КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (33), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.33.1128
УДК 004.89:005.52

Школьніков Владислав Ігорович
доктор філософії в галузі права, доцент,
завідувач кафедри кримінології та інформаційних технологій
Національна академія внутрішніх справ, Київ, Україна
ORCID: 0000-0003-2041-9450
shkolnikov.v.i@navs.edu.ua

Лисов Богдан Сергійович
аспірант 2го року навчання,
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна
ORCID: 0009-0007-7963-6958
bogukraine@gmail.com

Халигов Артем Азимович
аспірант 3го року навчання,
Інститут телекомунікацій і глобального інформаційного простору
Національної академії наук України, Київ, Україна
0009-0006-5465-4650
khalygovartem@gmail.com

Гуськова Віра Геннадіївна
PhD, доцентка кафедри штучного інтелекту
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна
ORCID: 0000-0001-7637-201X
guskovavera2009@gmail.com

ГІБРИДНА АРХІТЕКТУРА СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ВИЯВЛЕННЯ ТА ОЦІНЮВАННЯ КІБЕРРИЗИКІВ В ОБ'ЄКТАХ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

Анотація. У статті досліджено проблему інтеграції нейромережевих методів виявлення аномалій із формалізованими механізмами оцінювання ризику та підтримки прийняття рішень у середовищі об'єктів критичної інфраструктури. У вступі обґрунтовано актуальність забезпечення кіберстійкості критичної інфраструктури в умовах зростання складності атак, розширення поверхні атаки та обмеженості традиційних систем. У розділі постановки проблеми визначено ключові обмеження існуючих нейромережевих підходів, зокрема відсутність контекстуалізації, ризик-орієнтованої інтерпретації та механізмів автоматизованого формування рекомендацій. В аналітичному огляді літератури систематизовано сучасні підходи до LSTM-, autoencoder- та transformer-based детекції аномалій, а також методи динамічного оцінювання кіберризику. Сформульовано мету дослідження – розроблення гібридної архітектури, що поєднує поведінкову детекцію, агрегування anomaly score у інтегральний показник ризику та формування сценаріїв реагування. Запропонована багаторівнева архітектура включає Data Layer, Neural Detection Layer (LSTM Autoencoder), Risk Aggregation Layer (top-k агрегування з урахуванням коефіцієнта критичності активу) та Decision Support Layer з пороговою моделлю рекомендацій. Формалізовано обчислення anomaly score, інтегрального ризику та функції управлінських дій. Експериментальна перевірка виконана на датасеті BETH із застосуванням двох режимів оцінювання: Normal-only та Mixed. У режимі Mixed отримано ROC-AUC = 0.874 та PR-AUC = 0.828 на рівні часових вікон, Session ROC-AUC = 0.8235 після агрегування ризику. Показник Action Precision = 0.9333 підтверджує ефективність механізму пріоритизації інцидентів. Низька латентність (~0.35 мс) засвідчує придатність до застосування, близького до реального часу. У висновках доведено, що інтеграція нейромережевого детектора з ризик-орієнтованою моделлю підтримки прийняття рішень

© В. І. Школьніков, Б. С. Лисов, А. А. Халигов, В. Г. Гуськова, 2026

забезпечує підвищення інтерпретованості результатів, зменшення хибних ескалацій та адаптивність до різних режимів даних.

Ключові слова: критична інфраструктура; кіберризик; LSTM Autoencoder; anomaly score; агрегування ризику; система підтримки прийняття рішень.

ВСТУП

Критична інфраструктура (КІ) охоплює системи та об'єкти, функціонування яких є важливими для національної безпеки, економічної стабільності та суспільної життєдіяльності. До них належать енергетичні, транспортні, інформаційні, фінансові та медичні системи, що характеризуються високим рівнем цифровізації та взаємозалежності. Порушення роботи одного елемента може спричиняти каскадні ефекти в суміжних секторах.

Інтеграція кіберфізичних систем, промислових мереж (SCADA/ICS), хмарних технологій та IoT-рішень суттєво розширила поверхню атаки. Сучасні кіберзагрози набувають складного багаторівневого характеру, використовують механізми прихованого проникнення та імітацію легітимної поведінки, що ускладнює їх своєчасне виявлення.

Традиційні системи виявлення вторгнень (Intrusion Detection Systems, IDS) є обмеженими в умовах постійної еволюції атак та появи раніше невідомих загроз. У зв'язку з цим дедалі більшого значення набувають методи поведінкового аналізу та машинного навчання, зокрема нейромережеві моделі, здатні виявляти аномальні патерни у часових послідовностях подій.

Разом із тим, для об'єктів критичної інфраструктури недостатньо лише детекції аномалій. Необхідні механізми оцінювання ризику, пріоритизації інцидентів та формування рекомендацій, що дозволяють інтегрувати результати аналізу у систему підтримки прийняття рішень та підвищити рівень кіберстійкості.

Постановка проблеми. Сучасні нейромережеві підходи до виявлення аномалій у кібербезпекових системах демонструють високу ефективність у задачах класифікації та реконструкції поведінкових патернів. Моделі на основі автоенкодерів, LSTM або трансформерів здатні ідентифікувати відхилення від нормальної поведінки шляхом аналізу реконструкційної похибки або латентних представлень послідовностей подій. Проте більшість таких підходів мають низку концептуальних обмежень, що є критичними для застосування в середовищах об'єктів критичної інфраструктури.

По-перше, нейромережеві моделі переважно орієнтовані на виявлення аномалій як статистичних відхилень, не враховуючи операційний контекст активу. Однакова величина показника аномальності може мати різні наслідки залежно від типу системи (енергетичний вузол, транспортний сегмент, IT-інфраструктура) та рівня її критичності. Таким чином, відсутність механізму контекстуалізації знижує практичну релевантність результатів.

По-друге, існуючі рішення, як правило, не трансформують показники аномальності у ризик-орієнтовану інтерпретацію, придатну для прийняття управлінських рішень. Аномалія не є тотожною ризику: ризик передбачає оцінку потенційних наслідків, ймовірності розвитку інциденту та впливу на функціонування об'єкта. Відсутність формалізованого механізму переходу від детекції до оцінювання ризику створює розрив між аналітичним модулем і системою підтримки прийняття рішень.

По-третє, більшість нейромережевих інформаційних систем функціонують як ізольовані модулі детекції, не інтегровані з механізмами автоматизованого формування рекомендацій. У результаті оператори систем безпеки змушені самостійно інтерпретувати числові показники та визначати подальші дії, що збільшує когнітивне навантаження та час реагування.

По-четверте, відсутня цілісна архітектура, яка б поєднувала поведінкове моделювання; агрегування аномальних сигналів у стійку оцінку ризику; урахування критичності активів; генерацію сценаріїв реагування. У середовищах критичної інфраструктури така інтеграція є принципово важливою, оскільки навіть локальна аномалія може мати системні наслідки через міжсекторальні залежності та каскадні ефекти.

Отже, виникає науково-практична проблема розроблення інтегрованої нейромережевої моделі виявлення аномалій з механізмом формалізованого оцінювання ризику та формування рекомендацій, адаптованої до умов функціонування об'єктів критичної інфраструктури.

Аналіз останніх досліджень і публікацій. Сучасні дослідження у сфері кібербезпеки активно застосовують нейромережеві підходи для виявлення аномалій у логах та мережевих подіях. Malhotra et al. [1] запропонували LSTM-encoder–decoder для детекції відхилень у часових рядах, а Du et al. [2] у моделі DeepLog використали LSTM для прогнозування послідовностей системних подій. Подальший розвиток цього напряму пов'язаний із застосуванням transformer-архітектур: Chourasiya et al. [3] поєднали LSTM і Transformer для аналізу логів, Nasirzadeh et al. [4] запропонували модель CoLog для виявлення колективних аномалій, а Kummerow et al. [5] реалізували transformer-based autoencoder з механізмами пояснюваності.

Паралельно розвиваються ризик-орієнтовані підходи. Poolsappasit et al. [6] застосували байєсівські графи атак для динамічного оцінювання ризику, а Feng et al. [7] формалізували ризик як функцію ймовірності інциденту та його впливу. Також було розглянуто задачі пріоритизації та кореляції сповіщень у IDS, проте без повної інтеграції з глибокими моделями детекції. У сфері підтримки прийняття рішень для критичної інфраструктури досліджуються інтелектуальні DSS, тоді як [4] демонструють застосування ML у кіберфізичних системах.

Отже, література демонструє значний прогрес у нейромережевій детекції, оцінюванні ризику та автоматизованому реагуванні окремо. Водночас інтеграція цих компонентів у єдину архітектуру для критичної інфраструктури залишається недостатньо дослідженою.

Метою статті є розроблення та наукове обґрунтування гібридного підходу, що дає змогу інтегрувати поведінкове виявлення аномалій із формалізованим механізмом оцінювання кіберризику та автоматизованим формуванням рекомендацій для об’єктів критичної інфраструктури. Запропонований підхід спрямований на забезпечення переходу від ізольованої детекції відхилень до контекстно-залежної risk-aware моделі підтримки прийняття рішень, яка враховує критичність активів, агрегує аномальні сигнали у стійку оцінку ризику та формує управлінські дії відповідно до рівня загрози. Реалізація цієї мети передбачає побудову нейромережевої моделі аналізу подій, розроблення алгоритму агрегування anomaly score у ризиковий показник та інтеграцію отриманих результатів у систему рекомендацій для підвищення рівня кіберстійкості критичної інфраструктури.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Запропонований підхід реалізується у вигляді багаторівневої архітектури, що забезпечує послідовний перехід від аналізу подій до формування управлінських рішень. Архітектура складається з чотирьох взаємопов’язаних рівнів: підготовки даних, нейромережевого виявлення аномалій, агрегування ризику та формування рекомендацій.

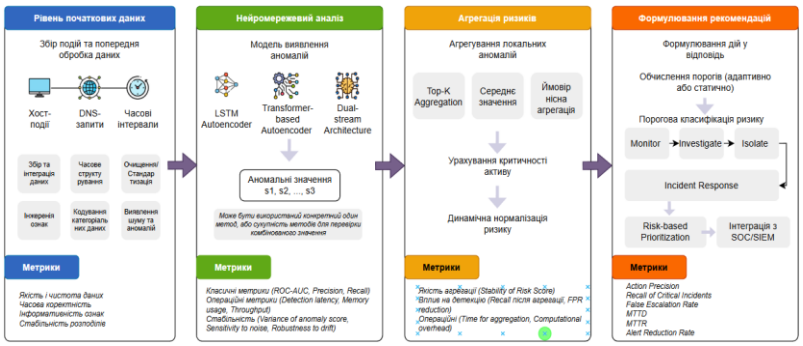

Рис. 1. Архітектура гібридної системи підтримки прийняття рішень

На першому рівні здійснюється формування та структурування вхідних даних. Події організовуються у часові послідовності або сесії, виконується нормалізація числових ознак і кодування категоріальних характеристик. Метою цього етапу є побудова уніфікованого вхідного представлення для подальшого моделювання.

Другий рівень реалізує нейромережеву модель поведінкового аналізу. Послідовності подій подаються на вхід архітектури автоенкодера, яка навчається відтворювати нормальні патерни функціонування системи. Відхилення від очікуваної реконструкції інтерпретуються як локальні аномалії.

Третій рівень передбачає агрегування локальних показників аномальності у стійку оцінку ризику. Для цього використовується механізм селективного узагальнення найбільш значущих відхилень з урахуванням коефіцієнта критичності активу.

Четвертий рівень трансформує кількісну оцінку ризику у набір рекомендацій, що відповідають різним сценаріям реагування. Таким чином, забезпечується інтеграція аналітичного модуля з системою підтримки прийняття рішень.

МЕТОДИКА ДОСЛІДЖЕННЯ

Набір даних для експерименту. Для експериментальної перевірки запропонованої гібридної архітектури використовується відкритий датасет BETH (Kaggle), що містить часові журнали подій рівня операційної системи (таблиця 1). Набір даних відображає поведінку процесів у середовищі хоста та призначений для дослідження задач поведінкового моделювання та виявлення аномалій. Датасет

складається з послідовностей системних подій, кожна з яких характеризується часовою міткою та набором атрибутів, що описують виконання процесів та системні виклики. Така структура дозволяє формувати поведінкові профілі вузла та аналізувати відхилення у часовій динаміці.

Таблиця 1

Характеристику набору даних BETH

Поле	Опис
timestamp	Часова мітка події
processId	Ідентифікатор процесу
parentProcessId	Ідентифікатор батьківського процесу
threadId	Ідентифікатор потоку
userId	Ідентифікатор користувача
processName	Назва процесу
eventId	Ідентифікатор системного виклику
eventName	Назва події/системного виклику
argsNum	Кількість аргументів виклику
returnValue	Код повернення
args	Параметри системного виклику
sus	Ознака підозрілої активності
evil	Ознака шкідливої активності

У межах роботи датасет використовується для реалізації та тестування чотирирівневої архітектури:

- *Data Layer* – формування часових послідовностей та інженерія ознак;
- *Neural Detection Layer* – обчислення anomaly score на основі поведінкового моделювання;
- *Risk Aggregation Layer* – агрегування локальних відхилень у інтегральний ризик;
- *Decision Support Layer* – відображення ризику у керовані управлінські рішення.

Формалізоване представлення. Нехай представлена послідовність подій, де кожен x_i є вектором ознак системного виклику.

$$X = \{x_1, x_2, \ldots, x_T\} \quad (1)$$

Завдання полягає у обчисленні локального показника аномальності s_i, агрегуванні його у ризик R та побудові функції рекомендацій $A(R)$.

Нейромережевий аналіз.

На першому етапі реалізується нейромережевий модуль виявлення аномалій, метою якого є моделювання нормальної поведінки системи та визначення відхилень у часових послідовностях подій. З огляду на те, що події в середовищах критичної інфраструктури формують складні темпоральні залежності (послідовності системних викликів, поведінкові патерни процесів, взаємодію користувачів і сервісів), класичні статичні методи класифікації є недостатніми.

Нейромережевий модуль функціонує у режимі unsupervised / semi-supervised anomaly detection, де модель навчається виключно на нормальних даних (evil = 0), формуючи внутрішній профіль типової поведінки системи. Будь-яке суттєве відхилення від цього профілю інтерпретується як потенційна аномалія.

LSTM Autoencoder. Для моделювання часових залежностей використовується автоенкодер на основі рекурентної нейронної мережі типу LSTM. Нехай:

$$X = \{x_1, x_2, \ldots, x_T\} \quad (2)$$

– послідовність подій для певного процесу або користувача. Модель складається з Encoder, який стискає послідовність у латентне представлення z та Decoder, який відновлює послідовність $\hat{X}$. Показник аномальності обчислюється як реконструкційна похибка:

$$s_i = ||x_i - \hat{x}_i|| \quad (3)$$

Для сесії обчислюється узагальнений anomaly score:

$$S = \frac{1}{T}\sum_{i=1}^{T} s_i \quad (4)$$

Навчання здійснюється на нормальних подіях (evil = 0), після чого модель тестується на повному наборі даних.

Dual-Stream Architecture (Host). Для підвищення точності детекції може бути використана двопотокова архітектура, у якій перший потік обробляє host-level події, а інший потік - dns-level події. Кожен потік має окремий encoder, після чого латентні представлення об'єднуються:

$$z = f(z_{host}, z_{dns}) \tag{5}$$

Об'єднаний латентний простір використовується для обчислення інтегрального anomaly score. Перевага такого підходу полягає у здатності враховувати як локальну системну поведінку, так і зовнішню мережеву активність.

Модель агрегування ризику. Оскільки окремі аномалії можуть мати випадковий або шумовий характер, наступним етапом є агрегування локальних показників аномальності у стійку ризик-орієнтовану оцінку. Нейромережевий модуль формує числові значення anomaly score для кожної події або часової сесії, однак ізольоване відхилення не завжди свідчить про реальну загрозу. Тому застосовується механізм узагальнення найбільш значущих відхилень у межах часових вікон, що дозволяє зменшити вплив статистичного шуму та підвищити стабільність оцінки.

Додатково враховується критичність активу, оскільки однакова за величиною аномалія може мати різний рівень небезпеки залежно від функціональної ролі об'єкта в системі критичної інфраструктури. У результаті формується інтегральний показник ризику, який поєднує інтенсивність аномальної поведінки та значущість активу. Такий підхід забезпечує перехід від статистичної детекції відхилень до контекстно-залежної оцінки ризику, придатної для подальшої інтеграції у систему підтримки прийняття рішень. Нехай

$$S = \{s_1, s_2, \ldots, s_T\} \tag{6}$$

– множина локальних anomaly score для сесії.

Top-k Aggregation. Для зменшення впливу шуму застосовується:

$$R_{session} = mean(topK(s_i)) \tag{7}$$

де обираються k найбільших значень.

Урахування критичності активу. Для систем критичної інфраструктури вводиться коефіцієнт критичності $C \in [0,1]$. Фінальний ризик визначається як:

$$R_{final} = R_{session}(1 + \alpha C) \tag{8}$$

де α – ваговий коефіцієнт. Таким чином однакова аномалія може мати різну вагу залежно від важливості активу.

Рівень рекомендацій. Останній етап трансформує інтегральний показник ризику у конкретні управлінські дії, орієнтовані на мінімізацію потенційних наслідків інциденту. На цьому рівні кількісна оцінка ризику, отримана після агрегування аномальних сигналів та врахування критичності активу, інтерпретується у вигляді сценаріїв реагування. Таким чином забезпечується перехід від аналітичної оцінки до операційного управління. Механізм рекомендацій може базуватися на пороговій моделі, де різні діапазони ризику відповідають різним рівням реагування (моніторинг, поглиблена перевірка, ізоляція вузла, ініціювання процедури реагування на інцидент). Пороги можуть бути статичними або адаптивними, залежно від історичних даних, профілю активу та поточного навантаження системи. Додатково може формуватися індекс пріоритетності, який враховує як величину ризику, так і системну важливість активу, що дозволяє оптимізувати черговість обробки інцидентів у середовищі SOC або SIEM.

Реалізується порогова модель, де A(R) визначається у наступному діапазоні:

$$\begin{gathered} Monitor, R < \tau_1 \\ Investigate, \tau_1 < R < \tau_2 \\ Isolate, \tau_2 < R < \tau_3 \\ Incident\ Response, \geq \tau_3 \end{gathered} \tag{9}$$

Пороги можуть бути статичними та адаптивними (на основі $\mu R + \beta\sigma R$). Додатково формується показник пріоритетності:

$$P_{priority} = R_{final} \cdot Impact_{sector} \tag{10}$$

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Експерименти. Після формування теоретичних засад та опису запропонованої архітектури було здійснено покрокову реалізацію експериментального дослідження. На першому етапі виконано підготовку та попередню обробку даних, формування часових послідовностей і побудову наборів для навчання та тестування. Далі реалізовано нейромережевий модуль виявлення аномалій, проведено його навчання на нормальних подіях та обчислено показники якості детекції.

На наступному кроці локальні anomaly score було агреговано у інтегральні показники ризику на рівні сесій, після чого застосовано механізм формування управлінських рекомендацій. Завершальним етапом стало порівняння результатів у різних режимах даних (за наявності та відсутності шкідливих подій), що дозволило оцінити адаптивність і практичну придатність запропонованого підходу.

Крок 1. Підготовка даних та перевірка придатності. На першому етапі завантажуються один або декілька CSV-файлів host-level датасету BETH. Виконується первинний контроль якості даних:

- перевірка наявності ключових колонок (timestamp, processId, eventId, argsNum, returnValue, evil);
- перевірка коректності часових міток;
- обчислення частки шкідливих подій evil_rate=mean(evil)

Залежно від результату визначається режим експерименту:

- Режим A (Normal-only) – якщо evil_rate=0;
- Режим B (Mixed) – якщо evil_rate>0.

Цей крок є критично важливим для коректної інтерпретації метрик, оскільки показники ROC-AUC та PR-AUC можуть бути обчислені лише за наявності обох класів.

Крок 2. Формування ознак та врахування часової структури.

Формуються такі групи ознак:

- числові: eventId, argsNum, returnValue;
- часові: dt=Δtimestamp та ts_mod=timestamp mod60;
- категоріальні (за потреби): хешоване кодування processName, eventName.

Дані впорядковуються за часовими мітками для збереження послідовності подій.

Крок 3. Формування часових послідовностей (sliding windows).

Для кожної сутності (наприклад, processId) формується набір послідовностей фіксованої довжини W зі зсувом S. Кожне вікно представляє собою послідовність X={x1,x2,…,xW}. Мітка вікна визначається як $ywindow$=max($evili$ у межах вікна). Такий підхід дозволяє застосовувати рекурентні нейронні мережі для моделювання поведінки у часі.

Крок 4. Навчання Neural Detection Layer (LSTM Autoencoder).

Модель LSTM Autoencoder навчається виключно на нормальних вікнах (ywindow=0).

Архітектура складається з Encoder (LSTM), що формує латентне представлення; Decoder, що відновлює початкову послідовність. Аномальність оцінюється через реконструкційну похибку:

$$S = \frac{1}{W}\sum_{i=1}^{W} \quad ||x_i - \hat{x}_i|| \tag{11}$$

Крок 5. Оцінювання якості детекції (два режими).

Режим A: відсутність шкідливих подій (Normal-only)

У випадку, коли у вибірці відсутні події з evil=1, метрики класифікації не обчислюються. Натомість виконується sanity-check у вигляді разрахунку середнього значення anomaly score; дисперсії score; частки тривог при фіксованому порозі; латентності (часу обробки одного вікна). Це дозволяє перевірити стабільність та операційну придатність моделі.

Режим B: наявність шкідливих подій (Mixed) з evil=1. У цьому режимі обчислюються ROC-AUC; PR-AUC; Recall@FPR=1% (строгий режим мінімізації хибних спрацювань); латентність.

Крок 6. Модель агрегування ризику (Risk Aggregation Layer).

Локальні anomaly score агрегуються на рівні сесії або процесу. Застосовується стратегія top-k: Rsession=mean(top-k(s)). За необхідності враховується критичність активу: Rfinal=Rsession(1+αC), де C – коефіцієнт критичності активу.

Крок 7. Decision Support Layer (формування рекомендацій).

На основі інтегрального ризику формується рекомендація: Monitor or Investigate or Isolate or Incident Response. Оцінюються такі показники:

- Action Precision – точність критичних ескалацій;
- Recall Critical – частка атак, переведених у критичний режим;
- False Escalation Rate – рівень хибної ескалації;
- Underreaction Rate – рівень недореагування.

Таблиці результатів. Таблиця 2 відображає логіку адаптивного експериментального протоколу залежно від структури вхідних даних. Виділено два основні режими: Normal-only, коли у вибірці відсутні події з міткою evil=1, та Mixed, коли дані містять як нормальні, так і шкідливі події. У першому випадку класичні метрики класифікації (ROC-AUC, PR-AUC) не можуть бути коректно обчислені через відсутність позитивного класу, тому оцінювання зосереджується на перевірці стабільності детектора, частці потенційних тривог і операційній придатності моделі (sanity-check). У другому режимі здійснюється повноцінна кількісна оцінка якості детекції та ефективності рівня підтримки прийняття рішень. Такий поділ забезпечує методологічну коректність аналізу та дозволяє застосовувати запропонований підхід незалежно від наявності розмічених атак у даних.

Таблиця 2

Режими даних та доступні метрики

Режим даних	Наявність evil	ROC/PR метрики	Основна мета
Normal-only	Ні	Ні	Sanity-check, стабільність
Mixed	Так	Так	Повна оцінка детекції та рішень

У Таблиці 3а наведено результати роботи нейромережевого детектора в режимі Normal-only, тобто за відсутності подій з міткою evil=1. У цьому випадку класичні метрики класифікації (ROC-AUC, PR-AUC, Recall@FPR) не визначаються, оскільки позитивний клас відсутній. Натомість основна увага приділяється операційним характеристикам моделі: швидкості обробки (Latency) та частці потенційних тривог при заданому пороговому значенні. Такий підхід дозволяє виконати sanity-check експерименту, перевірити стабільність реконструкційної похибки та оцінити схильність моделі до генерації хибних аномалій на повністю нормальному трафіку. Отримані результати підтверджують коректність навчання автоенкодера та його придатність до використання у середовищах без явної розмітки атак.

Таблиця 3а

Результати детекції (window-level). Режим A (без атак)

Модель	ROC_AUC	PR_AUC	Recall@FPR=1%	Latency	Частка тривог
LSTM_AE	—	—	—	обчислюється	оцінюється

У Таблиці 3б представлено результати оцінювання нейромережевого детектора в режимі Mixed, коли у вибірці наявні як нормальні, так і шкідливі події. Модель LSTM Autoencoder продемонструвала високий рівень розділення класів на рівні часових вікон: ROC-AUC = 0.874 та PR-AUC = 0.828. Це свідчить про здатність моделі ефективно відокремлювати аномальні поведінкові послідовності від нормальних. Показник Recall@FPR=1% = 0.103 відображає здатність виявляти частину атак у строгому режимі обмеження хибних спрацювань, що є важливим для систем кіберзахисту критичної інфраструктури. Низька латентність (~0.35 мс на вікно) підтверджує можливість використання підходу у режимі близькому до реального часу.

Таблиця 3б

Результати детекції. Режим B (з атаками)

Модель	ROC_AUC	PR_AUC	Recall@FPR=1%	Latency
LSTM_AE	0.874	0.828	0.103	~0.35 ms

Отримані значення ROC-AUC та PR-AUC свідчать про високу якість поведінкової детекції аномалій. Показник Recall@FPR=1% демонструє компроміс між чутливістю та мінімізацією хибних тривог у строгому операційному режимі. Загалом, результати підтверджують ефективність LSTM Autoencoder як базового модуля виявлення аномалій у запропонованій гібридній архітектурі.

На рисунку представлено ROC-криву (Receiver Operating Characteristic), що відображає залежність між часткою істинно-позитивних спрацювань (True Positive Rate, TPR) та часткою хибно-позитивних спрацювань (False Positive Rate, FPR) для моделі LSTM Autoencoder на рівні часових вікон. Площа під кривою (AUC = 0.874) свідчить про високий рівень розділення нормальних та шкідливих послідовностей.

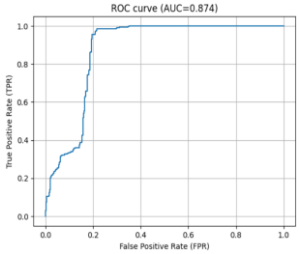

Рис. 2. ROC-крива залежності стинно-позитивних та хибно-позитивних спрацювань

Крива має виражений вигин у верхньому лівому куті, що означає здатність моделі досягати високого значення TPR при відносно низькому FPR. Зокрема, при FPR ≈ 0.2 модель вже забезпечує TPR, близький до 1.0, що демонструє ефективне відокремлення атак від нормальної поведінки.

Отримане значення AUC значно перевищує 0.5 (випадковий класифікатор), що підтверджує доцільність застосування LSTM Autoencoder для поведінкового виявлення аномалій у задачах кібербезпеки. Результат узгоджується з кількісними показниками, наведеними у таблиці результатів, і підтверджує стабільність детекційного шару запропонованої гібридної архітектури.

На рисунку 3 представлено криву Precision-Recall для моделі LSTM Autoencoder на рівні часових вікон. Площа під кривою (AUC = 0.821) свідчить про високий рівень точності виявлення аномалій в умовах дисбалансу класів, що є характерним для задач кібербезпеки.

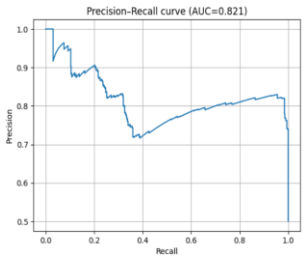

Рис. 3. Precision-Recall крива на рівні часових вікон

Графік демонструє, що при невеликих значеннях recall модель забезпечує дуже високий precision (близький до 1), тобто більшість виявлених аномалій є справді шкідливими. Із зростанням recall, точність поступово знижується, що відображає типовий компроміс між повнотою виявлення атак та кількістю хибних спрацювань. Отримані результати підтверджують здатність моделі ефективно працювати в умовах значного класового дисбалансу та забезпечувати надійну поведінкову детекцію.

У Таблиці 4 наведено результати оцінювання запропонованої архітектури після застосування механізму агрегування ризику та формування управлінських рішень на рівні сесій. Значення Session ROC-AUC = 0.8235 підтверджує, що агрегування локальних anomaly score (top-k стратегія) зберігає високу здатність до розділення нормальних і шкідливих сесій та зменшує вплив шумових відхилень на рівні окремих вікон.

Таблиця 4

Risk + Decision (session-level)

Модель	Session ROC-AUC	Action Precision	Recall Critical	False Escalation	Underreaction
LSTM_AE	0.8235	0.9333	0.8235	0.0556	0.1687

Показник Action Precision = 0.9333 свідчить про те, що більшість критичних ескалацій (Isolate / Incident Response) є обґрунтованими, що мінімізує кількість необґрунтованих втручань. Значення Recall Critical = 0.8235 демонструє, що понад 82% атак коректно переводяться у критичний режим реагування. Водночас False Escalation = 0.0556 відображає низький рівень хибних критичних рішень, а Underreaction = 0.1687 характеризує частку атак, які не були ескаловані до найвищого рівня реагування. Сукупність цих показників підтверджує ефективність інтеграції нейромережевого детектора з ризик-орієнтованим механізмом підтримки прийняття рішень.

Запропонований експериментальний протокол є адаптивним і коректно працює у двох режимах – як за відсутності, так і за наявності шкідливих подій. У режимі Normal-only система виконує sanity-check та підтверджує стабільність детектора без генерації неконтрольованої кількості хибних тривог.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У межах проведеного дослідження було реалізовано та експериментально перевірено гібридну архітектуру, орієнтовану на виявлення аномалій та підтримку прийняття рішень у середовищі критичної інфраструктури. Запропонований підхід поєднує поведінкове моделювання на основі LSTM Autoencoder, механізм агрегування локальних аномалій у інтегральний показник ризику та порогову модель формування управлінських рекомендацій.

Отримані результати підтверджують ефективність запропонованої архітектури на декількох рівнях. На рівні часових вікон модель продемонструвала високу здатність до розділення нормальних та шкідливих послідовностей (ROC-AUC = 0.874; PR-AUC = 0.828), що свідчить про адекватність обраного підходу до поведінкового аналізу. Перехід до рівня сесій через механізм top-k агрегування дозволив зберегти якість детекції (Session ROC-AUC = 0.8235) та водночас підвищити інтерпретованість результатів.

Особливо важливим є те, що Decision Support Layer забезпечив високий рівень точності критичних ескалацій (Action Precision = 0.9333) при збереженні значної частки виявлених атак (Recall Critical = 0.8235). Це свідчить про те, що інтеграція нейромережевого детектора з ризик-орієнтованим механізмом прийняття рішень дозволяє зменшити кількість необґрунтованих критичних тривог і водночас не втрачати значну частину реальних інцидентів.

Додатково підтверджено адаптивність експериментального протоколу. У випадку відсутності шкідливих подій у вибірці система коректно переходить до режиму sanity-check, що дозволяє оцінити стабільність детектора, рівень потенційних хибних ескалацій та операційну придатність моделі без спотворення метрик класифікації. Таким чином, запропонований підхід є придатним як для повністю розмічених, так і для частково або повністю нерозмічених даних.

Низька латентність обчислення (~0.35 мс на вікно) демонструє можливість застосування розробленої архітектури у near real-time сценаріях моніторингу, що є критично важливим для об’єктів критичної інфраструктури.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2016). LSTM-based encoder–decoder for multi-sensor anomaly detection. *arXiv*. https://arxiv.org/abs/1607.00148
2. Du, M., Li, F., Zheng, G., & Srikumar, V. (2017). DeepLog: Anomaly detection and diagnosis from system logs through deep learning. https://www.researchgate.net/publication/320678676_DeepLog_Anomaly_Detection_and_Diagnosis_from_System_Logs_through_Deep_Learning
3. Chourasiya, V., Kumar, A., & Singh, P. (2025). Advanced system log analyzer for anomaly detection and cyber forensic investigations using LSTM and transformer networks. https://www.researchgate.net/publication/396999897_Advanced_system_log_analyzer_for_anomaly_detection_and_cyber_forensic_investigations_using_LSTM_and_transformer_networks
4. Nasirzadeh, M., Tahmoresnezhad, J., & Rashidi-Khazaee, P. (2025). A unified framework for detecting point and collective anomalies in operating system logs via collaborative transformers. *Scientific Reports, 15*, Article 45698. https://doi.org/10.1038/s41598-025-27693-4
5. Kummerow, M., Müller, T., & Freiling, F. (2024). Explainable transformer-based autoencoders for anomaly detection. *arXiv*. https://arxiv.org/abs/2404.06517
6. Poolsappasit, N., Dewri, R., & Ray, I. (2012). Dynamic security risk management using Bayesian attack graphs.https://www.researchgate.net/publication/224244206_Dynamic_Security_Risk_Management_Using_Bayesian_Attack_Graphs
7. Feng, C., Li, T., & Chana, I. (2018). Multi-level anomaly detection in industrial control systems via deep learning. *IEEE Access, 6*, 701–715. https://doi.org/10.1109/ACCESS.2017.2784915
8. Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2016). Long short-term memory networks for anomaly detection in time series. In *Proceedings of the 23rd European Symposium on Artificial Neural Networks (ESANN 2016)* (pp. 89-94).
9. Poolsappasit, N., Dewri, R., & Ray, I. (2012). Dynamic security risk management using Bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing, 9*(1), 61-74. https://doi.org/10.1109/TDSC.2011.34
10. Saha, S., Deb, S., & Das, S. (2020). Machine learning-based intrusion detection system for critical infrastructure protection. *Future Generation Computer Systems, 108*, 121-134. https://doi.org/10.1016/j.future.2020.02.049

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (33), 2026
ISSN 2663 – 4023

Vladyslav Shkolnikov
Doctor of Philosophy in Law, Associate Professor,
Head of the Department of Criminology and Information Technologies
National Academy of Internal Affairs, Kyiv, Ukraine
ORCID: 0000-0003-2041-9450
shkolnikov.v.i@navs.edu.ua

Bohdan Lysov
2nd-year PhD student
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
ORCID: 0009-0007-7963-6958
bogukraine@gmail.com

Artem Khalygov
3rd-year PhD student
Institute of Telecommunications and Global Information Space
National Academy of Sciences of Ukraine, Kyiv, Ukraine
ORCID: 0009-0006-5465-4650
khalygovartem@gmail.com

Vira Huskova
PhD, Associate Professor of the Department of Artificial Intelligence
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
ORCID: 0000-0001-7637-201X
guskovavera2009@gmail.com

HYBRID ARCHITECTURE OF A DECISION SUPPORT SYSTEM FOR DETECTION AND ASSESSMENT OF CYBER RISKS IN CRITICAL INFRASTRUCTURE OBJECTS

Abstract. The article investigates the problem of integrating neural network-based anomaly detection methods with formalized risk assessment mechanisms and decision support models within critical infrastructure environments. The introduction substantiates the relevance of ensuring cyber resilience of critical infrastructure under conditions of increasing attack complexity, expanding attack surfaces, and the limitations of traditional security systems. The problem statement identifies key limitations of existing neural network approaches, including the lack of contextualization, risk-oriented interpretation, and automated recommendation generation mechanisms. The literature review systematizes contemporary approaches to LSTM-, autoencoder-, and transformer-based anomaly detection, as well as methods for dynamic cyber risk assessment. The research objective is formulated as the development of a hybrid architecture that integrates behavioral anomaly detection, aggregation of anomaly scores into an integral risk indicator, and automated response scenario generation. The proposed multi-layer architecture includes a Data Layer, Neural Detection Layer (LSTM Autoencoder), Risk Aggregation Layer (top-k aggregation with asset criticality coefficient), and a Decision Support Layer with a threshold-based recommendation model. The computation of anomaly scores, integral risk, and management action functions is formally defined. Experimental validation was conducted on the BETH dataset using two evaluation modes: Normal-only and Mixed. In the Mixed mode, the model achieved ROC-AUC = 0.874 and PR-AUC = 0.828 at the window level, and Session ROC-AUC = 0.8235 after risk aggregation. The Action Precision metric of 0.9333 confirms the effectiveness of the incident prioritization mechanism. Low latency (~0.35 ms) demonstrates suitability for near real-time application. The conclusions demonstrate that integrating a neural anomaly detector with a risk-oriented decision support model improves interpretability, reduces false escalations, and ensures adaptability to different data regimes.

Keywords: critical infrastructure; cyber risk; LSTM Autoencoder; anomaly score; risk aggregation; decision support system.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2016). LSTM-based encoder–decoder for multi-sensor anomaly detection. *arXiv*. https://arxiv.org/abs/1607.00148
2. Du, M., Li, F., Zheng, G., & Srikumar, V. (2017). DeepLog: Anomaly detection and diagnosis from system logs through deep learning.
3. Chourasiya, V., Kumar, A., & Singh, P. (2025). Advanced system log analyzer for anomaly detection and cyber forensic investigations using LSTM and transformer networks.
4. Nasirzadeh, M., Tahmoresnezhad, J., & Rashidi-Khazaee, P. (2025). A unified framework for detecting point and collective anomalies in operating system logs via collaborative transformers. *Scientific Reports, 15*, Article 45698. https://doi.org/10.1038/s41598-025-27693-4
5. Kummerow, M., Müller, T., & Freiling, F. (2024). Explainable transformer-based autoencoders for anomaly detection. *arXiv*. https://arxiv.org/abs/2404.06517
6. Poolsappasit, N., Dewri, R., & Ray, I. (2012). Dynamic security risk management using Bayesian attack graphs.
7. Feng, C., Li, T., & Chana, I. (2018). Multi-level anomaly detection in industrial control systems via deep learning. *IEEE Access, 6*, 701–715. https://doi.org/10.1109/ACCESS.2017.2784915
8. Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2016). Long short-term memory networks for anomaly detection in time series. In *Proceedings of the 23rd European Symposium on Artificial Neural Networks (ESANN 2016)* (pp. 89-94).
9. Poolsappasit, N., Dewri, R., & Ray, I. (2012). Dynamic security risk management using Bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing, 9*(1), 61-74. https://doi.org/10.1109/TDSC.2011.34
10. Saha, S., Deb, S., & Das, S. (2020). Machine learning-based intrusion detection system for critical infrastructure protection. *Future Generation Computer Systems, 108*, 121-134. https://doi.org/10.1016/j.future.2020.02.049

Отримано редакцією журналу / Received: 12.02.26
Прорецензовано / Revised: 25.02.26
Схвалено до друку / Accepted: 25.06.26

(cc) BY-NC-SA
This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.