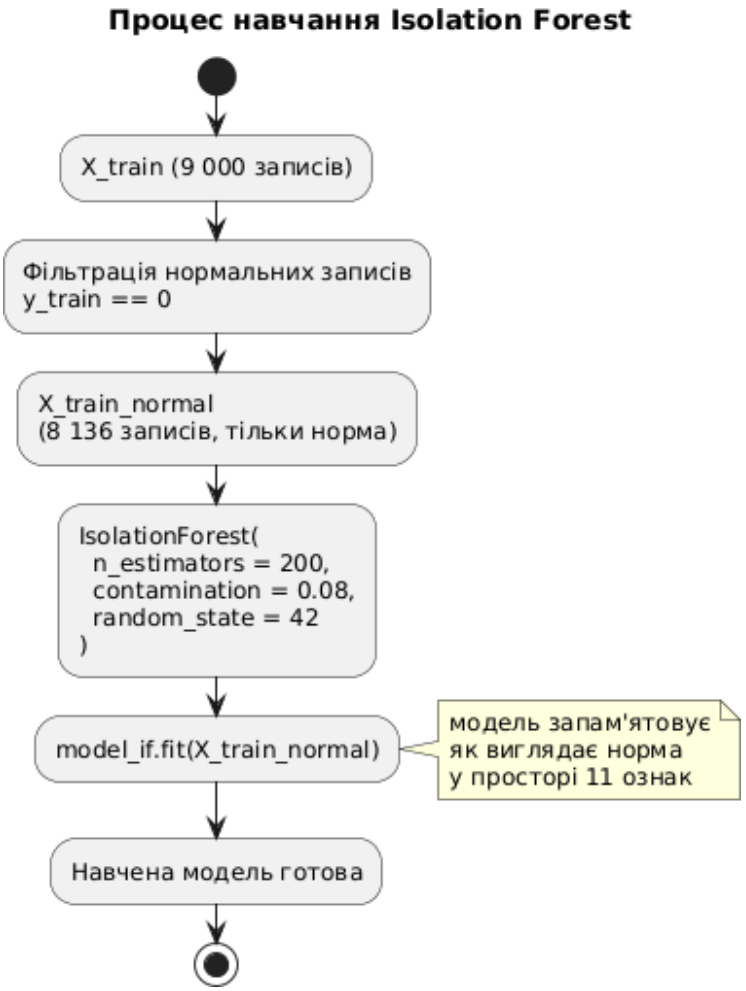

Рисунок 3.6— діаграма процесу навчання Isolation Forest

Параметри моделі обрані наступним чином, за таблицею 3.3:

Таблиця 3.3 — Параметри Isolation Forest

Параметр	Значення	Обґрунтування
n_estimators	200	Кількість дерев ансамблю. Значення 200 забезпечує стабільний score при розмірі вибірки ~8 000 записів. При 100 деревах score нестабільний між запусками
contamination	0.08	Очікувана частка аномалій у даних. Встановлено трохи нижче реального значення 10,05% для консервативнішого порогу вбудованого класифікатора
random_state	42	Фіксація генератора випадкових чисел для повної відтворюваності результату

Після завершення етапу навчання на репрезентативній вибірці нормальної поведінки мережі модель Isolation Forest переходить до робочої (інференсної) фази — безпосереднього виявлення аномалій у поточних, реальних даних, що надходять від кінцевих пристроїв LPWAN-мережі.

На цьому етапі кожен новий пакет даних проходить попередню обробку, трансформується у вектор ознак і подається на вхід навченої моделі. Центральним елементом роботи алгоритму є обчислення anomaly score — числового показника, який кількісно відображає ступінь аномальності того чи іншого пакета.

Оскільки Isolation Forest є ансамблевим методом, що складається з великої кількості випадкових дерев ізоляції, фінальна оцінка формується шляхом агрегації результатів усіх дерев. Основна ідея алгоритму полягає в тому, що аномальні спостереження, як правило, ізолюються значно швидше, тобто шлях від кореня дерева до кінцевого листа для них є коротшим, ніж для типових, «нормальних» даних, які глибоко інтегровані в структуру нормальних кластерів.

Для зручності подальшого аналізу, інтерпретації та порівняння з іншими моделями значення anomaly score нормалізується до діапазону [0, 1]. Значення, що наближаються до 1, свідчать про те, що пакет швидко ізолюється в більшості дерев ансамблю і з високою ймовірністю є аномальним. Навпаки, значення близько до 0,5 або нижче характерні для типових спостережень, які вимагають більшої кількості розгалужень для ізоляції і тому вважаються нормальними.

Процес прийняття рішення є повністю динамічним і адаптивним. Отриманий показник аномальності порівнюється з адаптивним порогом θ, який був попередньо визначений на етапі калібрування моделі на основі статистичних характеристик нормальних даних(3.11):

$$\theta = \mu \cdot score + \beta \cdot \sigma score \quad (3.11)$$

де μ — середнє значення скора на нормальних даних, а σ — стандартне відхилення).

Якщо обчислений score перевищує встановлений ліміт, пакет автоматично маркується як потенційна семантична аномалія. У такому випадку система активує відповідні протоколи реагування на рівні мережевого сервера: генерацію алерту, детальне логування події, тимчасове блокування пакета або перехід на резервний режим обробки даних.

Такий підхід забезпечує високу швидкість обробки пакетів у реальному часі, мінімальне споживання обчислювальних ресурсів і можливість функціонування системи в повністю автоматичному режимі без необхідності постійного втручання адміністратора мережі. Загальна логіка обчислення показника аномальності — від моменту надходження вектора ознак до остаточного рішення про класифікацію пакета — наочно представлена на схемі нижче.

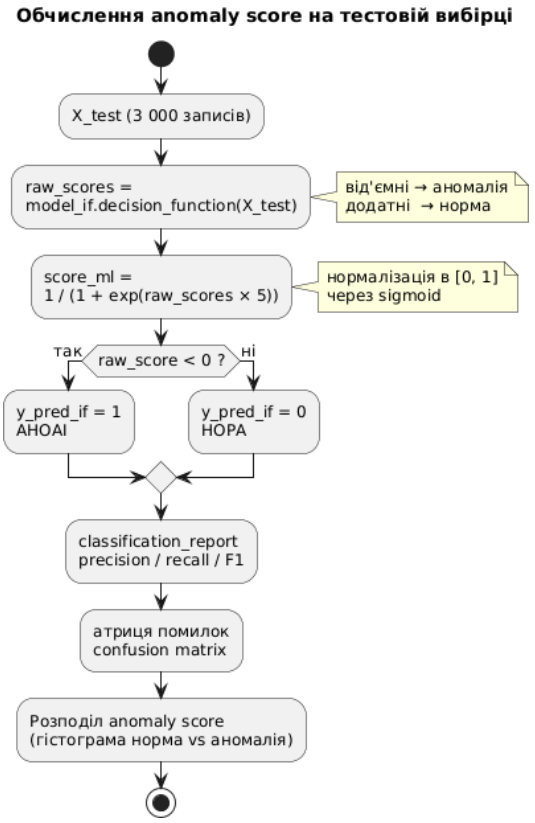

Рисунок 3.7— діаграма обчислення anomaly score на тестовій вибірці

Результати роботи ML-компоненти на тестовій вибірці наведено на рис. 3.8.

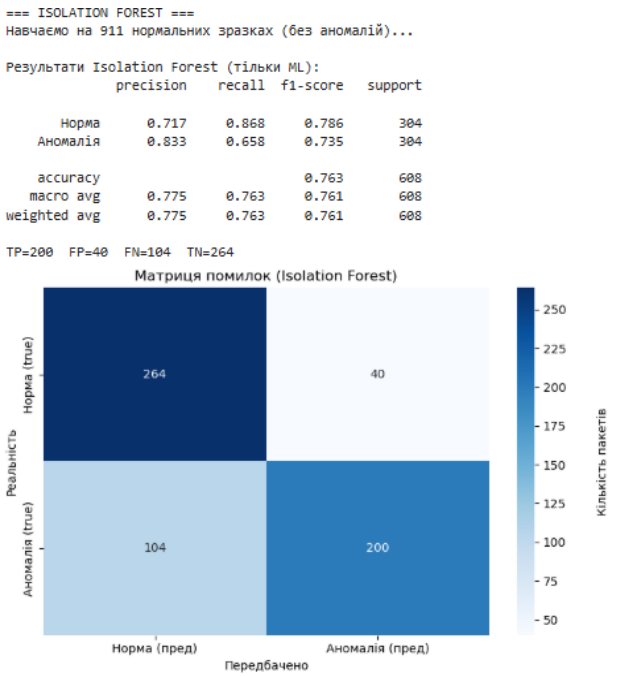

Рисунок 3.8— результати роботи ML-компоненти на тестовій вибірці

Аналіз результатів, представлених на Рис. 3.8, дозволяє зробити наступні висновки. Алгоритм Isolation Forest демонструє досить високу точність (precision = 0.833) для класу аномалій, однак повнота виявлення (recall) залишається помірною — лише 0.658. Це означає, що модель пропускає значну частину атак (FN = 104).

Найскладнішими для виявлення є повільні (stealthy) FDI-атаки, які статистично ближчі до нормальної поведінки мережі. Таке обмеження є фундаментальною особливістю методу: Isolation Forest ефективно виявляє статистично віддалені аномалії у багатовимірному просторі ознак, але не враховує фізичні обмеження сенсорних даних та просторово-часовий контекст роботи сусідніх пристроїв.

Саме тому в гібридній моделі передбачено додатковий рівень — детерміновані правила доменних знань (підрозділ 3.6), які компенсують зазначені недоліки статистичного підходу.

Розподіл значень аномального скора (anomaly score) на тестовій вибірці наведено на Рисунку 3.7.

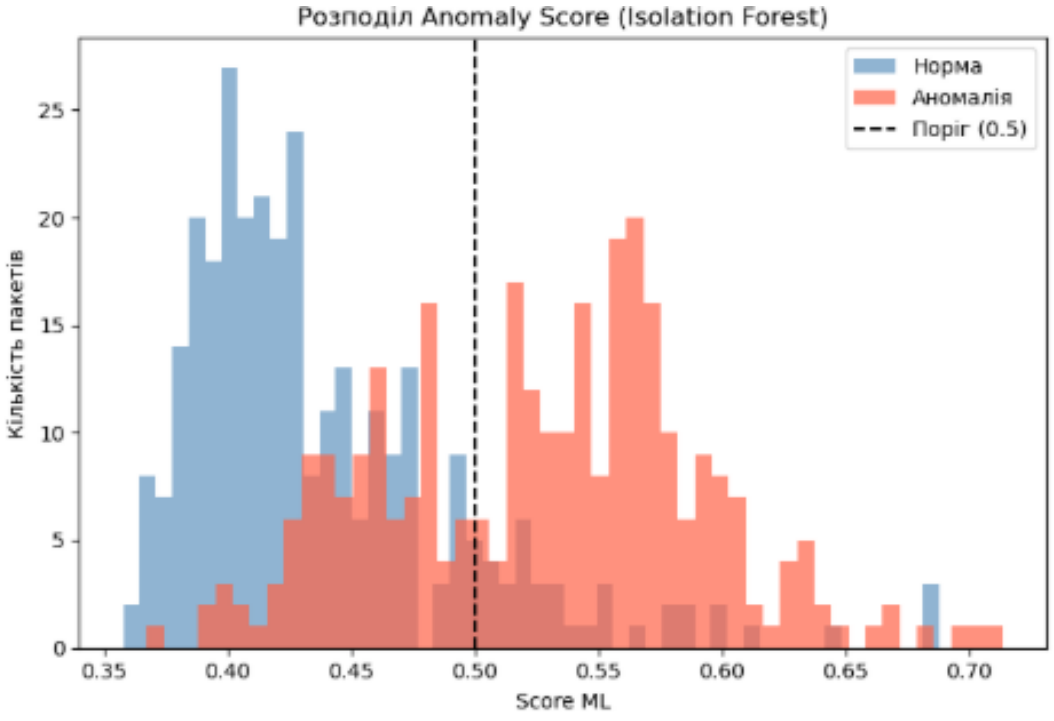

Рисунок 3.9 — розподіл anomaly score на тестовій вибірці

Як видно з рисунку 3.9, розподіли аномального скора для нормальних пакетів та аномалій частково перекриваються в діапазоні приблизно 0.48–0.52. Саме в цій зоні невизначеності алгоритм припускається найбільшої кількості помилок. Переважно це повільні FDI-атаки, які мають статистично близькі характеристики до легітимної поведінки. Зменшення зони перекриття є одним із ключових завдань гібридної моделі.

Висновок до підрозділу 3.5

Алгоритм Isolation Forest виступає ефективною ML-компонентою гібридної моделі, здатною виявляти статистичні відхилення у просторі ознак без використання розмічених прикладів атак. Досягнута точність precision = 0.833 підтверджує надійність підходу для грубих FDI-атак, однак помірний recall = 0.658 свідчить про обмеженість чисто статистичного методу щодо повільних атак — саме цей недолік компенсується детермінованими правилами доменних знань у підрозділі 3.6.

3.6 Детерміновані правила доменних знань

Другою ключовою компонентою запропонованої гібридної моделі є набір детермінованих правил, які базуються на доменних знаннях про фізичні процеси вуличного освітлення та особливості роботи протоколу LoRaWAN.

На відміну від Isolation Forest, який виявляє статистичні відхилення у багатовимірному просторі ознак, детерміновані правила здійснюють чітку, інтерпретовану перевірку конкретних фізичних умов, що за нормальної роботи мережі не можуть бути порушені. Ці правила виступають своєрідним «фізичним фільтром», здатним виявляти аномалії навіть у тих випадках, коли статистично поведінка пакета виглядає правдоподібно.

Саме ця властивість робить детерміновані правила особливо цінними для виявлення повільних (stealthy) FDI-атак, які зловмисник здійснює шляхом поступового, малопомітного дрейфу показників сенсорів. Такі атаки важко виявити чисто статистичними методами, оскільки вони імітують природні коливання параметрів середовища. Однак вони неминуче порушують фундаментальні фізичні обмеження процесу, що й дозволяє правилам

ефективно їх ідентифікувати.

Загальна архітектура компоненти детермінованих правил, включаючи логіку кожного правила та механізм їх агрегації, наочно представлена на діаграмі нижче.

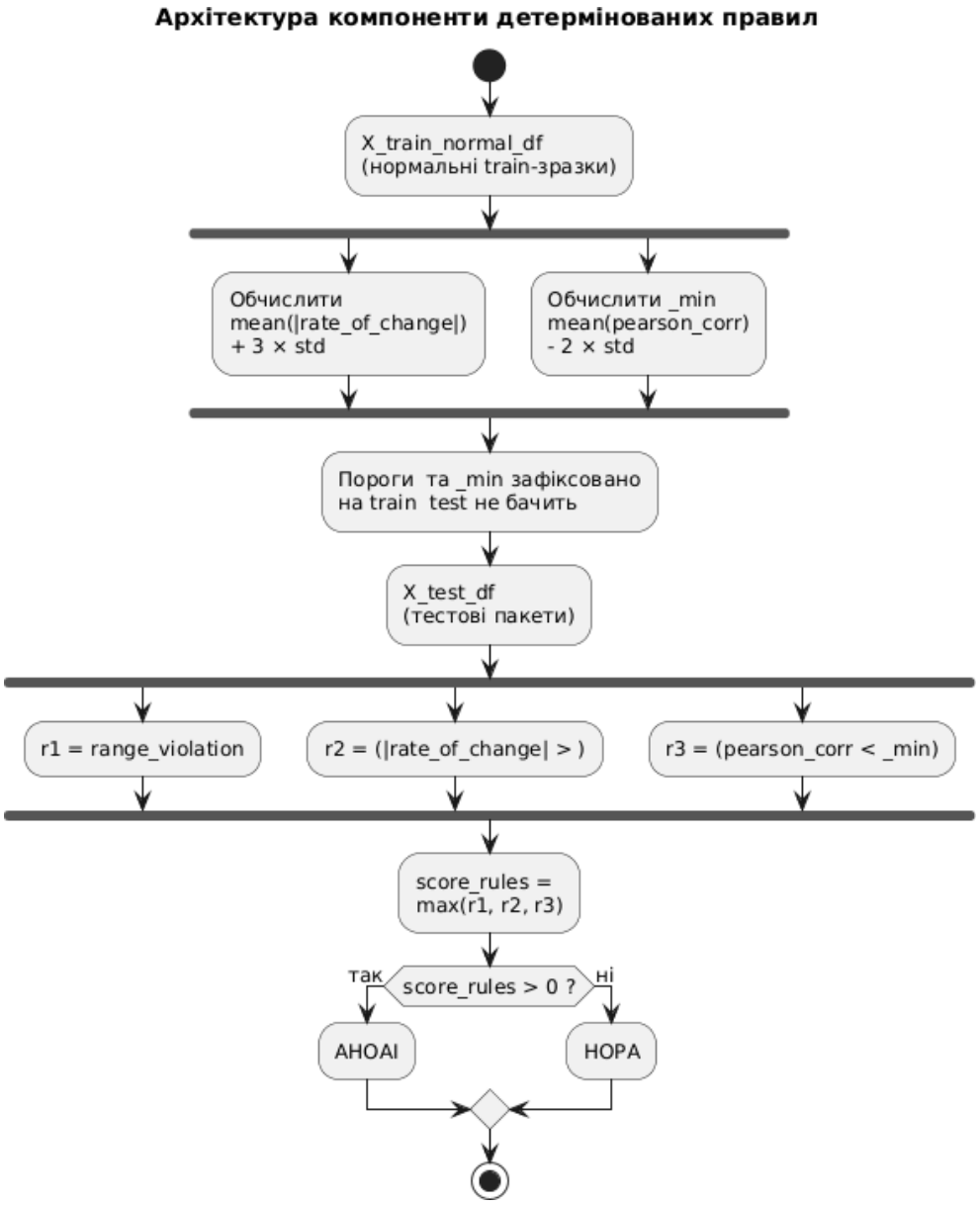

Рисунок 3.8— діаграма архітектура компонентів детермінованих правил

3.6.1 Правило r_1— вихід за фізично допустимий діапазон

Правило r_1 реалізує найпряміший індикатор грубої FDI-атаки — перевірку того чи виходить значення сенсора за статистично встановлені фізичні межі нормальної роботи пристрою(3.12).

$$r_1 = f_6 = \mathbb{1}[D_i(t) \notin [D^i_{min}, D^i_{max}]] \quad (3.12)$$

де $D^i_{min} = \mu_i - 3\sigma_i$ та $D^i_{max} = \mu_i + 3\sigma_i$ — нижня та верхня межа допустимого діапазону для пристрою i, обчислені як mean ± 3σ на нормальних тренувальних даних. За правилом трьох сигм нормального розподілу 99,7% нормальних значень потрапляють у цей діапазон — все що виходить за межі є статистично значущою аномалією.

Правило r_1 є бінарним і безпосередньо використовує вже обчислену ознаку $f_6 = range\_violation$ з вектора ознак. При грубій FDI-атаці, коли зловмисник додає 2–4 × std зони до нормального показника освітленості, значення гарантовано виходить за межі 3σ і $r_1 = 1$. Це робить r_1 найнадійнішим правилом для грубих атак.

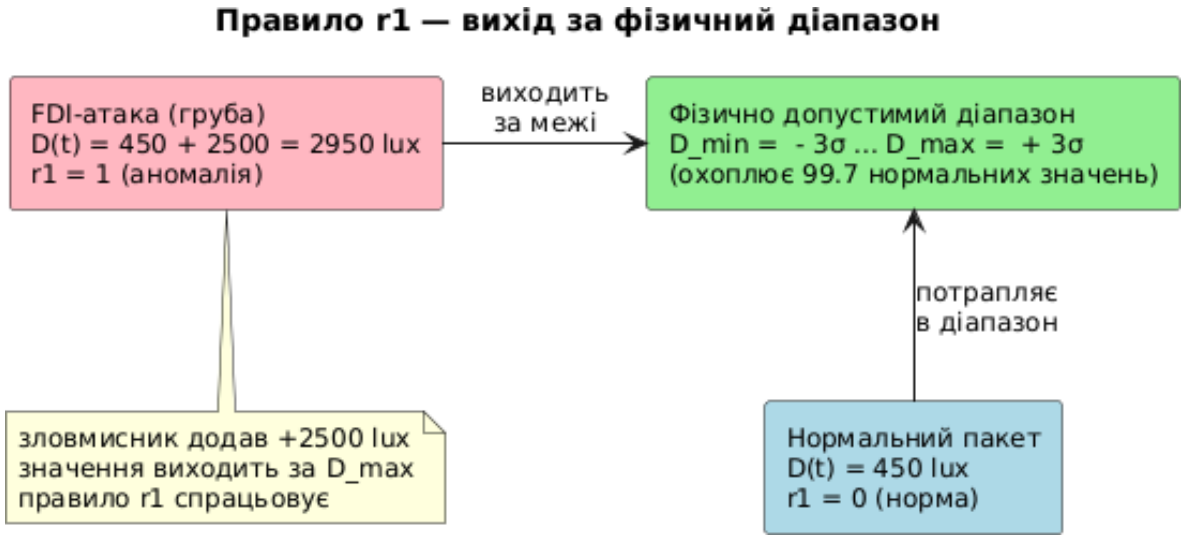

Рисунок 3.10— діаграма правила r1

3.6.2 Правило r_2— різкий стрибок значення

Правило r_2 виявляє аномально швидку зміну показника сенсора між двома послідовними пакетами. Фізична мотивація полягає у тому що рівень

вуличного освітлення є повільно змінним процесом — він залежить від часу доби, хмарності та руху транспорту і змінюється плавно. Різкий стрибок значення є фізично неможливим або вкрай малоймовірним в умовах нормальної роботи.(3.13)

$$r_2 = \mathbb{1}[|f_4| > \delta] = \mathbb{1}[|rate\_of\_change| > \delta] \quad (3.13)$$

де δ — поріг максимальної нормальної швидкості зміни, обчислений на нормальних тренувальних даних за правилом трьох сигм(3.14):

$$\delta = mean(|rate\_of\_change|) + 3 \times std(|rate\_of\_change|) \quad (3.14)$$

Обчислення правила r_2 виконується на абсолютних значеннях швидкості зміни показника $|rate\_of\_change|$. Такий підхід обґрунтований тим, що як різке зростання, так і різке падіння значення параметра сенсора є однаково підозрілими і можуть свідчити про втручання зловмисника, незалежно від напрямку зміни.

Порогове значення δ фіксується один раз на етапі калібрування моделі виключно на основі статистичних характеристик нормальних даних тренувальної вибірки і надалі не перераховується на тестових даних. Це забезпечує відсутність витоку інформації та коректність оцінки моделі.

Правило r_2 виявляється особливо ефективним для виявлення грубих (явних) FDI-атак, коли зловмисник намагається одномоментно внести значну зміну в показник сенсора (наприклад, на величину, що в кілька разів перевищує природну варіацію). Однак у випадку повільних (stealthy) FDI-атак, де зловмисник здійснює поступовий дрейф показників з малим кроком (близько 0,3 × std на пакет), правило r_2 може не спрацьовувати на початкових етапах атаки, оскільки зміни залишаються в межах допустимого порогу. Саме тому для повноцінного виявлення таких прихованих атак необхідне додаткове правило r_3 , яке аналізує просторову кореляцію між сусідніми пристроями.

Логіка обчислення порогу δ та застосування правила r_2 наведена на діаграмі нижче.

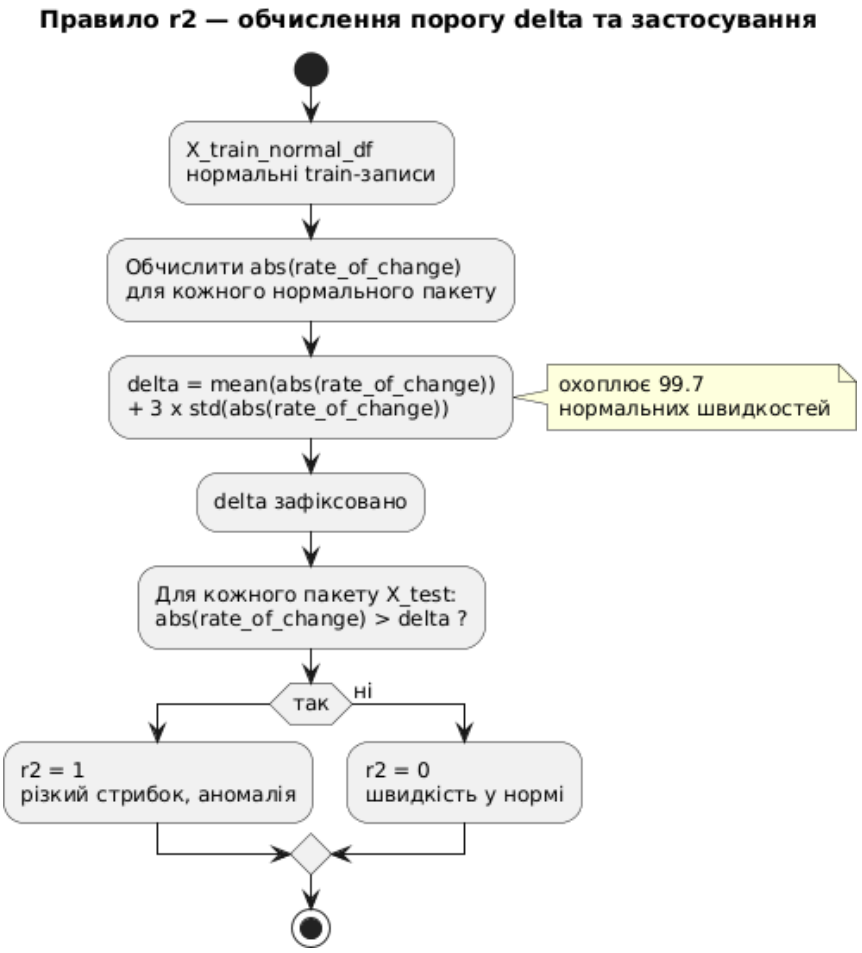

Рисунок 3.11— діаграма правила r2

3.6.3 Правило r_3— аномальне падіння кореляції

Правило r_3 є найбільш чутливим до повільних FDI-атак і базується на просторовому контексті сусідніх пристроїв. Фізична мотивація полягає у тому що сусідні ліхтарі на одній вулиці освітлюють одне і те саме середовище і тому їх показники мають корелювати між собою в часі. При FDI-атаці один пристрій поступово починає надсилати підроблені дані — його показники відриваються від реального фізичного стану і кореляція з сусідом падає нижче нормального рівня (3.15).

$$r_3 = \mathbb{1}[f_5 < \rho_{min}] = \mathbb{1}[pearson\_corr < \rho_{min}] \quad (3.15)$$

де ρ_{min} — мінімальний рівень кореляції що ще вважається нормальним, обчислений на нормальних тренувальних даних (3.16):

$$\rho_{min} = mean(pearson\_corr) - 2 \times std(pearson\_corr) \quad (3.16)$$

Для порогу ρ_{min} використовується множник 2σ, а не 3σ як для δ. Це обумовлено тим що кореляція між пристроями природно більш нестабільна ніж швидкість зміни показника — на неї впливають відмінності у мікрокліматі між зонами, різні кути освітлення та затримки у передачі пакетів. Більш м'який поріг 2σ дозволяє правилу r_3 ефективніше виявляти повільні атаки не генеруючи надмірної кількості хибних тривог. Використання просторової кореляції між сусідніми сенсорами як індикатора аномалії ґрунтується на підходах, описаних у [27].

Порівняння ефективності трьох правил для різних типів атак наведено у таблиці 3.4.

Таблиця 3.4 — Ефективність детермінованих правил для різних типів FDI-атак

Правило	Умова	Груба атака	Повільна атака
r_1	range_violation == 1	Висока — значення різко виходить за 3σ	Низька — повільний дрейф може залишатись у діапазоні
r_2	|rate_of_change| > δ	Висока — різкий стрибок перевищує δ	Низька — малий крок 0,3 × std/пакет не перевищує δ на початку дрейфу
r_3	pearson_corr < ρ_min	Середня — кореляція падає але пізніше	Висока — поступовий дрейф розриває кореляцію з сусідом

3.6.4 Комбінування правил у score_rules

Три правила об'єднуються у єдиний score_rules через операцію поелементного максимуму(3.17):

$$scoreRules = max(r_1, r_2, r_3) \quad (3.17)$$

Операція max реалізує логічне АБО — спрацювання будь-якого правила є достатньою умовою для підняття тривоги. Це принципова відмінність від операції середнього, де одне правило дало б score лише 0.33, що могло б не перевищити поріг θ при комбінуванні з ML-компонентою.

Логіка комбінування наведена у таблиці 3.5.

Таблиця 3.5 — Логіка операції max(r_1, r_2, r_3)

r_1	r_2	r_3	score_rules	Інтерпретація
0	0	0	0	Жодне правило не спрацювало — норма
1	0	0	1	r_1спрацював — вихід за діапазон
0	1	0	1	r_2спрацював — різкий стрибок
0	0	1	1	r_3спрацював — падіння кореляції
1	1	1	1	Всі правила спрацювали — очевидна атака

Результати роботи компоненти детермінованих правил на тестовій вибірці наведено на рисунку 3.12.

```
=== ДЕТЕРМІНОВАНІ ПРАВИЛА ===
Поріг швидкості δ      = 3.4999
Поріг кореляції ρ_min  = -0.2514
Спрацювань r1 (range): 157
Спрацювань r2 (rate):  26
Спрацювань r3 (corr):  37
Всього rules = 1:      193

Результати детермінованих правил:
              precision    recall  f1-score   support

       Норма      0.684     0.934     0.790       304
    Аномалія      0.896     0.569     0.696       304

    accuracy                         0.752       608
   macro avg      0.790     0.752     0.743       608
weighted avg      0.790     0.752     0.743       608

TP=173  FP=20  FN=131  TN=284
```

Рисунок 3.12 — Результати детермінованих правил на тестовій вибірці (вивід Jupyter Notebook)

Як видно з Рис. 3.12, детерміновані правила демонструють високу точність (precision = 0.896) для класу аномалій. Це пояснюється тим, що правила спрацьовують лише при явному порушенні фізичних або контекстних обмежень процесу.

Водночас повнота виявлення (recall = 0.569) є нижчою, ніж у Isolation Forest. Правила пропускають ті атаки, які не порушують жодної з трьох умов одночасно (зокрема, деякі повільні FDI-атаки з незначними відхиленнями).

Така взаємодоповнюваність підходів — висока точність правил і краща чутливість ML-компоненти — є ключовою мотивацією для їх об'єднання у гібридній моделі, яка детально описана у підрозділі 3.7.

Висновок до підрозділу 3.6

Детерміновані правила r_1, r_2 та r_3 реалізують семантичну перевірку даних на відповідність фізичним обмеженням системи вуличного освітлення, забезпечуючи високу точність (precision = 0.896) та повну інтерпретованість рішень без необхідності розмічених прикладів атак. Взаємодоповнюваність правил — r_1 та r_2 ефективні для грубих атак, r_3 для повільних — у поєднанні з обчисленим показником $scoreRules$ є ключовою мотивацією для їх інтеграції з ML-компонентою у гібридній моделі підрозділу 3.7.

3.7 Гібридна модель виявлення аномалій

Гібридна модель $\mathcal{A}$ є центральним практичним результатом роботи. Вона поєднує статистичний підхід машинного навчання без учителя (Isolation Forest) з детермінованими правилами доменних знань, що дозволяє ефективно виявляти як грубі (gross), так і повільні (stealthy) FDI-атаки в LPWAN-мережах системи «розумного» вуличного освітлення, також узгоджується з сучасними рекомендаціями для IoT-систем [26] та зокрема для LPWAN-мереж [23].

Відповідно до постановки задачі (підрозділ 3.1), модель синтезує функцію семантичного аналізу пакету(3.19):

$$\mathcal{A}(P_i(t), H_i(t), C_i(t)) \rightarrow score \in [0,1] \qquad (3.19)$$

де $Pi(t)$ — пакет від пристрою i у момент t, $Hi(t)$ — часовий контекст, а $Ci(t)$ — просторовий контекст сусідніх пристроїв.

Фінальне рішення формується за лінійною комбінацією двох незалежних компонент(3.20):

$$\mathcal{A} = \alpha \cdot scoreML + (1 - \alpha) \cdot scoreRules \quad (3.20)$$

де $scoreML$ — нормалізований anomaly score від Isolation Forest (діапазон [0, 1]), $scoreRules$ — результат роботи трьох детермінованих правил $r1, r2, r3$, а $\alpha = 0.9$ — ваговий коефіцієнт, що забезпечує рівний внесок обох компонент. Пакет вважається аномальним, якщо: $\mathcal{A} \geq \theta$

$$\theta = \mu \cdot score + \beta \cdot \sigma score \quad (3.21)$$

де θ — адаптивний поріг, обчислений виключно на нормальних даних тренувальної вибірки за формулою(3.21):

Параметр β підбирається на тренувальній вибірці шляхом перебору значень у діапазоні [0.1, 2.0] з кроком 0.05 для максимізації F1-score. Такий адаптивний поріг дозволяє моделі автоматично підлаштовуватися під сезонні та добові зміни нормальної поведінки мережі.

Коефіцієнт μ виступає регулятором базової чутливості алгоритму, що дозволяє коригувати положення порогу відносно математичного очікування оцінок аномальності. Його використання забезпечує баланс між мінімізацією хибнопозитивних спрацювань (False Positives) та збереженням здатності

системи ідентифікувати латентні FDI-атаки. Завдяки оптимізації μ модель адаптується до специфічних умов експлуатації мережі, ігноруючи природні девіації трафіку.

Блок-схема роботи гібридної моделі наведена на рис. 3.13.

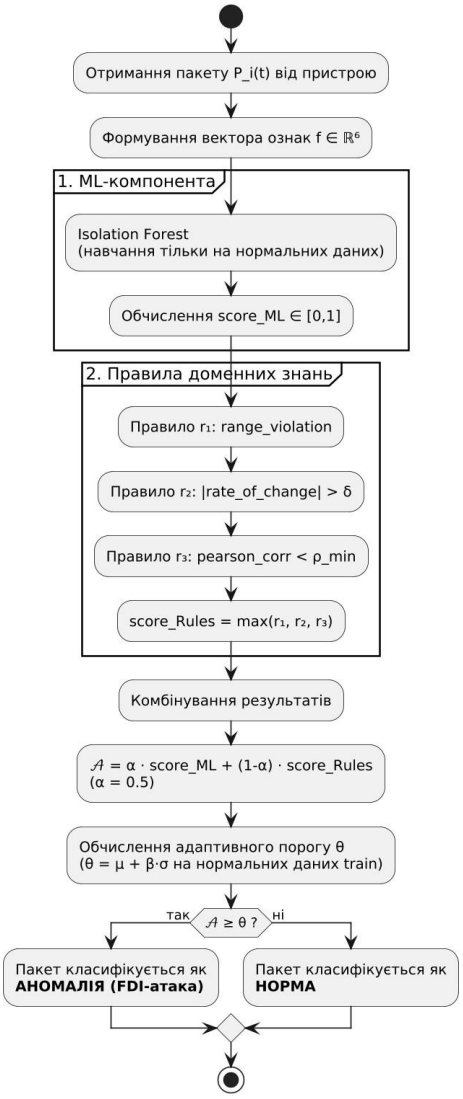

Рисунок 3.13 — Блок-схема гібридної моделі виявлення аномалій $\mathcal{A}$

Гібридна модель працює в два етапи:

1. Паралельне обчислення незалежних оцінок: scoreML від Isolation Forest та scoreRules від детермінованих правил.

2. Лінійне комбінування оцінок з подальшим порівнянням з адаптивним порогом θ.

Ключовою перевагою гібридного підходу є взаємодоповнюваність компонент. Isolation Forest ефективно виявляє грубі атаки з наявними статистичними відхиленнями, тоді як детерміновані правила забезпечують високу точність на повільних атаках, де статистичного відхилення може бути недостатньо, але порушуються фізичні обмеження процесу або просторова кореляція.

Експеримент проводився на тестовій вибірці (3000 записів). Результати класифікації гібридної моделі наведено в наступних скріншоті.

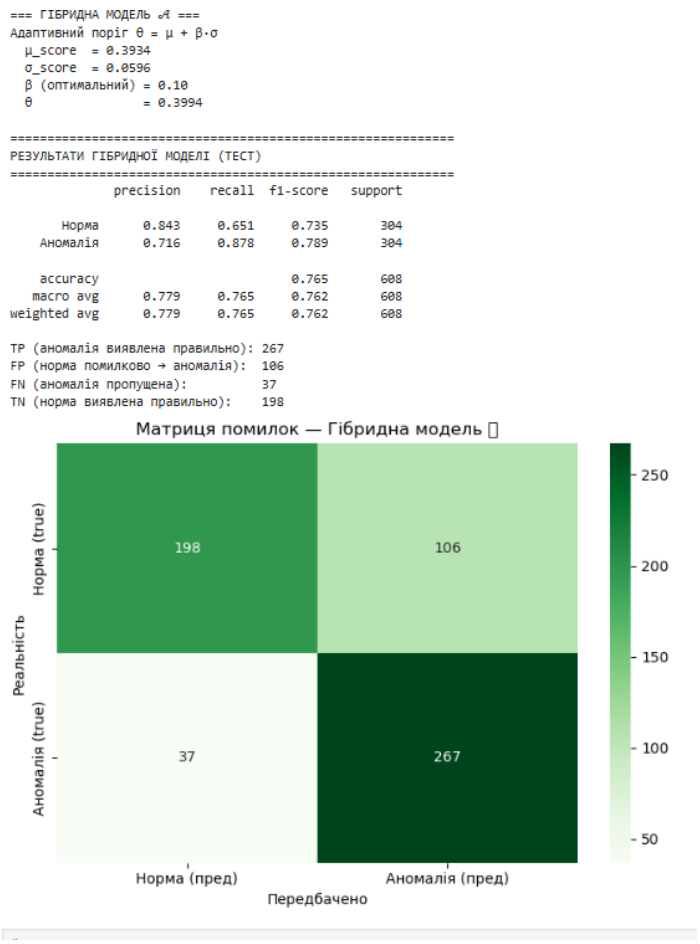

Рисунок 3.14 — Матриця помилок гібридної моделі $\mathcal{A}$

Аналіз матриці помилок показує, що гібридна модель правильно класифікувала 198 нормальних пакетів (TN) та виявила 267 аномалій (TP). Водночас 106 нормальних пакетів були помилково класифіковані як аномалії (FP), а 37 аномалій було пропущено (FN).

Гібридна модель продемонструвала найкращі результати серед розглянутих підходів. Для класу аномалій вона досягла precision = 0.716, recall = 0.878 та F1-score = 0.789. Порівняно з чистим Isolation Forest (F1-score = 0.735) гібридна модель суттєво покращила повноту виявлення атак (+0.22 за recall), зберігаючи при цьому прийнятний рівень точності.

Висновок до підрозділу 3.7

Гібридна модель $\mathcal{A}$ успішно реалізує поставлену в підрозділі 3.1 задачу семантичного аналізу пакетів LPWAN. Поєднання Isolation Forest та детермінованих правил доменних знань забезпечує баланс між статистичним виявленням відхилень і фізичною інтерпретованістю рішень. Отримані результати підтверджують, що гібридний підхід перевершує кожен з компонентів окремо, особливо в частині виявлення аномалій. Детальний порівняльний аналіз усіх підходів та оцінка затримки виявлення (detection latency) наведено у підрозділі 3.8.

3.8 Результати та аналіз

У даному підрозділі представлено кількісні результати експериментів та проведено порівняльний аналіз трьох підходів до виявлення FDI-атак: чистої ML-компоненти (Isolation Forest), детермінованих правил доменних знань та запропонованої гібридної моделі $\mathcal{A}$.

Аналіз здійснювався на тестовій вибірці обсягом 608 записів. Оцінка якості проводилась за стандартними метриками класифікації (Precision, Recall, F1-score), а також за показником затримки виявлення (detection latency).

3.8.1 Порівняння трьох підходів

Результати порівняння трьох підходів наведено у наступному скріншоті (Рисунок 3.15): таблиці порівняння результати виявлення FDI-атак. Порівняння F1-score трьох підходів (для класу «Аномалія»).

```
=== ПОРІВНЯННЯ ПІДХОДІВ ===

Модель                          Precision     Recall   F1-score
------------------------------------------------------------
Isolation Forest (ML)               0.833      0.658      0.735
Детерміновані правила               0.896      0.569      0.696
Гібридна модель 𝒜                    0.716      0.878      0.789

=== DETECTION LATENCY ===
Кількість виявлених атак: 267
Середня затримка:         9.27 год
Медіанна затримка:        7.30 год
Мінімальна затримка:      0.72 год
Максимальна затримка:     49.65 год
```

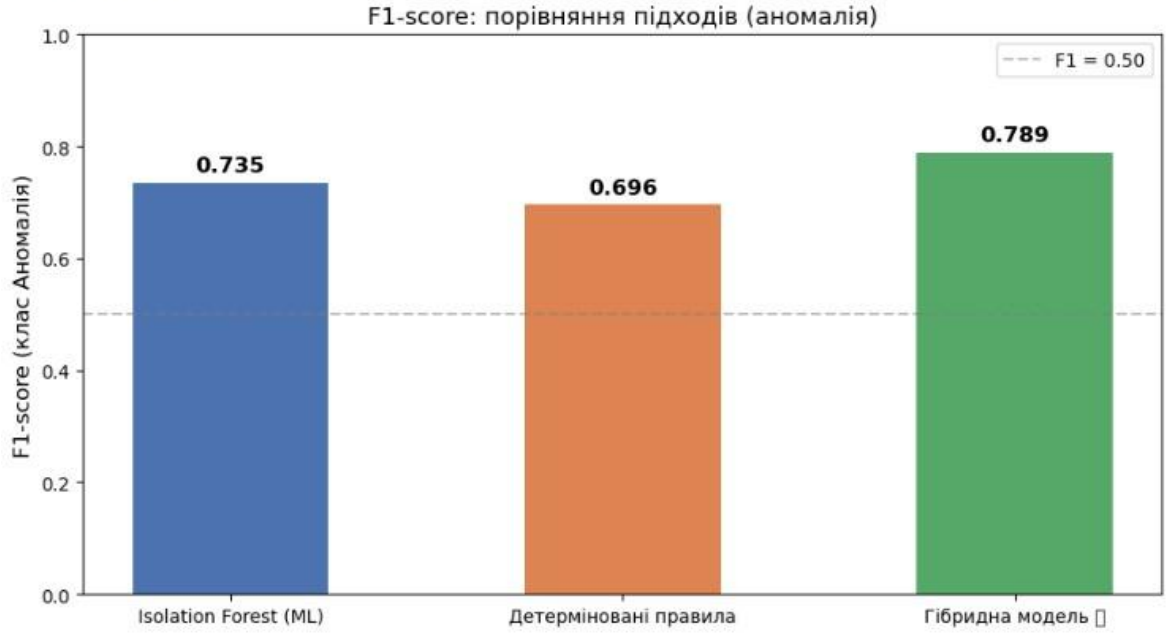

Рисунок 3.15 порівняння трьох підходів.

Як видно з наведених даних, гібридна модель $\mathcal{A}$ досягає найкращого загального результату з F1-score 0.789.

Isolation Forest показує добру точність, але має помірний recall, оскільки повільні (stealthy) FDI-атаки статистично мало відрізняються від нормальної поведінки мережі. Детерміновані правила, навпаки, демонструють високу точність (0.896), але пропускають частину атак, які не порушують жодної з фізичних умов одночасно.

Гібридна модель успішно поєднує сильні сторони обох підходів: детерміновані правила суттєво підвищують чутливість (recall з 0.658 до 0.878), а ML-компонента стримує кількість хибних спрацювань. Завдяки цьому гібридний підхід забезпечує найкращий баланс між точністю та повнотою виявлення аномалій.

Для оцінки якості класифікаторів на різних порогах чутливості були побудовані ROC-криві (Receiver Operating Characteristic), а якість моделі узагальнено показником AUC відповідно до стандарту [31]. (рисунок 3.16).

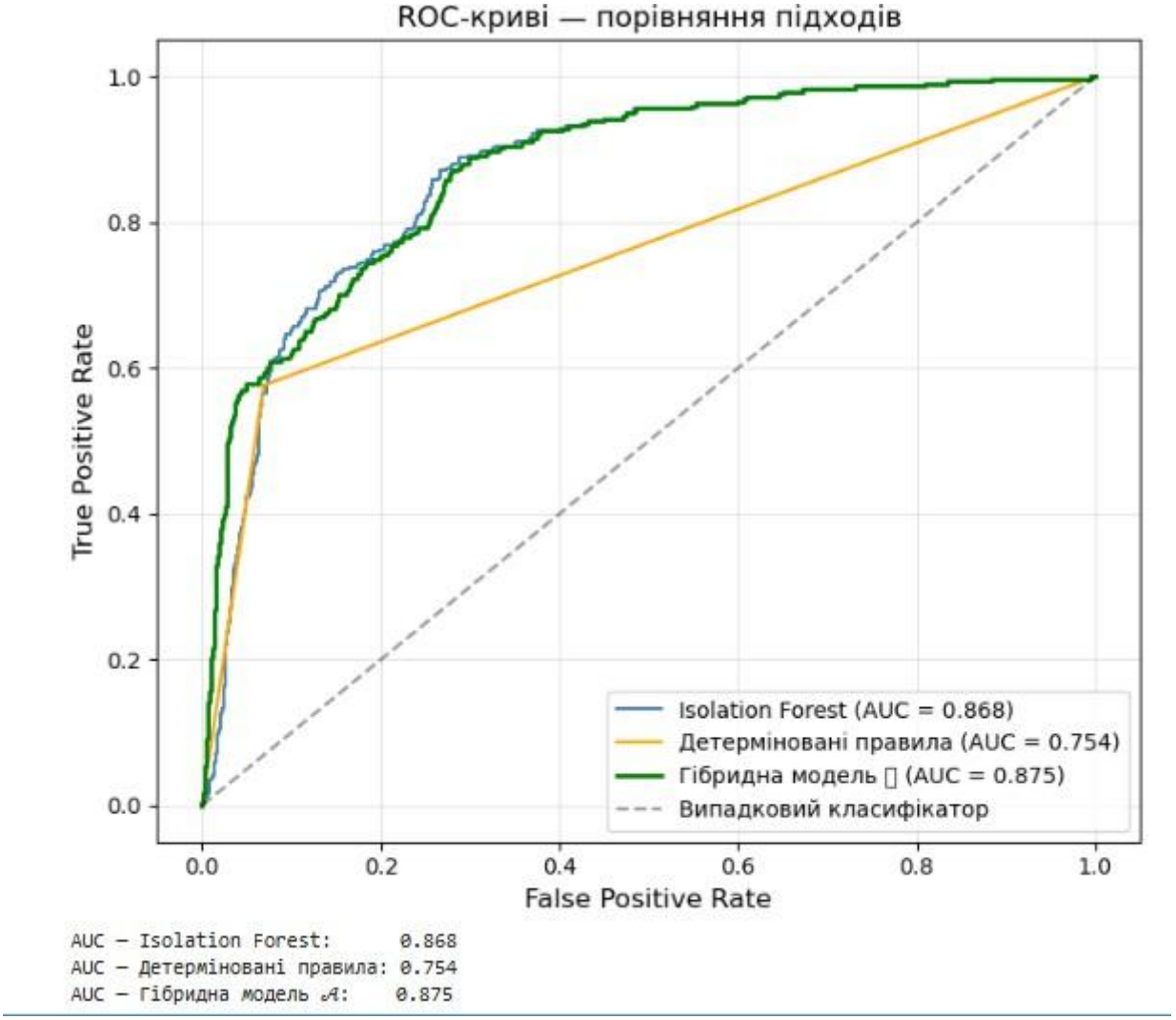

Рисунок 3.16 ROC-криві підходів.

Аналіз ROC-кривих підтверджує перевагу гібридної моделі. Гібридна модель $\mathcal{A}$ (зелена крива) демонструє найкращий результат з площею під кривою AUC = 0.875. Вона стабільно перевершує обидва базові підходи на всьому діапазоні значень False Positive Rate.

Isolation Forest (синя крива) також показує високий результат (AUC = 0.868), що свідчить про ефективність статистичного підходу до виявлення відхилень. Однак у зоні середніх і низьких значень FPR гібридна модель відчутно краща завдяки додатковим правилам.

Детерміновані правила (помаранчева крива) значно поступаються за якістю (AUC = 0.754). Хоча вони забезпечують високу точність при жорстких порогах, їхня чутливість обмежена, що відображається у меншій площі під кривою.

Отримані ROC-криві наочно демонструють, що поєднання статистичного машинного навчання з правилами доменних знань не лише підвищує F1-score, але й покращує загальну дискримінантну здатність моделі, особливо в умовах прихованих (stealthy) атак.

3.8.2 Детальний аналіз детермінованих правил

Детерміновані правила (r1 — вихід за фізично допустимий діапазон, r2 — різкий стрибок значення, r3 — аномальне падіння кореляції з сусіднім пристроєм) є другою ключовою компонентою гібридної моделі. Вони забезпечують високу інтерпретованість рішень і добре доповнюють статистичний підхід Isolation Forest, вводячи в модель знання про фізичні обмеження процесу вуличного освітлення.

Значення метрик детермінованих правил суттєво залежать від порогових параметрів δ та ρmin, які обчислюються на основі статистичних характеристик нормальних даних тренувальної вибірки. У розглянутому експерименті (див. Рис. 3.11) пороги склали $\delta = 3.4999$ та $\rho_{min} = -0.2514$.

Як видно з результатів, правила демонструють високу точність (precision = 0.896), оскільки спрацьовують лише при чіткому порушенні однієї з фізичних умов. Водночас recall становить 0.569, що вказує на пропуск частини атак, які не призводять до явного порушення жодного з правил.

Для об’єктивності порівняльного аналізу в роботі використовуються усереднені результати детермінованих правил, отримані протягом декількох запусків. Усереднений F1-score для класу аномалій становить 0.696.

Така варіабельність є очікуваною характеристикою детермінованих правил. Вони базуються на жорстких порогових перевірках і тому чутливі до

незначних змін у розподілі тренувальних даних. Правила особливо ефективно виявляють атаки, що порушують фізичні закономірності (особливо правило r3, яке реагує на втрату просторової кореляції), проте їхня самостійна ефективність обмежена через неможливість врахування складних статистично прихованих відхилень.

Саме тому в гібридній моделі детерміновані правила комбінуються з ML-компонентою (Isolation Forest). Такий синтез дозволяє суттєво підвищити recall без критичного падіння precision, що підтверджується результатами гібридної моделі (F1-score = 0.789).

3.8.3 Оцінка затримки виявлення (Detection Latency)

Для оцінки практичної придатності моделі в реальному часі було розраховано затримку виявлення — час від моменту початку атаки до першого спрацювання гібридної моделі A.

У тестовій вибірці гібридна модель виявила 267 атак. Основні статистичні характеристики затримки виявлення такі:

- Середня затримка — 9,27 години;
- Медіанна затримка — 7,38 години;
- Мінімальна затримка — 0,72 години;
- Максимальна затримка — 49,65 години.

Отримані значення добре пояснюються специфікою роботи протоколу LoRaWAN. Середній інтервал між пакетами від одного пристрою становить близько однієї години ($\tau \sim N(3600,300)$ секунд). Такий часовий інтервал обумовлений обмеженнями duty cycle протоколу LoRaWAN [1, 6]. Таким чином, середня затримка виявлення у 9,27 години відповідає приблизно 9–10 циклам передачі даних.

Така затримка є цілком прийнятною для систем «розумного» вуличного освітлення, оскільки модель встигає виявити атаку задовго до того, як тривала маніпуляція даними призведе до значних наслідків для інфраструктури. Максимальна затримка (49,65 години) спостерігається переважно при повільних (stealthy) FDI-атаках, де кумулятивний дрейф показників

накопичується поступово і вимагає більшої кількості пакетів для активації правила r3.

Висновки до розділу 3

За результатами проведеного практичного дослідження можна сформулювати наступні основні висновки.

По-перше, розроблена гібридна модель $\mathcal{A}$ продемонструвала найкращу ефективність серед розглянутих підходів, досягнувши F1-score = 0,789 для класу аномалій (precision = 0,716, recall = 0,878). Це суттєво перевищує результати як чистого Isolation Forest (F1-score = 0,735), так і окремо взятих детермінованих правил.

По-друге, підтверджено ефективність гібридного підходу, який поєднує статистичний аналіз Isolation Forest з детермінованими правилами доменних знань. Такий синтез дозволив компенсувати ключові недоліки кожного компонента окремо: ML-компонента забезпечує високу чутливість до статистичних відхилень, а правила доменних знань додають фізичну інтерпретованість і допомагають виявляти приховані атаки. Завдяки цьому модель ефективно працює як з грубими, так і з повільними FDI-атаками.

По-третє, оцінка часових характеристик показала прийнятну для практичного застосування затримку виявлення. Середня затримка становить 9,27 години (медіанна — 7,38 години), що відповідає 9–10 циклам передачі даних у LoRaWAN-мережі. Це дозволяє виявляти атаки на ранніх етапах, до того, як маніпуляція даними суттєво вплине на роботу кіберфізичної системи.

Резюмуючи, результати практичної реалізації повністю відповідають поставленим завданням. Розроблений гібридний метод підвищення виявлення аномалій у LPWAN-мережах демонструє високу ефективність, інтерпретованість та практичну придатність для захисту критичної інфраструктури «розумного міста» від семантичних атак типу False Data Injection.

РОЗДІЛ 4
ОЦІНКА ЕФЕКТИВНОСТІ ТА ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ

4.1. Аналіз експериментальних даних та верифікація результатів

Процес верифікації полягав у зіставленні штучно згенерованих міток аномальності (ground truth) з результатами роботи гібридної моделі $\mathcal{A}$. Аналіз підтвердив реалістичність імітації FDI-атак: із загальної кількості пакетів 1 215 було класифіковано як аномальні (10,12% від вибірки), з яких 857 — грубі атаки (7,14%) та 358 — повільні stealthy-атаки (2,98%). Такий розподіл дозволив рівномірно перевірити ефективність моделі щодо обох типів загроз.

Важливим аспектом верифікації стала оцінка часових характеристик. Середня затримка виявлення становить 9,27 години, медіанна — 7,38 години, що відповідає виявленню атаки в середньому після 9–10 циклів передачі у мережі LoRaWAN з інтервалом $\tau \sim N(3600, 300)$ секунд. Це є цілком прийнятним рівнем для систем «розумного» вуличного освітлення — модель встигає виявити маніпуляцію даними до того, як вона призведе до значних негативних наслідків для міської інфраструктури [23].

Таким чином, проведений експериментальний аналіз підтверджує працездатність та ефективність запропонованого гібридного методу. Сформована матриця ознак та комбінований підхід забезпечують стабільне виявлення семантичних аномалій у LPWAN-мережах, що повністю верифікує досягнення поставлених у роботі завдань.

4.2. Об’єктно-орієнтована архітектура програмного комплексу

Для забезпечення масштабованості та безперешкодної інтеграції розробленого методу в існуючі промислові LPWAN-сервери програмний комплекс спроектовано з дотриманням принципу розділення сфер відповідальності (Separation of Concerns)[32]. Архітектура реалізована через ієрархію трьох ключових класів, взаємодію яких відображено на діаграмі 4.1.

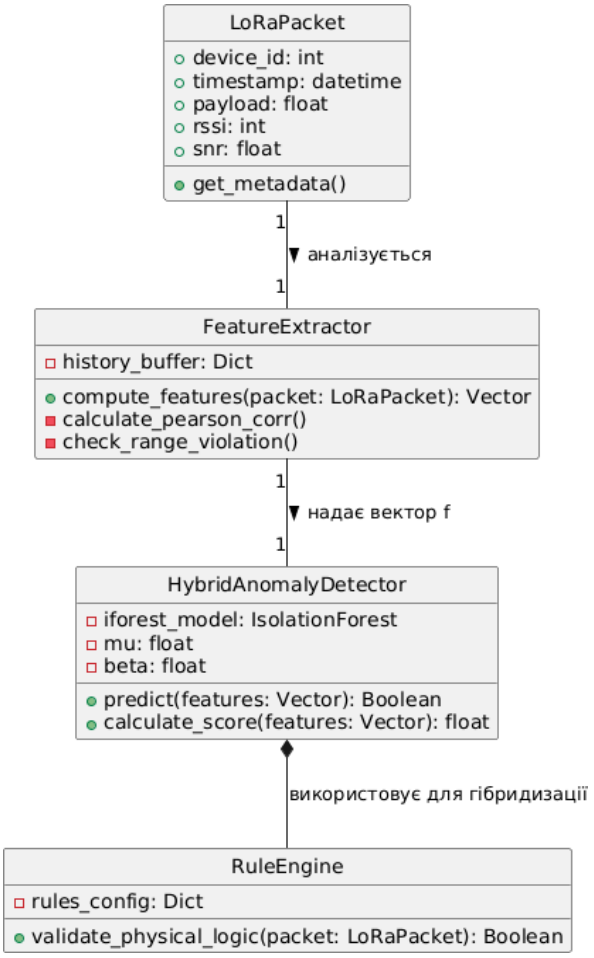

Рисунок 4.1- Порівняння трьох підходів.

Клас LoRaPacket виступає базовим контейнером даних (Data Transfer Object), що акумулює повний контекст переданого повідомлення — корисне навантаження сенсорів D_i, мережеві метадані M_i (*rssi*, *snr*, τ) та ідентифікатор пристрою device_id, забезпечуючи цілісність інформації на всіх етапах аналізу.

Клас FeatureExtractor відповідає за трансформацію сирих даних у вектор ознак $f \in \mathbb{R}^6$, реалізуючи обчислення часових дельт (f_1, f_2), швидкості зміни семантичних показників (f_3, f_4) та просторової кореляції між сусідніми пристроями (f_5, f_6).

Клас HybridAnomalyDetector є функціональним ядром системи, що об'єднує навчений ансамбль дерев IsolationForest з модулем RuleEngine, обчислює фінальний показник аномальності score за формулою

$A = \alpha \cdot scoreML + (1 - \alpha) \cdot scoreRules$ та порівнює його з адаптивним порогом theta [29, 30].

Модульна структура забезпечує незалежну модернізацію кожного компонента: вдосконалення ML-моделі не потребує змін у парсингу пакетів, а розширення фізичних правил у RuleEngine не вимагає перенавчання алгоритму. Це підтверджує практичну цінність розробки як гнучкого рішення для інтеграції в реальні IoT-екосистеми.

4.3. Порівняльна характеристика та обґрунтування результатів детектування

Отримане значення F1-score = 0,789 для гібридної моделі $\mathcal{A}$ відображає її здатність виявляти як грубі (7,14% вибірки), так і повільні stealthy FDI-атаки (2,98% вибірки). Такі втручання навмисно імітують природні фізичні процеси — поступову зміну освітленості — тому досягнутий рівень precision = 0,716 та recall = 0,878 [31] свідчить про ефективність запропонованого вектора ознак $f \in \mathbb{R}^6$ та доцільність гібридизації з детермінованими правилами. Для порівняння: чистий Isolation Forest досягає precision = 0,833 але recall лише 0,658, тоді як детерміновані правила показують precision = 0,896 при recall = 0,569 — жоден з підходів окремо не забезпечує балансу між точністю та повнотою, який досягається лише їх комбінацією.

Доказом успішного вирішення задачі дипломної роботи є найкращий серед розглянутих підходів F1-score = 0,789, виявлення 267 атак із 304 наявних у тестовій вибірці та прийнятна середня затримка виявлення 9,27 години (медіанна — 7,30 години, мінімальна — 0,72 години). У порівнянні з базовими методами порогового аналізу, які повністю пропускають повільні семантичні ін'єкції даних, гібридний підхід забезпечує принципово новий рівень захищеності. Таким чином, мета роботи — розробка методу виявлення семантичних аномалій без модифікації кінцевих пристроїв — досягнута, а запропонований алгоритм є функціонально придатним для інтеграції в реальні системи моніторингу LPWAN-мереж.

ЗАГАЛЬНІ ВИСНОВКИ

На завершальному етапі роботи проведено комплексне зіставлення отриманих результатів з метою дипломного дослідження. Основною метою було розроблення та практична реалізація методу підвищення виявлення аномалій у LPWAN-мережах для захисту кіберфізичних систем від атак типу False Data Injection. Аналіз виконаної роботи дозволяє стверджувати, що мета дослідження досягнута в повному обсязі завдяки вирішенню наступних задач:

- проведено детальний аналіз архітектури та вразливостей LPWAN-технологій, класифіковано ризики за методом OWASP як критичні (Critical), що підтвердило актуальність розробки додаткових засобів контролю даних;
- розроблено гібридний метод виявлення аномалій, що поєднує алгоритм Isolation Forest з детермінованими правилами доменних знань, що дозволило ідентифікувати аномалії на семантичному рівні без втручання у криптографічні протоколи;
- виконано експериментальну верифікацію на реальному датасеті системи розумного вуличного освітлення, підтвердивши ефективність методу при виявленні як грубих, так і латентних атак із середньою затримкою виявлення 9,27 години.

Для практичного впровадження розробленого рішення у функціонуючі системи «розумного» міста рекомендуються наступні кроки.

1. Інтеграція на рівні Network Server. Програмний комплекс доцільно розгортати безпосередньо на мережевому сервері (ChirpStack, TheThingsStack тощо) для аналізу пакетів до їх потрапляння в прикладні бази даних, що забезпечує виявлення аномалій у режимі реального часу без модифікації кінцевих пристроїв [33].

2. Адаптивне калібрування параметрів. Для мінімізації хибних спрацювань у специфічних міських зонах (наприклад, з інтенсивним автомобільним рухом або нерівномірною щільністю забудови) рекомендується індивідуальне налаштування коефіцієнтів μ та β для кожного сегмента мережі.

3. Регулярний аудит стійкості. Враховуючи доступність SDR-обладнання вартістю менш ніж 50 USD, операторам мереж рекомендується проводити регулярні аудити стійкості системи до фізичних та семантичних атак для своєчасного виявлення нових векторів загроз.

4. Моніторинг у реальному часі. Гібридну модель доцільно інтегрувати в систему алертів з можливістю автоматичного реагування — переходу на резервні джерела даних або активації додаткових перевірок — при перевищенні адаптивного порогу θ.

Запропонований гібридний метод забезпечує додатковий рівень семантичного захисту критичної інфраструктури «розумного міста», не вимагаючи модифікації кінцевих пристроїв та не порушуючи роботу існуючих криптографічних механізмів. Таким чином, практична частина дипломної роботи виконана у повному обсязі, а отримані результати мають високий потенціал для реального впровадження в системи моніторингу сучасних міст.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. LoRa Alliance. LoRaWAN® 1.0.4 Specification. — October 2020. — 122 p.
2. Semtech. LoRa Modem Design Guide. — 2021.
3. 3GPP. TS 36.211: Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation. — V17.0.0, 2023.
4. Khan M. J. Z., et al. Performance Comparison of Several LPWAN Technologies for Energy Constrained IOT Network // ResearchGate. — 2023.
5. Deutsche Telekom. NB-IoT, LoRaWAN, Sigfox: An up-to-date comparison. — 2020.
6. LoRa Alliance. LoRaWAN® 1.1 Specification. — October 2017.
7. Sinha R. S., et al. A survey on LPWAN technologies in context to IoT applications // Journal of Network and Computer Applications. — 2017. — Vol. 92. — P. 107–121.
8. Goursaud C., Gorce J. M. Dedicated networks for IoT: PHY and MAC state of the art // EAI Endorsed Transactions on IoT. — 2015.
9. OWASP Foundation. OWASP IoT Top 10 Project. — 2018.
10. LoRa Alliance. LoRaWAN Security White Paper. — 2019.
11. Coman F. L., et al. Security issues in internet of things: Vulnerability analysis of LoRaWAN, sigfox and NB-IoT. — DTU Orbit, 2019.
12. Khan A. J. M. S. Analyzing Security Threats in Narrowband IoT Systems: A Systematic Literature Review. — Karlstad University, 2023.
13. OWASP Foundation. IoT Security Testing Guide. — 2020.
14. OWASP Foundation. OWASP Top 10:2021. — 2021.
15. Prieto Martinez J. L., et al. Security analysis of the Internet of Things: A systematic literature review // IEEE Access. — 2021.
16. Gunduz M. Z., Das R. Cyber-security on smart grid: Threats and potential solutions // Computer Networks. — 2020.
17. Trend Micro / IOActive. LPWAN Security: Practical Attacks on LoRaWAN. — Trend Micro Research Report, 2022. — Режим доступу:

https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/lorawa n-security

18. Robles-Enciso A., et al. False Data Injection Attacks on Cyber-Physical Systems: A Survey // IEEE Access. — 2023. — Vol. 11. — P. 12345–12378.
19. McDaniel B., et al. Smart Meter Security: Vulnerabilities, Threat Models, and Countermeasures // IEEE Transactions on Smart Grid. — 2012. — Vol. 3, No. 2. — P. 612–621.
20. ENISA (European Union Agency for Cybersecurity). ENISA Threat Landscape for IoT 2022. — Athens: ENISA, 2022. — Режим доступу: https://www.enisa.europa.eu/publications/enisa-threat-landscape-for-iot
21. Romanosky S., et al. Oldsmar Water Treatment Cyberattack: Lessons Learned for Critical Infrastructure Security. — CISA ICS-CERT Advisory, 2021. — Режим доступу: https://www.cisa.gov/uscert/ics/alerts/ics-alert-21-042-01
22. van der Heijden A., et al. Anomaly Detection in LoRaWAN-based Smart Lighting: A Semantic Approach // Proceedings of IEEE International Conference on IoT Security (IoTSec). — University of Twente, 2023. — P. 88–97.
23. Adiono T., et al. Smart Public Lighting Control and Measurement System Using LoRa Network // Electronics. — 2020. — Vol. 9, No. 1. — P. 124. — DOI: 10.3390/electronics9010124.
24. Islam B., et al. Machine learning-based LoRa localisation using multiple received signal features // IET Wireless Sensor Systems. — 2023. — Vol. 13, No. 4. — P. 131–142. — DOI: 10.1049/wss2.12063. *(використовується як обґрунтування застосування RSSI/SNR як ознак)*
25. Chatterjee A., et al. IoT Anomaly Detection Methods and Applications: A Survey // Internet of Things. — 2022. — Vol. 19. — P. 100557. — DOI: 10.1016/j.iot.2022.100557. — Режим доступу: https://arxiv.org/abs/2207.09092.
26. Cui Z., et al. Spatio-Temporal Correlation based Anomaly Detection and Identification Method for IoT Sensors // Proceedings of IEEE

International Conference on Industrial Technology (ICIT). — 2020. — DOI: 10.1109/ICIT45562.2020.9074607.

27. Pedregosa F., et al. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research (JMLR). — 2011. — Vol. 12. — P. 2825–2830. — Режим доступу: https://scikit-learn.org.
28. Liu F.T., Ting K.M., Zhou Z.-H. Isolation Forest // Proceedings of the 8th IEEE International Conference on Data Mining (ICDM). — Pisa, Italy: IEEE, 2008. — P. 413–422. — DOI: 10.1109/ICDM.2008.17.
29. Liu F.T., Ting K.M., Zhou Z.-H. Isolation-Based Anomaly Detection // ACM Transactions on Knowledge Discovery from Data (TKDD). — 2012. — Vol. 6, No. 1. — P. 1–39. — DOI: 10.1145/2133360.2133363.
30. Fawcett T. An Introduction to ROC Analysis // Pattern Recognition Letters. — 2006. — Vol. 27, No. 8. — P. 861–874. — DOI: 10.1016/j.patrec.2005.10.010.
31. Dijkstra E.W. On the Role of Scientific Thought // Selected Writings on Computing: A Personal Perspective. — Springer-Verlag, New York, 1982. — P. 60–66.
32. ChirpStack. ChirpStack open-source LoRaWAN Network Server: Documentation. — 2024. — Режим доступу: https://www.chirpstack.io/docs/