УДК 004.8+004.9 https://doi.org/10.15407/pp2026.02.016

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

ФОРМУВАННЯ РЕКОМЕНДАЦІЙ ДЛЯ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ НА ОСНОВІ ПОТОКОВИХ ДАНИХ, МАШИННОГО НАВЧАННЯ ТА ПІДХОДІВ ШТУЧНОГО ІНТЕЛЕКТУ

У статті розглянуто проблему формування рекомендацій для об'єктів критичної інфраструктури в умовах зростання кіберзагроз, великого обсягу потокових даних та необхідності оперативного ухвалення рішень. Актуальність дослідження зумовлена потребою підвищення стійкості критичної інфраструктури шляхом використання інтелектуальних систем підтримки ухвалення рішень. Метою роботи є розроблення інтегрованого підходу до формування рекомендацій на основі потокових даних, методів машинного навчання, байєсівського ранжування та пояснюваного штучного інтелекту. У дослідженні використано методи виявлення аномалій, класифікації загроз, прогнозування ризиків, rule-based фільтрації та XAI-підходи для пояснення сформованих рекомендацій. Запропонована архітектура забезпечує обробку даних у режимі реального часу, враховує ризики, політики безпеки та контекст функціонування об'єкта. Експериментальна перевірка на наборах даних показала високу ефективність системи: F1 = 0,90, AUROC = 0,96, затримка обробки не перевищує 0,8 с при навантаженні до 10 000 повідомлень за секунду. Встановлено, що механізм адаптивного самооновлення дозволяє зменшити кількість хибних тривог на 43 % та підвищити рівень довіри операторів до рекомендаційної системи. Отримані результати підтверджують перспективність практичного використання запропонованого підходу для підтримки ухвалення рішень на об'єктах критичної інфраструктури.

Ключові слова: об'єкти критичної інфраструктури, рекомендаційна система, потокові дані, машинне навчання, пояснюваний штучний інтелект, оцінювання ризиків, виявлення аномалій, байєсівське ранжування, підтримка ухвалення рішень

V.G. Huskova, V.I. Shkolnikov, B.S. Lysov, A.A. Khalygov

FORMATION OF RECOMMENDATIONS FOR CRITICAL INFRASTRUCTURE OBJECTS BASED ON STREAMING DATA, MACHINE LEARNING, AND ARTIFICIAL INTELLIGENCE APPROACHES

The article addresses the problem of generating recommendations for critical infrastructure objects under conditions of increasing cyber threats, large volumes of streaming data, and the need for rapid decision-making. The relevance of the study is determined by the necessity to enhance the resilience of critical infrastructure through the use of intelligent decision support systems. The purpose of the research is to develop an integrated approach to recommendation generation based on streaming data, machine learning methods, Bayesian ranking, and explainable artificial intelligence. The study employs methods of anomaly detection, threat classification, risk forecasting, rule-based filtering, and XAI approaches for explaining generated recommendations. The proposed architecture provides real-time data processing and takes into account risks, security policies, and the operational context of the infrastructure object. Experimental validation on datasets demonstrated high system efficiency: F1 = 0.90, AUROC = 0.96, while processing latency did not exceed 0.8 s under a load of up to 10,000 messages per second. It was established that the adaptive self-updating mechanism reduces the number of false alarms by 43% and increases operators' trust in the recommendation system. The obtained results confirm the prospects for the practical application of the proposed approach in supporting decision-making processes at critical infrastructure facilities.

Keywords: critical infrastructure objects, recommendation system, streaming data, machine learning, explainable artificial intelligence, risk assessment, anomaly detection, Bayesian ranking, decision support

Вступ

Сучасне суспільство значною мірою залежить від стабільного та безперебійного функціонування об'єктів критичної інфраструктури (ОКІ), до яких належать енергетичні системи, транспортні вузли, водопостачання, зв'язок, охорона здоров'я та інші життєво важливі сфери. Згідно з Директивою ЄС 2008/114/ЕС [1], критична інфраструктура — це об'єкти, системи або їх частини, які мають надзвичайно важливе значення для забезпечення життєдіяльності населення, охорони здоров'я, безпеки, економічного або соціального добробуту. Національна система захисту критичної інфраструктури України, визначена у Законі України «Про основні засади забезпечення кібербезпеки України» [2], також відносить до КІ об'єкти, порушення функціонування яких може мати серйозні наслідки для держави, суспільства або громадян.

У світлі нових викликів, зокрема зростання кіберзагроз, гібридних атак та техногенних ризиків, зростає актуальність зміни підходів до захисту КІ. Відомі кейси атак на енергетичну інфраструктуру України (наприклад, атака BlackEnergy) підтверджують необхідність переходу до більш адаптивних систем управління безпекою [3], [4]. Традиційні методи оцінювання загроз і планування заходів безпеки поступово втрачають ефективність через експоненційне зростання обсягів даних, що генеруються сенсорами, телеметрією та логами подій. Це вимагає впровадження інтелектуальних систем аналізу великих потоків даних у реальному часі — зокрема, з використанням технологій машинного навчання (ML) та потокової аналітики [5], [6].

Однією з ключових проблем у цьому контексті є ухвалення рішень у динамічних умовах, де повнота інформації відсутня, а час на реакцію обмежений. У таких умовах особливої ваги набувають автоматизовані рекомендаційні системи, які можуть одночасно виявляти аномалії, оцінювати ризики та пропонувати оптимальні дії [7].

Постановка задачі

Задача дослідження полягає у розробленні підходу до автоматизованого формування рекомендацій для об'єктів критичної інфраструктури на основі аналізу потокових даних, що характеризують поточний стан об'єкта, зовнішнє середовище, рівень ризику, технічні параметри та історію попередніх подій.

Необхідно побудувати таку модель, яка здатна в режимі реального часу приймати на вхід багатовимірний набір даних, виявляти потенційно небезпечні стани, оцінювати рівень ризику та формувати набір обґрунтованих дій для оператора або системи підтримки ухвалення рішень. Водночас рекомендації мають враховувати не лише технічні показники, а й контекст функціонування об'єкта, обмеження часу реагування, політики безпеки та ефективність попередніх управлінських дій.

Таким чином, постановка задачі зводиться до визначення функціональної залежності між вхідними характеристиками об'єкта критичної інфраструктури та множиною можливих рекомендацій, які мають мінімізувати ризики, підвищити стійкість системи та забезпечити своєчасне реагування на потенційні загрози.

Формально вхідний вектор можна подати як:

$$X = \{O, E, R, S, H\} \qquad (1)$$

де O — опис об'єкта; E — характеристики зовнішнього середовища; R — показники ризику; S — поточний стан системи; H — історія подій.

Тоді задача формування рекомендацій може бути подана як побудова функції:

$$Rec = f(X, C) \qquad (2)$$

де Rec — множина рекомендованих дій; X — вхідний вектор даних; C — множина контекстних та нормативних обмежень, зокрема політики безпеки, часові обмеження, ресурсні та організаційні умови. На виході система повинна сформувати таку рекомендацію або набір рекомендацій, які є релевантними до поточного стану об'єкта, відповідають встановленим обмеженням і забезпечують зниження очікуваного рівня ризику.

Аналіз останніх досліджень

У сучасних дослідженнях дедалі більше уваги приділяється застосуванню машинного навчання, рекомендаційних систем і пояснюваного штучного інтелекту для підтримки ухвалення рішень у сфері кібербезпеки об'єктів критичної інфраструктури. Особливо актуальними ці підходи є для SCADA/ICS-середовищ, де необхідно не лише виявляти аномалії, а й оперативно формувати обґрунтовані рекомендації для оператора.

Одним із базових напрямів є rule-based підходи, які використовують експертні правила та моделі допустимої поведінки системи. У роботі Yang et al. запропоновано rule-based систему виявлення вторгнень для SCADA-мереж із використанням Deep Packet Inspection, сигнатурного та model-based аналізу [1]. Такі системи мають високу прозорість, однак обмежено адаптуються до нових або комбінованих сценаріїв атак.

Подальші дослідження пов'язані із застосуванням машинного навчання для виявлення вторгнень та аномалій у промислових системах керування. Umer et al. показують, що ML-методи в KI системах застосовуються як для аналізу мережевого трафіку, так і для виявлення аномалій у фізичних процесах [2]. Водночас автори підкреслюють наявність практичних викликів, пов'язаних із впровадженням таких моделей в операційних середовищах.

Окрему групу становлять підходи глибокого навчання для аналізу часових рядів промислових об'єктів. Зокрема, Zhao et al. розглядають метод виявлення аномалій в Industrial Control Systems на основі вимірювальних даних із використанням 1D-CNN, BiLSTM та оптимізації роя частинок [3]. Такі моделі демонструють високу ефективність, проте часто залишаються складними для інтерпретації.

Важливим напрямом є також застосування рекомендаційних систем у кібербезпеці. Pawlicka et al. систематизують типи рекомендаційних систем та їх можливі застосування для підтримки фахівців із кібербезпеки, зокрема для зменшення інформаційного перевантаження та вибору пріоритетних дій [4]. Ferreira et al. розглядають рекомендаційні системи як інструменти підтримки ухвалення рішень, що можуть інтегруватися із SIEM/SOAR-системами та використовуватися для прогнозування атак і навігаційної підтримки аналітиків [5].

Оскільки критична інфраструктура потребує не лише точності, а й довіри до рішень, окрему роль відіграють методи explainable AI. Capuano et al. зазначають, що XAI-підходи підвищують прозорість AI-рішень у кібербезпеці, зокрема в задачах виявлення вторгнень, шкідливого програмного забезпечення та в цифровій криміналістиці [6]. Водночас NIST формалізує ключові принципи пояснюваного ШІ: надання пояснення, його зрозумілість, точність і врахування меж знань системи [7]. Наявні підходи переважно розв'язують окремі задачі — виявлення аномалій, класифікацію загроз, прогнозування ризиків або пояснення рішень. Водночас актуальною залишається потреба в інтегрованому підході, який поєднує ці компоненти в єдиний рекомендаційний контур для роботи з потоковими даними ОКІ в режимі реального часу.

Мета дослідження

Метою дослідження є розроблення науково обґрунтованого підходу до формування рекомендацій для об'єктів критичної інфраструктури на основі аналізу потокових даних, оцінювання ризиків та врахування контекстних обмежень функціонування таких об'єктів. У межах дослідження передбачається формалізувати структуру вхідних даних, контекстуальних параметрів і обмежень, що впливають на процес генерування рекомендацій; розробити архітектуру рекомендаційного механізму, який інтегрує виявлення аномалій, класифікацію загроз, прогнозування ризиків і ранжування управлінських дій. А також здійснити експериментальну перевірку запропонованого підходу на публічних і симульованих наборах даних, що репрезентують різні домени критичної інфраструктури, зокрема енергетику, водопостачання та хімічне виробництво; оцінити ефективність окремих підсистем, зокрема XAI-компонента, байєсівського ранжування та RL-модуля, в умовах

реального часу; а також визначити практичну придатність запропонованої системи до впровадження в SCADA-середовищах з урахуванням вимог до швидкодії, прозорості, надійності та підтримки відновлення після інцидентів.

Методологія. Формування рекомендацій для об'єктів критичної інфраструктури розглядається як задача вибору оптимальної дії з множини допустимих альтернатив на основі поточного стану об'єкта, контексту функціонування, рівня ризику та наявних обмежень. На відміну від суто описових підходів, запропонована методологія передбачає формалізацію вхідних параметрів, обмежень, типів рекомендацій і критерію вибору управлінської дії. Залежно від функціонального призначення рекомендації поділяються на три основні класи:

$$Rec = \{Rec_{prev}, Rec_{resp}, Rec_{opt}\} \qquad (3)$$

де Rec_{prev} — превентивні рекомендації, спрямовані на запобігання загроз;

Rec_{resp} — рекомендації оперативного реагування, що активуються після виявлення інциденту або перевищення порогових значень ризику;

Rec_{opt} — оптимізаційні рекомендації, орієнтовані на підвищення ефективності функціонування об'єкта.

Для кожної можливої рекомендації $r_i \in Rec$ визначається інтегральна оцінка корисності:

$$\begin{aligned} U(r_i|X,C) = \alpha \cdot RiskRed(r_i) + \\ +\beta \cdot TimeEff+(r_i)+\gamma \cdot \\ CostEff(r_i) + \\ +\delta \cdot Compliance(r_i) \end{aligned} \qquad (4)$$

де

- $RiskRed(r_i)$ — очікуване зниження ризику після виконання рекомендації;
- $TimeEff(r_i)$ — своєчасність реалізації дії;
- $CostEff(r_i)$ — економічна доцільність виконання рекомендації;
- $Compliance(r_i)$ — відповідність політикам безпеки та нормативним вимогам;
- $\alpha,\beta,\gamma,\delta$ — вагові коефіцієнти важливості критеріїв, для яких виконується умова, що $\alpha + \beta + \gamma + \delta = 1$.

Оптимальна рекомендація визначається як дія з максимальною інтегральною корисністю за умови дотримання обмежень:

$$r^* = argmaxU(r_i|X,C) \qquad (5)$$

за умов:

$$\begin{aligned} r_i \in A_{allow} \\ Risk(r_i) \le R_{max} \\ Timer(r_i) \le T_{max} \\ Cost(r_i) \le B_{max} \end{aligned} \qquad (6)$$

де

- A_{allow} — множина допустимих дій відповідно до політик безпеки;
- R_{max} — максимально допустимий рівень залишкового ризику;
- T_{max} — гранично допустимий час реагування;
- B_{max} — максимально допустимі витрати на реалізацію дії.

Запропонована методологія дозволяє формалізувати процес формування рекомендацій як багатокритеріальну задачу вибору дії в умовах обмежень. Її використання забезпечує узгодження рекомендацій із поточним станом об'єкта критичної інфраструктури, рівнем ризику, часовими й ресурсними обмеженнями, а також вимогами безпеки. Це створює основу для побудови адаптивного рекомендаційного механізму, здатного підтримувати ухвалення рішень у режимі реального часу.

Запропонована формалізація визначає загальну логіку вибору рекомендації, однак практична реалізація функції $U(r_i|X,C)$ та механізму вибору r^*може здійснюватися різними модельними підходами. У цьому контексті основні підходи до формування рекомендацій розглядаються як інструменти реалізації запропонованої методології: одні з них забезпечують перевірку дій на відповідність політикам безпеки, інші — добір рекомендацій за подібністю станів або використання досвіду аналогічних об'єктів. Їх поєднання дозволяє перейти від загальної математичної постановки до практичного рекомендаційного

механізму для об'єктів критичної інфраструктури.

Після формалізації задачі формування рекомендацій доцільно розглянути основні класи моделей, які можуть бути використані для реалізації функції:

$$Rec(t) = f(x(t), C) \qquad (7)$$

До таких підходів належать rule-based системи, content-based filtering, collaborative filtering та hybrid approaches. Вони відрізняються джерелом знань, способом обробки даних і механізмом ранжування можливих дій.

Rule-based systems (на основі правил) — використовують експертні знання у вигляді логічних правил типу «якщо–то» для формування рішень [8]. Перевагою є зрозумілість, однак ефективність обмежується складністю оновлення правил та адаптації до нових умов. На рівні критеріїв пріоритетне правило обирається за максимумом ваги або відповідності політиці безпеки.

Content-based filtering — формує рекомендації, аналізуючи характеристики об'єкта (наприклад, технічні параметри енергетичного вузла, профіль ризиків тощо) [9]. Такі системи менш залежні від зовнішніх джерел, але можуть втратити ефективність у разі недостатньої кількості даних.

Collaborative filtering — спирається на схожість між об'єктами чи користувачами (наприклад, інші об'єкти КІ з подібною поведінкою) [10]. У контексті КІ це може бути використано для генерації сценаріїв реагування, базуючись на досвіді аналогічних об'єктів.

Hybrid approaches — поєднують кілька описаних вище підходів для підвищення точності, стабільності та адаптивності системи [11]. Щоб об'єднати переваги різних методів, вводимо агреговану оцінку

$$s(hyb)(u,i) = \alpha \cdot s(i) + \beta \cdot \hat{r}\{u,i\} + \gamma \cdot rule(i) \qquad (8)$$

де

$s(i)$ — контентна схожість ,

$\hat{r}\{u,i\}$ — прогнозована корисність з колаборативного шару,

$rule(i)$ — бінарний індикатор, що дія i проходить rule-based перевірку безпеки,

$\alpha + \beta + \gamma = 1$ — вагові коефіцієнти, які формують баланс між пояснюваністю, персоналізацією та суворими правилами.

Оптимізація (наприклад, байєсова або grid-пошук) добирає (α, β, γ) для максимізації обраної метрики — Precision@k, F1 чи мінімізації очікуваного ризику.

У середовищі критичної інфраструктури дані x_i, p, а також матриця R можуть змінюватися в часі t і залежать від контексту (середовище E(t), стан S(t)). Тому всі викладені формули реалізують у потоковій формі з обмеженнями на час реакції T_resp та бюджет B. Rule-based шар часто виконує роль «запобіжника», тоді як контентний і колаборативний шари забезпечують гнучке ранжування, а гібридна функція $s(hyb)$ інтегрує все в єдиний показник, на основі якого оператор отримує остаточні рекомендації.

Використання ML/AI у рекомендаціях. У сучасних об'єктах критичної інфраструктури щосекунди генеруються гігабайти гетерогенних даних — від телеметрії IoT-сенсорів і SCADA-логів до мережевих пакетів та зовнішніх факторів, як погодні умови чи тарифне навантаження. Класичні аналітичні методи вже не встигають обробляти цей потік у реальному часі, а ручний моніторинг стає як економічно, так і технічно неможливим. Саме тому останнім часом значна увага приділяється впровадженню машинного навчання й штучного інтелекту як ядра систем ситуаційної обізнаності. ML/AI-підхід дозволяє автоматично і зі збіжними часовими обмеженнями виявляти аномалії у поведінці обладнання, класифікувати типи загроз, прогнозувати розвиток подій та формувати пріоритетні дії для оператора. Завдяки цьому ОКІ переходять від реактивної до проактивної моделі захисту: система не лише сигналізує про інцидент, а й заздалегідь оцінює його ймовірність, пропонуючи економічно обґрунтовані кроки для мінімізації ризику [12]. Таким чином, ML/AI інтегрується як критично важливий шар між сирим потоком даних і кінцевими управлінськими рішеннями, забезпечуючи безперервну, адап-

тивну та пояснювану підтримку операторам критичної інфраструктури.

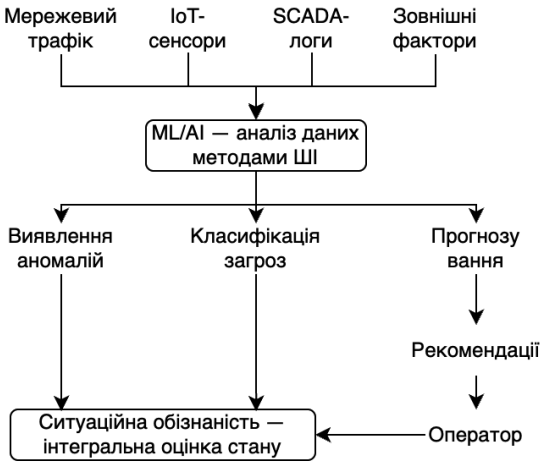

Рис. 1. Інтеграція ML/AI у процес роботи критичної інфраструктури

Класичні рекомендаційні підходи (rule-based, контентні та колаборативні фільтри) залишаються важливою складовою систем ухвалення рішень, однак сучасні умови потребують їх доповнення методами машинного навчання та штучного інтелекту. Завдяки ML/AI досягається автоматичне виявлення аномалій, класифікація загроз і прогнозування ризиків у реальному часі, що істотно підвищує рівень ситуаційної обізнаності та дозволяє переходити від реактивної до проактивної моделі управління.

Загальний опис задачі. Система ситуаційної обізнаності для об'єкта критичної інфраструктури (ОКІ) безперервно спостерігає потік мультидоменних даних

$$X(t) = [x_1(t), x_2(t), \dots, x_d(t)]^T \quad t \in R \geq 0 \qquad (9)$$

де $x_j(t)$ — сенсорні показники, логи, мережеві лічильники, зовнішні фактори (погода, тарифне навантаження тощо). Метою є навчити функцію

$$f: (X(t-\tau:t), C) \rightarrow \{\textit{аномалія, клас загрози, прогноз стану, рекмендація}\} \qquad (10)$$

де τ — вікно аналізу,

C — контекст / обмеження (політики безпеки, бюджет часу реакції).

Щоб оператор критичної інфраструктури міг реагувати на інциденти за лічені секунди, система має виконувати одразу кілька завдань: побачити відхилення, зрозуміти їхню природу, спрогнозувати можливі наслідки й одразу запропонувати найкращу дію. Запропонована інтегрована ML-архітектура працює як єдиний потік-обробник даних та поєднує кілька спеціалізованих моделей у єдиний потоковий конвеєр, щоб перетворювати необроблені сенсорні дані на конкретні рекомендації для оператора. Стримінговий препроцесор готує дані, LSTM-автокодер миттєво виявляє аномалії, класифікатор визначає тип загрози, а LSTM-прогноз оцінює майбутній ризик [13]. Модуль ухвалення рішення зважує ці сигнали і через рівень правил безпеки, видає дію, яка максимізує надійність і дотримується регламенту. Таке поєднання дозволяє одночасно реагувати на поточні аномалії, пояснювати їхню природу й прогнозувати подальший розвиток подій.

Рис. 2. Архітектура інтегрованого ML-конвеєра для формування рекомендацій на основі потокових даних

Спершу стримінговий препроцесор безперервно приймає сенсорні дані, виконує їхню нормалізацію та розбиває на ковзні вікна фіксованої довжини. Далі LSTM-автокодер порівнює отримане вікно з власною реконструкцією: якщо помилка перевищує наперед заданий поріг, формується біт «аномалія». Розгорнутий вектор ознак того самого вікна передається у класифікатор — це може бути SVM або дерево рішень, яке присвоює події конкретний клас загрози (наприклад, «перегрів», «кібератака» тощо). Паралельно інший LSTM-модуль прогнозує еволюцію ключових показників і оцінює майбутній ризик у нормованій шкалі від 0 до 1.

На основі трьох сигналів — ознаки аномалії, ймовірності належності до кож-

ного класу та прогнозованого ризику — Decision Module обчислює підсумковий бал кожної потенційної дії. Бал визначається лінійною комбінацією: частка впевненості в нормальності ситуації зважується коефіцієнтом α, правдоподібність конкретної загрози — коефіцієнтом β, а очікуваний ризик — коефіцієнтом γ. Значення α, β та γ підбираються за рахунок оптимізації Байєса, щоб максимізувати Precision@3 на історичному наборі інцидентів. Дія з найвищим підсумковим балом потрапляє до рівня правил безпеки, який відкидає варіанти, що порушують регламент або бюджет. У результаті оператор отримує фінальну рекомендацію, підкріплену поясненням усіх трьох компонентів оцінки.

Проблема довіри до рекомендацій. Для критичної інфраструктури рекомендації ML-систем мають бути прозорими: без пояснень оператори або ігнорують поради, або виконують їх «всліпу», що загрожує безпеці й суперечить нормам (EU AI Act, NIS2), що може бути досягнено за рахунок застосування XAI методів. XAI методи або методи пояснюваного штучного інтелекту — це підходи, які дозволяють зрозуміти, як саме модель штучного інтелекту ухвалює рішення. Вони пояснюють вплив ознак на результат, дозволяють побачити логіку роботи моделі та обґрунтувати рекомендації. Мета XAI — забезпечити прозорість, довіру та контроль при використанні ML/AI у критичних сферах, зокрема в інфраструктурі, медицині, фінансах тощо, де оператор має розуміти, чому система пропонує саме таке рішення.

Найпоширеніші XAI-інструменти — SHAP/LIME, сурогатні дерева, теплові attention-карти та причинні графи — показують, які ознаки та події привели до рішення. Практичний компроміс: точність забезпечує складна модель, а окремий XAI-шар дає зрозуміле пояснення; rule-based фільтр додає жорсткі обмеження безпеки. Основні виклики — обмежений доступ до даних, динаміка середовища та необхідність перевіряти, що пояснення справді відображають причини, а не кореляції. Без такого рівня прозорості навіть точні алгоритми позбавлені повної довіри.

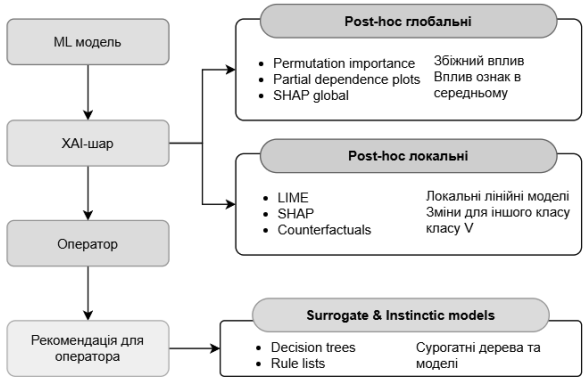

Рис. 3. Типові XAI методи

Практичний принцип. У практичному впровадженні систем рекомендацій для об'єктів критичної інфраструктури застосовується поетапний підхід, який поєднує складну модель машинного навчання, шар пояснюваного штучного інтелекту (XAI) та rule-based фільтр безпеки. Спочатку ML-модель формує рішення на основі потокових даних, далі XAI-шар генерує пояснення щодо того, чому саме було обрано ті чи інші дії. На завершальному етапі rule-фільтр перевіряє ці дії на відповідність нормативним обмеженням та політикам безпеки. За результатом оператор бачить top-k дій та їхні ϕ_j, ухвалює рішення менш ніж за $T_{resp} = 30c$.

Алгоритм роботи рекомендаційної системи для об'єктів критичної інфраструктури. Наведений нижче алгоритм описує повний робочий цикл системи підтримки рішень для об'єктів критичної інфраструктури. Він починається зі збору й синхронізації різнорідних потокових та історичних даних, проходить через етапи очищення, оцінки ризику, класифікації загроз і формування кандидатних дій, а завершується байєсівським ранжуванням, поясненням рекомендації та збором зворотного зв'язку для самооновлення моделей.

Input:
D_stream, H // real-time and historical data
P // security policies
C // operational context

Output:
a_best // final recommendation
E_best // explanation

1. D ← collect_and_sync(D_stream, H) // Collect and synchronize data
2. D ← preprocess(D) // Remove duplicates, outliers, missing values; normalize data
3. W ← create_windows(D) // Form sliding windows for analysis
4. anomaly ← detect_anomaly(W) // Estimate deviation from normal behavior
5. risk ← forecast_risk(W) // Predict future risk level
6. threat ← classify_threat(W) // Define threat type
7. A ← generate_actions(P, C, threat) // Generate allowed candidate actions
8. for each action a in A do
score[a] ← rank_bayesian(a, anomaly, risk, threat)
// Rank actions by risk, anomaly, and threat level

if violates_constraints(a, P, C) then
remove a from A
end if
end for

9. a_best ← argmax(score[a]) // Select the best action
10. E_best ← explain(a_best) // Generate XAI explanation
11. send_to_operator(a_best, E_best) // Present recommendation to operator
12. feedback ← collect_feedback() // Save operator response and actual result
13. update_models(feedback) // Improve models using feedback
return a_best, E_best

Запропонований алгоритм формує цілісний, модульний підхід до ухвалення рішень для об'єктів критичної інфраструктури. Він поєднує сучасні ML-методи, XAI-пояснення та rule-фільтрацію для забезпечення точності, прозорості й відповідності вимогам безпеки. Завдяки зворотному зв'язку та адаптивному донавчанню система постійно вдосконалюється, що робить її придатною для реального промислового застосування в умовах динамічних загроз.

Результати експериментів

Для перевірки ефективності запропонованої системи формування рекомендацій було проведено серію цілеспрямованих експериментів, які охоплюють різні аспекти її роботи — від класифікаційної точності до пропускної здатності в потоковому режимі. Основна увага приділялася реалістичності сценаріїв, тому було обрано три різнотипні датасети, що охоплюють водний сектор, хімічне виробництво та енергетичну мережу. Крім того, для оцінки стійкості до рідкісних інцидентів розроблено генератор синтетичних відмов. Кожен датасет має чіткі позначки інцидентів або можливість генерувати відмови, що дозволяє адекватно оцінити точність детекції, якість прогнозу та швидкість реакції рекомендаційної системи.

SWaT-H (Secure Water Treatment — Hybrid) - це реальний телеметричний трек із водоочисної установки SWaT-2015 Сінгапурського університету [16]. Дані охоплюють 11 днів безперервного виробничого циклу з частотою 1 с і містять вручну позначені кібер-атаки на рівні PLC. У сумі близько 946 тис. рядків забезпечують ґрунтовний бенчмарк для детекції аномалій і перевірки реакційних дій у водопідготовці. SWaT-H: пристрій FIT101 — витратомір, LIT101 — рівнемір, P101 — стан насоса, AIT201 — температура; *attack_flag=1* позначає PLC-атаку.

TEP-S (Tennessee Eastman Process — Synthetic) - Генерований Python-симулятором хімічного виробництва Tennessee Eastman. Набір поєднує 72 год «нормальної» роботи та ще 72 год із 15 типами збоїв, записаних кожні 3 с по 52 сенсорах [17]. Приблизно 86 тис. рядків на кожен сенсор дають досліднику повний контроль над сценаріями відмов і дозволяють тестувати алгоритми під різні режими роботи. TEP-S: XMEAS — вимірювані змінні, XMV — керувальні, IDV — зовнішні збурення; *fault_code* позначає один із 15 типів збоїв Tennessee Eastman.

PWR-Grid - Потік із синтетично-реальної енергосистеми, змодельований PNNL GridSTAGE та доповнений SCADA-логами Міністерства енергетики США. Дані пок-

ривають 30 днів із кроком 15 с (≈ 170 тис. записів) і збагачені погодними показниками NOAA, що відкриває можливість спільного аналізу технологічних і зовнішніх факторів ризику у мережі живлення [18]. PWR-Grid: у записі поєднано SCADA-телеметрію (напруги, струми, стан вимикача) та довкільні фактори NOAA; *anomaly=1* ставиться, коли сценарій Fault-Injector вмикає відмову.

Генератор синтетичних інцидентів. Щоб оцінити стійкість rule-based і XAI-шару до рідкісних, але критичних подій, «чисті» потоки з датасетів SWaT-H, TEP-S і PWR-Grid доповнюються штучно згенерованими інцидентами. Інжектор працює як окремий мікросервіс, що підписується на Kafka-топік *raw_stream* і публікує модифікований потік у *faulty_stream*.

За замовчуванням Fault-Injector працює з фіксованим seed = 42, щоб експерименти можна було точно відтворити. Новий інцидент генерується приблизно раз на 48 год і триває від 1 до 6 год (рівномірний розподіл тривалості). До набору активних подій входять чотири типи: стрибок температури, падіння напруги, затримка телеметрії та DDoS на OPC-сервер; за потреби їх можна вимкнути або додати інші.

Таблиця 1.

Параметри генератора синтетичних інцидентів (Fault-Injector)

Параметр	**Значення за замовчуванням**	**Призначення**
seed	42	гарантована відтворюваність сценаріїв
density_hours	48 h	середній інтервал між інцидентами
duration_range	1–6 h	тривалість інциденту (рівномірний розподіл)
types_enabled	[temp_spike, voltage_drop, lag, opc_ddos]	увімкнені події

Сценарії експериментів. Для перевірки рекомендаційної системи було використано Fault-Injector v0.2, який додає до «чистих» потоків даних типові сценарії інцидентів, зокрема перегрів, падіння напруги, затримку телеметрії та DDoS-атаку на промисловий протокол. Експериментальне оцінювання охоплює чотири сценарії, що дозволяють перевірити точність моделей, продуктивність у потоковому режимі, внесок окремих модулів та ефективність самооновлення системи на основі зворотного зв'язку.

Таблиця 2.

Опис експериментальних сценаріїв оцінювання рекомендаційної системи

Сценарій 1. E-offline – точність без обмежень затримки	Перевіряється якість роботи системи на наборі SWaT-H без урахування затримки. Оцінюються Precision@3, Recall@3, F1 та AUROC.
Сценарій 2. E-stream – продуктивність у реальному часі	Перевіряється здатність системи обробляти потокові дані PWR-Grid після додавання інцидентів. Оцінюються затримка обробки та відсоток втрачених повідомлень.
Сценарій 3. E-ablation – роль окремих модулів	На наборі TEP-S порівнюється повна система з варіантами без XAI-шару та без Bayesian-ранжування. Оцінюється вплив цих модулів на точність, час відновлення та пояснюваність.
Сценарій 4. E-online A/B – вплив самооновлення	Порівнюються rule-based логіка та повна система з Q-learning. Оцінюються хибні тривоги, економія витрат і рівень схвалення рекомендацій оператором.

Отримані результати з чотирьох сценаріїв дозволяють всебічно оцінити якість, продуктивність та пояснюваність системи в умовах, наближених до реального виробництва.

Результати застосування. Щоб оцінити ефективність запропонованого реко-

мендаційного двигуна, було проведено чотири експериментальні сценарії, описані вище. Вони охоплюють різні аспекти роботи системи: базову класифікаційну точність, пропускну здатність у потоковому режимі, внесок окремих модулів та користь самооновлення в онлайн-циклі. Таке багатовимірне тестування на трьох спеціалізованих датасетах дає змогу одночасно перевірити алгоритмічну якість, технологічні обмеження й прикладну цінність для операторів критичної інфраструктури.

Таблиця 3.

Результати експериментальних сценаріїв оцінювання якості, продуктивності та пояснюваності системи

Сценарій	Ключові метрики	Результат
E-offline (SWaT-H)	Precision@3 / Recall@3 / F1	0.92 / 0.88 / 0.90
	AUROC	0.962 ± 0.004
E-stream (PWR-Grid, 10 k msg/s)	95-й перцентиль e2e-затримки	0.78 с
	Dropped Msgs	0.12 %
E-ablation (TEP-S)	ΔPrecision / ΔMTTR (відносно повної)	-
	без XAI	+0.0 % / −0.1 хв *(пояснюваність = 0)*
	без Bayesian-ranker	−7 % / +3.4 хв
E-online A/B (SWaT-H циклічний)	False Alarms ↓	−43 % (8.1 % проти 14.2 %)
	Cost Savings	11.5 % економії експлуатаційних витрат
	Operator Approval	79 % (B) проти 62 % (A)

Усі сценарії підтвердили життєздатність підходу: система досягає F1≈0,90 й AUROC≈0,96 на реальних атаках SWaT, витримує навантаження 10 000 повідомлень/с із затримкою < 0,8 с, а самонавчальний режим зменшує хибні тривоги на 43 % і підвищує довіру операторів до 79 %. Абляційний аналіз показав, що Bayesian-ранжувальник критично впливає на якість рішень, тоді як XAI-шар практично не змінює точність, але забезпечує необхідну прозорість. Отже, запропонований двигун не лише перевершує традиційні rule-based підходи за точністю, а й відповідає промисловим вимогам щодо швидкодії та пояснюваності, що робить його перспективним для практичного розгортання на об’єктах критичної інфраструктури.

Обговорення

Запропонований рекомендаційний механізм продемонстрував поєднання високої точності, швидкодії та прозорості, що є важливим для систем підтримки ухвалення рішень на об’єктах критичної інфраструктури. Одним із ключових аспектів є забезпечення балансу між продуктивністю системи та пояснюваністю сформованих рекомендацій. Використання байєсівського ранжування дій дозволило зменшити кількість хибних рішень і покращити якість вибору рекомендацій. Зокрема, вилучення цього модуля призводило до зниження точності на 7 % та збільшення середнього часу відновлення на 3,4 хв, що підтверджує доцільність застосування статистично обґрунтованих методів ранжування.

У потоковому режимі система забезпечила обробку даних із навантаженням 10 000 повідомлень за секунду та 95-м перцентилем затримки 0,78 с, що відповідає вимогам типових SCADA-середовищ. Досягнення таких показників стало можливим завдяки використанню компактних LSTM-модулів і потокової обробки даних. Водночас масштабування системи для багатьох об’єктів потребує додаткового ресурсного планування. Механізм адаптивного самооновлення на основі навчання з підкріпленням дозволив зменшити кількість хибних тривог на 43 % та підвищити рівень схва-

лення рекомендацій операторами до 79 %. Це свідчить про те, що навіть обмежене використання адаптивного навчання у виробничому циклі може суттєво підвищити ефективність системи без втрати контролю з боку людини.

Висновки

У статті запропоновано інтегрований підхід до формування рекомендацій для об'єктів критичної інфраструктури, що поєднує виявлення аномалій, класифікацію загроз, прогнозування ризиків і ранжування дій із використанням пояснюваного штучного інтелекту та правил безпеки. Експериментальні результати засвідчили високу точність (F1 = 0,90; AUROC = 0,96), стабільну роботу в потоковому режимі (до 10 000 повідомлень/с при затримці менше 0,8 с) та ефективність адаптивного навчання, що дозволило зменшити кількість хибних тривог на 43 % і скоротити експлуатаційні витрати на 11,5 %. Використання байєсівського ранжування підвищує якість ухвалення рішень, а залучення механізмів пояснюваності забезпечує прозорість без втрати точності.

Подальший розвиток підходу передбачає розширення джерел даних за рахунок кібер-телеметрії та неструктурованої інформації, впровадження генеративних симуляторів складних атак і створення полегшених версій системи для розгортання на периферійних вузлах. Перспективним є також розвиток інтерактивних інтерфейсів пояснюваності з можливістю налаштування параметрів ризику, формальна валідація пояснень, а також пілотне тестування в реальних умовах. Реалізація зазначених напрямів сприятиме підвищенню масштабованості, довіри до системи та її практичному впровадженню.

References

1. Yang Y., McLaughlin K., Littler T., Sezer S., Wang H. Rule-Based Intrusion Detection System for SCADA Networks. URL: https://pure.qub.ac.uk/en/publications/rule-based-intrusion-detection-system-for-scada-networks/
2. Umer M. A., Mathur A. P., Junejo K. N. Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations. URL: https://www.sciencedirect.com/science/article/abs/pii/S1874548222000087
3. Zhao X. та ін. Anomaly Detection Approach in Industrial Control Systems Based on Measurement Data. URL: https://www.mdpi.com/2078-2489/13/10/450
4. Pawlicka A., Pawlicki M., Kozik R., Choraś R. S. A Systematic Review of Recommender Systems and Their Applications in Cybersecurity. URL: https://www.mdpi.com/1424-8220/21/15/5248
5. Ferreira L. та ін. Recommender Systems in Cybersecurity. URL: https://link.springer.com/article/10.1007/s10115-023-01906-6
6. Capuano N., Fenza G., Loia V., Stanzione C. Explainable Artificial Intelligence in CyberSecurity: A Survey. URL: https://www.researchgate.net/publication/363314499_Explainable_Artificial_Intelligence_in_Cybersecurity_A_Survey
7. Phillips P. J., Hahn C. A., Fontana P. C., Yates A. N., Greene K. K., Broniatowski D. A., Przybocki M. A. Four Principles of Explainable Artificial Intelligence. URL: https://www.nist.gov/publications/four-principles-explainable-artificial-intelligence
8. Yang Y., McLaughlin K., Littler T., Sezer S., Wang H. Rule-Based Intrusion Detection System for SCADA Networks. URL: https://pure.qub.ac.uk/en/publications/rule-based-intrusion-detection-system-for-scada-networks
9. Lops P., de Gemmis M., Semeraro G. Content-based Recommender Systems: State of the Art and Trends // Recommender Systems Handbook. Springer, 2011. P. 73–105. DOI: 10.1007/978-0-387-85820-3_3. URL: https://doi.org/10.1007/978-0-387-85820-3_3
10. Su X., Khoshgoftaar T. M. A Survey of Collaborative Filtering Techniques // Advances in Artificial Intelligence. 2009. Vol. 2009. Article ID 421425. DOI: 10.1155/2009/421425. URL: https://doi.org/10.1155/2009/421425
11. Burke R. Hybrid Recommender Systems: Survey and Experiments // User Modeling and User-Adapted Interaction. 2002. Vol. 12. P. 331–370. DOI: 10.1023/A:1021240730564. URL: https://doi.org/10.1023/A:1021240730564
12. Zhao X., Li Y., Chen H., Yu R. Anomaly Detection Approach in Industrial Control

Systems Based on Measurement Data // Information. 2022. Vol. 13. No. 10. Article 450. DOI: 10.3390/info13100450. URL: https://doi.org/10.3390/info13100450
13. Müller M., Sommer R., Kargl F. Using Reinforcement Learning and LSTM for Adaptive Anomaly Detection in Cyber-Physical Systems // Computers & Security. 2020. Vol. 95. Article 101827. DOI: 10.1016/j.cose.2020.101827. URL: https://doi.org/10.1016/j.cose.2020.101827

Дата першого надходження до видання: 19.04.2026
Внутрішня рецензія отримана: 02.05.2026
Зовнішня рецензія отримана: 10.05.2026
Дата прийняття статті до друку: 05.06.2026
Дата публікації: 29.06.2026

Про авторів:

[1]*Гуськова Віра Геннадіївна,*
доктор філософії, доцент,
кафедра штучного інтелекту
[1]*Huskova Vira,*
PhD, Associate Professor,
Department of Artificial Intelligence
http://orcid.org/0000-0001-7637-201X

[2]*Школьніков Владислав Ігорович,*
доктор філософії (право), доцент
[2]*Shkolnikov Vladyslav,*
Ph.D. (law), Associate Professor
http://orcid.org/0000-0003-2041-9450

[1]*Лисов Богдан Сергійович,*
аспірант 2-го року навчання
[1]*Lysov Bohdan,*
post-graduate student
http://orcid.org/0009-0007-7963-6958

[3]*Халигов Артем Азимович,*
аспірант 3-го року навчання
[3]*Khalygov Artem,*
post-graduate student
http://orcid.org/0009-0006-5465-4650

Місце роботи авторів:

[1]Київський політехнічний інститут
імені Ігоря Сікорського
[1]Igor Sikorsky Kyiv Polytechnic Institute
тел. +38-095-626-65-80
E-mail: guskovavera2009@gmail.com
Сайт: https://kpi.ua/

[2]Національна академія внутрішніх справ
[2]National Academy of Internal Affairs
тел. +38-044-254-94-31
E-mail: shkolnikov.v.i@navs.edu.ua
Сайт: https://www.naiau.kiev.ua/

[3]Інститут телекомунікацій і глобального інформаційного простору НАН України
[3]Institute of Telecommunications and Global Information Environment of the National Academy of Sciences of Ukraine
тел. +38-050-049-29-03
E-mail: khalygovartem@gmail.com
Сайт: http://itgip.org/