https://doi.org/10.31891/2219-9365-2026-86-24
УДК 004.056

ІЛЬЄНКО Анна
Державний університет «Київський авіаційний інститут»
https://orcid.org/0000-0001-8565-1117
e-mail: anna.ilienko@npp.kai.edu.ua
ТЕЛЮЩЕНКО Валентина
Державний університет «Київський авіаційний інститут»
https://orcid.org/0000-0001-6026-5105
e-mail: valentyna.teliushchenko@npp.kai.edu.ua

ГІБРИДНИЙ МЕТОД ОЦІНЮВАННЯ КІБЕРРИЗИКІВ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

У статті запропоновано гібридний підхід до оцінювання кіберризиків об'єктів критичної інфраструктури, який поєднує багатокритеріальну модель зваженої суми (WSM), методи машинного навчання та елементи нечіткої логіки. Актуальність дослідження зумовлена зростанням складності кіберзагроз у сучасних кіберфізичних системах, де традиційні експертні методи оцінювання ризику не забезпечують достатньої адаптивності до динамічних змін середовища. Запропонований підхід орієнтований на інтеграцію даних оперативного кібермоніторингу із класичними механізмами оцінювання ризиків для підвищення точності, адаптивності та інтерпретованості результатів. Базою моделі виступає метод зваженої суми, який забезпечує прозору структуру оцінювання та можливість експертного налаштування критеріїв і вагових коефіцієнтів. На відміну від традиційних моделей, у роботі використано розширену структуру критеріїв, яка включає не лише ймовірність реалізації загрози та критичність впливу, а й рівень експозиції та рівень вразливості. Критерій вразливості представлено як агреговану композицію підкритеріїв, що враховують CVSS-оцінки, доступність експлойтів, стан оновлень і конфігураційну безпеку системи. У межах дослідження формалізовано механізм інтеграції ML-модулів у модель оцінювання ризику. Для аналізу телеметричних даних SIEM/XDR, журналів подій та мережевого трафіку використано методи класифікації, кластеризації, виявлення аномалій, регресії та аналізу часових рядів. Результати роботи ML-модулів агрегуються у вигляді інтегрального ML-індикатора, який використовується для адаптивного коригування оцінки ймовірності реалізації загрози. Для забезпечення контрольованої інтеграції машинного навчання введено коефіцієнт довіри до ML-шару, що дозволяє балансувати між експертними оцінками та автоматизованим аналізом.

Практична цінність запропонованого підходу полягає у можливості побудови адаптивних систем підтримки прийняття рішень у сфері кібербезпеки, здатних поєднувати експертні знання, технічні характеристики вразливостей та результати аналізу телеметричних даних.

Ключові слова: кіберризик, критична інфраструктура, машинне навчання, SIEM, оцінювання ризиків, виявлення аномалій, кібербезпека, гібридна модель.

ILIENKO Anna, TELIUSHCHENKO Valentyna
State University "Kyiv Aviation Institute"

HYBRID METHOD FOR ASSESSING CYBERRISKS OF CRITICAL INFRASTRUCTURE OBJECTS USING MACHINE LEARNING

The article proposes a hybrid approach to assessing cyber risks of critical infrastructure facilities, which combines a multi-criteria weighted sum model (WSM), machine learning methods, and elements of fuzzy logic. The relevance of the study is due to the increasing complexity of cyber threats in modern cyber-physical systems, where traditional expert risk assessment methods do not provide sufficient adaptability to dynamic changes in the environment. The proposed approach is focused on integrating operational cyber monitoring data with classical risk assessment mechanisms to increase the accuracy, adaptability, and interpretability of results. The model is based on the weighted sum method, which provides a transparent assessment structure and the possibility of expert adjustment of criteria and weighting coefficients. Unlike traditional models, the work uses an extended criteria structure that includes not only the probability of threat realization and the criticality of the impact, but also the level of exposure and the level of vulnerability. The vulnerability criterion is presented as an aggregated composition of subcriteria that take into account CVSS scores, exploit availability, update status, and system configuration security. The study formalizes the mechanism for integrating ML modules into the risk assessment model. Classification, clustering, anomaly detection, regression, and time series analysis methods are used to analyze SIEM/XDR telemetry data, event logs, and network traffic. The results of ML modules are aggregated in the form of an integral ML indicator that is used to adaptively adjust the assessment of the probability of threat implementation. To ensure controlled integration of machine learning, a confidence factor is introduced for the ML layer, which allows balancing between expert assessments and automated analysis.

The practical value of the proposed approach lies in the possibility of building adaptive decision support systems in the field of cybersecurity that are capable of combining expert knowledge, technical characteristics of vulnerabilities, and the results of telemetry data analysis.

Keywords: cyber risk, critical infrastructure, machine learning, SIEM, risk assessment, anomaly detection, cybersecurity, hybrid model.

Стаття надійшла до редакції / Received 27.03.2026
Прийнята до друку / Accepted 22.04.2026
Опубліковано / Published 31.05.2026

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

Сучасні об'єкти критичної інфраструктури функціонують у середовищі постійного зростання складності кіберзагроз, що зумовлено цифровізацією технологічних процесів, інтеграцією IT- та OT-систем, широким використанням віддаленого доступу та збільшенням кількості взаємопов'язаних сервісів. Особливої актуальності ця проблема набуває для енергетики, транспорту, авіаційної галузі, промислових систем керування та інших кіберфізичних систем, порушення роботи яких може призвести не лише до економічних втрат, а й до критичних наслідків для безпеки держави та населення. Традиційні методи оцінювання ризиків, що базуються переважно на статичних експертних оцінках, не забезпечують достатньої адаптивності до динамічних змін кіберзагроз та не враховують оперативні дані моніторингу безпеки, отримані із SIEM/XDR-платформ, систем OT-телеметрії та засобів виявлення інцидентів.

Існуючі підходи до оцінювання кіберризиків, зокрема ISO/IEC 27005, NIST SP 800-30 та інші ризик-орієнтовані методики, забезпечують формалізовану основу управління ризиками, однак мають низку обмежень у контексті сучасних кіберфізичних систем. [1]. Більшість моделей використовують фіксовані критерії ризику та недостатньо враховують невизначеність, властиву процесам аналізу кіберзагроз і вразливостей. Крім того, класичні моделі ризику здебільшого не інтегрують результати машинного навчання, які дозволяють автоматизовано виявляти аномальні події, приховані закономірності та тенденції зміни стану безпеки на основі великих обсягів телеметричних даних. Це ускладнює побудову адаптивних систем підтримки прийняття рішень у сфері кібербезпеки критичної інфраструктури.

У зв'язку з цим актуальним науково-прикладним завданням є розроблення гібридного підходу до оцінювання кіберризиків, який поєднує переваги багатокритеріальних моделей та методів машинного навчання. Такий підхід повинен забезпечувати одночасно інтерпретованість результатів, можливість експертного контролю, адаптивність до змін загрозового середовища та врахування невизначеності оцінок. Особливо важливим є створення механізмів інтеграції ML-індикаторів у структуру ризику без втрати прозорості моделі, що дозволить формувати більш об'єктивні та динамічні оцінки кіберризиків для об'єктів критичної інфраструктури.

ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є удосконалення та формалізація розширеної гібридної моделі оцінювання кіберризиків, яка поєднує інтерпретовану багатокритеріальну основу методу зваженої суми (WSM) із динамічними індикаторами машинного навчання. На відміну від традиційних підходів, що розглядають ризик переважно як функцію ймовірності та впливу, запропонована модель інтегрує додатковий критерій рівня вразливості як самостійну компоненту та передбачає механізм адаптивного коригування оцінки ризику на основі даних кіберзахисної телеметрії. Такий підхід дозволяє враховувати як статичні експертні оцінки, так і динамічні прояви загроз у реальному або наближеному до реального часу.

Наукова новизна роботи полягає у удосконаленні формалізованої гібридної моделі та методу оцінювання кіберризиків, у яких WSM використовується як базова структура агрегації критеріїв, а її функціональні можливості розширено за рахунок: інтеграції ML-індикаторів у компоненту оцінки ймовірності реалізації загрози; введення окремого критерію рівня вразливості з можливістю його декомпозиції на підкритерії; використання механізму гібридизації експертних та ML-оцінок на основі коефіцієнта довіри до ML-шару.

На відміну від існуючих моделей, де машинне навчання застосовується переважно як окремий інструмент класифікації або прогнозування, у запропонованому підході ML-індикатори інтегруються безпосередньо у формулу оцінювання ризику, що забезпечує узгодження між експертним і дано-орієнтованим підходами оцінювання та підвищує адаптивність моделі до динамічних змін кіберзагроз.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Одним із ключових обмежень класичних моделей оцінювання ризику, зокрема методу зваженої суми (WSM), є їх статичний характер, який передбачає використання фіксованих або експертно заданих значень критеріїв. Такий підхід забезпечує високу інтерпретованість, однак не дозволяє враховувати динамічні зміни у середовищі кіберзагроз. З іншого боку, методи машинного навчання здатні ефективно аналізувати потоки даних у реальному часі, проте часто позбавлені прозорості та складні для інтерпретації у контексті управління ризиками. У зв'язку з цим у даному дослідженні запропоновано гібридний підхід, що поєднує переваги обох методів.

Базою інтеграції виступає модель WSM, яка використовується для формування інтегральної оцінки ризику на основі множини критеріїв. При цьому результати роботи ML-модулів не замінюють класичні критерії, а використовуються для їх уточнення, зокрема для динамічного коригування оцінки ймовірності реалізації загрози. Такий підхід дозволяє зберегти інтерпретовану структуру моделі та одночасно підвищити її чутливість до змін у системі.

Запропонований метод оцінювання кіберризиків базується на інтеграції багатокритеріального підходу зваженої суми (WSM), методів машинного навчання та розширеної моделі вразливості, що дозволяє поєднати інтерпретованість класичних моделей із адаптивністю сучасних дано-орієнтованих підходів. Архітектурно модель побудована за модульним принципом, де кожен компонент відповідає окремому етапу обробки даних і формування інтегрального показника ризику.

На першому етапі формується множина сценаріїв ризику $S = S_i$ кожен з яких описує конкретну комбінацію активу, загрози та вразливості. Для кожного сценарію визначається вектор критеріїв $Q = (L_i, I_i, E_i, V_i)$, що включає як класичні параметри ризику (ймовірність і вплив), так і розширені характеристики (експозицію та рівень вразливості). Значення критеріїв формуються на основі комбінації експертних оцінок, історичних даних та технічної інформації з систем кіберзахисту.

Далі виконується нормалізація критеріїв з метою приведення їх до єдиної шкали [0,1], що забезпечує можливість коректного порівняння та агрегування. Нормалізовані значення формують матрицю $\bar{Q}$, яка використовується як вхід для багатокритеріального аналізу. На цьому етапі також визначається множина вагових коефіцієнтів W, що відображають відносну важливість кожного критерію. Значення ваг можуть задаватися експертно або визначатися на основі аналітичних методів.

Наступним кроком є формування матриці зважених компонентів p, яка отримується шляхом множення нормалізованих значень критеріїв на відповідні вагові коефіцієнти. Саме ця матриця є основою для обчислення інтегрального ризику за моделлю WSM, яка виступає центральним елементом запропонованої архітектури.

Адитивна структура цієї моделі забезпечує прозорість та інтерпретованість результатів, що є критично важливим для практичного застосування у системах управління ризиками.

Для підвищення адаптивності моделі до динамічного середовища кіберзагроз до архітектури інтегровано модуль машинного навчання. Цей модуль включає сукупність моделей, орієнтованих на аналіз поведінкових патернів, виявлення аномалій та прогнозування інцидентів на основі телеметричних даних (SIEM, XDR, журнали подій тощо). Результати роботи ML-моделей агрегуються у вигляді інтегрального індикатора Z^{ML}, який використовується для корекції оцінки ймовірності реалізації загрози. Таким чином, формується гібридна ймовірність, що поєднує експертну оцінку та дано-орієнтовану інформацію.

Окремим важливим компонентом архітектури є модель вразливості, яка дозволяє деталізувати технічний стан системи. На відміну від традиційних підходів, де вразливість враховується узагальнено, у даному дослідженні вона представлена як зважена композиція підкритеріїв, що включають CVSS-оцінки, доступність експлойтів, стан оновлень та конфігураційну безпеку. Такий підхід забезпечує більш точне відображення експлуатованості слабкостей та дозволяє інтегрувати дані з різних джерел кібербезпеки.

З метою підвищення аналітичної глибини оцінювання в архітектурі також передбачено використання альтернативних форм агрегування критеріїв — мультиплікативної та геометричної. Вони не замінюють базову модель WSM, а виконують допоміжну функцію, дозволяючи аналізувати взаємозалежність критеріїв та перевіряти стійкість отриманих результатів. Це особливо важливо у випадках, коли ризик визначається складними нелінійними залежностями між факторами.

Загалом архітектура запропонованої моделі забезпечує поєднання трьох ключових властивостей: інтерпретованості, адаптивності та масштабованості. Інтерпретованість досягається за рахунок використання WSM як базової моделі, адаптивність — завдяки інтеграції машинного навчання, а масштабованість — через модульну структуру, яка дозволяє розширювати набір критеріїв та джерел даних без зміни загальної логіки моделі. Такий підхід створює основу для побудови сучасних систем оцінювання кіберризиків, здатних ефективно функціонувати в умовах постійно змінюваного загрозового середовища.

Вибір критеріїв оцінювання кіберризику є ключовим етапом побудови адекватної та інтерпретованої моделі, оскільки саме вони визначають, які аспекти ризику будуть враховані та яким чином буде сформовано інтегральний показник.

У даному дослідженні за основу взято чотири критерії: ймовірність реалізації загрози L, критичність впливу I, рівень експозиції E та рівень вразливості V. Такий вибір зумовлений необхідністю поєднання класичних підходів оцінювання ризику з сучасними вимогами до врахування технічних характеристик кіберзагроз.

Критерії ймовірності та впливу традиційно використовуються у більшості методик оцінювання ризику (зокрема, ISO/IEC 27005, NIST, OWASP) і забезпечують базове уявлення про можливість реалізації загрози та масштаби її наслідків. Вони дозволяють формалізувати ризик як функцію двох основних факторів: частоти виникнення події та тяжкості її наслідків. Разом з тим, у сучасних кіберфізичних системах, зокрема в об'єктах критичної інфраструктури, така двокомпонентна модель є недостатньою, оскільки не враховує ані технічної природи вразливостей, ані умов їх експлуатації. [2,3].

Для усунення цього обмеження до моделі введено критерій експозиції E, який відображає ступінь доступності активу для потенційного зловмисника та рівень його захищеності. Цей критерій дозволяє врахувати вплив архітектури системи, наявність механізмів захисту, сегментацію мережі, політики доступу та інші фактори, що визначають, наскільки легко реалізувати атаку в конкретному середовищі. Таким чином, експозиція характеризує зовнішній контекст функціонування системи та виступає своєрідним "фільтром", який модифікує потенціал реалізації загрози.

Окремо у моделі виділено критерій рівня вразливості V, який відображає внутрішні технічні характеристики слабкостей системи. На відміну від експозиції, що описує середовище та умови доступу, вразливість визначає потенціал самої слабкості, зокрема її критичність, складність експлуатації та наявність інструментів для атаки. Включення цього критерію дозволяє перейти від абстрактної оцінки ризику до більш деталізованого опису технічного стану системи. При цьому вразливість моделюється як агрегований показник, сформований на основі підкритеріїв, таких як CVSS-оцінка, доступність експлойтів, стан оновлень та конфігураційна безпека. [4,5].

Важливим аспектом є те, що критерії експозиції та вразливості, незважаючи на їх взаємозв'язок, описують різні виміри ризику і не є дублюючими. Вразливість характеризує потенціал слабкості незалежно від контексту її використання, тоді як експозиція визначає можливість реалізації цієї слабкості у конкретному середовищі. Таке розділення дозволяє більш точно моделювати реальні умови функціонування системи, зокрема ситуації, коли наявність критичної вразливості не призводить до високого ризику через її ізоляцію, або навпаки — відносно незначна вразливість стає критичною за умов високої доступності.

Запропонований набір критеріїв забезпечує баланс між інтерпретованістю, яка є характерною для класичних моделей, та здатністю враховувати технічні та динамічні аспекти кіберзагроз. Крім того, така структура є придатною для інтеграції з методами машинного навчання, що дозволяє використовувати дані кіберзахисної телеметрії для уточнення оцінки окремих компонентів ризику. Таким чином, вибір критеріїв є обгрунтованим як з точки зору теоретичних підходів до оцінювання ризику, так і з позицій практичного застосування у сучасних інформаційно-комунікаційних системах.

Разом з тим, слід зазначити, що ефективність запропонованої системи критеріїв залежить від якості вхідних даних та коректності їх інтерпретації. Зокрема, визначення вагових коефіцієнтів та нормалізація показників можуть впливати на результати оцінювання, що потребує використання експертних знань або додаткових методів калібрування. Незважаючи на це, обрана структура критеріїв створює основу для побудови гнучкої, розширюваної та адаптивної моделі оцінювання кіберризиків.

Представимо розширену модель оцінювання кіберризику кортежем:

$$\Re^{ext} = \left\{ S, K, W, Q, \bar{Q}, P, R^{class}, M_{ML}, A, \Lambda, Z^{ML}, \bar{V}, R^{hyb} \right\} \tag{1}$$

де S - множина сценаріїв ризику, K - множина критеріїв, W - множина ваг критеріїв, Q - матриця оцінок, $\bar{Q}$ - матриця нормованих оцінок, P - матриця зважених компонентів, R^{class} — класичний ризик, — сукупність ML-модулів, A - множина ваг ML-модулів, Λ - множина коефіцієнтів довіри до ML, Z^{ML} — агрегований ML-індикатор, R^{hyb} — гібридний ризик.

Така декомпозиція узгоджується з NIST-логікою сценаріїв «джерело загрози → подія → вразливість/передумови → наслідок». [1,2].

Перший компонент моделі оцінювання ризиків є множина сценаріїв S кіберризику, яка визначається як:

$$S = \{S\}_{i=1}^{n},\ S_i = \langle A_i, T_i, U_i \rangle \tag{2}$$

де S_i - *i*-ий сценарій ризику; A_i — актив або група активів, для яких розглядається сценарій; T_i - загроза/подія або клас загроз; U_i - множина вразливостей або конкретна вразливість, релевантна для сценарію S_i. n - кількість сценаріїв.

Другий компонент K - множина критеріїв оцінювання для кожного сценарію S_i задається у вигляді:

$$K = \{K\}_{j=1}^{m} = \{L, I, E, V\},\ m = 4 \tag{3}$$

де, K_j - j-й критерій оцінювання ризику; $j = \{1, ..., m\}$ — індекс критерію; m - кількість критеріїв.

Причому: $K_1 = L, K_2 = I, K_3 = E, K_4 = V$ де, L - апріорна (експертна) чи гібридна ймовірність реалізації загрози. I - рівень критичності впливу. E - експозиція (рівень розриву контролів безпеки) V - рівень вразливості активу або компонента ІКС.

Третій компонент моделі включає в себе множину вагових коефіцієнтів критеріїв оцінювання W, визначається як

$$W = \{W_j\}_{j=1}^{m} = \{W_L, W_I, W_E, W_V\} \tag{4}$$

де W_i — ваговий коефіцієнт j-го критерію K_j; j — індекс критерію оцінювання; W_L — вага критерію ймовірності реалізації загрози L; W_I — вага критерію критичності впливу I; W_E — вага критерію експозиції або розриву контролів безпеки E; W_V — вага критерію рівня вразливості V

Як правило, ваги нормуються так, що:

$$\sum_{j=1}^{m} W_j = 1, \quad W_j \geq 0 \ \ \forall_j \in \{1,...,m\} \tag{5}$$

Для зручності подаємо ілюстративний вектор ваг, але позначимо його як приклад, а не як нормативно фіксований вибір,

Для кожного сценарію ризику S_i формується четвертий компонент - вектор оцінок, Q_i який містить значення всіх критеріїв, що використовуються у розширеній моделі оцінювання кіберризику і визначається як:

Для кожного сценарію формуємо вектор оцінок за критеріями для сценарію S_i визначається як:

$$Q = \{q_{ij}\}_{j=1}^{m} = \{L_i, I_i, E_i, V_i\}, \tag{6}$$

де, q_{ij} — значення j-го критерію для сценарію S_i; i — індекс сценарію ризику j — індекс критерію оцінювання; m — кількість критеріїв оцінювання, у даній моделі m =4; L_i — значення критерію ймовірності реалізації загрози для сценарію S_i I_i — значення критерію критичності впливу для сценарію S_i; E_i — значення критерію експозиції або розриву контролів безпеки для сценарію S_i; V_i — значення критерію рівня вразливості для сценарію S_i

Усі оцінки сценаріїв об'єднуються в одну матрицю Q, де кожен рядок відповідає окремому сценарію ризику, а кожен стовпець - окремому критерію оцінювання і це відображається за формулою:

$$Q = \{q_{ij}\}_{n \times m}, \ i = \overline{1,n}, \ j = \overline{1,m} \tag{7}$$

Оскільки в запропонованій моделі враховуються чотири критерії, такі як ймовірність реалізації загрози, вплив, експозиція та рівень вразливості, вектор Q_i має чотирикомпонентну структуру.

Для забезпечення сумісності з ML-індикаторами оцінки цих параметрів лінійно нормалізуємо до інтервалу [0; 1]: Так як критерії надходять із різних шкал, вводимо нормовану матрицю оцінок :

$$\overline{Q} = \{\overline{q_{ij}}\} \in [0,1]^{n \times m} \tag{8}$$

де, $\overline{Q}$ — матриця нормалізованих оцінок критеріїв; $\overline{q_{ij}}$ — нормалізоване значення j-го критерію для i-го сценарію ризику S_i; i — індекс сценарію ризику; j — індекс критерію оцінювання; n — кількість сценаріїв ризику; m — кількість критеріїв оцінювання; Нормалізована матриця $\overline{Q}$ формується з початкової матриці оцінок Q_i шляхом приведення всіх значень критеріїв до єдиної шкали $[0,1]$.

Кожен елемент $\overline{q_{ij}}$ відображає відносний рівень прояву відповідного критерію: значення, близькі до 0, відповідають низькому рівню ризику за даним критерієм, тоді як значення, близькі до 1, — високому рівню.

Нормалізовані значення знаходимо за формулою (використовуємо мінімаксний підхід щодо нормування), яка має вигляд:

$$\overline{q_{ij}} = \frac{q_{ij} - q_j^{\min}}{q_j^{\max} - q_j^{\min}}, \ i = \overline{1,n}, \ j = \overline{1,m} \tag{9}$$

Після формування нормалізованої матриці оцінок критеріїв $\overline{Q}$ наступним етапом є врахування відносної важливості кожного критерію. Для цього використовується множина вагових коефіцієнтів W, які відображають ступінь впливу відповідного критерію на інтегральну оцінку ризику.

На основі нормалізованих значень критеріїв та їх ваг формується матриця зважених компонентів P, яка є базою для подальшого обчислення класичного ризику.

Матриця зважених компонентів визначається виразом:

$$P = \left\{ p_{ij} \right\} \ , \ p_{ij} = W_j \overline{q_{ij}} \tag{10}$$

де, p_{ij} - зважений компонент за критерієм K_j W_j - ваговий коефіцієнт критерію K_j; $\overline{q_{ij}}$ - нормалізоване значення критерію K_j для сценарію S_i

У рамках WSM **класичний** (інтегральний) ризик визначається як сума зважених нормалізованих значень критеріїв

$$R_i^{class,WSM} = \sum_{j=1}^{m} W_j \overline{q_{ij}} = W_L \overline{L_i} + W_I \overline{I_i} + W_E \overline{E_i} + W_V \overline{V_i} \tag{11}$$

де, $R_i^{class,WSM}$ — інтегральна оцінка ризику для сценарію S_i за методом зваженої суми; W_j — ваговий коефіцієнт j -го критерію; $\overline{q_{ij}}$ — нормалізоване значення відповідного критерію; W_L, W_I, W_E, W_V — ваги критеріїв ймовірності, впливу, експозиції та вразливості відповідно; $\overline{L_i}, \overline{I_i}, \overline{E_i}$ $\overline{V_i}$ — значення критеріїв для сценарію S_i.

Такий підхід дозволяє отримати лінійну інтерпретовану модель, у якій зміна будь-якого параметра пропорційно впливає на підсумковий результат. [8].

У мультиплікативній моделі ризик визначається як добуток значень критеріїв:

$$R_i^{class,\times} = \overline{L}_i \cdot \overline{I}_i \cdot \overline{E}_i \cdot \overline{V}_i \tag{12}$$

де, $R_i^{class,\times}$ — оцінка ризику у мультиплікативній формі; $\overline{L_i}, \overline{I_i}, \overline{E_i}, \overline{V_i}$ — значення відповідних критеріїв.

З огляду на зазначені особливості, у дослідженні також використовується зважене геометричне середнє як узагальнюючий підхід, що поєднує властивості адитивної та мультиплікативної моделей. На відміну від мультиплікативної форми, геометричне середнє дозволяє зберегти ваговий внесок кожного критерію, що є важливим для багатокритеріального аналізу.

Зважене геометричне середнє визначається як:

$$R_i^{class,\times} = \overline{L}_i^{W_L} \cdot \overline{I}_i^{W_I} \cdot \overline{E}_i^{W_E} \cdot \overline{V}_i^{W_V} \tag{13}$$

де, $R_i^{class,\times}$ — оцінка ризику у формі зваженого геометричного середнього; W_L, W_I, W_E, W_V — вагові коефіцієнти критеріїв; $\overline{L_i}, \overline{I_i}, \overline{E_i}, \overline{V_i}$ — значення критеріїв для сценарію S_i.

Після визначення класичних форм оцінювання ризику доцільно детальніше розглянути критерії, що формують їх значення, зокрема критерій рівня вразливості, який відображає технічну можливість реалізації загрози та ступінь експлуатованості слабкостей системи. На відміну від класичних підходів, де вразливість розглядається неявно або включається до ймовірності реалізації загрози, у цьому дослідженні вона виділяється як окремий компонент моделі для підвищення точності та інтерпретованості оцінювання ризику. Оскільки вразливість має чітке технічне підґрунтя, базується на об'єктивних даних і є багатовимірним поняттям, її доцільно представляти як інтегральну характеристику, сформовану з кількох взаємодоповнюючих підкритеріїв.

Формально критерій рівня вразливості V_i для сценарію ризику S_i визначається як зважена композиція підкритеріїв::

$$V_i = \sum_{r=1}^{g} \beta_r v_{ir} \ , \ B = \left\{ \beta_r \right\}_{r=1}^{g} \tag{14}$$

Умова нормування ваг:

$$\sum_{r=1}^{g} \beta_r = 1 \ , \ \beta_r \geq 0 \tag{15}$$

le, V_i — інтегральний показник рівня вразливості для сценарію ризику S_i; v_{ir} — значення r-го підкритерію вразливості для сценарію S_i; r — індекс підкритерію вразливості, $r \in \{1,...,g\}$; g — кількість підкритеріїв, що використовуються для оцінювання вразливості; β_r — ваговий коефіцієнт r-го підкритерію; B — множина ваг підкритеріїв.

Така формалізація дозволяє гнучко налаштовувати модель залежно від доступності даних та специфіки досліджуваної системи.

У рамках запропонованого підходу доцільно використовувати чотири базові підкритерії, що відображають ключові аспекти вразливості:

$$(v_{i1}, v_{i2}, v_{i3}, v_{i4}) = (v_i^{cvss}, v_i^{\exp l}, v_i^{patch}, v_i^{cfg}) \tag{17}$$

де v_i^{cvss} — нормована оцінка CVSS, характеризує технічну критичність вразливості та базується на стандартизованій оцінці CVSS, яка враховує такі параметри, як складність атаки, необхідні привілеї та потенційний вплив; $v_i^{\exp l}$ — нормована ознака доступності експлойта / Exploit Maturity, відображає доступність експлойта або рівень Exploit Maturity і характеризує реальну можливість використання вразливості зловмисником. v_i^{patch} — патч-лаг або зворотна патч-комплаєнс, описує стан оновлень і дозволяє врахувати затримки у встановленні виправлень (patch lag), що безпосередньо впливає на рівень ризику — вразливість у системі secure-configuration, відображає наявність конфігураційних слабкостей, пов'язаних із недостатнім рівнем захисту, помилками налаштування або відсутністю механізмів hardening.

ML-модулі, індикатори та джерела даних.

У сучасних умовах зростання кіберзагроз традиційні експертні методи оцінювання ризику не забезпечують достатньої адаптивності до динамічних змін середовища, що обумовлює необхідність використання методів машинного навчання для аналізу телеметричних даних і формування кількісних індикаторів ризику. У запропонованому підході ML-модулі виступають додатковим аналітичним шаром, який на основі даних SIEM, XDR, журналів подій і мережевого трафіку формує динамічні індикатори ризику з урахуванням поведінкових аномалій і трендів загроз. [9]. Для цього використано п'ять груп ML-модулів — класифікацію, кластеризацію, виявлення аномалій, регресію та аналіз часових рядів, що дозволяє одночасно враховувати статичні й динамічні аспекти кіберризику, включаючи сегментацію об'єктів, прогнозування загроз і оцінювання ймовірності інцидентів.

Як приклади алгоритмів доцільно використовувати усталені методи: k-means, DBSCAN, LOF, Isolation Forest, Random Forest, gradient boosting та LSTM. Вони є стандартними в літературі й добре переносяться на задачі кібертелеметрії та профілювання поведінки. [11].

Множина ML-модулів визначається як

$$M_{ML} = \{M_k\}_{k=1}^{s},\ M_{ML} = \{M_{cls}, M_{clu}, M_{anom}, M_{reg}, M_{ts}\},\ \text{при}\ s = 5 \tag{18}$$

де, M_{ML} — множина моделей машинного навчання; M_k — k-й ML-модуль; k — індекс ML-модуля, $k \in \{1,...,s\}$; s — кількість ML-модулів, у даному випадку $s = 5$; M_{cls} — модуль класифікації (оцінка ймовірності інциденту); M_{clu} — модуль кластеризації (групування подібних сценаріїв); M_{anom} — модуль виявлення аномалій (аналіз відхилень поведінки); M_{reg} — модуль регресії (прогнозування числових показників ризику); M_{ts} — модуль аналізу часових рядів (виявлення трендів та динаміки ризику/

Кожен із модулів вирішує окрему задачу, яка пов'язана з різними аспектами оцінювання ризику. Наприклад, моделі класифікації дозволяють оцінити ймовірність виникнення інциденту, моделі аномалій виявляють підозрілу поведінку, а моделі часових рядів визначають тенденції зростання або зниження ризику. Таким чином, використання декількох типів моделей забезпечує комплексний аналіз кіберзагроз.

Кожний модуль генерує нормований індикатор:

$$I_{k,i} \in [0,1],\ k \in \{cls, clu, anom, reg, ts\} \tag{19}$$

де, $I_{k,i}$ — нормований індикатор, сформований k-м ML-модулем для сценарію S_i; i — індекс сценарію ризику;

Індикатори $I_{k,i}$ відображають результати роботи ML-моделей у вигляді узагальнених числових значень. Нормалізація до інтервалу $[0,1]$ забезпечує їх сумісність із багатокритеріальною моделлю та дозволяє використовувати їх для подальшої агрегації. На відміну від класичних експертних оцінок,

агрегований ML-індикатор формується динамічно на основі телеметричних даних та результатів автоматизованого аналізу.

Для забезпечення комплексного аналізу у моделі використовуються декілька типів ML-модулів, що відповідають різним аналітичним задачам, зокрема класифікації, кластеризації, виявленню аномалій, регресії та аналізу часових рядів. Кожен модуль формує окремий ML-індикатор та використовує відповідні джерела даних [10]. Узагальнену характеристику ML-модулів, сформованих індикаторів та типових джерел даних наведено в табл. 1.

Агрегований ML-індикатор сценарію визначається як:

$$Z_i^{ML} = \sum_{k=1}^{s} a_k I_{k,i} = a_{cls} I_{cls,i} + a_{clu} I_{clu,i} + a_{anom} I_{amom,i} + a_{reg} I_{reg,i} + a_{ts} I_{ts,i} \quad (20)$$

У межах запропонованої моделі використовуються п’ять основних ML-модулів, кожен із яких відповідає окремому типу аналітичної задачі (табл.1).

Інтеграція ML-модулів у модель оцінювання ризику дозволяє перейти від статичного до динамічного аналізу, забезпечуючи врахування поведінкових характеристик системи, виявлення прихованих закономірностей та підвищення точності оцінювання кіберризиків. На відміну від традиційних підходів, що базуються виключно на експертних оцінках або фіксованих наборах правил, використання методів машинного навчання дає можливість адаптивно аналізувати великі обсяги телеметричних даних, автоматично виявляти аномальні події, прогнозувати розвиток загроз та оцінювати тенденції зміни рівня ризику у часі.

Таблиця 1

Типи ML-модулів, індикатори ризику та джерела даних у гібридній моделі оцінювання кіберризиків

Тип задачі ML	Приклади методів	Сформований ML-індикатор	Нормалізація індикатора	Типові джерела даних
Класифікація	логістична логістична регресія, Random Forest, градієнтний бустинг	$I_{cls,i}$ - ймовірність того, що подія/сценарій є інцидентом або належить до високого класу ризику	[0,1], за потреби — калібрування ймовірностей	SIEM/XDR-системи, журнали подій безпеки, телеметрія кінцевих пристроїв, журнали автентифікації, мережеві події
Кластеризація	k-means, DBSCAN, ієрархічна кластеризація	$I_{clu,i}$ - ризиковий профіль кластера або індикатор “noise/new pattern”	[0,1], min-max або нормалізація за щільністю кластера	база конфігурацій активів (CMDB), інвентаризація активів, мережева топологія, графи залежностей сервісів
Виявлення аномалій	Isolation Forest, LOF, One-Class SVM, автоенкодери	$I_{amom,i}$ - інтенсивність відхилення від нормальної поведінки	([0,1], percentile scaling або robust scaling	телеметрія OT/ICS-систем, IDS/IPS, журнали міжмережевих екранів, мережевий трафік, журнали процесів
Регресія	лінійна регресія, Random Forest Regressor, Gradient Boosting Regressor	$I_{reg,i}$ - нормований прогноз кількісного показників, таких як збиток, частота інцидентів, час простою	[0,1], відносно референтного масштабу	історичні інциденти, SOC/ticketing-системи, журнали простоїв, статистика збитків, база вразливостей
Часові ряди	ARIMA, LSTM	($I_{ts,i}$) - короткостроковий прогноз тренду ризику/інцидентів/вразливостей	[0,1], min-max або відносно критичного baseline	часові ряди подій безпеки, журнали оновлень, динаміка вразливостей, тренди експозиції та ризику

Після формування агрегованого ML-індикатора Z_i^{ML} наступним етапом є його інтеграція у базову WSM-модель. У запропонованому підході ML-шар не замінює експертну оцінку ризику повністю, а використовується для адаптивного уточнення окремих компонентів моделі. Насамперед це стосується критерію ймовірності реалізації загрози L_i, оскільки саме ймовірність найбільш чутлива до поточних змін у середовищі кіберзагроз, появи аномалій, зростання кількості інцидентів або активності зловмисника.

Для керування ступенем впливу ML-шару вводиться коефіцієнт довіри λ_i, який задається для кожного сценарію ризику:

$$\Lambda = \{\lambda_i\}_{i=1}^{n},\ 0 \le \lambda_i \le 1 \quad (21)$$

де Λ — множина коефіцієнтів довіри до ML-шару для всіх сценаріїв ризику; λ_i — коефіцієнт довіри до ML-індикатора для сценарію S_i; n — загальна кількість сценаріїв; $0 \le \lambda_i \le 1$ — умова, яка означає, що вплив ML-шару може змінюватися від повної відсутності впливу до повного домінування ML-оцінки.

Коефіцієнт довіри λ_i показує, яку частку в гібридній оцінці ймовірності займає ML-компонент. Якщо λ_i=0, то модель повністю спирається на експертну або базову оцінку L_i. Якщо λ_i=1, то ймовірність

визначається лише ML-індикатором Z_i^{ML}. У практичних умовах найчастіше використовується проміжне значення λ_i, що дозволяє поєднати стабільність експертного підходу з адаптивністю машинного навчання.

На основі цього коефіцієнта гібридна ймовірність реалізації загрози визначається як:

$$L_i^{hyb} = (1-\lambda_i)\overline{L_i} + \lambda_i Z_i^{ML} \tag{22}$$

де L_i^{hyb} — гібридна оцінка ймовірності реалізації загрози для сценарію S_i; $\overline{L_i}$ — нормалізована базова, експертна або статистична оцінка ймовірності реалізації загрози; Z_i^{ML} — агрегований ML-індикатор для сценарію (S_i); λ_i — коефіцієнт довіри до ML-шару; $(1-\lambda_i)$ — частка впливу базової експертної оцінки; $\lambda_i Z_i^{ML}$ — частка впливу ML-індикатора.

Ця формула є механізмом зваженого поєднання двох джерел інформації: експертного знання та поточної телеметрії. Її перевага полягає в тому, що вона не руйнує базову WSM-структуру, а лише уточнює один із її критеріїв — ймовірність. Такий підхід особливо важливий для кіберризиків, оскільки ймовірність атаки може швидко змінюватися залежно від поточних подій у системі.

Після отримання гібридної ймовірності вона підставляється у базову модель WSM. Тоді результуючий **гібридний ризик у формі WSM** визначається як:

$$R_i^{hyb,WSM} = W_L L_i^{hyb} + W_I \overline{I_i} + W_E \overline{E_i} + W_V \overline{V_i} \tag{23}$$

де, $R_i^{hyb,WSM}$ — гібридна оцінка ризику для сценарію, обчислена за методом WSM; W_L, W_I, W_E, W_V — ваги критеріїв ймовірності, впливу, експозиції та вразливості відповідно; $\overline{I_i}, \overline{E_i}, \overline{V_i}$ — нормалізовані значення критеріїв для сценарію S_i. L_i^{hyb} — гібридна ймовірність реалізації загрози.

У цій формулі WSM залишається базовою моделлю інтеграції критеріїв, а ML впливає лише на ту компоненту, яка справді потребує динамічного уточнення. Саме це забезпечує баланс між пояснюваністю та адаптивністю: результат можна інтерпретувати через окремі критерії та їхні ваги, але він уже враховує поточний стан кіберзагроз.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ У ДАНОМУ НАПРЯМКУ

У статті запропоновано гібридний підхід до оцінювання кіберризиків об'єктів критичної інфраструктури, який поєднує багатокритеріальну модель зваженої суми (WSM) та методи машинного навчання. Основою моделі виступає WSM-підхід, що забезпечує прозорість структури ризику, інтерпретованість результатів та можливість експертного налаштування критеріїв і вагових коефіцієнтів. На відміну від традиційних статичних моделей, запропонований підхід дозволяє враховувати динамічний стан кіберзагроз за рахунок інтеграції ML-індикаторів, сформованих на основі телеметричних даних систем моніторингу безпеки.

У межах дослідження формалізовано розширену структуру ризику, яка, крім класичних критеріїв ймовірності та впливу, включає критерії експозиції та рівня вразливості. Показано, що таке розширення дозволяє більш повно врахувати технічний стан системи, рівень доступності активів для потенційного зловмисника та реальні умови експлуатації вразливостей. Критерій вразливості представлено як агреговану композицію підкритеріїв, що включають CVSS-оцінки, доступність експлойтів, стан оновлень і конфігураційну безпеку.

У роботі також реалізовано механізм гібридизації WSM та ML, у межах якого результати роботи ML-модулів використовуються для адаптивного уточнення оцінки ймовірності реалізації загрози та, за необхідності, критичності впливу. Запропоновано механізм коефіцієнта довіри до ML-шару, який дозволяє збалансувати вплив експертних оцінок і результатів автоматизованого аналізу залежно від якості моделей та повноти телеметричних даних. Це забезпечує контрольовану інтеграцію машинного навчання без втрати пояснюваності моделі.

Додатково у статті формалізовано використання методів класифікації, кластеризації, виявлення аномалій, регресії та аналізу часових рядів для формування агрегованого ML-індикатора. Для кожного типу задач визначено математичні моделі, джерела даних та механізми нормалізації індикаторів, що забезпечує цілісність і узгодженість гібридної архітектури оцінювання ризику.

Практична цінність запропонованого підходу полягає у можливості побудови адаптивних систем оцінювання кіберризиків для критичної інфраструктури, здатних поєднувати експертні знання, технічні характеристики вразливостей та дані оперативного кібермоніторингу. Запропонована модель може бути використана як основа для систем підтримки прийняття рішень у сфері кібербезпеки, а також для подальшого розвитку інтелектуальних платформ управління ризиками.

Перспективами подальших досліджень є апробація моделі на реальних наборах телеметричних даних, автоматизоване налаштування вагових коефіцієнтів критеріїв і ML-модулів, а також розширення моделі за рахунок інтеграції графових методів аналізу атак, методів зрозумілий AI та механізмів онлайн-навчання для адаптації до нових типів кіберзагроз.

Література

1. National Institute of Standards and Technology. *Guide for Applying the Risk Management Framework to Federal Information Systems and Organizations : SP 800-37 Rev. 2* [Електронний ресурс]. – Gaithersburg, MD: NIST, 2018. – URL: https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final.
2. National Institute of Standards and Technology. *Security and Privacy Controls for Information Systems and Organizations: SP 800-53 Rev. 5* [Електронний ресурс]. – Gaithersburg, MD: NIST, 2020. – URL: https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final.
3. OWASP Foundation. *OWASP Risk Rating Methodology* [Електронний ресурс]. – URL: https://owasp.org/www-community/OWASP_Risk_Rating_Methodology.
4. *Common Vulnerability Scoring System Version 4.0: Specification Document* [Електронний ресурс]. – 2023. – Режим доступу: https://www.first.org/cvss/v4-0/cvss-v40-specification.pdf.
5. Scikit-learn developers. *Probability calibration* [Електронний ресурс]. – URL: https://scikit-learn.org/stable/modules/calibration.html.
6. Microsoft. *Connect data sources to Microsoft Sentinel* [Електронний ресурс]. – Режим доступу: https://learn.microsoft.com/en-us/azure/sentinel/connect-data-sources.
7. Zadeh L. A. Fuzzy sets / L. A. Zadeh // *Information and Control*. – 1965. – Vol. 8, No. 3. – P. 338–353. – Режим доступу: https://www.sciencedirect.com/science/article/pii/S001999586590241X.
8. Ільєнко А., Телющенко В. Методи оцінювання ризиків кіберзагроз в інформаційно-комунікаційних системах об'єктів цивільної авіації // *Безпека інформаційних систем і технологій*. – 2025. – № 2(10). – С. 5–15. – DOI: https://doi.org/10.17721/ISTS.2025.10.5-15.
9. Ільєнко А., Кривокульська О., Яковенко О., Телющенко В. Інтелектуальні технології у кібербезпеці: аналіз потенціалу та викликів застосування штучного інтелекту // *Кібербезпека: освіта, наука, техніка*. – 2026. – № 4(32). – С. 711–723. – DOI: https://doi.org/10.28925/2663-4023.2026.32.1139.
10. Ільєнко А., Ільєнко С., Телющенко В., Маляренко С. Теоретичний підхід застосування методів машинного навчання для оцінки ризиків об'єктів критичної інфраструктури // *Наукоємні технології*. – 2026. – Т. 69, № 1. – С. 35–47. – DOI: https://doi.org/10.18372/2310-5461.69.20945/
11. Bellman R. E. Decision-making in a fuzzy environment / R. E. Bellman, L. A. Zadeh // *Management Science*. – 1970. – Vol. 17, No. 4. – P. B-141–B-164. – Режим доступу: https://projecteuclid.org/download/pdf_1/euclid.bsmsp/1200512992.

References

1. National Institute of Standards and Technology. *Guide for Applying the Risk Management Framework to Federal Information Systems and Organizations : SP 800-37 Rev. 2* [Електронний ресурс]. – Gaithersburg, MD: NIST, 2018. – URL: https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final.
2. National Institute of Standards and Technology. *Security and Privacy Controls for Information Systems and Organizations: SP 800-53 Rev. 5* [Електронний ресурс]. – Gaithersburg, MD: NIST, 2020. – URL: https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final.
3. OWASP Foundation. *OWASP Risk Rating Methodology* [Електронний ресурс]. – URL: https://owasp.org/www-community/OWASP_Risk_Rating_Methodology.
4. *Common Vulnerability Scoring System Version 4.0: Specification Document* [Електронний ресурс]. – 2023. – Режим доступу: https://www.first.org/cvss/v4-0/cvss-v40-specification.pdf (дата звернення: 14.08.2026).
5. Scikit-learn developers. *Probability calibration* [Електронний ресурс]. – URL: https://scikit-learn.org/stable/modules/calibration.html.
6. Microsoft. *Connect data sources to Microsoft Sentinel* [Електронний ресурс]. – Режим доступу: https://learn.microsoft.com/en-us/azure/sentinel/connect-data-sources.
7. Zadeh L. A. Fuzzy sets / L. A. Zadeh // *Information and Control*. – 1965. – Vol. 8, No. 3. – P. 338–353. – Режим доступу: https://www.sciencedirect.com/science/article/pii/S001999586590241X.
8. Ilienko A., Teliushchenko V. Methods for Assessing the Risks of Cyber Threats in Information and Communication Systems of Civil Aviation Facilities // Information Systems and Technologies Security. – 2025. – Vol. 2, No. 10. – P. 5–15. – DOI: https://doi.org/10.17721/ISTS.2025.10.5-15.
9. Ilienko A., Kryvokulska O., Yakovenko O., Teliushchenko V. Intelligent Technologies in Cybersecurity: Analysis of the Potential and Challenges of the Application of Artificial Intelligence // Cybersecurity: Education, Science, Technique. – 2026. – Vol. 4, No. 32. – P. 711–723. – DOI: https://doi.org/10.28925/2663-4023.2026.32.1139.
10. Ilienko A., Ilienko S., Teliushchenko V., Maliarenko S. Theoretical Approach to the Application of Machine Learning Methods for Assessing the Risks of Critical Infrastructure Facilities // Science-Based Technologies. – 2026. – Vol. 69, No. 1. – P. 35–47. – DOI: https://doi.org/10.18372/2310-5461.69.20945/
11. Bellman R. E. Decision-making in a fuzzy environment / R. E. Bellman, L. A. Zadeh // *Management Science*. – 1970. – Vol. 17, No. 4. – P. B-141–B-164. – Режим доступу: https://projecteuclid.org/download/pdf_1/euclid.bsmsp/1200512992.