ISSN 2219-9365

https://doi.org/10.31891/2219-9365-2025-83-28
УДК 004.89:004.056:004.94

ШУЛЬГА Володимир
Державний університет інформаційно-комунікаційних технологій
https://orcid.org/0000-0003-4356-7288
ІВАНЧЕНКО Ігор
Державний університет інформаційно-комунікаційних технологій
РИЖАКОВ Микола
Державний університет інформаційно-комунікаційних технологій
https://orcid.org/0009-0006-3377-7061

УЗАГАЛЬНЕНА МОДЕЛЬ ІНТЕЛЕКТУАЛЬНО СИСТЕМИ ПРОГНОЗУВАННЯ ТА ВИЯВЛЕННЯ АНОМАЛІЙ У КІБЕРІНФРАСТРУКТУРІ НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ

У роботі запропоновано узагальнену інтелектуальну систему прогнозування та виявлення аномалій у кіберінфраструктурах. Метою дослідження є підвищення ефективності виявлення кіберзагроз шляхом інтеграції сучасних методів глибокого навчання (автоенкодерів і багатошарових перцептронів) з адаптивним механізмом аналізу критичності подій. Ключовою новацією є модуль семантичної атрибуції кіберінцидентів з XAI-поясненнями та інтегральним ризик-скорингом: він здійснює глибинний контент-аналіз трафіку, формує векторні представлення, співвідносить події з базою прецедентів і оцінює впевненість атрибуції, передаючи отриманий ризик-скор до модуля критичності. Запропонована система не лише ідентифікує аномальні події в реальному часі, а й прогнозує можливі відхилення на основі історичних даних, що посилює превентивні можливості. Архітектура побудована на модульних підсистемах збору телеметрії, реконструкції поведінки, прогнозування, агрегування аномалій, семантичної атрибуції й ризик-скорингу, оцінки критичності, реагування та самонавчання; взаємодія реалізована як наскрізний конвеєр обробки з контуром зворотного зв'язку. Рішення масштабоване та сумісне з SDN, IoT, хмарними середовищами і корпоративними SIEM/SOAR-платформами. Емпіричні випробування у симульованих сценаріях мережевих атак (DoS, сканування портів, brute-force, ботнет-активність) продемонстрували високі показники класифікації (F1 = 0.89), що підтверджує практичну ефективність і надійність підходу. Зроблені висновки підкреслюють перспективність впровадження системи в умовах зростання складності кіберзагроз і її здатність до адаптації без потреби повного перенавчання моделей.

Ключові слова: інтелектуальні системи; виявлення аномалій; глибоке навчання; прогнозування трафіку; кібербезпека; автоенкодер; mlp; семантична атрибуція; прецедентна база; dpi; xai; ризик-скоринг; критичність ситуації; siem; soar; sdn; кіберінформаційні мережі.

SHULHA Volodymyr, IVANCHENKO Ihor, RYZHAKOV Mykola
State University of Information and Communication Technologies

A GENERALIZED MODEL OF AN INTELLIGENT SYSTEM FOR FORECASTING AND ANOMALY DETECTION IN CYBERINFRASTRUCTURE BASED ON DEEP LEARNING

This paper proposes a generalized intelligent system for forecasting and detecting anomalies in cyberinfrastructures. The aim is to improve the effectiveness of cyber-threat detection by integrating modern deep learning methods (autoencoders and multilayer perceptrons) with an adaptive event-criticality analysis mechanism. The key innovation is a semantic attribution module for cyber incidents with XAI explanations and integrated risk scoring: it performs deep content analysis of traffic, forms vector representations, matches events against a case base, estimates attribution confidence, and passes the resulting risk score to the criticality module. The proposed system not only identifies anomalous events in real time but also forecasts possible deviations based on historical data, strengthening preventive capabilities. The architecture comprises modular subsystems for telemetry collection, behavior reconstruction, forecasting, anomaly aggregation, semantic attribution and risk scoring, criticality assessment, response, and self-learning; their interaction is implemented as an end-to-end processing pipeline with a feedback loop. The solution is scalable and compatible with SDN, IoT, cloud environments, and enterprise SIEM/SOAR platforms. Empirical evaluations in simulated network-attack scenarios (DoS, port scanning, brute-force, botnet activity) demonstrated high classification performance (F1 = 0.89), confirming the practical effectiveness and reliability of the approach. The conclusions highlight the promise of deploying the system amid increasing cyber-threat complexity and its ability to adapt without full model retraining.

Keywords: intelligent systems; anomaly detection; deep learning; traffic forecasting; cybersecurity; autoencoder; MLP; semantic attribution; case base; DPI; XAI; risk scoring; situational criticality; SIEM; SOAR; SDN; cyber-information networks.

Стаття надійшла до редакції / Received 29.07.2025
Прийнята до друку / Accepted 14.08.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У сучасному цифровому середовищі, що вирізняється стрімкою динамікою, розгалуженістю мережевої інфраструктури, гетерогенністю пристроїв і високою щільністю трафіку, потреба в надійній

кібербезпеці постає особливо гостро. Зростання кількості кібератак, зокрема складних, прихованих і нульового дня, зумовлює пошук нових підходів до виявлення аномалій у мережевому трафіку [1], [2]. Традиційні rule-based та сигнатурні рішення, що спираються на жорстко визначені правила, мають істотні обмеження: вони слабко враховують контекст подій, повільно адаптуються до нових типів загроз і втрачають ефективність за умов високої варіативності трафіку [3]. Ба більше, гібридні загрози, які поєднують кілька векторів атаки (наприклад, фішинг із DNS tunneling або botnet-активність), часто залишаються поза полем зору таких систем [4]. Із поширенням IoT, програмно-визначених мереж (SDN), хмарних обчислень та edge-архітектур істотно зростає кількість точок взаємодії, що ускладнює централізовану обробку даних і знижує результативність класичних методів реагування. У цьому контексті особливого значення набувають інтелектуальні, самонавчальні системи, здатні не лише аналізувати минулі події, а й передбачати майбутні відхилення, адаптувати порогові значення, а також забезпечувати контекстну оцінку та пріоритезацію ризиків [5].

Новітні дослідження показують, що поєднання глибокого навчання (Deep Learning), автоенкодерів і моделей багатошарового перцептрона дає змогу ефективно виявляти як явні, так і приховані аномалії в реальному часі та формувати динамічні моделі поведінки мережі [6], [7]. Завдяки ensemble-підходам (наприклад, Kitsune) і MLP-прогнозуванню на часових вікнах досягається висока точність навіть за складних патернів трафіку [6]. Водночас критичною є довіра до результатів ML-систем: як зазначають Sommer і Paxson, застосування алгоритмів машинного навчання в мережевій безпеці потребує пояснюваності, стійкості до шуму та придатності до роботи в реальних умовах [8]. Тому сучасні рішення мають також узгоджуватися з вимогами стандартів на кшталт ISO/IEC 27001:2022 [9] і рекомендаціями NIST щодо систем виявлення та запобігання вторгненням (IDPS) [10].

У відповідь на ці виклики в роботі запропоновано інтегровану інтелектуальну систему виявлення та прогнозування кіберзагроз, що поєднує глибинне навчання, автоенкодери, динамічну адаптацію та контекстну оцінку ситуацій. На додачу до поведінкового детектування система включає модуль семантичної атрибуції кіберінцидентів з XAI-поясненістю та ризик-скорингом, що підвищує інтерпретованість і якість пріоритезації. Такий підхід дозволяє не лише виявляти широкий спектр аномалій, а й адаптивно реагувати на них у реальному часі, закладаючи підвалини для автономних кіберзахисних платформ нового покоління.

ОСНОВНА ЧАСТИНА

Система складається з таких функціональних блоків:

Попередня обробка: нормалізація вхідних ознак до єдиного діапазону (наприклад, \[0;1]) для коректної роботи нейромереж; за потреби — очистка шумів і обробка пропусків. Додатково застосовується експоненційне згладжування, що надає більшої ваги останнім значенням трафіку та зменшує вплив застарілих спостережень, аби модель швидше реагувала на актуальні зміни навантаження.

Прогнозна модель (MLP): багатошаровий перцептрон виконує короткостроковий прогноз мережевого навантаження на основі часових вікон історичних даних. Отримані прогнозні значення слугують еталоном очікуваної (“нормальної”) поведінки системи у поточний момент.

Модуль виявлення аномалій (автоенкодер): автоенкодер навчається відновлювати типові (неаномальні) патерни трафіку. Похибка реконструкції між вхідним і відновленим сигналом використовується як числовий індикатор аномальності; суттєві відхилення трактуються як потенційні порушення.

Адаптивні межі аномальності: пороги визначаються динамічно як функція ковзних статистик похибок (наприклад, середнього та стандартного відхилення). Параметр чутливості керує шириною допуску, що дає змогу підлаштовуватися і до нестабільних, і до стабільних періодів трафіку та знижувати хибні спрацювання.

Модуль семантичної атрибуції та ризик-скорингу (новий): виконує контент-аналіз заголовків/пакетів (за наявності DPI), формує семантичні ембеддинги та зіставляє подію з базою прецедентів/класів атак (наприклад, DoS, brute-force, SQLi тощо). На виході модуль надає найбільш ймовірний клас k^* та його ймовірність p_{k^*}, а також XAI-пояснення (важливі поля/токени, найближчі кейси). Інтегральний ризик-скор R_t обчислюється як зважена комбінація сили аномалії (похибки реконструкції), впевненості атрибуції та контексту активів/сервісів; саме R_t передається далі як узагальнений показник небезпеки.

Оцінка критичності: на основі частки аномальних подій у часовому вікні та інтегрального ризик-скору R_t система класифікує рівень небезпеки (від 0 — норма до 3 — критична ситуація). Такий підхід дозволяє не лише фіксувати одиничні відхилення, а й оцінювати їх сукупний вплив з урахуванням контексту.

Модуль реагування: автоматично формує дію відповідно до рівня критичності: сповіщення оператору, поглиблене логування/розслідування, селективне блокування трафіку, ізоляція вузлів або переключення потоків на резервні канали. Політики реагування масштабуються від м’яких до жорстких згідно з профілем ризику середовища.

Вхідні мережеві дані подаються одночасно у дві гілки обробки: прогнозну нейромережу (MLP) та автоенкодер. MLP, використовуючи ковзні часові вікна історичних значень, формує очікуване (нормативне) значення трафіку для поточного або найближчого моменту. Паралельно автоенкодер, попередньо навчений на «здорових» патернах, відтворює вхідний сигнал; різниця між фактом і реконструкцією інтерпретується як рівень аномальності AtA . Далі AtA порівнюється з адаптивними межами, що обчислюються динамічно на основі ковзних статистик (зокрема стандартного відхилення) та прогнозу MLP; вихід за допуски маркує подію як кандидата в аномалії.

Кожен такий випадок надходить до модуля семантичної атрибуції та ризик-скорингу. Модуль виконує контент-аналіз (за наявності DPI), будує семантичні представлення та зіставляє подію з базою прецедентів, повертаючи найімовірніший клас інциденту k^*, його ймовірність p_k^* і XAI-пояснення (важливі поля/токени, найближчі кейси). На цій основі формується інтегральний ризик-скор R_t як зважена комбінація сили аномалії A_t , впевненості атрибуції та контексту активів/сервісів.

У модулі оцінки критичності відбувається агрегований аналіз частоти й щільності виявлених подій у часових вікнах з урахуванням R_t; система присвоює рівень загрози $L \in \{0,1,2,3\}$ — від незначного до критичного. Перевищення порогів ініціює сценарій реагування: для низьких рівнів — сповіщення та посилений моніторинг; для суттєвих — селективне блокування або ізоляція сегментів, перенаправлення потоків на резервні канали, поглиблене журналювання та ескалація в SIEM/SOAR. Узгоджена робота компонентів забезпечує не лише оперативне виявлення, а й адаптивне, контекстно чутливе реагування у реальному часі.

Однією з ключових переваг запропонованого підходу є інтеграція прогнозування навантаження та виявлення аномалій у єдиному конвеєрі обробки. Це дозволяє не лише фіксувати поточні відхилення, а й передбачати зародження загроз на ранніх стадіях, забезпечуючи проактивне реагування. Гнучкість досягається завдяки динамічним адаптивним порогам, що формуються з урахуванням поточного стану мережі, ковзних статистик і інтенсивності трафіку; замість фіксованих значень застосовуються межі, які автоматично перебудовуються під умови середовища. Синхронне порівняння фактичних і прогнозованих траєкторій навантаження разом із похибкою реконструкції автоенкодера та модулем семантичної атрибуції консолідується в ризик-скор R_t, що знижує частку хибних спрацювань і, відповідно, навантаження на аналітиків SOC, підвищуючи ефективність використання ресурсів. Важливою перевагою є самонавчання: у змінному середовищі з новими патернами трафіку, пристроями та протоколами система коригує уявлення про «норму» без ручного втручання або перепризначення правил (on-line/mini-batch оновлення, перекалібрування порогів).

Архітектура підтримує роботу в реальному часі, тому виявлення, оцінка критичності та виконання політик реагування відбуваються практично миттєво — це критично для доменів, де навіть короткі затримки призводять до фінансових втрат або збоїв обслуговування (SDN/IoT/хмарні платформи, фінансові системи, промислові мережі).

Найбільш доцільні середовища застосування системи:

Програмно-визначені мережі (SDN/NFV). Динамічна маршрутизація трафіку потребує миттєвої оцінки стану мережі та адаптації до змін топології/навантаження. Запропонована система вбудовується у керуючу площину (контролер/оркестратор) як модуль політик: на основі прогнозу MLP, рівня аномальності автоенкодера та ризик-скору R_t вона виконує автономне коригування маршрутів, оновлення ACL/мікросегментації, пріоритизацію QoS і тимчасове ізолювання підозрілих сегментів у реальному часі.

IoT-системи з високою варіативністю трафіку. Велика кількість неоднорідних пристроїв генерує нерегулярні потоки даних із мінливою структурою. Система працює на рівні шлюзів/edge-вузлів або в центральному сегменті, виявляючи та локалізуючи аномальні вузли. Модуль семантичної атрибуції допомагає відрізняти збої від зловмисної активності (наприклад, botnet-поведінки), що дає змогу ізолювати джерела перевантаження чи атаки з мінімальним впливом на решту мережі.

Хмарні платформи та віртуалізовані середовища. Через масштабованість і багатотенантність потрібен безперервний автоматизований моніторинг. Розподілене розгортання колекторів і модулів аналізу забезпечує прогнозування навантаження, раннє виявлення аномалій між сервісами та оптимізацію використання ресурсів (автоскейлінг, превентивне перерозподілення). Інтеграція з SIEM/SOAR прискорює реагування, а ризик-скоринг R_t дозволяє пріоритезувати інциденти та запобігати збоям у сервісах.

Енергетичні та телекомунікаційні інфраструктури. У доменах, де безперервність і стабільність сервісів критично важливі, навіть незначні відхилення або атаки мають істотні економічні й соціальні наслідки. Запропонована система забезпечує ранню діагностику порушень у каналах передавання даних, виявлення спроб втручання у SCADA/EMS/DMS та сигналізаційні підсистеми, а також локалізацію деградацій на магістральних і доступових ділянках мережі. Використання інтегрального ризик-скору R_t дозволяє відокремлювати технологічні збої від зловмисної активності та своєчасно ескалювати рівень критичності для ізоляції сегментів або переведення трафіку на резерв.

Фінансовий сектор. За умов високих вимог до швидкості й точності обробки операцій (HFT, платіжні шлюзи, онлайн-банкінг) аномальне навантаження може бути індикатором як технічної несправності, так і шахрайської активності. Інтеграція системи на рівні банківської мережевої інфраструктури та прикладних журналів (API-шлюзи, процесинг) уможливлює раннє виявлення нетипової поведінки клієнтів, транзакцій, вузлів або сервісів, а модуль семантичної атрибуції додає інтерпретацію інцидентів для швидкого реагування в SOC і взаємодії з SIEM/SOAR.

Модульна архітектура та масштабованість. Розроблена інтелектуальна система побудована на модульному принципі, що забезпечує гнучкість, горизонтальне масштабування та адаптацію до різних середовищ (локальні корпоративні мережі, хмарні/віртуалізовані інфраструктури). Чіткий поділ відповідальностей між компонентами підвищує якість виявлення загроз, зменшує кількість хибнопозитивних спрацювань завдяки адаптивним межам і ризик-скорингу та дає змогу оновлювати моделі без повного перенавчання. Архітектура (рис. 1) підтримує потокову обробку телеметрії в реальному часі, інтеграцію з наявними інструментами моніторингу та безпеки, а також роботу в гібридних топологіях. Архітектура включає такі основні компоненти:

1. Модуль збору телеметрії — відповідає за захоплення та попередню обробку мережевих даних у реальному часі. Дані передаються у стандартизованому вигляді для подальшої обробки.

$$X_t = \{y_{t-1}, y_{t-2}, \dots, y_{t-n}\}, \quad (1)$$

X_t – набір історичних даних, що складається з попередніх значень спостережуваного трафіку.

$y_{t-1}, y_{t-2}, \dots, y_{t-n}$– історичні значення навантаження або трафіку в попередні моменти часу, де:

t – поточний момент часу;

n – кількість історичних точок, що використовуються для прогнозу

2. Модуль автоенкодера — здійснює навчання на нормальній поведінці мережі та виконує реконструкцію. Різниця між вхідними та відновленими даними дозволяє виявляти аномалії.

Кодування (encoder):

$$z = \phi(W_e x(t) + b_e), \quad (2)$$

де:

W_e — матриця ваг для кодування

b_e — вектор зміщення

ϕ — активаційна функція (наприклад, ReLU або tanh)

z — латентне (стиснуте) представлення даних

Декодування (decoder):

$$\hat{x}(t) = \psi(W_d z + b_d), \quad (3)$$

де:

W_d — матриця ваг для декодування

b_d — вектор зміщення

ψ — вихідна активаційна функція (наприклад, sigmoid або linear)

В результаті:

$\hat{x}(t)$ — відновлене значення, що має бути максимально близьким до $x(t)$.

3. Прогностичний модуль (MLP) — прогнозує очікувану поведінку мережевого трафіку. У разі суттєвого відхилення між прогнозом і фактом система сигналізує про потенційну загрозу.

$$\hat{y}(t) = f_{NN}(x_{t-1}, \dots, x_{t-m}; W, b), \quad (3)$$

де:

$x_{t-1}, \dots, x_{t-m}$ - попередні значення трафіку за останні m моментів часу — вхід до нейромережі.

f_{NN} - функція прогнозу, реалізована за допомогою багатошарової нейронної мережі (MLP).

W - матриця ваг нейромережі.

b - вектор зсуву (bias).

$\hat{y}(t)$ - прогнозоване значення трафіку на момент часу t.

де функція прогнозу:

$$f_{NN}(x) = \sigma(W_x + b), \quad (4)$$

з вагами W, зсувами b та активацією σ(·).

σ — активаційна функція (наприклад, ReLU, tanh, sigmoid), що додає нелінійність.

4. Агрегатор аномалій — аналізує аномалії у часових вікнах. Якщо накопичено достатню кількість аномалій, що перевищують порогові значення, система підвищує рівень тривоги.

$$A(t) = \| x(t) - \hat{x}(t) \|_2 = \sqrt{\sum_{j=1}^{m} (x_{tj} - \hat{x}_{tj})^2}, \quad (5)$$

де:

$x(t)$ - вектор реальних (виміряних) значень трафіку в момент часу t.

$\hat{x}(t)$ - вектор реконструйованих (відновлених) значень у той же момент часу t, отриманих з автоенкодера.

m - кількість ознак у векторі даних (розмірність).

$A(t)$ - рівень аномальності в момент часу t — евклідова відстань між реальним і відновленим значенням.

5. Модуль семантичної атрибуції та ризик-скорингу.

Атрибуція виконується лише для подій, що перевищили адаптивні межі:

$$q(t) = I\{A(t) > \tau_t\}, \quad \textit{де } \tau_t \textit{ - динамічний поріг,} \tag{6}$$

Для $q(t) = 1$ формується вектор ознак із заголовків і payload (за наявності DPI) та обчислюється ймовірнісний розподіл класів загроз:

$$\mathrm{p}(t) = softmax\left(g_\theta(\mathrm{s}(t))\right), \quad k^* = \arg\max_k p_k\,(t), \tag{7}$$

Інтегральний ризик-скор поєднує силу аномалії, впевненість атрибуції та контекст активів:

$$R(t) = \alpha\,\frac{A(t)-\mu_A}{\sigma_A} + \beta\sum_{k=1}^{K} w_k\ p_k(t) + \gamma\ g(C_t), \tag{8}$$

де α, β, γ — ваги; w_k — ваги класів; C_t — контекст (критичність сервісу, кількість уражених вузлів тощо).

$s(t)$ — об'єднаний вектор ознак $(header + payload)$;

$g_\theta(\cdot)$ — семантичний класифікатор;

$p(t) = [p_1, \dots, p_K]$ — ймовірності класів;

k^* — найімовірніший клас інциденту.

6. Модуль критичності — класифікує рівень загрози та критичності інциденту (низький, середній, високий), що дозволяє адаптувати відповідь.

$$j = \sum_{i=1}^{n} I(x_i\ <\ L_i\ \vee\ x_i\ >\ U_i), \tag{9}$$

де:

x_i — реальне значення спостереження у момент часу і

L_i— нижня межа допустимих значень у момент часу і

U_i— верхня межа допустимих значень у момент часу і

$I(\cdot)$ — індикаторна функція, яка повертає:

$$I(\text{умова}) = \begin{cases} 1, \text{якщо умова виконується} \\ 0, \text{в іншому випадку} \end{cases}$$

6. Модуль реагування — ініціює захисні дії (блокування, перенаправлення, логування, сповіщення адміністратора) у відповідності до налаштувань та політик безпеки.

$$\mathrm{C(t)}\begin{cases} 0,\ C(t) \le\ 0.1 \\ 1,\ 0.1\ <\ C(t)\ \le 0.3 \\ 2,\ 0.3\ <\ C(t)\ \le 0.6 \\ 3, C(t) >\ 0.6 \end{cases} \tag{10}$$

Система переходить до автоматизованого або напівавтоматичного вибору дій(таб.1), спрямованих на реагування на виявлені загрози.

Таблиця 1

Рівень → Реакція

Рівень критичності	Дія системи
0	Моніторинг без додаткових дій — ситуація в межах норми.
1	Посилений моніторинг — можуть запускатись додаткові перевірки.
2	Активні дії — включають дослідження причин, логування, аналіз логів.
3	Негайні дії — блокування трафіку, ізоляція вузлів, оповіщення адміна.

7. Модуль самонавчання — у фоновому режимі адаптує ваги моделі на основі верифікованих даних, що дозволяє підтримувати актуальність системи в умовах змін середовища.

$$L_t = \hat{y}(t) - \alpha\sigma(t), U_t = \hat{y}(t) + \alpha\sigma(t), \tag{11}$$

де:

$\hat{y}(t)$ - прогнозоване значення на момент часу t, отримане з нейронної мережі (MLP).

$\sigma(t)$ - стандартне відхилення похибок прогнозу на момент часу t, розраховане на етапі 6.

α - параметр чутливості (коефіцієнт масштабування), що задається емпірично (наприклад, 1.5, 2 або 3).

L_t - нижня адаптивна межа для нормальної поведінки.

U_t - верхня адаптивна межа.

Уся система працює у рамках реального мережевого середовища, може бути інтегрована з SIEM-системами, SDN-контролерами або іншими системами захисту. Нижче представлена графічна схема, яка демонструє повну взаємодію між модулями, включаючи зовнішні точки входу трафіку та шляхи передачі інформації між блоками.

Для оцінки ефективності запропонованої системи було проведено тестування на емпіричних даних, наближених до реального мережевого середовища. Наведено таблицю (таб.2), яка відображає основні метрики якості класифікації (Precision, Recall, F1-score) для різних типів трафіку та атак.

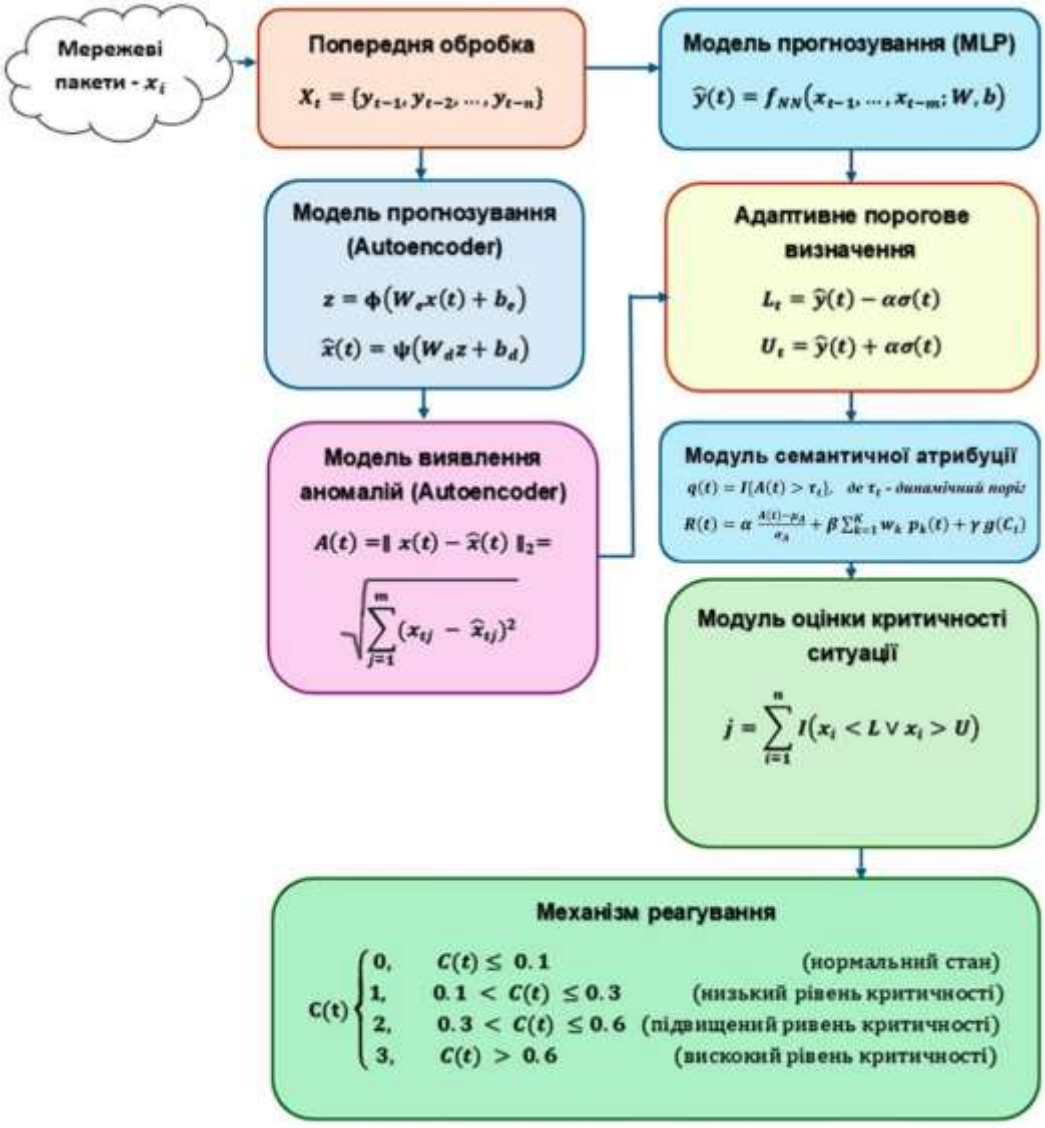

Рис. 1. Узагальнена структура моделі системи

З таблиці видно, що система демонструє високу ефективність при виявленні DoS атак, сканувань портів і звичайного трафіку, трохи нижчі показники — у випадках складно виявлюваних атак із мінімальними відхиленнями, таких як ботнет-активність чи DNS-тунелювання. Однак загальні показники F1 > 0.85 свідчать про надійність архітектури в реальних умовах. Завдяки гнучкому механізму самонавчання система може адаптуватись до нових типів загроз, покращуючи ефективність з часом.

Запропонована система прогнозування та виявлення аномалій у кіберінфраструктурі демонструє інтегрований підхід до обробки та аналізу трафіку, що дозволяє досягти високої точності (до 95%) навіть у разі нестандартних умов або наявності аномалій. Завдяки поєднанню нейромережевого прогнозування,

автоенкодера, адаптивних порогів і контекстної оцінки критичності — система ефективно виявляє потенційні загрози й автоматично адаптується до змін у поведінці середовища.

Таблиця 2

Ввідображення основних метрик

Тип трафіку	**Precision**	**Recall**	**F1-score**	**Пояснення результату**
Нормальний трафік	0.98	0.94	0.96	Висока точність і повнота у розпізнаванні звичайної активності.
DoS атаки	0.94	0.97	0.95	Система ефективно виявляє перевантаження і надлишковий трафік.
Brute-force (SSH/HTTP)	0.91	0.88	0.89	Висока precision, але не всі випадки виявляються.
Port scanning	0.89	0.93	0.91	Завдяки агрегації частоти аномалій вдається добре ідентифікувати сканування.
Botnet-активність	0.86	0.79	0.82	Мережевий трафік ботів часто маскується під нормальний, однак модель демонструє прийнятні результати.
Phishing / DNS tunneling	0.83	0.76	0.79	Складні для виявлення через низький рівень відхилення, але система частково справляється.
Загальна середня оцінка	0.91	0.88	0.89	Інтегрована система забезпечує надійні результати у більшості сценаріїв.

На графіку(рис.2) зображено значення основних метрик якості класифікації — Precision, Recall та F1-score — для кожного типу мережевого трафіку, який розглядалася у дослідженні.

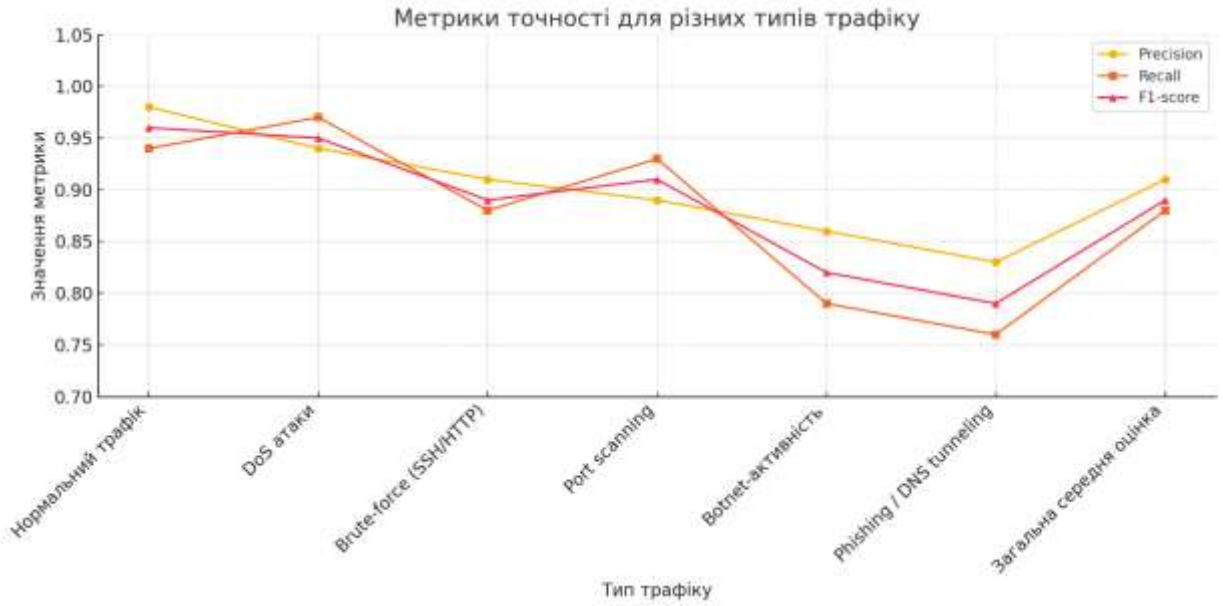

Рис. 2. Порівняльний аналіз метрик точності для різних типів мережевого трафіку

Зокрема:

- Нормальний трафік демонструє найвищі показники точності та повноти (Precision = 0.98, Recall = 0.94), що свідчить про високу здатність системи коректно ідентифікувати звичайну активність без помилкових спрацювань.
- DoS-атаки також добре виявляються системою, особливо за рахунок високої Recall (0.97), що є критично важливим для запобігання перевантаженням.
- Для Brute-force атак та Port scanning модель демонструє збалансовану ефективність, хоча деякі спроби можуть не виявлятись на ранніх етапах.
- Виявлення ботнет-активності і фішингових атак / DNS tunneling є менш точним через схожість таких атак із легітимним трафіком. Однак система все ж забезпечує прийнятну якість розпізнавання, що підтверджується F1-метрикою на рівні 0.79–0.82.
- Загальна середня оцінка (F1 = 0.89) свідчить про стабільну роботу інтегрованої системи на різних типах загроз.

Система продемонструвала здатність ефективно виявляти як очевидні, так і приховані аномалії, що підтверджує її доцільність для впровадження у середовища з підвищеними вимогами до інформаційної безпеки.

Структурно запропонована система включає шість основних модулів, тісно інтегрованих між собою: (1) попередню обробку даних (з нормалізацією та експоненційним згладжуванням), (2) модуль прогнозування (MLP), (3) модуль виявлення аномалій (автоенкодер), (4) блок адаптивного визначення меж аномальності, (5) оцінку критичності ситуації та (6) модуль реагування. Усі компоненти функціонують у реальному часі та забезпечують безперервну взаємодію як між собою, так і з зовнішніми інформаційними системами — такими як SIEM, IDS/IPS, засоби резервування або SDN-оркестратори.

В основі реалізації лежить наскрізна обробка мережевого трафіку. Система інтегрується в точку збору телеметрії — через NetFlow, SNMP, sFlow або спеціалізовані агенти, після чого дані передаються до модуля попередньої обробки. Там відбувається нормалізація показників, видалення шуму та застосування згладжування для посилення значущості актуальних змін. Далі нейромережевий модуль прогнозування (MLP) формує очікувану норму функціонування, яку паралельно аналізує автоенкодер, відтворюючи типовий шаблон поведінки. Якщо похибка реконструкції перевищує адаптивні межі, розраховані динамічно, система реєструє потенційно аномальну подію.

Оцінка критичності ситуації проводиться агреговано в межах часових вікон, а отриманий індекс визначає сценарій реагування — від простого сповіщення адміністратора до автоматичного переналаштування політик SDN-контролера, блокування сегментів або активації резервних каналів. Уся взаємодія відбувається у безперервному режимі, з можливістю зворотного навчання моделі на основі актуальних інцидентів.

Така структура дозволяє забезпечити прогностичну кіберстійкість навіть у динамічних та високонавантажених мережах, адаптуючись до змін у поведінці трафіку, нових типів загроз та умов експлуатації. У результаті, система не лише виявляє аномалії, але й активно протидіє їм, трансформуючи класичний підхід до кіберзахисту в сторону автономного ризик-менеджменту. Система придатна до використання в умовах змінного трафіку, мультисегментних топологій, у гібридних середовищах (онлайн/офлайн), та забезпечує не лише ретроспективну, але й прогностичну кіберстійкість. Це робить її перспективною для впровадження у великомасштабних інфраструктурах, зокрема в телекомі, енергетиці, фінансах, оборонному секторі.

Нижче представлено типову архітектуру взаємодії модулів(рис.3) у реальному мережевому середовищі:

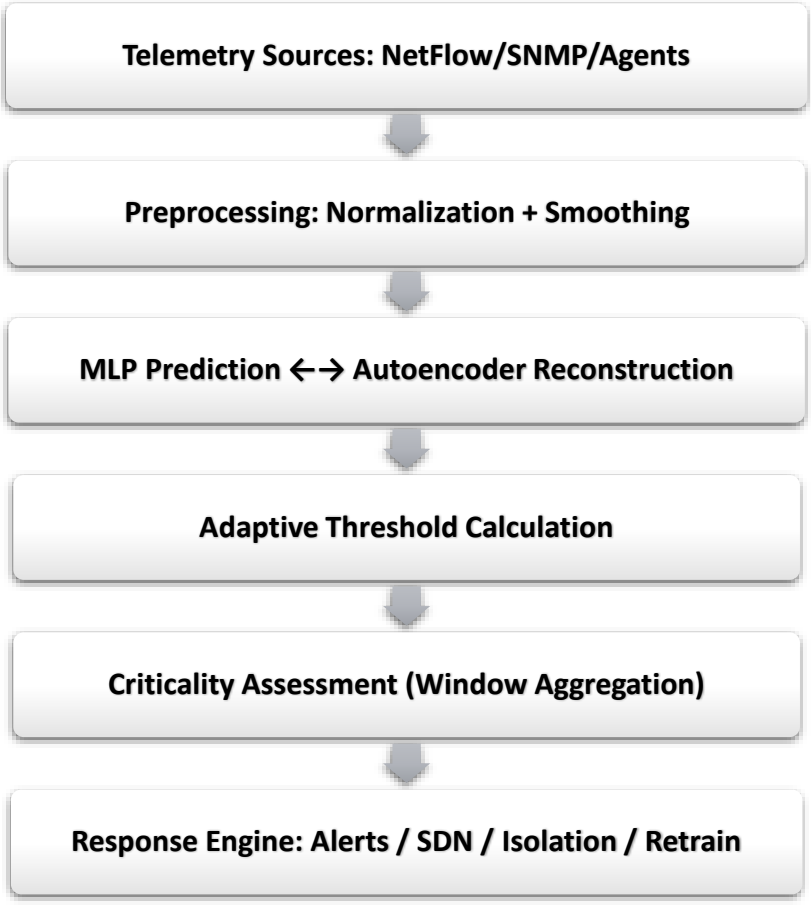

Рис. 3. Типову архітектура взаємодії модулів

Розроблена інтелектуальна система відіграє ключову роль не лише як засіб виявлення аномалій, а як повноцінна архітектура управління інформаційною безпекою. Її функціональність виходить далеко за межі традиційного моніторингу: вона поєднує в собі механізми збору телеметрії, прогнозування поведінки, динамічної оцінки ризиків, а також адаптивного реагування на виявлені загрози в реальному часі. Модульна архітектура дозволяє інтегрувати систему в широкий спектр кіберінфраструктур — від корпоративних мереж і дата-центрів до промислових систем SCADA та IoT-рішень. Вона підтримує обробку великих обсягів трафіку, адаптується до змін у шаблонах поведінки користувачів та інфраструктури, самонавчається на основі нових подій, а також формує зворотний зв'язок для автоматичної корекції порогів спрацювання. Завдяки використанню гібридного підходу, що поєднує автоенкодери, багатошарові нейронні мережі та модулі оцінки критичності ситуацій, забезпечується глибокий рівень розуміння контексту загроз.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Запропонована система поєднує прогнозування навантаження (MLP), поведінкове виявлення (autoencoder) та модуль семантичної атрибуції з XAI і ризик-скорингом R(t). Така інтеграція забезпечує роботу в реальному часі, зниження хибнопозитивних спрацьовувань і коректну пріоритезацію інцидентів. Емпіричні випробування на сценаріях DoS, сканувань портів, botnet-активності та прихованих загроз на кшталт DNS tunneling підтвердили стабільно високі значення Precision/Recall і загальний F1 (до ≈0,89), що свідчить про практичну придатність підходу. Агрегування подій у часових вікнах дозволяє відстежувати не лише поодинокі відхилення, а й стійкі патерни атак. Модульна архітектура спрощує масштабування обробки потоків і оновлення моделей без повного перенавчання, підтримує розгортання у SDN/NFV, IoT та хмарних середовищах і інтегрується з наявними SIEM/SOAR-процесами, дотримуючись сучасних вимог управління безпекою. Таким чином, система формує основу для автономного, адаптивного й проактивного кіберзахисту.

Подальші напрями розвитку включають: попереднє навчання на доменних вибірках і continual/online-learning для боротьби з concept drift; використання reinforcement learning для оптимізації політик реагування; підвищення стійкості до adversarial впливів і розширення бази прецедентів; стандартизацію профілів впровадження на рівні державних і корпоративних інфраструктур.

References

1. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys (CSUR), 41(3), 1–58.
2. Zhang, Y., Meratnia, N., & Havinga, P. (2010). Outlier detection techniques for wireless sensor networks: A survey. IEEE Communications Surveys & Tutorials, 12(2), 159–170.
3. Hodge, V. J., & Austin, J. (2004). A survey of outlier detection methodologies. Artificial Intelligence Review, 22(2), 85–126.
4. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690–1700.
5. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2019). Applying deep learning approaches for network traffic prediction and classification. Computer Communications, 141, 72–85.
6. Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: An ensemble of autoencoders for online network intrusion detection. Network and Distributed Systems Security (NDSS) Symposium.
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
8. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. IEEE Symposium on Security and Privacy, 305–316.
9. ISO/IEC 27001:2022. Information Security Management Systems — Requirements.
10. NIST Special Publication 800-94. Guide to Intrusion Detection and Prevention Systems (IDPS).