КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (29), 2025
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2025.29.948
УДК 004.056:004.8:004.75(045)

Гречанінов Віктор Федорович
кандидат технічних наук, старший дослідник,
завідувач науково-дослідного відділу
Інститут проблем математичних машин і систем НАН України, м. Київ, Україна
ORCID ID: 0000-0001-6268-3204
vgrechaninov@gmail.com

МОДЕЛІ ТА ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ КРИТИЧНОЇ ІНФРАСТРУКТУРИ ДЛЯ ПІДВИЩЕННЯ СТІЙКОСТІ

Анотація. У статті обґрунтовано доцільність використання сучасних інформаційних технологій для захисту стійкого функціонування об’єктів критичної інфраструктури (ОКІ), із фокусом на забезпечення кіберзахисту їхніх інформаційних систем, що є ключовим чинником національної безпеки та стійкості держави до гібридних загроз. Показано, що підвищення рівня захищеності й відновлюваності систем ОКІ можливе завдяки створенню розгалуженої мережі кризових центрів, інтегрованих із платформами моніторингу, виявлення та реагування на кіберінциденти у режимі реального часу. Особливу увагу приділено сценарному моделюванню, яке забезпечує прогнозування варіантів розвитку кібератак, формування моделей управління безпекою та підтримку процесу прийняття рішень. Такий підхід дає змогу завчасно визначати ймовірні канали впливу на системи, оцінювати наслідки їх порушення та генерувати оптимальні стратегії нейтралізації загроз. Запропоновано архітектуру багаторівневої системи управління захистом і відновленням роботи інформаційних систем ОКІ, що враховує як фізичні, так і кібернетичні ризики. Вона базується на інтеграції інтелектуальних технологій, здатних забезпечувати адаптивне реагування на зміни середовища та автоматизовану підтримку процедур резервування й відновлення. Обґрунтовано доцільність використання штучного інтелекту в кризових центрах, зокрема агентних систем, що підвищують ефективність аналізу великих масивів даних, виявлення аномалій у трафіку, оцінки ризиків і вироблення управлінських рекомендацій. Застосування інтелектуальних агентів забезпечує швидкість і точність у локалізації кіберзагроз, що суттєво підвищує стійкість інформаційних систем критичної інфраструктури та формує основу для проактивних механізмів кіберзахисту.

Ключові слова: критична інфраструктура, інформаційні системи, кіберзахист, стійке функціонування, кризові (ситуаційні) центри, сценарне моделювання, штучний інтелект, інтелектуальні агенти, системи підтримки прийняття рішень.

ВСТУП

Актуальність проблематики захисту інформаційних систем об’єктів критичної інфраструктури (ОКІ) зумовлена високим рівнем залежності державних і приватних секторів від їх безперервного та надійного функціонування. У сучасному світі інформаційні системи забезпечують управління технологічними процесами, моніторинг стану виробничих об’єктів, підтримку комунікацій та координацію діяльності у сфері безпеки, тому їхня вразливість стає фактором стратегічних ризиків.

Особливістю сучасних викликів є зростання кількості кібератак на ОКІ, які призводять до порушення роботи енергетичних систем, транспортної інфраструктури, медичних установ та фінансового сектору [1, 3]. Атаки на інформаційні системи відрізняються комплексністю, швидкістю поширення та здатністю викликати каскадні

наслідки [2, 5]. Це визначає потребу у формуванні стійких механізмів протидії, що базуються не лише на класичних засобах захисту, але й на інтелектуальних методах аналізу, прогнозування та реагування.

Науково-практичним завданням у цій сфері стає розробка моделей та технологій, орієнтованих на захист інформаційних систем критичної інфраструктури з урахуванням сучасних кіберзагроз [8-10]. Йдеться про побудову адаптивних систем, здатних підтримувати стійкість функціонування ІС навіть за умов інтенсивного інформаційного та фізичного впливу. Важливе значення при цьому мають кризові (ситуаційні) центри, які забезпечують оперативну координацію дій, моніторинг інцидентів та інтеграцію рішень штучного інтелекту.

Серед сучасних підходів до забезпечення кіберзахисту особливу роль відіграє сценарне моделювання, що дозволяє здійснювати імітацію можливих варіантів розвитку кризових ситуацій. Використання агентів штучного інтелекту в таких моделях створює умови для прогнозування динаміки кібератак, оптимізації управлінських рішень та формування стратегій відновлення функціонування систем [4-6]. Таким чином, інтелектуальні технології стають основою нових методів захисту, спрямованих на мінімізацію ризиків та підвищення рівня стійкості ОКІ.

Згідно з положеннями Закону України «Про критичну інфраструктуру» (2021), а також міжнародних стандартів ISO/IEC 27001, ISO/IEC 27005 та NIST SP 800-30 [2, 7, 13], питання забезпечення стійкого функціонування інформаційних систем є пріоритетним напрямом національної та глобальної безпеки. Це визначає потребу в інтеграції нормативно-правових вимог, організаційних заходів і сучасних науково-технологічних рішень у сфері кіберзахисту.

У статті вперше запропоновано інтегровану архітектуру інтелектуального захисту інформаційних систем критичної інфраструктури, яка поєднує кризово-ситуаційні центри, сценарне ядро та агентні технології штучного інтелекту [1, 3, 9, 11]. Ключовою особливістю є те, що сценарне прогнозування безпосередньо пов’язане з керуванням ресурсами відновлення завдяки використанню цифрових двійників процесів, що забезпечує адаптивність та відтворюваність прийнятих рішень.

Розроблено єдиний контур керування стійкістю, де узгоджені метрики MTTD, MTTR та ILP інтегруються у композитний показник ρ. На відміну від класичних підходів, ці показники виконують функцію не лише постфактум-індикаторів, але й регуляторів у MDP/RL-політиках, що робить систему керованою та навчальною у режимі реального часу.

Наукова новизна також полягає у формальній зшивці різнорідних моделей — часових мереж Петрі для інцидент-менеджменту [12], черг із пріоритетами для операційної логістики [2] та нечіткого скорингу ризику для пріоритизації інцидентів [5] — в один вимірюваний контур SLA. Такий підхід створює прозорий механізм контролю стійкості та обґрунтованого розподілу ресурсів реагування [6, 10, 13]. Додатковим внеском є метод сценарного моделювання, який враховує міжзалежності між ОТ- і ІТ-сегментами та впливи з боку ланцюгів постачання. Використання цифрових двійників дозволяє валідувати політики реагування та логістику відновлення у симуляційному середовищі, зменшуючи ризики хибних рішень та підвищуючи готовність систем до динамічних багатоканальних загроз. Таким чином, запропоновані моделі й технології формують науково новий підхід до проактивного, метрико-керованого управління стійкістю інформаційних систем критичної інфраструктури.

Постановка проблеми. Захист інформаційних систем об’єктів критичної інфраструктури (ОКІ) в умовах сучасних загроз є одним із ключових завдань державної

безпеки та стійкого розвитку суспільства [7, 13]. Вразливість інформаційних систем, які керують технологічними процесами, координують роботу енергетики, транспорту, медицини чи фінансового сектору, створює потенційну можливість масштабних збоїв, економічних втрат і негативних соціальних наслідків.

Аналіз міжнародного досвіду та практики України показує, що традиційні підходи до забезпечення кіберзахисту ОКІ здебільшого орієнтовані на реактивні заходи – виявлення та ліквідацію наслідків інцидентів [8]. Проте сучасні кібератаки характеризуються високим рівнем координації, застосуванням складних сценаріїв і використанням уразливостей у SCADA/ICS-системах [3], що робить звичайні засоби захисту малоефективними [2, 10]. Додатковим ускладнюючим чинником є зростання кількості взаємозалежних ланцюгів постачання, які можуть стати джерелом каскадних збоїв при порушенні роботи одного з елементів.

Існує також низка організаційних та технологічних проблем. По-перше, відсутність інтегрованої системи управління захистом інформаційних систем у кризових ситуаціях знижує ефективність прийняття рішень [8]. По-друге, обмеженість ресурсів і недостатній рівень автоматизації створюють ризик неузгоджених дій під час реагування на атаки [7, 12]. По-третє, збереження залежності від людського фактора в умовах дефіциту часу підвищує ймовірність управлінських помилок.

У зв’язку з цим виникає потреба у створенні моделей і технологій інтелектуального захисту, які б поєднували можливості сценарного моделювання [11], інтелектуальних агентів [4] і технологій штучного інтелекту [1, 5-6] для моніторингу, прогнозування та запобігання загрозам [7-9, 13]. Такий підхід дозволяє забезпечити проактивність у протидії кібератакам, оптимізувати розподіл ресурсів і підвищити рівень стійкості функціонування інформаційних систем критичної інфраструктури [14]. Крім того, він створює основу для інтеграції механізмів автоматизованого прийняття рішень у кризових центрах, що сприяє скороченню часу реагування на інциденти. Запровадження таких рішень формує нову парадигму управління кіберзахистом, орієнтовану на динамічну адаптацію до змін середовища та мінімізацію наслідків атак.

Таким чином, проблема дослідження полягає у відсутності комплексної моделі та технологічної архітектури інтелектуального захисту інформаційних систем ОКІ, яка б враховувала динамічний характер сучасних кіберзагроз, забезпечувала підтримку прийняття рішень у кризових умовах і сприяла підвищенню стійкості функціонування критично важливих об’єктів.

Аналіз останніх досліджень і публікацій. Упродовж останніх років спостерігається стале зростання наукового інтересу до застосування інтелектуальних технологій у захисті інформаційних систем об’єктів критичної інфраструктури, що відображає перехід від реактивних до проактивних парадигм кібербезпеки. Репрезентативним прикладом є робота Амджада Рехмана та співавт. [1], де запропоновано гібридну рамку, що поєднує нечітку логіку та федеративне навчання для підвищення захисту у високодинамічних, розподілених середовищах (IoT-орієнтовані транзакції в метапросторі). Вказане поєднання дає змогу адаптувати моделі без централізованого збирання даних, зменшуючи як ризики конфіденційності, так і затрати на синхронізацію, та є релевантним для ОКІ з територіально розподіленою інфраструктурою.

Вагомий внесок у формалізацію ризик-менеджменту для складних, багатовекторних загроз запропоновано Сергієм Зибіним, Олександром Корченком, Олександром Користіним та співавт. [2], де на основі теорії нечітких множин обґрунтовано метод оцінювання гібридних загроз, здатний інтегрувати експертні

судження та неоднорідні дані. Отримані результати є методологічною основою для побудови адаптивних механізмів підтримки прийняття рішень у системах кіберзахисту ОКІ. На рівні прикладних протоколів інтернету речей Чайон Кім та співавт. [3] демонструють ефективність нечіткої логіки у задачах виявлення вторгнень, пропонуючи схему FLSec-RPL проти атак на RPL-мережі; ці результати ілюструють доцільність використання нечітких правил у мережевих сценаріях з обмеженими ресурсами, типовими для сегментів ОКІ.

Паралельно розвивається напрям агентного штучного інтелекту як технологічна основа автономного виявлення та реагування. Нір Кшетрі [4] концептуалізує трансформаційний потенціал агентного ШІ в протидії емерджентним загрозам, акцентуючи на здатності агентів до координації дій, адаптації політик і самостійного планування у реальному часі. Комплементарно Алі Каракая [5] показує, що поєднання ансамблевого навчання з нечіткою логікою підвищує стійкість систем оцінювання безпеки IoT до шуму та концептуального дрейфу даних, що є критично важливим для гетерогенних середовищ ОКІ. Практичну реалізацію агентних підходів у вигляді відкритої, модульної архітектури представляє Віктор Майораль-Вільчес та співавт. [6]: система CAI орієнтована на інтеграцію інструментів перевірки безпеки та автоматизоване виконання складених «ланцюжків дій», демонструючи перспективність агентів для застосувань від CTF до елементів оперативного моніторингу.

Український контекст проблематики, включно з організаційно-правовими та операційними аспектами забезпечення стійкості під час воєнного стану, ґрунтовно висвітлено у праці Андрія Гавриса, Валентини Філіппової та Наталії Тур [7]. Автори узагальнюють профілі загроз для ОКІ та окреслюють потреби модернізації систем захисту, що створює емпіричне тло для інтеграції інтелектуальних технологій у національні моделі кіберзахисту. У цьому контексті особливої ваги набуває питання синхронізації національних рішень із міжнародними стандартами та рамками безпеки, що забезпечує узгодженість підходів у сфері протидії гібридним загрозам. Таким чином, результати дослідження розкривають не лише теоретичну, а й практичну значущість запропонованих моделей для підвищення стійкості інформаційних систем України в умовах воєнних викликів.

Сукупність проаналізованих наукових публікацій демонструє, що нечіткі підходи, федеративне навчання, ансамблеві методи та агентний ШІ формують доведено ефективні компоненти інтелектуального захисту ІС ОКІ. Проте на рівні системної інженерії все ще бракує комплексної інтегрованої архітектури, яка б поєднувала ці методи з інструментарієм сценарного моделювання, механізмами кризових (ситуаційних) центрів та процедурами відновлення, забезпечуючи узгоджене управління ризиками і ресурсоорієнтоване реагування в умовах багатоканальних та мультиагентних загроз.

Мета статті. Метою статті є розробка моделей та технологій інтелектуального захисту інформаційних систем об’єктів критичної інфраструктури, що забезпечують підвищення їх стійкого функціонування в умовах сучасних загроз. У межах досягнення цієї мети передбачено обґрунтування доцільності застосування інтелектуальних методів у сфері кіберзахисту інформаційних систем критичної інфраструктури, аналіз можливостей сценарного моделювання для прогнозування й нейтралізації кібератак, а також формування архітектурних підходів до побудови систем управління захистом і відновленням критично важливих інформаційних ресурсів. Окрему увагу приділено визначенню ролі кризових (ситуаційних) центрів та агентів штучного інтелекту, здатних підвищити ефективність моніторингу, аналізу ризиків і підтримки прийняття

управлінських рішень у режимі реального часу, що у комплексі створює основу для підвищення стійкості інформаційних систем об'єктів критичної інфраструктури.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

У дослідженні використано взаємодоповнювальний набір методів системної інженерії та прикладної математики [2, 7], що разом утворюють єдиний контур моделювання, вимірювання і оптимізації стійкості ІС ОКІ. Спершу виконуємо системний аналіз і декомпозицію функцій захисту на OT/IT-шарах: це дає формальну карту активів, потоків даних і точок контролю, від яких далі залежать метрики MTTD, MTTR та ILP [9-11-13]. На цій основі проводимо сценарний аналіз і дискретно-подієве імітаційне моделювання для відтворення каскадних ефектів і деградаційних режимів; таким чином перевіряємо, як різні класи загроз впливають на часові вікна виявлення/відновлення та втрати продуктивності.

Робочі потоки інцидент-менеджменту формалізуємо часовою мережею Петрі [8, 12]:

$$N = (P, T, F, W, M_0, \tau), \quad (1)$$

де P – множина місць (стани процесу: $Det, Tri, Iso, Fix, Restore$), T – множина переходів (події: $t_{det}, t_{tri}, t_{iso}, t_{fix}, t_{close}$), $F \subseteq (P \times T) \cup (T \times P)$ – дуги, $W: F \to N$ – ваги дуг, $M_0: P \to N$ – початкова розмітка (токени відповідають активним інцидентам), $\tau: T \to Dist(R_{>0})$ – часові затримки переходів (експоненційні або емпіричні). Перехід дозволений (enabled), якщо на всіх його вхідних місцях кількість токенів не менша за ваги відповідних дуг, спрацьовування переводить токени на вихідні місця згідно з F та W. Досяжність місця $Restore$ з будь-якої досяжної розмітки гарантує завершуваність плейбука, T-інваріанти контролюють збереження потоку (усі гілки завершуються), а P-інваріанти – відсутність «витоків» токенів (інциденти не «зависають» поза процесом). Часові параметри безпосередньо зв'язуються з метриками: затримка на t_{det} інтерпретує MTTD, затримки на t_{fix} та t_{close} (або $t_{restore}$) – компонент T_s у складі MTTR, накопичення токенів у місцях Tri та Fix відображає чергу очікування та узгоджується з оцінкою W_q у моделі черг (отже, $MTTR \approx W_q + T_s$) [11]. Структурний аналіз мережі (обмеженість, живучість, відсутність блокувань) дозволяє виявляти «вузькі місця» (наприклад, брак ресурсу на Tri або слабку паралелізацію на Fix) і синтезувати коригувальні політики: додавання виконавців/плейбуків на критичних місцях, введення паралельних гілок Fix, або пріоритетних дуг для категорій A/B/C (за потреби – кольорові токени з атрибутом критичності та SLA-порогів) [13, 16]. Така формалізація робить життєвий цикл інциденту прозорим для аналізу, а параметри мережі – керованими важелями досягнення цільових порогів MTTD/MTTR/ILP.

Якщо для мережі $N = (P, T, F, W, M_0, \tau)$ виконується обмеженість розмітки, існують T-інваріанти, що покривають переходи $t_{det}, t_{tri}, t_{iso}, t_{fix}, t_{close}$, та немає заблокованих підмереж, тоді місце *Restore* досяжне з будь-якої розмітки, а життєвий цикл інциденту завершується скінченно [2-3]. Це гарантує відсутність нескінченної циркуляції токенів у часткових підмережах і забезпечує прогрес процесу до відновлення [9]. Формально це означає, що система інцидент-менеджменту є завершуваною і володіє властивістю гарантованої досяжності стану відновлення (*Restore*).

Для будь-якої часової мережі Петрі інваріанти забезпечують консервативність потоку токенів, що формально виражається рівнянням [11, 14-15]:

$$M = M_0 + C \cdot \sigma, \tag{2}$$

де M – поточна розмітка, M_0 – початкова розмітка, C – інцидентна матриця мережі, σ – вектор кількостей спрацьовувань переходів. Це співвідношення гарантує збереження балансу токенів у системі та виключає їх втрату чи неконтрольоване накопичення [12]. Обмеженість розмітки означає, що кількість токенів у будь-якому місці завжди залишається скінченною, а відсутність заблокованих підмереж виключає «зависання» інцидентів у недосяжних станах [7, 10]. Таким чином, усі дозволені переходи поступово приводять систему до місця *Restore*, що забезпечує завершуваність життєвого циклу інциденту та його коректну обробку.

Рис. 1. відображає послідовність переходів від моменту виявлення проблеми до її повного усунення й відновлення роботи системи [13, 15-16]. Інцидент фіксується у стані Det, після чого через перехід t_{det} з урахуванням затримки MTTD відбувається перехід до Tri (тріаж). Далі перехід t_{tri} переводить систему в Iso (ізоляція), що локалізує інцидент, а t_{iso} – у Fix (усунення), де враховується час MTTR. Завершальний перехід t_{fix}, приводить до стану *Restore*, що означає відновлення сервісу відповідно до SLO. Передбачено також пряму ескалацію з тріажу до фіксу при високому пріоритеті [1, 4-6]. Модель підтверджує, що за умови обмеженості розмітки, покриття переходів T-інваріантами та відсутності заблокованих підмереж життєвий цикл інциденту гарантовано завершується у стані *Restore*.

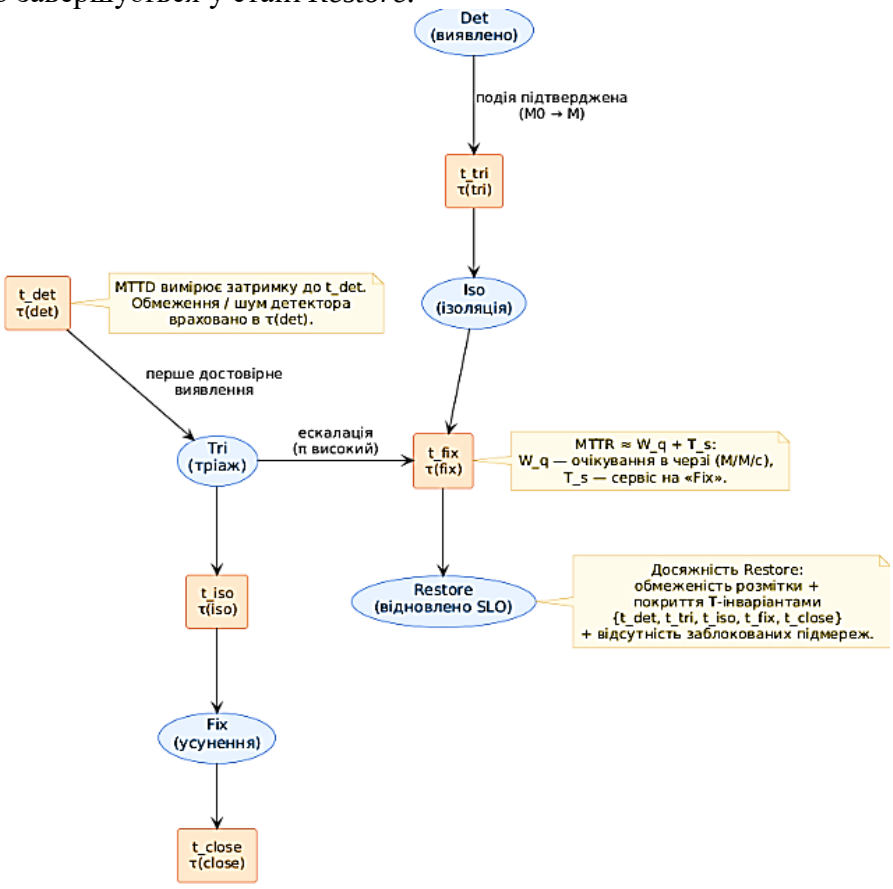

Рис. 1. Часова мережа Петрі життєвого циклу інциденту

ISSN 2663 - 4023

Канал реагування моделюємо системою $M/M/c$ з пріоритетами (preemptive-resume для A, non-preemptive для B і C). Середній час відновлення для категорії $k \in \{A, B, C\}$:

$$MTTR_k \approx W_{q,k} + T_{s,k}, \quad (3)$$

де $W_{q,k}$ – середній час очікування в черзі, що залежить від інтенсивності надходження інцидентів λ_k, числа «серверів» c (чергові бригади/плейбуки SOAR) та швидкості обслуговування $\mu_k = \frac{1}{T_{s,k}}$, а $T_{s,k}$ – середній час сервісу (відновлення). Для стабільності системи необхідно $\sum_k \lambda_k < c\mu$. Обрані правила пріоритетів та параметри c, μ_k прямо визначають досяжність порогів MTTR у SLA, а отже – впливають на ρ.

На рис. 2 показано роботу пріоритетної черги M/M/c, у якій інциденти трьох класів (A – критичні, B – важливі, C – звичайні) надходять із відповідними інтенсивностями λ_k у спільну чергу та розподіляються між c серверами [7-9, 11-12, 14]. Класи обробляються з різним рівнем пріоритету: спершу критичні, потім важливі, далі звичайні. Середній час відновлення $MTTR_k$ визначається співвідношенням інтенсивностей потоків λ_k і швидкості обслуговування μ на доступних серверах.

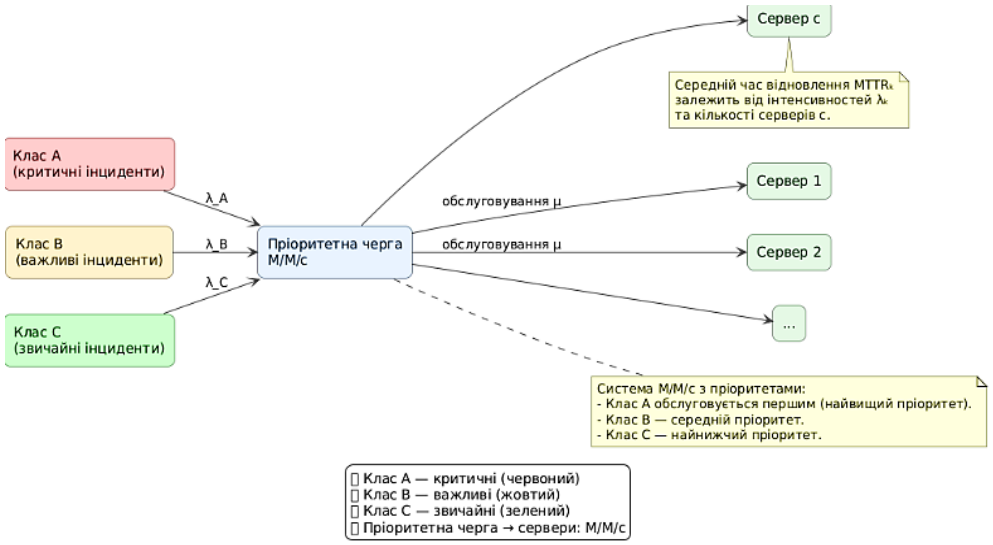

Рис. 2. Пріоритетна черга M/M/c і вплив параметрів на $MTTR_k$

У модулі ризик-скорингу інтегруємо неповні й гетерогенні ознаки інцидентів за допомогою нечіткої системи типу Мамдані [1-3, 7], яка працює в термінах експертних правил «якщо–то» та забезпечує прозоре, відтворюване прийняття рішень. Як вхідні лінгвістичні змінні використовуємо *Severity* (потенційний вплив), *Confidence* (достовірність сигналу), *Exposure* (ступінь експозиції активу) та *Criticality* (критичність об'єкта A/B/C). Кожну змінну проєцюємо у простір належностей $\mu_{low}, \mu_{med}, \mu_{high}$ (для *Criticality* – у множину $\{A, B, C\}$) за трикутними або трапецієподібними функціями, параметри яких калібруємо на історичних даних. База правил формує нечіткий висновок *Risk*, наприклад: «*якщо Severity = high і Exposure = high та Criticality = A, тоді Risk = very_high*». Логічні операції реалізуємо як $And = min, OR = max$, імплікацію – «обрізанням» наслідку правил, а агрегований вихід дефазифікуємо методом центроїда, отримуючи скаляр $Risk \in [0,1]$.

Отримане значення Risk перетворюємо на вагу пріоритизації $\pi_i = f(Risk)$, яка визначає місце інциденту в пріоритетній черзі та вибір SOAR-плейбуків [6]; додатково вводимо пороги ескалації θ_k для категорій $k \in \{A, B, C\}$ (наприклад, для A: $Risk \geq \theta_A$ – негайне втручання) [4, 9]. Значення Risk також використовуємо як штрафний коефіцієнт у інтегральному показнику стійкості ρ, що узгоджує скоринг з метриками MTTD/MTTR/ILP і SLA. Калібрування функцій належності та правил виконуємо за ROC-кривими з мінімізацією хибних ескалацій при фіксованій імовірності пропусків; стабільність контролюємо перехресною валідацією, моніторингом дрейфу даних і періодичним оновленням бази правил [11-13, 15]. Журналювання спрацювань правил забезпечує аудит і дає навчальні дані для подальшої адаптації агентних політик, завдяки чому нечіткий скоринг безпосередньо замикає контур «детектування–пріоритизація–реагування» та підсилює досяжність порогів A/B/C у межах міжвідомчих SLA.

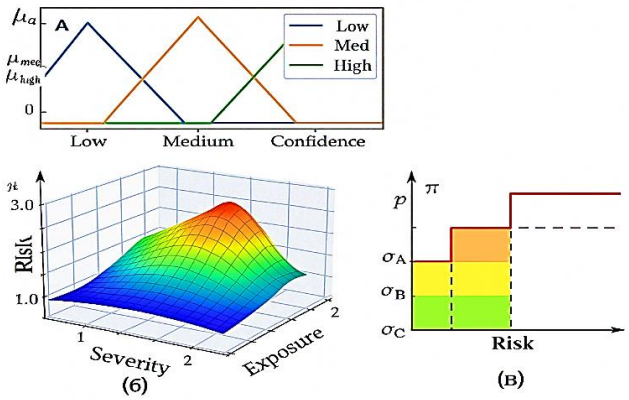

Рис. 3. Нечіткий скоринг і пріоритизація інцидентів: (а) функції належності; (б) поверхня інтегрального ризику; (в) правило Risk→π з політичними порогами ескалації

На рис. 3 показано: (а) функції належності для критеріїв *Severity*, *Exposure* та *Confidence* з рівнями *Low, Medium, High*; (б) тривимірну поверхню інтегрального ризику $Risk = f(Severity, Exposure)$ при *Criticality* = *A*, яка демонструє взаємодію двох параметрів; (в) кусочно-задане правило перетворення *Risk→π* із порогами ескалації $\theta_A > \theta_B > \theta_C$, що визначає рівень пріоритету інциденту. Таким чином, рисунок візуалізує повний процес: від нечіткої оцінки ризику до його трансформації у пріоритет дій реагування.

Невизначеність і неповнота спостережень обробляються засобами нечіткої логіки: будуємо ризик-скоринг, який узгоджує експертні судження та гетерогенні телеметричні ознаки [2-3, 5, 16], а результат використовується як ваги у пріоритизації інцидентів і виборі дій. Далі формулюємо багатокритеріальну оптимізацію розподілу ресурсів відновлення під обмеженнями бюджету та SLA (мінімізуємо MTTD/MTTR/ILP і максимізуємо інтегральний показник ρ-стійкості з урахуванням критичності A/B/C) [11, 14]. Операційні політики узгоджуємо через агентний підхід у постановці MDP/RL: стан включає телеметрію і беклог, дії – ізоляція, ескалація, активація резервів, а функція винагороди спеціально побудована для зменшення MTTD/MTTR/ILP і зростання ρ; це переводить метрики з постфактум-індикаторів у регулятори поведінки агентів.

Значення *Risk* також використовуємо як штрафний коефіцієнт у інтегральному показнику стійкості ρ, що узгоджує скоринг з метриками MTTD/MTTR/ILP і SLA [2, 11].

Функцію перетворення ризику у пріоритет інциденту можна формалізувати у вигляді правила:

$$\pi_i = \begin{cases} 3, & Risk \geq \theta_A, \\ 2, & \theta_B \leq Risk < \theta_A, \\ 1, & \theta_C \leq Risk < \theta_B, \\ 0, & Risk < \theta_C, \end{cases} \quad (4)$$

де 3 – найвищий пріоритет, $\theta_A > \theta_B > \theta_C$ задаються політикою управління.

Задачу ресурсно-обмеженого підвищення стійкості формулюємо як оптимізацію під SLA-обмеженнями. Для об'єктів $i \in O$, сценаріїв $s \in S$, і дій $d \in D$, (пакети реагування/відновлення) розв'язуємо:

$$max \sum_{i \in O} w_i p_i \;\; s.t. \;\; MTTD_i \leq MTTD_i^{SLA}, \;\; MTTR_i \leq MTTR_i^{SLA}, \;\; ILP_i \leq ILP_i^{SLA}, \quad (5)$$

$$\sum c_d \, z_{i,d} \leq B, \;\; z_{i,d} \in \{0,1\}, \quad (6)$$

де w_i – ваги критичності (A/B/C), c_d – вартість дії d, B – бюджет/ресурсний ліміт, $z_{i,d}$ – бінарні рішення застосування пакета дій d до об'єкта i. Постановка має структуру задачі «багатократного рюкзака» (multiple knapsack problem, MKP) і може розв'язуватися за допомогою ILP-розв'язувачів (наприклад, CPLEX, Gurobi, CBC) або жадібних евристик з гарантіями наближення. Штрафи за порушення SLA доцільно враховувати методом лагранжевої релаксації, що переводить задачу у форму максимізації лагранжіана.

Насамкінець результати верифікуємо статистично: оцінюємо бутстреп-довірчі інтервали для змін MTTD/MTTR/ILP і ρ, застосовуємо непараметричні тести для перевірки значущості відмінностей між базовими та запропонованими політиками, а також проводимо ablation-аналіз, щоб кількісно відокремити внесок сценарного ядра, черг із пріоритетами та агентів ШІ [11-12, 14-16]. Така послідовність кроків забезпечує логічний перехід від архітектурної декомпозиції до формального моделювання, від оптимізації до впроваджуваних політик, і – головне – робить вимірювані поліпшення стійкості відтворюваними та порівнюваними для різних категорій об'єктів.

У межах запропонованої парадигми інтелектуального захисту інформаційних систем критичної інфраструктури сформовано інтегровану архітектуру, що поєднує мережу кризово-ситуаційних центрів як координуючий контур, сценарно-модельний апарат для аналізу динаміки загроз і стійкості, а також агентні підходи штучного інтелекту для проактивного моніторингу, кореляції подій і напівавтоматизованого реагування [8, 10]. Концептуально архітектура реалізує багатошарову модель: у нижньому шарі здійснюється уніфікація телеметрії з OT/IT-сегментів, нормалізація журналів подій, підтримка таксономії активів і уразливостей [7, 9, 12]. У середньому шарі функціонують механізми кореляції, оцінювання ризику та пріоритизації інцидентів; у верхньому шарі працює сценарно-модельне ядро з цифровими двійниками критичних процесів, у яке вбудовано правила керування ресурсами відновлення [6, 10]. Структурну схему взаємодії рівнів критичності A/B/C з органами управління та ресурсним контуром наведено на рис. 4.

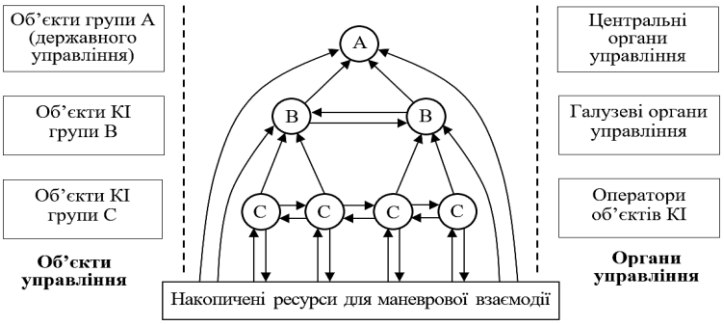

Рис. 4. Архітектура структури системи управління відновленням функціонування ОКІ та взаємодія рівнів критичності A/B/C з органами управління і «накопиченими ресурсами» для маневрової взаємодії

Схема на рис. 4 показує ієрархію об'єктів критичності A/B/C та їх взаємодію з органами управління (праворуч) і «накопиченими ресурсами для маневрової взаємодії» (внизу). Стрілки відображають інформаційні та ресурсні потоки між рівнями: зверху вниз – ескалація/делегування, знизу вгору – звітність/зворотний зв'язок; горизонтальні зв'язки ілюструють координацію між однорівневими об'єктами.

Щоб формалізувати взаємодію архітектурних шарів і процедур реагування, розглянемо керовану динамічну систему [2, 9]:

$$\dot{x}(t) = f\big(x(t), u(t), w(t)\big),\ y(t) = g(x(t)), \tag{7}$$

де $x(t)$ – вектор станів інформаційної системи (стани сервісів, цілісність даних, доступність каналів), $u(t)$ – керувальні дії (ізоляція сегментів, ротація ключів, перемикання на резерви), $w(t)$ – збурення/атаки, $y(t)$ – спостережувані OT/IT-показники. Функціонування описується множиною режимів $S = \{OK, DEG, FAIL\}$ з переходами, індукованими $w(t)$ та $u(t)$. Потоки інцидентів моделюємо множиною подій $E = \{e_1, \dots, e_K\}$ з часовими мітками та TTP-ознаками, сценарій Scn – впорядкована підмножина E з ресурсними обмеженнями та міжзалежностями.

На перетині OT/IT застосовано принципи Zero Trust: мінімально необхідні привілеї, верифікація кожної сесії, мікросегментація технологічних мереж, контроль схеми East–West трафіку, апаратні корені довіри для шлюзів телеметрії та обов'язкова атестація програмно-конфігураційних змін за моделлю «двох рук». Координація між шарами забезпечується через стандартизовані інтерфейси обміну даними та політиками, узгодженими з національними регуляторними вимогами й міжнародними практиками управління ризиками, що усуває фрагментацію оперативної картини та мінімізує часові лаги між виявленням, аналізом і дією. Обмін інцидентами реалізується через уніфіковані таксономії подій, узгоджені з державними вимогами та міжвідомчими регламентами, із визначеними часовими вікнами ескалації та взаємними SLA (Service Level Agreement) для операторів і кризово-ситуаційними центрами (КСЦ).

Ключовим елементом є формалізація «стійкісної динаміки» інформаційних систем, у межах якої стійкість трактується як здатність підтримувати критичні функції за наявності обмежень і збурень із мінімізацією часових та функціональних втрат. Для кількісної оцінки введено систему індикаторів: середній час виявлення інциденту та відновлення, частку попереджених подій до етапу впливу на сервіс, індекс втрати продуктивності при переході до деградаційних режимів, а також інтегральний показник ρ-стійкості, що агрегує доступність, цілісність і своєчасність виконання критичних

бізнес-процесів [9, 11-13, 15]. Запропоновані метрики дозволяють порівнювати альтернативні політики захисту, локалізовувати «вузькі місця» у ланцюзі виявлення-реагування-відновлення та обґрунтовувати розподіл обмежених ресурсів між об'єктами різних категорій критичності.

У подальшому викладі використовуємо взаємопов'язані метрики й позначення, які формують єдиний вимірювальний контур керування стійкістю: MTTD (Mean Time To Detect) інтерпретується як середній інтервал від виникнення чи активації інциденту до його першого достовірного детектування засобами моніторингу або аналітиком, тоді як MTTR (Mean Time To Restore/Recover) визначає середній інтервал від моменту підтвердження інциденту до відновлення цільового рівня надання послуги (SLO, service level objective), тобто повернення критичних функцій у погоджений стан працездатності; доповнює їх ILP (Index of Lost Productivity), що кількісно відображає відсоткову частку втраченої продуктивності в деградаційних режимах відносно номіналу за тривалість інциденту, а рамкові зобов'язання між операторами, КСЦ і підрядниками фіксуються в SLA, де закріплюються цільові значення MTTD/MTTR, вікна ескалації та розподіл відповідальності.

З урахуванням цієї класифікації встановлено об'єкти за критичністю на категорії A/B/C так, що A охоплює об'єкти загальнодержавного значення з високими міжсекторальними залежностями й потенціалом каскадних наслідків, B – об'єкти, збої яких мають значні регіональні або галузеві наслідки, а C – об'єкти з переважно локальними наслідками й більшими допустимими вікнами реагування; відповідно в кожній категорії встановлюємо диференційовані пороги, які безпосередньо пов'язують рівень критичності з потрібною оперативністю: для A – MTTD ≤ 5 хв, MTTR ≤ 30 хв, ILP ≤ 5%; для B – MTTD ≤ 15 хв, MTTR ≤ 2 год, ILP ≤ 10%; для C – MTTD ≤ 1 год, MTTR ≤ 8 год, ILP ≤ 15%. Оскільки коректність порогів залежить від стабільності вимірів, на практиці MTTD оцінюємо як середній інтервал між часовою міткою виникнення події та моментом її достовірного детектування [3], MTTR – як середній інтервал між підтвердженням інциденту та повним відновленням узгодженого SLO, тоді як ILP обчислюємо як нормовану частку втрати продуктивності за весь період інциденту, що забезпечує порівнюваність показників у різних доменах OT/IT.

Політики агентів тренують у MDP-постановці з винагородою, узгодженою з ρ-критерієм, що гарантує спрямованість дій на зменшення MTTD/MTTR/ILP [4, 6, 9]. Саме така уніфікація визначень і методик вимірювання логічно замикає контур управління: зафіксовані значення MTTD/MTTR/ILP одночасно слугують підставою для пріоритизації ресурсів відновлення і встановлення узгоджених міжвідомчих SLA (Service Level Agreement), а також повертаються до сценарно-модельного ядра для адаптації політик і безперервного донавчання інтелектуальних агентів, завдяки чому метрики перестають бути лише індикаторами постфактум і перетворюються на активні регулятори операційної стійкості.

Для кількісної оцінки стійкості використовуємо метрики MTTD, MTTR та ILP. Індекс втрати продуктивності оцінюємо як нормовану площу дефіциту продуктивності:

$$ILP = \frac{1}{T}\int_0^T \left(1 - \frac{p(t)}{p_0}\right) dt \cdot 100\%, \quad (8)$$

де p_0 – номінальний (еталонний) рівень продуктивності, $p(t)$ – фактична продуктивність системи у момент часу t, dt – крок дискретизації вимірювання, T – загальна тривалість спостереження. Менші значення MTTD і MTTR відображають скорочення «вікна атаки»

й швидше повернення сервісів до працездатності, тоді як нижчий *ILP* засвідчує менші операційні втрати в деградаційних режимах.

Для практичного застосування з дискретними вимірюваннями, коли значення p_j фіксуються через інтервал Δt, використовуємо дискретну форму:

$$ILP = \frac{1}{T}\int_{j=1}^{T/\Delta t}\left(1 - \frac{p_j}{p_0}\right)\cdot 100\%, \quad (9)$$

Ця форма зручна для обчислень у реальних умовах експлуатації, оскільки дозволяє інтегрувати показники системи моніторингу продуктивності та оцінювати втрати у відсотках відносно номінальної продуктивності.

На рис. 5 показано індекс втрати продуктивності ILP як нормовану площу між еталонним рівнем продуктивності p_0 та фактичною кривою $p(t)$. Заштрихована область відображає втрату продуктивності, а сині точки ілюструють дискретні вимірювання p_j, що відповідають дискретній формі розрахунку.

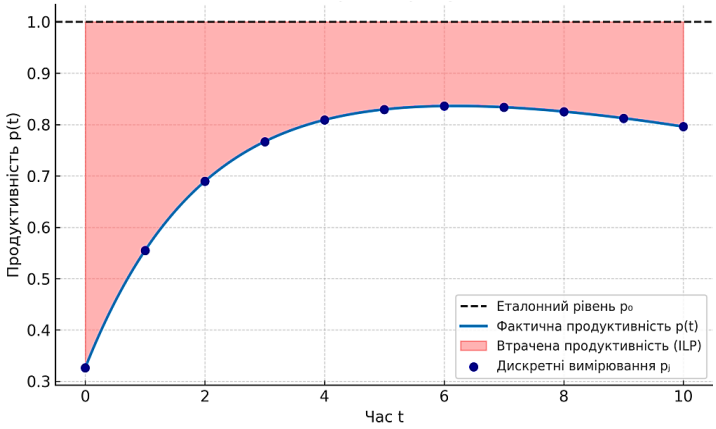

Рис.5. Індекс втрати продуктивності ILP: нормована площа дефіциту продуктивності (безперервна та дискретна форми)

Інтегральний показник ρ-стійкості будуємо як лінійно-фракційну агрегацію індикаторів доступності, цілісності та своєчасності з урахуванням штрафів за часові втрати й деградацію [9-12, 14]:

$$\rho = a\cdot\bar{A} + \beta\cdot\bar{I} + \gamma\cdot\bar{T} - \delta\cdot norm(MMTTD) - \varepsilon\cdot norm(MMTTR) - \xi\cdot ILP, \quad (10)$$

де $\bar{A}, \bar{I}, \bar{T}$ – нормовані середні показники доступності, цілісності та своєчасності критичних функцій, $norm(\cdot)$ – нормування до [0,1] за погодженими порогами, ILP – індекс втрати продуктивності в деградаційних режимах, $a, \beta, \gamma = 1$, $\delta, \varepsilon, \xi > 0$ визначаються політикою ризиків для категорій A/B/C. Така форма дозволяє калібрувати ваги під галузеві вимоги та прозоро порівнювати альтернативні політики.

У практичній реалізації середній час виявлення інциденту обчислюється як [3]:

$$MTTD = \frac{1}{N}\sum_{i=1}^{N}(t_{detect}^{(i)} - t_{onset}^{(i)}), \quad (11)$$

де N – кількість інцидентів у вибірці, $t_{onset}^{(i)}$ – час фактичного виникнення/активації i-го інциденту, $t_{detect}^{(i)}$ – час першого достовірного детектування цього інциденту засобами моніторингу або аналітиком. Використання показника MTTD дозволяє кількісно оцінити

ефективність механізмів виявлення аномалій та визначати вплив людського фактора у процесі реагування. Порівняння значень MTTD між різними сценаріями забезпечує можливість корекції алгоритмів моніторингу та оптимізації процесів кіберзахисту.

Середній час відновлення визначаємо як [11-13]:

$$MTTR = \frac{1}{N}\sum_{i=1}^{N}(t_{restore}^{(i)} - t_{confirm}^{(i)}), \quad (12)$$

де $t_{confirm}^{(i)}$ – час підтвердження i-го інциденту (момент фіксації події як інциденту із запуском процедур реагування), $t_{restore}^{(i)}$ – час відновлення до узгодженого рівня надання послуги (SLO) для цього інциденту. Наведені формули є операціоналізацією метрик для подальшої оптимізації ρ та перевірки досяжності порогів SLA за категоріями A/B/C. Додатково отримані значення MTTR дозволяють здійснювати порівняльний аналіз ефективності різних політик реагування та виявляти «вузькі місця» у процесах відновлення. Це забезпечує можливість інтеграції метрик у систему безперервного моніторингу та динамічного коригування планів забезпечення безперервності бізнес-процесів.

Рекомендовано поетапне впровадження у три хвилі: $W1$ – уніфікація телеметрії, каталог активів і базові агенти детектування; $W2$ – сценарне ядро з цифровими двійниками та оркестрація SOAR; $W3$ – розширення агентів (compliance/симуляції/постачання), міжвідомчі тренування й перегляд ваг $\alpha \dots \zeta$ у моделі ρ. Операційні рішення моделюємо як MDP $\langle S, A, P, R, \gamma \rangle$, де стан s включає агрегати телеметрії, беклог інцидентів, доступні ресурси та прогноз $\hat{\lambda}$, дії a – вибір плейбука (ізоляція сегмента, ротація ключів, переключення на резерв), ескалація або відкладення; перехідні ймовірності P апроксимуються на цифровому двійнику. Функція винагороди орієнтована на метрики стійкості [2, 9, 14]:

$$R(s,a) = -(w_1\,\Delta MTTD + w_2\,\Delta MTTR + w_3\,ILP) + w_4\,\Delta p, \quad (13)$$

з обмеженнями SLA, інкорпорованими як штрафи або лагранжеві множники [11]. Навчання – off-policy (Q-learning / DQN) у симуляторі; перед продакшеном політики проходять «сухі прогони» в цифровому двійнику. Дистанційно застосовується безпечне донавчання з контролем відхилень від базової політики [4, 9, 12, 16]. Навчені політики інтегруються у SOAR-ланцюжки та верифікуються на цифрових двійниках перед розгортанням. Завдяки такій методиці вдається поєднати гнучкість адаптивних стратегій з формальною верифікацією їх безпечності, що мінімізує ризики некоректного реагування в реальних умовах. Це забезпечує стійкість інтелектуальних агентів до динамічних загроз і створює основу для масштабованого впровадження в системах критичної інфраструктури.

На рис. 6 показано єдиний контур керування стійкістю, у якому телеметрія OT/IT надходить до блоку обчислення метрик MTTD, MTTR та ILP. Ці показники агрегуються у композитний індикатор ρ, що використовується як регулятор у політиках MDP/RL. Далі система SOAR виконує автоматизовані дії – ізоляцію, ротацію ключів чи фейловер – після чого формується нова телеметрія [6, 9-10]. Замикання петлі забезпечують сценарне ядро та цифровий двійник, які адаптують політики, перекалібровують p і уточнюють методики вимірювання.

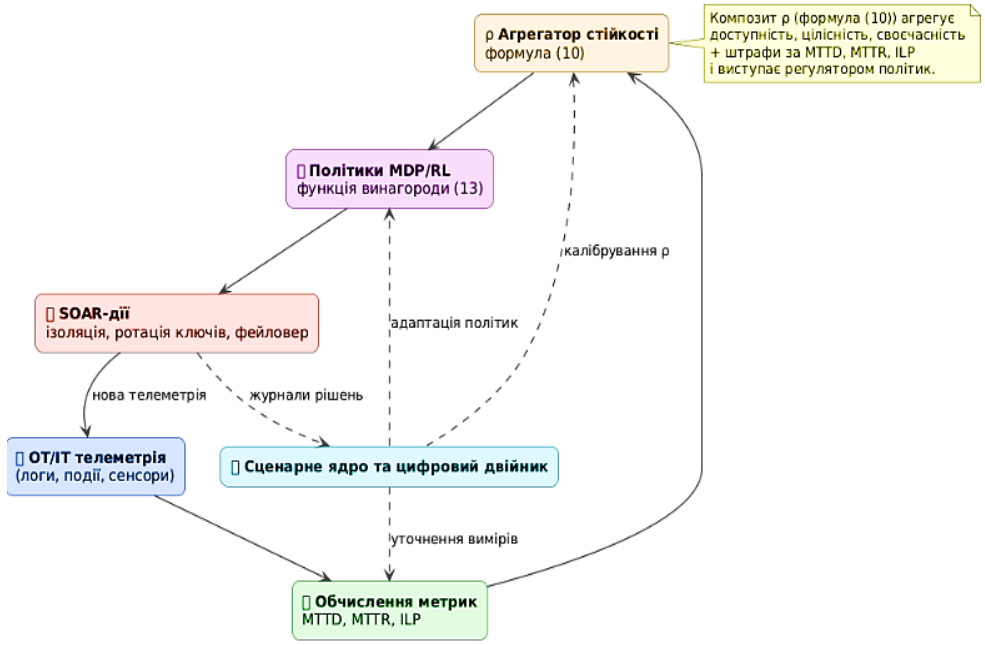

Рис. 6. Єдиний контур керування стійкістю (MTTD–MTTR–ILP → ρ → MDP/RL → SOAR)

Для забезпечення коректності та порівнюваності вимірів часові мітки синхронізуються (NTP/PTP), повторні або зв'язані події агрегуються в єдиний інцидент за узгодженими правилами дедуплікації, статистики стабілізуються робастними оцінками (медіана, усічене середнє) та контролем викидів, а для коректного розрахунку *ILP* фіксується профіль навантаження з метою адекватного вибору P_0 для різних режимів роботи [12, 15]. Така уніфікація визначень, порогів і процедур вимірювання логічно замикає контур управління: зафіксовані значення MTTD/MTTR/ILP одночасно слугують підставою для пріоритизації ресурсів відновлення і встановлення узгоджених міжвідомчих SLA, а також повертаються до сценарно-модельного ядра для адаптації політик і безперервного донавчання інтелектуальних агентів [4, 6, 8-9, 11]. У результаті метрики перестають бути лише постфактум-індикаторами й перетворюються на активні регулятори операційної стійкості.

Сценарне моделювання реалізовано як послідовність процедур побудови профілів загроз, синтезу базових і стресових сценаріїв з урахуванням міжзалежностей OT/IT і ланцюгів постачання, імітації каскадних ефектів із накладанням організаційних та ресурсних обмежень і, зрештою, післядієвого аналізу для ревізії правил реагування та планів відновлення [7-10]. Стандартна процедура включає: профілювання загроз (TTPs, карти уразливостей, міжзалежності), синтез базових і стрес-сценаріїв для A/B/C з комбінованими OT↔IT векторами та фізичними впливами, імітацію каскадних ефектів з ресурсними/організаційними обмеженнями, післядієвий аналіз із ревізією IRP/SOAR-процедур і політик ризику.

Імітаційні дослідження на репрезентативних кейсах – від комбінованих OT-ін'єкцій з компрометацією облікових даних до мережевих маніпуляцій і деградації каналів телеметрії – засвідчили, що включення сценарно-модельного шару до операційного циклу кризово-ситуаційних центрів забезпечує стабільне скорочення часу виявлення та відновлення, підвищує частку превентивно зірваних інцидентів і зменшує втрати продуктивності при переході до керованих деградаційних режимів [6, 9-10].

Наукова новизна полягає в ув'язці сценарного прогнозування з контролем ресурсів відновлення через цифрові двійники процесів, що дозволяє не лише оцінювати ймовірні траєкторії розвитку подій, а й валідовувати логістику постачання запасних вузлів, резервних каналів і кадрових змін.

Агентні технології ШІ у пропонованій моделі виконують роль операційних медіаторів між даними, правилами та діями. Агент телеметрії відповідає за адаптивну фільтрацію і виявлення аномалій із використанням поведінкових і контекстних ознак; агент кореляції будує причинно-наслідкові графи інцидентів, зшиваючи події з різних доменів; агент пріоритизації трансформує ризик у керовані черги реагування з урахуванням категорії об'єкта та крос-секторальних впливів; агент реагування оркеструє напівавтоматичні процедури SOAR – від ізоляції сегментів і ротації ключів до застосування тимчасових політик доступу; агент навчання формує цифровий журнал рішень для подальшої валідації правил і донавчання моделей. Еволюцію підходів, що мотивує перехід до «ери досвіду» агентів ШІ для операційної стійкості ОКІ, узагальнено на рис. 7. Графік із віссю часу (2014–2024+) та вертикальною віссю «увага на навчання з підкріпленням» ілюструє три етапи: ера симуляцій (Atari, AlphaGo, AlphaZero — зростання ролі RL), ера людських даних (GPT-3, ChatGPT – спад інтересу до RL), ера досвіду агентів ШІ (AlphaProof – нове зростання завдяки агентним підходам). Пунктиром позначено стратегічну ціль – «надлюдський розум».

Рис. 7. Схематична хронологія домінуючих парадигм ШІ (ера симуляцій – ера людських даних – ера досвіду агентів ШІ) [6]

Додатково передбачено: агент відповідності (compliance) для зіставлення дій реагування з регуляторними нормами та журналюванням аудиту [11]; агент симуляцій, який перед релізом політик виконує «сухі прогони» у цифровому двійнику; агент постачання, що корелює інциденти з логістикою запасних вузлів і графіками підрядників для скорочення «вузла відновлення» [7-8, 13]. Інтеграція агентів із процедурною логікою кризово-ситуаційних центрів забезпечує відтворюваність дій, скорочує інформаційне навантаження на операторів і зменшує вікно атаки, що є критеріально значущим для об'єктів категорій A/B.

Отримані результати демонструють, що синергія згаданих компонентів – уніфікованої телеметрії, сценарного прогнозування та агентного ШІ – створює кероване середовище підвищення стійкості інформаційних систем критичної інфраструктури [1-5, 9-10, 15]. Зокрема, при переході від реактивної моделі до інтегрованої «КСЦ + сценарії + агенти ШІ» спостерігається статистично значуще зниження середнього часу виявлення завдяки ранньому детектуванню відхилень, скорочення часу відновлення за рахунок

уніфікованих процедур і автоматизованої оркестрації дій, а також збільшення частки інцидентів, що нейтралізуються на докритичній стадії.

За результатами експертно-модельних оцінок на репрезентативних кейсах досягнуто: зменшення MTTD на 25–40%, MTTR на 20–35%, зростання частки превентивно зірваних інцидентів на 15–25% і зниження ILP на 10–20%, що еквівалентно приросту ρ на 12–18% залежно від категорії об'єкта.

Сукупний ефект зафіксовано інтегральним показником ρ-стійкості, який зростає завдяки синхронізації політик між доменами OT/IT, кращій видимості каскадних залежностей і формалізації «допустимих рівнів ризику» на етапі планування. Водночас процесний вимір – стандартизація ролей і відповідальностей, регламентація обміну інцидентами, ведення каталогів запасів і графіків готовності підрядників, регулярні міжвідомчі тренування зі стрес-сценаріями – є необхідною умовою матеріалізації технічних переваг у вимірювані покращення операційної стійкості.

Обмеження дослідження: частина параметрів моделювання калібрована на узагальнених даних і потребує доменної адаптації, ефекти залежать від зрілості вихідних процесів SOC/OT і готовності персоналу до процедур SOAR [6, 8, 11], цифрові двійники критичних процесів потребують періодичної валідації для коректного відтворення міжсекторальних залежностей.

Таким чином, розроблені моделі та технології інтелектуального захисту забезпечують науково обґрунтований перехід до проактивного, метрико-керованого управління безпекою інформаційних систем критичної інфраструктури. Запропонована архітектура узгоджує цілі виявлення, реагування й відновлення з об'єктивними обмеженнями ресурсу та часу, а також надає інструментарій для постійної адаптації політик на підставі даних і симуляцій. У підсумку отримано цілісний механізм підвищення стійкості, придатний до поетапного впровадження в національній мережі кризово-ситуаційних центрів і галузевих операторів, що відповідає сучасним викликам, нормативним вимогам і стратегічній меті забезпечення безпечного, безперервного функціонування критично важливих сервісів.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У статті запропоновано цілісну парадигму підвищення стійкості інформаційних систем об'єктів критичної інфраструктури на основі інтегрованої архітектури «кризово-ситуаційні центри (КСЦ) + сценарне ядро + агенти ШІ». Ключовим результатом є ув'язка сценарного прогнозування з керуванням ресурсами відновлення через цифрові двійники критичних процесів, що переводить управління з реактивної площини у проактивну та відтворювану. Запропоновано єдиний контур керування стійкістю, де узгоджені операційні метрики MTTD, MTTR і ILP агреговано в композит p, який виконує роль не лише підсумкового індикатора, а й регулятора в MDP/RL-політиках, що забезпечує адаптивність рішень у реальному часі.

Формально обґрунтовано зшивку різнорідних моделей – часових мереж Петрі для інцидент-менеджменту, пріоритетних черг M/M/c для операційної логістики та нечіткого скорингу ($Risk \to \pi$) – в один вимірюваний контур, синхронізований із вимогами SLA. Показано, що така інтеграція робить життєвий цикл інциденту прозорим для аналізу, а параметри процесів – керованими важелями досягнення цільових порогів. Сценарне моделювання з урахуванням міжзалежностей OT↔IT і впливів ланцюгів постачання,

підкріплене цифровими двійниками, дозволяє валідувати політики реагування та логістику відновлення до їх розгортання, знижуючи ризики хибних дій.

Отримані експертно-модельні оцінки свідчать про суттєві кількісні поліпшення: скорочення MTTD і MTTR, зниження ILP та підвищення p, що відтворюється на репрезентативних сценаріях і є релевантним для секторів енергетики, транспорту, медицини та фінансів. Практична значущість підходу полягає у створенні повторюваного, керованого середовища підвищення стійкості: від пріоритизації інцидентів і оркестрації SOAR-ланцюжків до планування запасів і синхронізації дій підрядників у межах міжвідомчих SLA та принципів Zero Trust.

Попри продемонстровані переваги, дослідження має низку обмежень. Частина параметрів симуляцій та нечітких правил калібрована на узагальнених наборах даних і потребує доменної адаптації; ефекти залежать від зрілості SOC/OT-процесів і готовності персоналу; цифрові двійники вимагають періодичної валідації для коректного відображення каскадних міжсекторальних впливів. Окремим викликом є забезпечення сталості даних (NTP/PTP-синхронізація), керування дрейфом даних і підтримання узгодженості таксономій інцидентів у багатосторонніх взаємодіях КСЦ-операторів.

Перспективи подальших досліджень бачимо у кількох напрямках. По-перше, формальна верифікація безпеки політик (control-invariant sets, перевірка властивостей мереж Петрі та марковських моделей) і розроблення процедур безпечного онлайн-навчання агентів із гарантованими обмеженнями на відхилення від базової політики. По-друге, підвищення достовірності цифрових двійників: ідентифікація параметрів із реальних потоків OT/IT-телеметрії, моделювання каскадів у ланцюгах постачання та врахування обмежень логістики запасних вузлів і підрядних ресурсів. По-третє, розвиток стійких до атак методів детектування (adversarial-robust ML), мультиагентної координації (договорні механізми/субординоване керування) та причинно-орієнтованих графів для кращої інтерпретованості рішень КСЦ. По-четверте, інтеграція економічних метрик у p (вартість простою, штрафи SLA, CAPEX/OPEX резервування) для ризик-орієнтованого бюджетування та порівняння альтернативних політик відновлення. По-п’яте, стандартизація інтерфейсів і форматів даних (інтероперабельність між галузями), а також створення відкритих бенчмарків і еталонних сценаріїв для незалежної оцінки методів сценарного моделювання й агентних підходів.

Таким чином, представлена архітектура та методологія задають практичну рамку переходу до проактивного, даними керованого управління стійкістю ІС ОКІ. Подальша наукова робота має зосередитися на формальних гарантіях безпеки, підвищенні вірогідності цифрових двійників, економічній інтеграції p та масштабуванні міжвідомчої взаємодії – щоб перетворити локальні впровадження на узгоджену національну систему операційної стійкості.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Rehman A., Awan K.A., Al-Rasheed A., Ara A., Alruwaili F., Alotaibi S., Saba T. (2025). A novel hybrid fuzzy logic and federated learning framework for enhancing cybersecurity and fraud detection in IoT-enabled metaverse transactions. *Egyptian Informatics Journal*, 30, 100668. https://doi.org/10.1016/j.eij.2025.100668
2. Zybin S., Korchenko O., Korystin O., Shulha V., Kazmirchuk S., Demediuk S. (2025). Method for the risk assessing of hybrid threats in cyber security based on fuzzy set theory. *SSRN Preprint*. https://ssrn.com/abstract=5143937 or http://dx.doi.org/10.2139/ssrn.5143937

3. Kim C., So-In C., Kongsorot Y., et al. (2024). FLSec-RPL: a fuzzy logic-based intrusion detection scheme for securing RPL-based IoT networks against DIO neighbor suppression attacks. *Cybersecurity*, 7, 27. https://doi.org/10.1186/s42400-024-00223-x
4. Kshetri N. (2025). Transforming cybersecurity with agentic AI to combat emerging cyber threats. *Telecommunications Policy*, 49, 102976. https://doi.org/10.1016/j.telpol.2025.102976
5. Karakaya A. (2025). A hybrid approach for IoT security: combining ensemble learning with fuzzy logic. *Sensors (Basel, Switzerland)*, 25(18), 5668. https://doi.org/10.3390/s25185668
6. Mayoral-Vilches V., Navarrete-Lozano L.J., Sanz-Gómez M., Espejo L.S., Crespo-Álvarez M., Oca-Gonzalez F., et al. (2025). CAI: An open, bug bounty-ready cybersecurity AI. arXiv preprint. arXiv:2504.06017
7. Havrys A., Filippova V., Tur N. (2024). Інформаційний аналіз систем захисту об'єктів критичної інфраструктури в період дії воєнного стану. *Вісник Львівського державного університету безпеки життєдіяльності*, 30, 173–187. https://doi.org/10.32447/20784643.30.2024.17
8. Гречанінов В.Ф. (2021). Особливості підтримки прийняття рішень у ситуаційних центрах органів влади з метою захисту критичної інфраструктури. *Реєстрація, зберігання і обробка даних*, 23(3), 80–90. URL: http://jnas.nbuv.gov.ua/article/UJRN-0001304909
9. Kostiuk Y., Skladannyi P., Samoilenko Y., Khorolska K., Bebeshko B., Sokolov V. (2025). A system for assessing the interdependencies of information system agents in information security risk management using cognitive maps. *In Cyber Hygiene & Conflict Management in Global Information Networks* (Vol. 3925, pp. 249–264).
10. Морозов А.О., Гречанінов В.Ф. (2025). Еволюційний розвиток та майбутнє технологій ситуаційного управління. *Математичні машини і системи*, 1, 3–11. URL: http://www.immsp.kiev.ua/publications/articles/2025/2025_1/01_25_Morozov.pdf
11. Kostiuk, Yu. V., Skladannyi, P. M., Hulak, H. M., Bebeshko, B. T., Khorolska, K. V., & Rzaieva, S. L. (2025). Information security systems. *[Textbook]* Kyiv: Borys Grinchenko Kyiv Metropolitan University.
12. Hulak, H. M., Zhyltsov, O. B., Kyrychok, R. V., Korshun, N. V., & Skladannyi, P. M. (2023). Enterprise information and cyber security. [Textbook] Kyiv: Borys Grinchenko Kyiv Metropolitan University.
13. Kostiuk Y., Skladannyi P., Sokolov V., Hulak H., Korshun N. (2025). Models and algorithms for analyzing information risks during the security audit of personal data information system. *In Proceedings of the Third International Conference on Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN'24)* (Vol. 3925, pp. 155–171).
14. Kostiuk Y., Skladannyi P., Sokolov V., Zhyltsov O., Ivanichenko Y. (2025). Effectiveness of information security control using audit logs. *In Proceedings of the Workshop on Cybersecurity Providing in Information and Telecommunication Systems (CPITS 2025)*, 524–538.
15. Abramov, V., Astafieva, M., Boiko, M., Bodnenko, D., Bushma, A., Vember, V., Hlushak, O., Zhyltsov, O., Ilich, L., Kobets, N., Kovaliuk, T., Kuchakovska, H., Lytvyn, O., Lytvyn, P., Mashkina, I., Morze, N., Nosenko, T., Proshkin, V., Radchenko, S., Yaskevych, V. (2021). Theoretical and practical aspects of the use of mathematical methods and information technology in education and science. https://doi.org/10.28925/9720213284km
16. Kostiuk, Yu. V., Skladannyi, P. M., Bebeshko, B. T., Khorolska, K. V., Rzaieva, S. L., & Vorokhob, M. V. (2025). Information and communication systems security. *[Textbook*] Kyiv: Borys Grinchenko Kyiv Metropolitan University.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (29), 2025
ISSN 2663 - 4023

Viktor Grechaninov
Associate Professor, Senior Researcher,
Head of the Research Department
Institute for Problems of Mathematical Machines and Systems of the NASU, Kyiv, Ukraine
ORCID ID: 0000-0001-6268-3204
vgrechaninov@gmail.com

MODELS AND TECHNOLOGIES OF INTELLIGENT PROTECTION OF INFORMATION SYSTEMS OF CRITICAL INFRASTRUCTURE FOR ENHANCING RESILIENCE

Abstract. The article substantiates the feasibility of using modern information technologies to ensure the sustainable functioning of critical infrastructure (CI) facilities, with a focus on protecting their information systems, which are a key factor in national security and the resilience of the state against hybrid threats. It is demonstrated that enhancing the protection and recoverability of CI systems is possible through the establishment of an extensive network of crisis centers integrated with platforms for monitoring, detection, and real-time response to cyber incidents. Particular attention is paid to scenario modeling, which enables forecasting possible developments of cyberattacks, designing security management models, and supporting the decision-making process. This approach makes it possible to identify likely channels of impact on systems in advance, assess the consequences of their disruption, and generate optimal strategies for threat neutralization. An architecture of a multi-level system for managing the protection and recovery of CI information systems is proposed, taking into account both physical and cyber risks. It is based on the integration of intelligent technologies capable of providing adaptive responses to environmental changes and automated support for backup and recovery procedures. The feasibility of applying artificial intelligence in crisis centers is substantiated, particularly through agent-based systems that enhance the efficiency of analyzing large datasets, detecting anomalies in traffic, assessing risks, and generating managerial recommendations. The use of intelligent agents ensures speed and accuracy in the localization of cyber threats, significantly increasing the resilience of critical infrastructure information systems and forming the foundation for proactive cybersecurity mechanisms.

Keywords: critical infrastructure, information systems, cybersecurity, resilient functioning, crisis (situational) centers, scenario modeling, artificial intelligence, intelligent agents, decision support systems.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Rehman, A., Awan, K.A., Al-Rasheed, A., Ara, A., Alruwaili, F., Alotaibi, S., & Saba, T. (2025). A novel hybrid fuzzy logic and federated learning framework for enhancing cybersecurity and fraud detection in IoT-enabled metaverse transactions. *Egyptian Informatics Journal*, 30, 100668. https://doi.org/10.1016/j.eij.2025.100668
2. Zybin, S., Korchenko, O., Korystin, O., Shulha, V., Kazmirchuk, S., & Demediuk, S. (2025). Method for the risk assessing of hybrid threats in cyber security based on fuzzy set theory. *SSRN Preprint*. https://ssrn.com/abstract=5143937 or http://dx.doi.org/10.2139/ssrn.5143937
3. Kim, C., So-In, C., Kongsorot, Y., et al. (2024). FLSec-RPL: a fuzzy logic-based intrusion detection scheme for securing RPL-based IoT networks against DIO neighbor suppression attacks. *Cybersecurity*, 7, 27. https://doi.org/10.1186/s42400-024-00223-x
4. Kshetri, N. (2025). Transforming cybersecurity with agentic AI to combat emerging cyber threats. *Telecommunications Policy*, 49, 102976. https://doi.org/10.1016/j.telpol.2025.102976
5. Karakaya, A. (2025). A hybrid approach for IoT security: combining ensemble learning with fuzzy logic. *Sensors (Basel, Switzerland)*, 25(18), 5668. https://doi.org/10.3390/s25185668
6. Mayoral-Vilches, V., Navarrete-Lozano, L.J., Sanz-Gómez, M., Espejo, L.S., Crespo-Álvarez, M., Oca-Gonzalez, F., et al. (2025). CAI: An open, bug bounty-ready cybersecurity AI. arXiv preprint. arXiv:2504.06017

7. Havrys, A., Filippova, V., & Tur, N. (2024). Information analysis of protection systems of critical infrastructure facilities during martial law. *Bulletin of the Lviv State University of Life Safety*, 30, 173–187. https://doi.org/10.32447/20784643.30.2024.17
8. Hrechaninov, V.F. (2021). Features of decision support in governmental situational centers aimed at protecting critical infrastructure. *Registration, Storage and Processing of Data*, 23(3), 80–90. URL: http://jnas.nbuv.gov.ua/article/UJRN-0001304909
9. Kostiuk, Y., Skladannyi, P., Samoilenko, Y., Khorolska, K., Bebeshko, B., & Sokolov, V. (2025). A system for assessing the interdependencies of information system agents in information security risk management using cognitive maps. *In Cyber Hygiene & Conflict Management in Global Information Networks* (Vol. 3925, pp. 249–264).
10. Morozov, A.O., & Hrechaninov, V.F. (2025). Evolutionary development and the future of situational management technologies. *Mathematical Machines and Systems*, 1, 3–11. URL: http://www.immsp.kiev.ua/publications/articles/2025/2025_1/01_25_Morozov.pdf
11. Kostiuk, Yu. V., Skladannyi, P. M., Hulak, H. M., Bebeshko, B. T., Khorolska, K. V., & Rzaieva, S. L. (2025). Information security systems. *[Textbook]* Kyiv: Borys Grinchenko Kyiv Metropolitan University.
12. Hulak, H. M., Zhyltsov, O. B., Kyrychok, R. V., Korshun, N. V., & Skladannyi, P. M. (2023). Enterprise information and cyber security. *[Textbook]* Kyiv: Borys Grinchenko Kyiv Metropolitan University.
13. Kostiuk Y., Skladannyi P., Sokolov V., Hulak H., Korshun N. (2025). Models and algorithms for analyzing information risks during the security audit of personal data information system. *In Proceedings of the Third International Conference on Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN'24)* (Vol. 3925, pp. 155–171).
14. Kostiuk Y., Skladannyi P., Sokolov V., Zhyltsov O., Ivanichenko Y. (2025). Effectiveness of information security control using audit logs. *In Proceedings of the Workshop on Cybersecurity Providing in Information and Telecommunication Systems (CPITS 2025)*, 524–538.
15. Abramov, V., Astafieva, M., Boiko, M., Bodnenko, D., Bushma, A., Vember, V., Hlushak, O., Zhyltsov, O., Ilich, L., Kobets, N., Kovaliuk, T., Kuchakovska, H., Lytvyn, O., Lytvyn, P., Mashkina, I., Morze, N., Nosenko, T., Proshkin, V., Radchenko, S., Yaskevych, V. (2021). Theoretical and practical aspects of the use of mathematical methods and information technology in education and science. https://doi.org/10.28925/9720213284km
16. Kostiuk, Yu. V., Skladannyi, P. M., Bebeshko, B. T., Khorolska, K. V., Rzaieva, S. L., & Vorokhob, M. V. (2025). Information and communication systems security. *[Textbook*] Kyiv: Borys Grinchenko Kyiv Metropolitan University.