

БЕЗПЕКА ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ

№ 2(10)/2025

НАУКОВИЙ ЖУРНАЛ

ISSN 2707-1758

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Засновано 2019 року

УДК 004.056:004.9

DOI: <https://doi.org/10.17721/ISTS.2025.10>

ROR: <https://ror.org/02aaqv166>

Представлено результати досліджень, що містять оригінальні наукові здобутки та висвітлюють теоретичні засади і практичні аспекти розвитку науки і техніки в галузі безпеки інформаційних систем і технологій. Тематика видання охоплює спеціальності: 125 (F5) – Кібербезпека та захист інформації, 122 (F3) – Комп'ютерні науки, 126 (F6) – Інформаційні системи та технології (кластер: Інформаційні технології та електроніка).

Для науковців, докторантів, аспірантів, фахівців відповідних галузей і студентів старших курсів.

ГОЛОВНИЙ РЕДАКТОР

ВІДПОВІДАЛЬНІ СЕКРЕТАРІ

РЕДАКЦІЙНА КОЛЕГІЯ

Бучик Сергій, д-р техн. наук, проф. (Україна)

Даков Сергій, канд. техн. наук (Україна)

Торошанко Олександр, канд. техн. наук (Україна)

Бернаш Марчін, д-р техн. наук, (Польща); Браїловський Микола, канд. техн. наук, доц. (Україна); Гопеєнко Віктор, д-р техн. наук, проф. (Латвія); Данієлієн Рената, д-р техн. наук, доц. (Литва); Дружинін Володимир, д-р техн. наук, проф. (Україна); Зюбіна Руслана, канд. техн. наук, доц. (Польща); Іларіонов Олег, канд. техн. наук, доц. (Україна); Казакова Надія, д-р техн. наук, проф. (Україна); Корнієнко Богдан, д-р техн. наук, проф. (Україна); Лаптев Олександр, д-р техн. наук, ст. наук. співроб. (Україна); Мілош Улман, д-р техн. наук, проф. (Чеська Республіка); Морозов Віктор, канд. техн. наук, проф. (Україна); Наконечний Володимир, д-р техн. наук, проф. (Україна); Пархоменко Іван, канд. техн. наук, доц. (Україна); Плєскач Валентина, д-р екон. наук, проф. (Україна); Рудзійоніс Вітаутас, д-р техн. наук, проф. (Литва); Субач Ігор, д-р техн. наук, проф. (Україна); Снитюк Віталій, д-р техн. наук, проф. (Україна); Толюпа Сергій, д-р техн. наук, проф. (Україна); Яновський Фелікс, д-р техн. наук, проф. (Нідерланди)

Адреса редколегії

вул. Богдана Гаврилишина, 24, м. Київ, 04116

☎ (38044) 481 44 07

e-mail: fit@univ.net.ua

Затверджено

вченою радою факультету інформаційних технологій
08.11.25 (протокол № 6)

Зареєстровано

Національною радою України з питань телебачення і радіомовлення
Рішення № 357 від 15.02.24
Ідентифікатор друкованого медіа: R30-02757

Атестовано

Міністерством освіти і науки України (категорія Б)
Наказ № 1415 від 02.10.24

Засновник та видавець

Київський національний університет імені Тараса Шевченка
Видавничо-поліграфічний центр "Київський університет"

Свідоцтво внесено до Державного реєстру
ДК № 1103 від 31.10.02

Адреса видавця

ВПЦ "Київський університет"

6-р Тараса Шевченка, 14, м. Київ, 01601

☎ (38044) 239 32 22, 239 31 58, 239 31 28

e-mail: vpc@knu.ua

INFORMATION SYSTEMS AND TECHNOLOGIES SECURITY

№ 2(10)/2025

SCIENTIFIC JOURNAL

ISSN 2707-1758

TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV

Established in 2019

UDC 004.056:004.9

DOI: <https://doi.org/10.17721/ISTS.2025.10>

ROR: <https://ror.org/02aaqv166>

The publication presents research findings that contain original scientific achievements and highlight the theoretical foundations and practical aspects of science and technology development in the field of information systems and technologies security. The scope of the publication covers the following specialties: 125 (F5) – Cybersecurity and Information Protection, 122 (F3) – Computer Science, 126 (F6) – Information Systems and Technologies (Cluster: Information Technologies and Electronics).

For scientists, doctoral students, postgraduate students, professionals in the relevant fields, and senior students.

EDITOR-IN-CHIEF

Buchykh Serhii, DSc (Engin.), Prof. (Ukraine)

EXECUTIVE SECRETARIES

Dakov Serhii, PhD (Engin.) (Ukraine)
Toroshanko Oleksandr, PhD (Engin.) (Ukraine)

EDITORIAL BOARD

Bernas Marcin, PhD (Engin.) (Poland); Brailovskyi Mykola, PhD (Engin.), Assoc. Prof. (Ukraine); Danieliene Renata, PhD (Engin.), Assoc. Prof. (Lithuania); Druzhynin Volodymyr, DSc (Engin.), Prof. (Ukraine); Gopejenko Viktor, DSc (Engin.), Prof. (Latvia); Ilarionov Oleh, PhD (Engin.), Assoc. Prof. (Ukraine); Kazakova Nadiia, DSc (Engin.), Prof. (Ukraine); Kornienko Bogdan, DSc (Engin.), Prof. (Ukraine); Laptiev Oleksandr, DSc (Engin.), Senior Researcher (Ukraine); Milos Ulman, DSc (Engin.), Prof. (Czech Republic); Morozov Viktor, PhD (Engin.), Prof. (Ukraine); Nakonechnyi Volodymyr, DSc (Engin.), Prof. (Ukraine); Parkhomenko Ivan, PhD (Engin.), Assoc. Prof. (Ukraine); Pleskach Valentyna, DSc (Econ.), Prof. (Ukraine); Rudzionis Vytautas, DSc (Engin.), Prof. (Lithuania); Subach Ihor, DSc (Engin.), Prof. (Ukraine); Snytyuk Vitaliy, DSc (Engin.), Prof. (Ukraine); Toliupa Serhii, DSc (Engin.), Prof. (Ukraine); Yanovsky Felix, DSc (Engin.), Prof. (Netherlands); Ziubina Ruslana, PhD (Engin.), Assoc. Prof. (Poland)

Address

24, Bohdan Hawrylyshyn str., Kyiv, 04116
☎ (38044) 481 44 07
e-mail: fit@univ.net.ua

Approved by

the Academic Council of the Faculty of Information Technology
08.11.25 (protocol № 6)

Registered by

the National Council of Ukraine on Television and Radio Broadcasting
Decision № 357 of 02.15.24
Identifier of printed media: R30-02757

Certified by

the Ministry of Education and Science of Ukraine (category B)
Order № 1415 dated 02.10.24

Founder and Publisher

Taras Shevchenko National University of Kyiv
Publishing and Polygraphic Center "Kyiv University"
Certificate of entry into the State Register
ДК № 1103 from 31.10.02

Address

PPC "Kyiv University"
14, Taras Shevchenko blvd., Kyiv, 01601
☎ (38044) 239 32 22, 239 31 58, 239 31 28
e-mail: vpc@knu.ua

КІБЕРБЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ

ІЛЬЄНКО Анна, ТЕЛЮЩЕНКО Валентина Методи оцінювання ризиків кіберзагроз в інформаційно-комунікаційних системах об'єктів цивільної авіації	5
ПАРХОМЕНКО Іван, СОКИРАН Юрій Методи оцінювання географічної концентрації та стабільності розробників у проєктах із відкритим кодом з OSINT-підходом	16
ПРОКОПОВИЧ-ТКАЧЕНКО Дмитро, ЧЕРКАСЬКИЙ Олександр, ЗУБЧЕНКО Назар, БІЛАН Максим, МАЦКО Володимир Компрометація ланцюга оновлень через SSL та SNMP як вектор атаки на телекомунікаційні вузли: технічне моделювання та заходи протидії	21
РАГІМОВА Ірада, ЕЮБ Абдул огли Аббасгулієв, СЄВРЮКОВА Єлизавета, ЛАПТЄВА Тетяна Експериментальне оцінювання ефективності криптографічних генераторів псевдовипадкових чисел.....	30
ЩЕБЛАНІН Юрій, СКИДАН Данило Оцінювання ефективності комбінованих криптографічних схем у захисті Wi-Fi-з'єднань.....	37
ОПІРСЬКИЙ Іван, ЯВОРІВСЬКА Соломія Дослідження особливостей використання SOAR як головного інструменту для прискорення реагування на інциденти	44
ЛУЦЕНКО Владислав, НАКОНЕЧНИЙ Володимир Децентралізовані механізми довіри та консенсусу для захисту конфіденційних даних у розподілених системах	54
КОБОЗЄВА Алла, ГРИГОРЕНКО Світлана Проблема синхронізації в селективній стеганографії	61
ДУДНІК Андрій, ВИГОВСЬКИЙ Сергій, ТОРОШАНКО Олександр, ФЕСЕНКО Андрій Метод інтеграції криптопроцесора у роботизовану сенсорну мережу системи розмінування	70

КОМП'ЮТЕРНІ НАУКИ

ПЛЕСКАЧ Валентина, ЖИЛЮК Ярослав Безпека поточкових даних у розподілених системах реального часу із застосуванням диференційної приватності та федеративного навчання	77
--	----

ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

КРАВЧЕНКО Юрій, ФІСУН Олена Метод забезпечення функціональної стійкості програмно-керованих мереж в умовах впливу кібератак	83
ДРУЖИНІН Володимир, БАТРАК Євгеній, ЦЬОПА Наталія, НІКОЛЕНКО Ольга Мультиджерельний алгоритмічний аналіз читачських і комерційних трендів книжкового ринку.....	96

CONTENTS

CYBERSECURITY AND INFORMATION PROTECTION

ILIENKO Anna, TELIUSHCHENKO Valentyna Methods for assessing the risks of cyber threats in information and communication systems of civil aviation facilities.....	5
PARKHOMENKO Ivan, SOKYRAN Yurii Methods for assessing geographic concentration and developer stability in open-source projects using an osint approach.....	16
PROKOPOVYCH-TKACHENKO Dmytro, CHERKASKYI Oleksandr, ZUBCHENKO Nazar, BILAN Maksym, MATSKO Volodymyr Compromise of the update chain via SSL and SNMP as an attack vector on telecommunication nodes: technical modeling and countermeasures	21
RAHIMOVA Irada, EYYUB Abdul oglu Abbasquliyev, SEVRIUKOVA Yelyzaveta, LAPTIEVA Tetiana Experimental evaluation of the efficiency of cryptographic generators of pseudo-random numbers	30
SHCHEBLANIN Yurii, SKYDAN Danylo Evaluating the effectiveness of combined cryptographic schemes in securing Wi-Fi connections	37
OPIRSKYI Ivan, YAVORIVSKA Solomiia Research on the features of using soar as a primary tool for accelerating incident response	44
LUTSENKO Vladyslav, NAKONECHNYI Volodymyr Decentralized trust and consensus mechanisms for protecting sensitive data in distributed systems	54
KOBOZIEVA Alla, HRYHORENKO Svitlana The synchronization problem in selective steganography	61
DUDNIK Andrii, VYHOVSKYY Sergiy, TOROSHANKO Olexandr, FESENKO Andrii Method of integration of cryptoprocessor into robotic sensor network of exchange system	70

COMPUTER SCIENCE

PLESKACH Valentyna, ZHYLIUK Yaroslav Privacy in streaming data: a hybrid approach with differential privacy and robust federated learning.....	77
---	----

INFORMATION SYSTEMS AND TECHNOLOGIES

KRAVCHENKO Yurii, FISUN Olena Method for ensuring functional resilience of software-defined networks under cyberattack influence	83
DRUZHYNIN Volodymyr, BATRAK Yevhenii, TSOPA Nataliia, NIKOLENKO OIha Multisource algorithmic analysis of reader and commercial trends in the book market	96



КІБЕРБЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ

УДК 004.056.5:656.7:519.86

DOI: <https://doi.org/10.17721/ISTS.2025.10.5-15>

Анна ІЛЬЄНКО, канд. техн. наук, доц.

ORCID ID: 0000-0001-8565-1117

e-mail: anna.ilienko@npp.kai.edu.uaДержавне некомерційне підприємство "Державний університет
«Київський авіаційний інститут»", Київ, Україна

Валентина ТЕЛЮЩЕНКО, асп.

ORCID ID: 0000-0001-6026-5105

e-mail: valentyna.teliushchenko@npp.kai.edu.uaДержавне некомерційне підприємство "Державний університет
«Київський авіаційний інститут»", Київ, Україна

МЕТОДИ ОЦІНЮВАННЯ РИЗИКІВ КІБЕРЗАГРОЗ В ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМАХ ОБ'ЄКТІВ ЦИВІЛЬНОЇ АВІАЦІЇ

Вступ. Цивільна авіація належить до критичної інфраструктури, а її наземні інформаційно-комунікаційні системи (ІКС) – аеропортові операційні платформи, інтеграційні шлюзи, мережі та служби підтримки – стикаються з різноманітним ландшафтом кіберзагроз. Для керованого вибору підходів до оцінювання ризиків потрібен прозорий і відтворюваний інструмент, узгоджений із вимогами ICAO/EASA/ISO й операційними обмеженнями аеропортів та авіаперевізників.

Методи. Запропоновано коротку модель багатокритеріального аналізу (МКА) з лінійним зваженням сумуванням (WSM). Як альтернативу взято NIST SP 800-30/CSF, CRAMM, OCTAVE Allegro, EBIOS RM, TARA. Критерії моделі: відповідність стандартам, глибина аналізу, гнучкість, легкість упровадження, вартість / ресурсомісткість. Ваги моделі нормовані (сума дорівнює 1), оцінки експертні за шкалою 1–5 з "оберненою" інтерпретацією вартості. Обчислено матрицю зважених компонентів, інтегральний бал і його нормовану форму. Введено чотирирівневу шкалу прийняття рішень із трьома порогоми та проведено аналіз чутливості через сценарне варіювання ваг.

Результати. Загальне ранжування підтвердило перевагу NIST SP 800-30/CSF, тоді як EBIOS RM продемонстрував результат, близький до лідера. У безпеково-критичному сценарії перші позиції посідають NIST і CRAMM, а у бюджетному та сценарії швидкого впровадження першість переходить до EBIOS RM, за яким іде NIST. OCTAVE Allegro доцільно розглядати як "умовно рекомендований" збалансований компроміс, а TARA – як рішення для швидкої пріоритизації за обмежених ресурсів. У типових умовах експлуатації, де ключовими є вартість і оперативність, CRAMM поступається через високу ресурсомісткість. Отримані результати загалом залишаються стійкими до помірних змін ваг.

Висновки. Модель МКА/WSM забезпечує прозоре, відтворюване й інтерпретоване порівняння фреймворків ризик-менеджменту для наземних ІКС цивільної авіації і може бути інтегрована у системи підтримки рішень. Підхід поєднує нормативну відповідність із практичною реалізованістю та підтримує сценарний вибір. Перспективи – автоматизація вибору ваг і критеріїв із застосуванням методів ШІ.

Ключові слова: цивільна авіація, ІКС, кіберризик, багатокритеріальний аналіз, метод зваженої суми, NIST.

Вступ

Цивільна авіація (ЦА) відіграє ключову роль у світовій інфраструктурі, поєднуючи країни та людей через економічні й соціальні зв'язки. Її функціонування все більше залежить від складних інформаційно-керуючих систем, які підтримують планування, управління, взаємодію та моніторинг у реальному часі. У цій роботі під "об'єктами цивільної авіації" розуміємо організації та інфраструктуру, залучені до забезпечення авіаційних перевезень і наземного обслуговування, зокрема й аеропорти, авіакомпанії, провайдери наземного обслуговування та суміжні служби. Інформаційні та керуючі системи, що забезпечують критичні бізнес-процеси цих суб'єктів, становлять інформаційно-комунікаційні системи (ІКС) ЦА. Подальший аналіз обмежено наземними ІКС, які обслуговують суб'єктів ЦА. Сюди належать системи аеропортових операцій (база операційних даних аеропорту, відображення рейсової інформації для пасажирів і операційних служб, реєстрація, оформлення багажу та посадка, транспортування, сортування та відстеження багажу тощо) та їхні підсистеми керування й моніторингу (SCADA/PLC, телеметрія), інтеграційні інтерфейси й операційні сервіси підтримки, корпоративні мережі та сервіси взаємодії, а також інженерні й охоронні системи керування об'єктів інфраструктури аеропорту. Бортіві комплекси авіоніки не входять до сфери, розглянутої в цій статті (Aryanti, Anwar, & Rahmawati, 2023).

Поширення цифрових технологій розширює сукупність точок входу та векторів доступу, під якими розуміємо набір цифрових активів, інтерфейсів, мережних вузлів, точок доступу, сервісів і процесів взаємодії, через які потенційний зломисник може реалізувати загрозу. Для об'єктів цивільної авіації це охоплює як прикладні ІТ-сервіси (електронна реєстрація, обмін повідомленнями, інформаційні портали для пасажирів і партнерів), так і технологічні компоненти наземної інфраструктури (системи керування багажем, енергетичні й інженерні підсистеми, доступ і відеоспостереження, мережні сегменти взаємодії з авіакомпаніями та підрядниками). З огляду на те, що об'єкти ЦА належать до критичної інфраструктури, наслідки кіберінцидентів можуть включати переривання бізнес-процесів, часткові або повні збої в наданні послуг аеропортом і суміжними службами, фінансові та репутаційні втрати, а в граничних випадках – ризики для безпеки пасажирів і персоналу (Dobrynin et al., 2018).



Міжнародні організації, зокрема IATA (International Civil Aviation Organization, 2025), систематично окреслюють виклики, пов'язані з новими кіберзагрозами для галузі, надають рекомендації та оновлюють стандарти для складових авіаційної екосистеми, що охоплює авіакомпанії, аеропорти та наземні служби. В Україні кіберзахист об'єктів критичної інфраструктури, включно з об'єктами цивільної авіації, регламентовано Постановою КМУ від 19.06.2019 № 518, яка вимагає впровадження політики управління ризиками інформаційної безпеки та застосування методик їх оцінювання й оброблення на всіх таких об'єктах (<https://zakon.rada.gov.ua/laws/show/518-2019-%D0%BF#Text>). Як методичну основу рекомендовано використовувати стандарт ДСТУ ISO/IEC 27005, що відповідає міжнародній практиці.

З огляду на зростання кіберзагроз і на нормативні вимоги, актуальним стає завдання обрати релевантні методи оцінювання ризиків саме для наземних ІКС цивільної авіації. У цій статті запропоновано формалізований підхід до порівняння таких методик на основі багатокритеріального аналізу з використанням методу зваженої суми, що забезпечує відтворюваний і інтерпретований вибір з урахуванням як відповідності стандартам, так і практичних обмежень упровадження в умовах реальної експлуатації.

Метою статті є розроблення математичної моделі багатокритеріального аналізу (МКА) для визначення найбільш релевантних фреймворків оцінювання кіберризиків в ІКС об'єктів ЦА на основі методу зваженої суми (WSM). Модель має забезпечувати формалізований і обґрунтований вибір методик управління ризиками шляхом порівняння різних альтернатив за системою кількісних критеріїв.

Огляд літератури. З кожним роком оцінювання ризиків кіберзагроз в ІКС стає дедалі важливішим для цивільної авіації і є одним із ключових напрямів сучасних досліджень у сфері авіаційної безпеки, і вимагає більш обґрунтованого і комплексного підходу, що поєднує міжнародні стандарти, національне законодавство, найкращі світові практики.

Науковці приділяють значну увагу розробленню методів, моделей для своєчасного виявлення кіберзагроз, використовують різні методики для кількісного та якісного оцінювання ризиків, удосконалюють існуючі методи.

Науковими дослідженнями окремих питань у сфері оцінювання ризиків кібербезпеки інформаційних систем на об'єктах критичної інфраструктури, зокрема і в авіаційній галузі присвячено роботи таких науковців як: А. Костін, О. Корченко, С. Гнатюк, А. Ільєнко, С. Ільєнко, І. Сіє, А. Алмараді, Е. Укванду, Ш. Фей та ін.

Дослідженнями А. Костіна запропоновано концепцію створення лабораторії, яка дозволяє як "тренажер" моделювати реалістичні кіберінциденти, створювати симуляції реальних кібератак і відпрацьовувати нові методи зниження ризиків.

Науковець Сунго Чо запропонував підхід, який передбачає розроблення тестового середовища для промислових систем, із використанням методології MITRE ATT&CK, як основи для побудови сценарію атак. Це дослідження є важливим для перевірки ефективності кіберзахисту систем управління аеропортами чи іншими важливими критичними об'єктами авіаційної інфраструктури, оскільки дає змогу відтворювати реалістичні умови атак і оцінювати готовність систем до потенційних загроз.

Останні наукові дослідження та публікації окреслюють актуальні виклики у сфері оцінювання ризиків кібербезпеки та пропонують різні шляхи їх подолання. Водночас ефективний захист ІКС цивільної авіації неможливий без комплексного поєднання міжнародних стандартів і гнучких методів, адаптованих до галузевих особливостей. Такий підхід передбачає не лише впровадження перевірених практик, але й їхнє систематичне вдосконалення. Із цією метою необхідно здійснити поглиблений аналіз наявних методик оцінювання ризиків ІКС, аби забезпечити формування більш цілісної та ефективної системи захисту цих критично важливих об'єктів.

Методи

Використано поєднання теоретичних і математичних методів. Здійснено аналіз і узагальнення наукових джерел і нормативних документів (систематизація стандартів ICAO/EASA/ISO), на підставі чого побудовано кортежну модель багатокритеріального аналізу (МКА) з визначенням критеріїв і ваг. Кількісне оброблення виконано методом зваженої суми із лінійним нормуванням min-max, ранжуванням альтернатив і пороговим правилом прийняття рішень.

Результати

Авіаційна галузь, що охоплює авіакомпанії, аеропорти та пов'язані з ними служби, функціонує як критична інфраструктура з багаторівневими цифровими взаємодіями. Інтенсивна цифровізація, використання інтелектуальних систем хмарних платформ, інтернету речей, аналітики великих даних і блокчейну формує складний ландшафт кіберзагроз, тобто множину потенційних векторів і каналів впливу на ІКС.

Захист цивільної авіації від кібератак здійснюється впровадженням Стандартів і рекомендованих практик кібербезпеки Міжнародною організацією цивільної авіації ICAO (International Civil Aviation Organization, 2018). Вона розробляє Стандарти та рекомендовану практику (SARPs, Standards and Recommended Practices), які є складовою Додатків до Чиказької конвенції 1944 р. Стандарти (Standards) є обов'язковими до виконання для держав-учасниць, тоді як рекомендована практика (Recommended Practices) установлює бажані заходи, які можуть адаптуватися відповідно до національних умов.

Розроблення та впровадження методик оцінювання ризиків кіберзагроз в авіації неможливі без урахування міжнародних стандартів і регламентів, які визначають вимоги до ІКС.

Огляд ключових міжнародних стандартів і методик управління ризиками. Управління ризиками кібербезпеки в ІКС авіаційної галузі ґрунтується на міжнародних стандартах і методиках, які формують цілісну систему оцінювання, контролю та зменшення ризиків.

Міжнародна практика управління ризиками інформаційної безпеки базується на ряді стандартів і фреймворків, серед яких ключову роль відіграють ISO/IEC 27005, ISO 31000, NIST SP 800-30, NIST CSF, а також галузеві підходи EASA Part-IS, ICAO Doc 10213, CRAMM, OCTAVE, EBIOS і TARA.

Розглянемо Міжнародний стандарт ISO/IEC 27005 "Інформаційні технології. Методи захисту. Управління ризиками інформаційної безпеки" (Information technology. Security techniques. Information security risk management), (ISO/IEC 27005: 2022. <https://www.iso.org/standard/80585.htm>) надає керівництво з управління ризиками інформаційної безпеки і є складовою сімейства стандартів ISO/IEC 27000. Процес включає етапи встановлення контексту, ідентифікації та аналізу ризиків, оброблення ризиків та моніторингу. Стандарт є гнучким, оскільки не встановлює конкретних методів оцінювання ризиків, що дає організаціям свободу у виборі інструментів.



ISO 31000 "Менеджмент ризиків. Принципи та настанови" (Risk management. Guidelines) (ISO 31000: 2018. <https://www.iso.org/standard/65694.html>) є загальним міжнародним документом, що визначає уніфіковані підходи до управління ризиками і містить принципи та настанови, придатні для будь-яких типів ризиків (фінансових, операційних, безпекових). Він охоплює етапи: встановлення контексту, оцінювання ризиків (ідентифікація, аналіз, оцінка), обробку, моніторинг та комунікацію ризиків. Перевага стандарту – універсальність і можливість інтеграції в будь-яку галузь.

Серед методик США особливе значення має NIST SP 800-30 "Guide for Conducting Risk Assessments" (NIST SP 800-30:2012. <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf>), який описує поетапний процес оцінювання ризиків інформаційних систем: від підготовки та збору даних до аналізу загроз, вразливостей і наслідків. Комплексна система NIST охоплює також пов'язані документи, серед яких:

- NIST SP 800-37 – упровадження Risk Management Framework у життєвий цикл систем безпеки;
- NIST SP 800-39 – багаторівневе управління ризиками на рівнях організації, місії та інформаційної системи;
- NIST SP 800-53 Rev.5 – каталог контролів безпеки й приватності;
- NIST SP 800-82 Rev.2 – керівництво з безпеки щодо промислових та авіаційних ІКС.

Ці стандарти взаємно доповнюють один одного, утворюючи цілісну модель, у якій оцінювання ризику (NIST SP 800-30) інтегрується з управлінськими процесами (NIST SP 800-37, 2018), впровадженням контролів (NIST SP 800-53, 2020) та специфікою промислових і авіаційних систем. Їх об'єднує NIST Cybersecurity Framework (CSF) – узагальнювальна "рамка" з п'яти базових функцій: ідентифікація (Identify), захист (Protect), виявлення (Detect), реагування (Respond), відновлення (Recover), що використовується як практичний орієнтир для критичної інфраструктури, зокрема й авіаційних ІКС.

Європейське агентство з авіаційної безпеки (European Union Aviation Safety Agency, EASA, 2023. <https://www.easa.europa.eu/en/document-library/regulations/commission-implementing-regulation-eu-2023203>) відповідає за встановлення стандартів безпеки польотів і сертифікацію авіаційних організацій.

Особливого значення набула регуляція EASA Part-IS (Information Security, 2023), що встановлює обов'язковість систем управління кіберризиками та проводить регулярне оцінювання загроз і вразливостей для авіаційних операторів, аеропортів і виробників. Регламент набуває повної чинності з лютого 2026 р. та вимагає інтеграції оцінювання кіберзагроз в операційні процеси.

Міжнародна асоціація повітряного транспорту IATA у своїх публікаціях (Aviation Cybersecurity Fact Sheet, 2025: <https://www.iata.org/contentassets/4c51b00fb25e4b60b38376a4935e278b/compilation-of-cyber-regs.pdf>) та (Compilation of Cyber Security Regulations, 2022: <https://www.iata.org/contentassets/4c51b00fb25e4b60b38376a4935e278b/compilation-of-cyber-regs.pdf>) узагальнює міжнародні вимоги, практики й рекомендації, наголошуючи на формуванні "спільнот довіри", стандартизації підходів і підвищенні обізнаності галузі щодо кіберризиків.

ICAO у своїх документах (ICAO Doc 10108, ICAO Doc 10213) визначають підходи до інтеграції кіберризиків у загальну авіаційну систему безпеки.

Відповідно до ICAO Doc 10213 (International Civil Aviation Organization, 2025) кількісне оцінювання ризику здійснюється за формулою

$$R = L \times I \times V, \quad (1)$$

де L – імовірність (Likelihood); I – вплив (Impact); V – рівень вразливості (Vulnerability).

Цей підхід дає змогу узгоджено оцінювати вплив кіберінцидентів на безпеку польотів і безперервність процесів у аеропортах.

В Україні нормативною основою управління кіберризиками є Постанова КМУ № 518 від 19.06.2019 р., яка встановлює загальні вимоги до кіберзахисту об'єктів критичної інфраструктури і зобов'язує операторів, зокрема й авіаційні організації, упроваджувати політику управління ризиками та здійснювати регулярне оцінювання відповідно до стандарту ДСТУ ISO/IEC 27005.

Поряд із нормативними документами, у практиці ризик-менеджменту застосовують відомі спеціалізовані методи, серед яких:

- CRAMM (CCTA Risk Analysis and Management Method) забезпечує детальний покроковий аналіз активів, загроз і вразливостей із формуванням каталогу контрзаходів. Ефективна для великих інфраструктурних об'єктів, але потребує значних ресурсів.
- OCTAVE (Operationally Critical Threat, Asset and Vulnerability Evaluation) орієнтована на організаційний рівень і поєднує технічні, людські й управлінські фактори, дає змогу будувати комплексну стратегію безпеки.
- EBIOS (Expression des Besoins et Identification des Objectifs de Securite), що розроблена агентством ANSSI, зосереджується на визначенні потреб безпеки й цілей захисту, забезпечуючи відповідність вимогам ISO 27005.
- TARA (Threat Assessment and Remediation Analysis) фокусується на аналізі актуальних загроз і пріоритизації ризиків з огляду на критичність функцій. Методика тісно пов'язана з даними розвідки загроз і застосовується як практичний інструмент "першої лінії" кіберзахисту.

Поєднання зазначених стандартів і методик формує основу інтегрованого ризик-орієнтованого підходу до управління кіберзагрозами в авіаційних ІКС, що відповідає вимогам міжнародної та національної нормативних баз і забезпечує узгодженість дій між авіаційними, технічними та кібернетичними підрозділами.

Отже, ефективна система оцінювання ризиків в авіаційних ІКС повинна ґрунтуватися на інтегрованому поєднанні нормативних (ISO, NIST, EASA, ICAO) та практичних методик (CRAMM, OCTAVE, EBIOS, TARA), що забезпечує структурованість, порівнюваність і відповідність міжнародним вимогам.

Проведений аналіз міжнародних стандартів і фреймворків (EASA Part-IS, ICAO Doc 10213, IATA Cybersecurity Framework, CRAMM, OCTAVE, EBIOS, TARA) показує, що всі вони пропонують різні рівні деталізації процесу управління ризиками, але не містять єдиного механізму кількісного порівняння та вибору найбільш релевантного підходу для конкретної організації цивільної авіації.



Зокрема, EASA Part-IS фокусується на нормативній відповідності та побудові системи управління інформаційною безпекою (СУІБ); ICAO Doc 10213 – на інтеграції кіберризиків у загальну систему авіаційних ризиків; CRAAM, OCTAVE, EBIOS, TARA провадять практичну реалізацію процесу оцінювання й управління ризиками, кожен із них має власні сильні сторони (глибина, гнучкість, структурованість) (Central Computer and Telecommunications Agency, 2001; International Civil Aviation Organization, 2018; International Civil Aviation Organization, 2025; National Agency for Information Systems Security, 2024).

У контексті авіаційної інфраструктури виникає потреба у формалізованому інструменті, який дозволяє одночасно враховувати кілька критеріїв ефективності підходів до управління ризиками (відповідність, гнучкість, вартість, масштабованість тощо) та забезпечує кількісне оцінювання їхньої придатності для конкретних ІКС.

Для цього доцільно застосовувати багатокритеріальний метод оцінювання (Multi-Criteria Analysis, MCA), що базується на методі зваженої суми (Weighted Sum Model, WSM). Такий підхід дозволяє об'єднати експертні оцінки за різними критеріями в єдину інтегральну метрику релевантності та здійснювати ранжування фреймворків за їхньою придатністю.

Метод WSM – це реалізація багатокритеріального аналізу, який дає змогу оцінювати альтернативи за сукупністю критеріїв, одночасно враховуючи відносну вагомість кожного з них. У результаті отримують інтегральний зважений бал, який використовують для кількісного порівняння варіантів і прийняття рішень (Potii et al., 2021).

Розглянемо математичну модель MCA методу визначення релевантних фреймворків для оцінювання ризиків в ІКС об'єктів цивільної авіації методом WSM.

Математична модель MCA для оцінювання ризиків в ІКС об'єктів цивільної авіації методом зваженої суми має такий узагальнений вигляд:

$$MCA = \{U_{i=1}^n MCA_i\} = \{MCA_1, MCA_2, MCA_3, \dots, MCA_n\}, \quad (2)$$

де $MCA_i \subseteq MCA$, $(1, n)$ – ключовий компонент кортежної моделі характеристик, що демонструє i -й ідентифікатор параметра оцінювання релевантного фреймворка, n – їхня кількість.

Наприклад, при $n = 8$ кортежна модель для узагальненої математичної моделі (2) виглядатиме як

$$MCA = \left\{ \bigcup_{i=1}^n MCA_i \right\} = \{MCA_1, MCA_2, MCA_3, \dots, MCA_n\} = \{M, K, W, Q, P, S, S^{norm}, D\}, \quad (3)$$

де $MCA_1 = M = \text{"Method"}$ – модуль, що містить множину альтернатив, тобто можливих фреймворків оцінювання кіберризиків; $MCA_2 = K = \text{"Criteria"}$ – модуль, що містить множину критеріїв оцінювання, за якими порівнюють альтернативи; $MCA_3 = W = \text{"Weight"}$ – модуль, який містить множину вагових коефіцієнтів критеріїв, що відображають відносну важливість кожного критерію з K в загальній оцінці; $MCA_4 = Q = \text{"Expert – assessments"}$ – модуль, який містить матрицю експертних оцінок; $MCA_5 = P = \text{"Weight – Sum"}$ – модуль матриці зважених компонентів; $MCA_6 = S = \text{"Integral – Sum"}$ – модуль сукупності підсумкових WSM-балів альтернатив (WSM – *Weighted Sum Method* – метод зваженої суми); $MCA_7 = S^{norm} = \text{"Integral – Sum – Normalized"}$ – модуль сукупності нормованих підсумкових WSM-балів альтернатив; $MCA_8 = D = \text{"Decision"}$ – модуль множин прийнятих рішень, тобто вибраних альтернатив-фреймворків, які визнані найкращими (релевантними) за результатами багатокритеріального аналізу.

Перший компонент узагальненої математичної моделі M – модуль сукупності альтернативних фреймворків оцінювання ризиків, допустимих для застосування в ІКС цивільної авіації визначається виразом

$$M = \{U_{i=1}^n M_i\} = \{M_1, M_2, M_3, \dots, M_n\}, \quad (4)$$

де $M_i \subseteq M$, $(1, n)$ – кожен елемент M_i репрезентує окремий фреймворк (методологію, стандарт чи підхід щодо оцінювання ризиків) із фіксованими властивостями й умовами застосування для оцінювання ризиків, релевантний для ІКС цивільної авіації. Тут n – загальна кількість таких альтернатив, надалі ці елементи підлягають порівнянню за множиною критеріїв K . Отже, M задає простір вибору для процедури ранжування та визначає перелік альтернатив, серед яких ухвалюється рішення.

Другий компонент узагальненої математичної моделі K – модуль структурованої сукупності оціночних критеріїв, за якими здійснюється зіставлення та ранжування альтернатив з M визначається виразом

$$K = \{U_{j=1}^m K_j\} = \{K_1, K_2, K_3, \dots, K_m\}, \quad (5)$$

де $K_j \subseteq K$, $(1, m)$, кожен елемент K_j описує окремий критерій (показник) якості або відповідності фреймворку (напр., охоплення технічних аспектів, відповідність міжнародним стандартам, простота впровадження, вартість застосування тощо); m – кількість вибраних критеріїв.

Третій компонент W – модуль сукупності вагових коефіцієнтів критеріїв, що відображають відносну важливість кожного критерію з K в загальній оцінці, визначається виразом

$$W = \{U_{j=1}^m W_j\} = \{W_1, W_2, W_3, \dots, W_m\}, \quad (6)$$

де $W_j \subseteq W$, $(1, m)$, кожна вага W_j зіставляється з критерієм K_j . Більше значення W_j означає, що критерій K_j більше впливає на підсумкову оцінку альтернативи, ніж критерії з меншими вагами.

Зазвичай, ваги нормуються так, що

$$\sum_{j=1}^m W_j = 1, \quad W_j \geq 0 \quad \forall j \in \{1, \dots, m\}. \quad (7)$$

Ваги є дробовими числами, сумарно рівними 1, що відображає розподіл важливості між критеріями.

Четвертий компонент Q – матриця експертних оцінок визначається виразом

$$Q := \{Q_i\}_{i=1}^n, \quad Q_i := \langle q_{i1}, \dots, q_{im} \rangle \in \mathbb{R}^m, \quad (8)$$

де $i = \{1, \dots, n\}$ – індекси альтернатив; $j = \{1, \dots, m\}$ – індекси критеріїв; q_{ij} – відображає значення показника альтернативи M_i за критеріями K_j , сформоване на підставі експертного оцінювання або аналітичного розрахунку.



П'ятий модуль P – модуль матриці зважених компонентів визначається виразом

$$P := \{P_i\}_{i=1}^n, P_i := \langle p_{i1}, p_{i2}, \dots, p_{im} \rangle, p_{ij} := W_j q_{ij}, \quad (9)$$

де p_{ij} – зважений компонент за критерієм K_j ; W_j – вага критерію K_j (з $\sum_{j=1}^m W_j = 1, W_j \geq 0$).

Матрицю P отримують шляхом почленного множення стовпців матриці оцінок Q на відповідні ваги з W . Отже, p_{ij} кількісно відображає, наскільки альтернатива M_i задовольняє критерій K_j з урахуванням критерію K_j .

Шостий модуль S – інтегральна оцінка (сукупність інтегральних підсумкових WSM-балів альтернатив) визначається виразом

$$S := \{S_i\}_{i=1}^n, S_i := \sum_{j=1}^m p_{ij} = \sum_{j=1}^m W_j q_{ij}, \quad (10)$$

де S_i – інтегральний бал альтернативи M_i ; W_j – вага критерію K_j .

Для кожної альтернативи M_i обчислюється сумарний бал S_i як сума всіх зважених оцінок за всіма критеріями. Модуль S містить інтегральні (узагальнені) оцінки всіх альтернатив і слугує основою для їх порівняння. Більше значення S_i означає кращу сукупну відповідність фреймворка M_i заданим критеріям.

Сьомий модуль S_i^{norm} нормованих інтегральних балів визначається виразом

$$S_i^{norm} = \{S_i^{norm}\}_{i=1}^n, S_i^{norm} = \frac{S_i - S_{min}}{S_{max} - S_{min}} \in [0, 1], I = \{1, \dots, n\}, \quad (11)$$

де S_i^{norm} – нормований інтегральний бал альтернативи M_i ; S_{max} – максимум підсумкових інтегральних балів; S_{min} – мінімум підсумкових інтегральних балів.

Восьмий модуль D – модуль прийняття рішень, тобто фреймворків-альтернатив, відібраних як оптимальні (релевантні) за підсумками багатокритеріального оцінювання, визначається виразом

$$D = \{M_i \in M | S_i^{norm} \geq \tau_{min}\}, \quad (12)$$

де S_i^{norm} – нормований інтегральний бал альтернативи M_i ; τ_{min} – поріг прийняття рішень.

Для прийняття рішень уведемо три впорядковані порогові значення інтегральної оцінки S_i^{norm} . $S_i^{norm} \in [0, 1]$:

$$0 \leq \tau_{acc} \tau_{rec} \tau_{min}, \quad (13)$$

де τ_{rec} – поріг рекомендації: рівень, з якого фреймворк вважається безумовно доцільним; τ_{acc} – поріг прийнятності: мінімальний рівень якості / релевантності, з якого варіант можна рекомендувати умовно; τ_{min} – поріг мінімальної доцільності: нижня межа, нижче якої варіант не рекомендується.

Задаємо чотири рівні рішення, у кожному з яких визначимо порогові значення. Задані рівні представлено в табл. 1.

Таблиця 1

Рівні прийняття рішень

Діпазон	Категорія	Опис
$D^4 = \{M_i S_i^{norm} \geq \tau_{rec}\}$	Безумовно рекомендовані	Фреймворк має дуже високий рівень безпеки, продуктивності й ефективності. Рекомендується для використання у критичних системах
$D^3 = \{M_i \tau_{acc} \leq S_i^{norm} < \tau_{rec}\}$	Умовно рекомендовані	Алгоритм загалом ефективний, але може мати окремі недоліки (напр., потребує додаткових налаштувань або оптимізації). Використання можливе, але бажано провести аудит або додаткове тестування
$D^2 = \{M_i \tau_{acc} \tau_{min} \leq S_i^{norm} < \tau_{acc}\}$	Обмежене використання	Алгоритм має помірну ефективність або окремі недоліки, які роблять його менш придатним у критичних середовищах. Використання можливе лише за відсутності кращих альтернатив
$D^1 = \{M_i S_i^{norm} < \tau_{min}\}$	Не рекомендовані	Алгоритм має низьку ефективність або потенційні вразливості, що роблять його небезпечним або малоефективним для використання у сучасних системах. Слід уникати його впровадження

Якщо дві чи більше альтернативи мають однакові нормовані інтегральні оцінки S_i^{norm} в тому самому рівні, вибір здійснюємо порівнянням векторів зважених компонент P_i за порядком важливості за критеріями.

Зазначений практичний підхід допоможе приймати рішення на основі конкретного значення S_i^{norm} , наприклад, чи фреймворк рекомендований, чи краще його уникати у критичних системах.

Приклад застосування. Розглянемо використання наведеної моделі на прикладі. Для обґрунтованого вибору найкращого підходу в оцінюванні ризиків кіберзагроз проведемо MCA методом WSM, що враховує найважливіші критерії ефективності методики у сфері ЦА.

Згідно з формулою (4) назначимо множину альтернативних фреймворків (M), де кожен елемент M_i репрезентує окремий фреймворк. Множину фреймворків зазначено в табл. 2.

Таблиця 2

Множина фреймворків

Множина M_i	Назва альтернативних фреймворків
M_1	NIST SP 800-30/NIST CSF
M_2	CRAMM
M_3	OCTAVE
M_4	EBIOS
M_5	TARA



Для порівняння фреймворків визначимо п'ять оціночних критеріїв (K_j), кожному з яких призначимо ваговий коефіцієнт W_j ($\sum_{j=1}^m W_j = 1$) згідно з формулами (5) і (6). Дані представлено в табл. 2.

В табл. 3 відображено вагові коефіцієнти для п'яти критеріїв, які визначають відносну важливість кожного критерію в інтегральній оцінці. Сума ваг дорівнює 1, що забезпечує коректну інтерпретацію підсумкових балів як зваженої "середньої корисності".

Таблиця 3

Розподіл вагових коефіцієнтів за критеріями

Критерій, K_j	Назва критерію	Пояснення (коротко)	Вагові коефіцієнти, W_j
K_1	Відповідність міжнародним стандартам	Ступінь сумісності / узгодження методики з міжнародно визнаними підходами (ICAO/EASA/ISO)	0,25
K_2	Гнучкість застосування	Адаптивність до масштабу / типів ІКС, сумісність з іншими практиками	0,20
K_3	Глибина аналізу ризиків	Рівень деталізації: загрози, вразливості, сценарії, контролю	0,25
K_4	Легкість упровадження	Складність запуску, потреба в експертизі / команді	0,15
K_5	Вартість / ресурсомісткість	Умовна інтегральна "ціна" (час, люди, бюджет)	0,15

У табл. 4 наведено результати оцінювання з урахуванням вагових коефіцієнтів, тобто подано вихідну матрицю Q – матрицю експертних оцінок, де q_{ij} – експертна оцінка методики M_i за критерієм K_j за шкалою від 1 до 5 балів за кожним критерієм (5 – найвищий рівень відповідності критерію, 1 – найнижчий). Усі критерії трактують як вигравні (чим більше, тим краще), а для K_5 застосовано "обернену" інтерпретацію (5 означають нижчі витрати / вища віддача, 1 – дорогий підхід).

Таблиця 4

Матриця експертних оцінок кожної методики

Фреймворк, M_i	q_{i1}	q_{i2}	q_{i3}	q_{i4}	q_{i5}
NIST SP 800-30 / CSF	3	4	5	4	3
CRAMM	2	3	5	1	1
OCTAVE (Allegro)	3	4	4	3	3
EBIOS RM	4	4	5	2	2
TARA	2	3	3	5	5

Скориставшись формулою (9), розрахували зважені компоненти p_{ij} за кожним фреймворком, що кількісно показують внесок кожного критерію у підсумкову оцінку відповідного фреймворка з урахуванням його ваги. Надалі ці компоненти підсумовують для отримання інтегрального бала S_i і порівняння методик між собою. Результати представлено в табл. 5.

Таблиця 5

Зважені компоненти (внесок критерію)

Фреймворк, M_i	p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}
NIST SP 800-30 / CSF	0,75	0,80	1,25	0,60	0,45
CRAMM	0,50	0,60	1,25	0,15	0,15
OCTAVE (Allegro)	0,75	0,80	1,00	0,45	0,45
EBIOS RM	1,00	0,80	1,25	0,30	0,30
TARA	0,50	0,60	0,75	0,75	0,75

Зважені компоненти демонструють, завдяки яким критеріям методики отримують перевагу. Для NIST максимальний внесок дає критерій K_3 (глибина аналізу) (1,25), що узгоджується з практичною орієнтацією SP 800-30/CSF. Для EBIOS RM провідними критеріями є K_1 (відповідність) і K_3 (глибина) (1,00 та 1,25), однак сумарний бал знижується через менші компоненти критеріїв K_4 (легкості) та K_5 (вартості) (по 0,30 кожен). OCTAVE (Allegro) демонструє збалансованість (усі компоненти близькі до середніх значень), що формує стабільний інтегральний результат.

TARA отримує найвищі компоненти за K_4 (легкості) та K_5 (вартість) (по 0,75 кожен), тому добре підходить для швидкої пріоритизації дій за обмежених ресурсів, але поступається за глибиною іншим фреймворкам. CRAMM має найбільший компонент глибини (1,25), проте дуже низькі компоненти K_4 (легкості) та K_5 (вартості) (по 0,150 кожен).

Компонентна подача полегшує аудит і аргументацію, тобто можна відразу бачити, чи відповідає домінування певної методики стратегічним пріоритетам (напр., $K_3 > K_1$ чи навпаки).

Скориставшись формулою (10), обрахували зважену суму S_i (інтегральний бал для кожної альтернативи M_i та згідно з формулою (11) розрахували нормований інтегральний бал S_i^{norm} альтернативи M_i . Результати підсумкових зважених WSM (інтегральних балів) і нормованих інтегральних балів представлено в табл. 6 ранжування методик. Ранжування виконано за спаданням інтегрального бала S_i .



Таблиця 6

Підсумкові та нормовано інтегральні бали (WSM)

Фреймворк, M_i	Інтегральний бал, S_i	Нормовано інтегральний бал, S_i^{norm}
NIST SP 800-30 / CSF	3,85	1,00
CRAMM	2,65	0,00
OCTAVE (Allegro)	3,45	0,67
EBIOS RM	3,65	0,83
TARA	3,35	0,58

Скориставшись формулою (12), введемо три впорядковані порогові значення.

Нехай $\tau_{i_{rec}} = 80$ – поріг рекомендації;

$\tau_{i_{acc}} = 0,55$ – поріг прийнятності;

τ_{min} – поріг мінімальної доцільності.

Тоді відповідно до табл. 1, задаємо чотири рівні рішення, у кожному з яких визначено порогові значення. Задані рівні прийняття рішень представлено в табл. 7.

Таблиця 7

Рівні прийняття рішень за заданих порогових значень

Діпазон, D^i	Категорія	Опис
$D^4 = \{M_i S_i^{norm} \geq 0,80\}$	Безумовно рекомендовані	Фреймворк має дуже високий рівень безпеки, продуктивності й ефективності. Рекомендується для використання у критичних системах
$D^3 = \{M_i 0,55 \leq S_i^{norm} < 0,80\}$	Умовно рекомендовані	Алгоритм загалом ефективний, але може мати окремі недоліки (напр., потребує додаткових налаштувань або оптимізації). Використання можливе, але бажано провести аудит або додаткове тестування
$D^2 = \{M_i 0,30 \leq S_i^{norm} < 0,55\}$	Обмежене використання	Алгоритм має помірну ефективність або окремі недоліки, які роблять його менш придатним у критичних середовищах. Використання можливе лише за відсутності кращих альтернатив
$D^1 = \{M_i S_i^{norm} < 0,30\}$	Не рекомендовані	Алгоритм має низьку ефективність або потенційні вразливості, що роблять його небезпечним або малоефективним для використання в сучасних системах. Слід уникати його впровадження

Скориставшись табл. 7, кожній з альтернатив визначили рівні прийняття рішень. Результати представлено в табл. 8.

Таблиця 8

Прийняття рішень щодо використання фреймворків

Фреймворк, M_i	Нормовано інтегральний бал, S_i^{norm}	Рівні прийняття рішень, D^i
NIST SP 800-30 / CSF	1,00	Безумовно рекомендовані
CRAMM	0,00	Не рекомендовані
OCTAVE (Allegro)	0,67	Умовно рекомендовані
EBIOS RM	0,83	Безумовно рекомендовані
TARA	0,58	Умовно рекомендовані

NIST SP 800-30/CSF має найвищий показник $S_i^{norm} = 1,00$ й віднесений до рівня "безумовно рекомендовані". Це зумовлено поєднанням максимальної глибини аналізу ризиків і високої відповідності вимогам і стандартам за збереження прийнятної легкості впровадження й витрат.

EBIOS RM має $S_i^{norm} = 0,83$ і також потрапляє до рівня "безумовно рекомендовані". Завдяки сильним компонентам відповідності та глибини він є близьким до лідера, однак поступається NIST через вищу ресурсомісткість / складність практичної реалізації (Intel Corporation, 2018; National Institute of Standards and Technology (NIST), 2018; National Institute of Standards and Technology (NIST), 2020).

OCTAVE (Allegro) віднесено до рівня "умовно рекомендовані". Методика демонструє стабільний, збалансований профіль за критеріями, однак величини відповідності і глибина недостатні, щоб перейти у верхній рівень за наявних ваг. Аналогічно,

TARA класифікується як "умовно рекомендовані". Попри те, що її перевагою є практичність і відносно низькі витрати / швидкість, упровадження обмеженості глибини та відповідності не дозволяє досягти порога рекомендації.

CRAMM отримує найнижчу нормовану оцінку та належить до рівня "не рекомендовані": попри високу глибину, метод є ресурсомістким і менш гнучким, що суттєво знижує інтегральний результат за зваженою сумою.



Оскільки WSM припускає адитивність і компенсаторність (слабкість за одним критерієм частково компенсується силою за іншим), для перевірки стійкості результатів (ранжування) до зміни пріоритетів виконано аналіз чутливості (варіювання ваг на $\pm 10\text{--}20\%$).

Аналіз чутливості виконано за рахунок перерозподілу ваг критеріїв відповідно до зміни пріоритетів (безпека / витрати / швидкість впровадження) для перевірки стійкості ранжування до змін. Для кожного сценарію перерозподілено ваги так, щоб $\sum_{j=1}^m W_j = 1$, і перераховано WSM-бали.

Виокремимо три сценарії, на які робиться практичний фокус: "безпеково-критичний", "бюджетно-орієнтований", "швидке впровадження". І кожному сценарію за нашими критеріями призначимо ваги. Дані представлено в табл. 9.

Таблиця 9

Альтернативні сценарії ваг критеріїв

Сценарій	W_1	W_2	W_3	W_4	W_5
A (безпеково-критичний)	0,25	0,15	0,35	0,15	0,10
B (бюджетно-орієнтований)	0,15	0,10	0,25	0,20	0,30
C (швидке впровадження)	0,15	0,10	0,20	0,30	0,25

Для безпеково-критичного сценарію підсилили вагу критерію K_3 (стійкість), для бюджетно-орієнтованого сценарію підсилили вагу критерію K_5 (вартість), для швидкого впровадження підсилили критерії K_4 (легкості) та K_5 (вартості) й повторно обчислили WSM-компоненти S_i для кожного сценарію і по кожному фреймворку M_i . Результати представлено в табл. 10–12. Для розрахунку матрицю експертних оцінок залишили незмінною, згідно з табл. 4.

Таблиця 10

WSM-компоненти, інтегральний і нормовано інтегральний бал для безпеково-критичного сценарію (ваги 0,25 / 0,15 / 0,35 / 0,15 / 0,10)

Фреймворк, M_i	p_{i1}	p_{i1}	p_{i2}	p_{i3}	p_{i4}	Інтегральний бал, S_i^{s-c}	Нормовано інтегральний бал, $S_i^{norm(s-c)}$
NIST SP 800-30 / CSF	0,75	0,60	1,75	0,60	0,30	4,00	1,00
CRAMM	1,00	0,60	1,75	0,30	0,20	3,85	0,86
OCTAVE (Allegro)	0,75	0,60	1,40	0,45	0,30	3,50	0,52
EBIOS RM	0,50	0,45	1,05	0,75	0,50	3,25	0,29
TARA	0,50	0,45	1,75	0,15	0,10	2,95	0,00

S_i^{s-c} – інтегральний бал по кожному фреймворку для безпеково-критичного сценарію;

$S_i^{norm(s-c)}$ – нормований інтегральний бал по кожному фреймворку для безпеково-критичного сценарію.

Таблиця 11

WSM-компоненти, інтегральний і нормовано інтегральний бал для бюджетно-орієнтованого сценарію (ваги 0,15 / 0,10 / 0,25 / 0,20 / 0,30)

Фреймворк, M_i	p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	Інтегральний бал, S_i^{b-or}	Нормовано інтегральний бал, $S_i^{norm(b-or)}$
NIST SP 800-30 / CSF	0,45	0,40	1,25	0,80	0,90	3,80	0,92
CRAMM	0,60	0,40	1,25	0,40	0,60	3,25	0,00
OCTAVE (Allegro)	0,45	0,40	1,00	0,60	0,90	3,35	0,17
EBIOS RM	0,30	0,30	0,75	1,00	1,50	3,85	1,00
TARA	0,45	0,40	1,25	0,80	0,90	3,80	0,92

S_i^{b-or} – інтегральний бал по кожному фреймворку для бюджетно-орієнтованого сценарію;

$S_i^{norm(b-or)}$ – нормований інтегральний бал по кожному фреймворку для бюджетно-орієнтованого сценарію.

Таблиця 12

WSM-компоненти, інтегральний і нормовано інтегральний бал для сценарію швидкого впровадження (ваги 0,15 / 0,10 / 0,20 / 0,30 / 0,25)

Фреймворк, M_i	p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	Інтегральний бал, S_i^{r-imp}	Нормовано інтегральний бал, $S_i^{norm(r-imp)}$
NIST SP 800-30 / CSF	0,45	0,40	1,00	1,20	0,75	3,80	0,92
CRAMM	0,60	0,40	1,00	0,60	0,50	3,10	0,53
OCTAVE (Allegro)	0,45	0,40	0,80	0,90	0,75	3,30	0,64
EBIOS RM	0,30	0,30	0,60	1,50	1,25	3,95	1,00
TARA	0,30	0,30	1,00	0,30	0,25	2,15	0,00



S_i^{r-imp} – інтегральний бал по кожному фреймворку для сценарію швидкого впровадження;

$S_i^{norm(r-imp)}$ – нормований інтегральний бал по кожному фреймворку для сценарію швидкого впровадження.

Скориставшись табл. 7, кожний з альтернатив визначили рівні прийняття рішень за трьома сценаріями: безпечно-критичним, бюджетно-орієнтованим, швидкого впровадження. Результати занесено в табл. 13–15 відповідно.

Таблиця 13

Прийняття рішень щодо використання фреймворків
для безпечно-критичного сценарію

Фреймворк, M_i	Нормовано інтегральний бал, $S_i^{norm(s-c)}$	Рівні прийняття рішень, D^i
NIST SP 800-30 / CSF	1,00	Безумовно рекомендовані
CRAMM	0,86	Безумовно рекомендовані
OCTAVE (Allegro)	0,52	Обмежене використання
EBIOS RM	0,29	Не рекомендовані
TARA	0,00	Не рекомендовані

Лідерами для безпечно-критичного сценарію є фреймворки NIST SP 800-30 / CSF і CRAMM. Вони мають найвищі ваги за критеріями глибини аналізу і відповідність стандартам. CRAMM втрачає у практичності, але це не критично за цього профілю ваг.

Таблиця 14

Прийняття рішень щодо використання фреймворків
для бюджетно-орієнтованого сценарію

Фреймворк, M_i	Нормований інтегральний бал, $S_i^{norm(b-or)}$	Рівні прийняття рішень, D^i
NIST SP 800-30 / CSF	0,92	Безумовно рекомендовані
CRAMM	0,00	Не рекомендовані
OCTAVE (Allegro)	0,17	Не рекомендовані
EBIOS RM	1,00	Безумовно рекомендовані
TARA	0,92	Безумовно рекомендовані

Лідерами для бюджетно-орієнтованого сценарію є фреймворки EBIOS RM, NIST SP 800-30 / CSF TARA. EBIOS очолює лідерство завдяки хорошему компромісу "глибина-практичність"; NIST SP 800-30 / CSF – завдяки стабільно високим таким критеріям: відповідність, гнучкість і глибина; TARA – через сильні критерії – легкість впровадження та вартість.

Таблиця 15

Прийняття рішень щодо використання фреймворків
для сценарію швидкого впровадження

Фреймворк, M_i	Нормовано інтегральний бал, $S_i^{norm(r-imp)}$	Рівні прийняття рішень, D^i
NIST SP 800-30 / CSF	0,92	Безумовно рекомендовані
CRAMM	0,53	Обмежене використання
OCTAVE (Allegro)	0,64	Умовно рекомендовані
EBIOS RM	1,00	Безумовно рекомендовані
TARA	0,00	Не рекомендовані

Лідерами сценарію швидкого впровадження є фреймворки NIST SP 800-30 / CSF, EBIOS RM. Ці методики забезпечують достатню глибину при прийнятній операційній "легкості".

Узагальнюючи питання вибору "найкращої методики", можна сказати, що NIST SP 800-30/CSF є стійким лідером (топівень у всіх сценаріях), коли в пріоритеті відповідність і глибина; оптимальний для безпечно-критичних систем і як базовий корпоративний підхід.

За потреби швидкого або економного розгортання у разі збереження достатньої глибини краще використовувати EBIOS RM. Він є лідером у сценаріях бюджет / швидкість.

Якщо потрібна операційна простота й економність, то раціонально використовувати методику TARA, але вона програє за глибиною та відповідністю.

Компромисним, переважно "умовним" вибором є OCTAVE (Allegro), CRAMM доречно використовувати у "вузьких" випадках, коли прийнятна висока ресурсомісткість заради формалізованої глибини.

Узагальнюючи результати, зазначимо, що застосування побудованої кортежної моделі MCA з обчисленням інтегрального бала за методом WSM продемонструвало свою практичну ефективність для вибору фреймворків оцінювання кіберризиків в ІКС авіації. На етапі обчислень підхід залишається простим і зрозумілим. Підсумковий бал S_i для кожної методики обраховують як суму добутків "оцінка × вага" за всіма критеріями, тож внесок кожного критерію легко побачити у векторі зважених компонент P_i . Для вхідних даних достатньо узгодженої експертної



шкали (напр., 1–5), а подальше нормування зводить результати до єдиної шкали [0,1]. Це спрощує інтерпретацію для керівництва, внутрішнього аудиту та регуляторних органів і робить аргументацію вибору прозорою.

Ключовою перевагою моделі є керована гнучкість. Критерії та їхні ваги W_i легко переналаштовуються під управлінський контекст: безпеково-критичний, бюджетно-орієнтований або сценарій швидкого впровадження. Саме це показано в отриманих нами результатах: за пріоритету відповідності та глибини (безпеково-критичний сценарій) найкращими виявилися NIST SP 800-30/CSF (лідер) і CRAMM, якщо спостерігаємо зсув ваг у бік вартості та швидкості імплементації, то верхні позиції посідають EBIOS RM і NIST, тоді як OCTAVE і TARA проявляють себе як раціональні компромісні варіанти залежно від обмежень.

Запроваджена чотирирівнева шкала прийняття з трьома порогами перетворює числові бали на чіткі управлінські висновки, такі як "безумовно рекомендовано", "умовно рекомендовано", "обмежене застосування", "не рекомендовано" й уможливорює єдине правило вибору переможця в межах рівня (Swan, 2019).

Дискусія і висновки

Застосування багатокритеріального аналізу за методом WSM продемонструвало ефективність як інструмент для обґрунтованого порівняння та ранжування методик оцінювання кіберризиків в ІКС ЦА. Метод забезпечує прозору логіку прийняття рішень, оскільки дозволяє кількісно врахувати внесок кожного критерію та вагового коефіцієнта у підсумкову оцінку.

Перевага підходу полягає у простоті реалізації, відсутності складних вимог до оброблення даних і можливості швидкої адаптації вагових коефіцієнтів до контексту завдання, наприклад, посилення критерію "глибина аналізу" для критичних систем безпеки або підвищення ваги "легкості впровадження" за потреби оперативного розгортання.

Проведення аналізу чутливості підвищує достовірність результатів, дозволяючи оцінити стабільність отриманого ранжування у разі зміни вагових коефіцієнтів. Такий підхід сприяє сценарному аналізу та мінімізує суб'єктивність у прийнятті управлінських рішень.

У контексті ІКС ЦА використання методу WSM є доцільним завдяки його узгодженості з вимогами міжнародних стандартів ISO/IEC 27005, ISO 31000 та NIST CSF. Отримані результати можуть слугувати основою для формування комбінованої системи управління ризиками, де WSM використовують як базовий інструмент для попередньої пріоритизації методик або сценаріїв ризику.

Отже, запропонований підхід поєднує простоту, гнучкість і аналітичну обґрунтованість, що робить його придатним для практичного застосування в системах кібербезпеки авіаційної інфраструктури, а також перспективним для подальшого розвитку, зокрема і шляхом інтеграції елементів штучного інтелекту для автоматизації вибору критеріїв і динамічного оновлення вагових коефіцієнтів.

Внесок авторів: Анна Ільєнко – концептуалізація, методологія; Валентина Телющенко – аналіз джерел, підготовка огляду літератури, теоретичних засад дослідження, проведення дослідження.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Aryanti, U., Anwar, M. T., & Rahmawati, T. (2023). Information security risk management using OCTAVE Allegro method at university. *International Journal of Engineering & Education Research (IJEER)*, 3(4). <https://doi.org/10.46336/ijeer.v3i4.506>
- Dobrynin, I., Radivilova, T., Maltseva, N., & Ageev, D. (2018). Use of approaches to the methodology of factor analysis of information risks for the quantitative assessment of information risks based on the formation of cause-and-effect links. <https://doi.org/10.1109/INFOCOMMST.2018.8632022>
- Central Computer and Telecommunications Agency (CCTA). (2001). *CRAMM User Guide Issue 2.0*. London: CCTA – The Stationery Office. <https://www.scribd.com/document/670786636/CRAMMGuide>
- International Civil Aviation Organization (ICAO). (2018). *Doc 10108: Global Risk Context Statement*. Montreal: ICAO. <https://www.scribd.com/document/688824651/DOC-10108-ICAO-RISK-CONTEXT-STATEMENT>
- International Civil Aviation Organization (ICAO). (2025). *Doc 10213: Global Cyber Risk Considerations (Advance Unedited Edition)*. Montreal: ICAO. <https://www.icao.int/Security/Pages/Cybersecurity.aspx>
- Intel Corporation. (2018). *Threat Assessment & Remediation Analysis (TARA)*. <https://www.intel.com/content/www/us/en/security/threat-assessment-remediation-analysis.html>
- National Agency for Information Systems Security. (2024). *EBIOS Risk Manager – The method (Version 1.5, English edition)*. Paris: ANSSI. Retrieved from https://cyber.gouv.fr/sites/default/files/document/20240901_Guide-EBIOS-RM-V1.5.pdf
- National Institute of Standards and Technology (NIST). (2018). *SP 800-37 Rev.2: Guide for Applying the Risk Management Framework to Federal Information Systems*. Gaithersburg, MD: NIST. <https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final>
- National Institute of Standards and Technology (NIST). (2020). *SP 800-53 Rev.5: Security and Privacy Controls for Information Systems and Organizations*. Gaithersburg, MD: NIST. <https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final>
- Potil, O. V., Horbenko, Yu. I., Zamula, O. A., & Isirova, K. V. (2021). Analysis of methods for assessing and managing cyber and information security risks. *Radiotekhnika*, 206, 5–24. <https://doi.org/10.30837/rt.2021.3.206.01>
- Swan, D. (2019). *Assessing Primary Cyber Threats to an International Airport's Critical Information Systems*. University of Maryland University College. Retrieved from https://www.academia.edu/1460287/Assessing_Primary_Cyber_Threats_to_an_International_Airport_s_Critical_Information_Systems

References

- Aryanti, U., Anwar, M. T., & Rahmawati, T. (2023). Information security risk management using OCTAVE Allegro method at university. *International Journal of Engineering & Education Research (IJEER)*, 3(4). <https://doi.org/10.46336/ijeer.v3i4.506>
- Dobrynin, I., Radivilova, T., Maltseva, N., & Ageev, D. (2018). Use of approaches to the methodology of factor analysis of information risks for the quantitative assessment of information risks based on the formation of cause-and-effect links. <https://doi.org/10.1109/INFOCOMMST.2018.8632022>
- Central Computer and Telecommunications Agency (CCTA). (2001). *CRAMM User Guide Issue 2.0*. London: CCTA – The Stationery Office. <https://www.scribd.com/document/670786636/CRAMMGuide>
- International Civil Aviation Organization (ICAO). (2018). *Doc 10108: Global Risk Context Statement*. Montreal: ICAO. <https://www.scribd.com/document/688824651/DOC-10108-ICAO-RISK-CONTEXT-STATEMENT>
- International Civil Aviation Organization (ICAO). (2025). *Doc 10213: Global Cyber Risk Considerations (Advance Unedited Edition)*. Montreal: ICAO. <https://www.icao.int/Security/Pages/Cybersecurity.aspx>
- Intel Corporation. (2018). *Threat Assessment & Remediation Analysis (TARA)*. <https://www.intel.com/content/www/us/en/security/threat-assessment-remediation-analysis.html>
- National Agency for Information Systems Security. (2024). *EBIOS Risk Manager – The method (Version 1.5, English edition)*. Paris: ANSSI. Retrieved from https://cyber.gouv.fr/sites/default/files/document/20240901_Guide-EBIOS-RM-V1.5.pdf



- National Institute of Standards and Technology (NIST). (2018). *SP 800-37 Rev.2: Guide for Applying the Risk Management Framework to Federal Information Systems*. Gaithersburg, MD: NIST. <https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final>
- National Institute of Standards and Technology (NIST). (2020). *SP 800-53 Rev.5: Security and Privacy Controls for Information Systems and Organizations*. Gaithersburg, MD: NIST. <https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final>
- Potii, O. V., Horbenko, Yu. I., Zamula, O. A., & Isirova, K. V. (2021). Analysis of methods for assessing and managing cyber and information security risks. *Radiotekhnika*, 206, 5–24. <https://doi.org/10.30837/rt.2021.3.206.01>
- Swan, D. (2019). *Assessing Primary Cyber Threats to an International Airport's Critical Information Systems*. University of Maryland University College. Retrieved from https://www.academia.edu/1460287/Assessing_Primary_Cyber_Threats_to_an_International_Airport_s_Critical_Information_Systems.

Отримано редакцією журналу / Received: 09.12.25

Прорецензовано / Revised: 20.12.25

Схвалено до друку / Accepted: 25.12.25

Anna ILIENKO, PhD (Engin.), Assoc. Prof.

ORCID ID: 0000-0001-8565-1117

e-mail: anna.ilienko@npp.kai.edu.ua

State Non-Commercial Company "State University «Kyiv Aviation Institute»", Kyiv, Ukraine

Valentyna TELIUSHCHENKO, PhD Student

ORCID ID: 0000-0001-6026-5105

e-mail: valentyna.teliushchenko@npp.kai.edu.ua

State Non-Commercial Company "State University «Kyiv Aviation Institute»", Kyiv, Ukraine

METHODS FOR ASSESSING THE RISKS OF CYBER THREATS IN INFORMATION AND COMMUNICATION SYSTEMS OF CIVIL AVIATION FACILITIES

Background. Civil aviation is a critical infrastructure, and its ground-based information and communication systems (ICS) – airport operating platforms, integration gateways, networks, and support services – face a diverse landscape of cyber threats. A transparent and reproducible tool aligned with ICAO/EASA/ISO requirements and the operational constraints of airports and air carriers is needed to guide the selection of risk assessment approaches.

Methods. A tuple model of multicriteria analysis (MCA) with linear weighted summation (WSM) is proposed. NIST SP 800-30/CSF, CRAMM, OCTAVE Allegro, EBIOS RM, TARA are taken as alternatives. Criteria: compliance with standards, depth of analysis, flexibility, ease of implementation, cost/resource intensity. The weights are normalized (the sum is 1), the expert estimates are on a scale of 1-5 with an "inverse" interpretation of the cost. The matrix of weighted components, the integral score and its normalized form are calculated. A four-level decision-making scale with three thresholds is introduced. The sensitivity analysis is carried out by means of scenario variation of weights.

Results. The overall ranking confirmed the superiority of NIST SP 800-30/CSF, while EBIOS RM demonstrated a close second. In the security-critical scenario, NIST and CRAMM take the top positions, while in the budget and rapid implementation scenarios, EBIOS RM takes the lead, followed by NIST. OCTAVE Allegro should be considered as a "conditionally recommended" balanced compromise, and TARA as a solution for rapid prioritization with limited resources. In typical operating conditions where cost and efficiency are key, CRAMM is inferior due to its high resource intensity. The results obtained generally remain robust to moderate changes in weights.

Conclusions. The MCA/WSM model provides a transparent, reproducible, and interpretable comparison of risk management frameworks for civil aviation ground-based ICS and can be integrated into decision support systems. The approach combines regulatory compliance with practical feasibility and supports scenario-based selection. Prospects include automating the selection of weights and criteria using AI methods.

Keywords: cybersecurity, civil aviation, ICS, risk assessment, multi-criteria analysis, WSM, frameworks.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

Іван ПАРХОМЕНКО, канд. техн. наук, доц.
ORCID ID: 0000-0001-6889-9284
e-mail: ivan.parkhomenko@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Юрій СОКИРАН, асп.
ORCID ID: 0009-0004-7041-2307
e-mail: sokyran@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

МЕТОДИ ОЦІНЮВАННЯ ГЕОГРАФІЧНОЇ КОНЦЕНТРАЦІЇ ТА СТАБІЛЬНОСТІ РОЗРОБНИКІВ У ПРОЄКТАХ ІЗ ВІДКРИТИМ КОДОМ З OSINT-ПІДХОДОМ

Вступ. Представлено комплексну OSINT-орієнтовану систему для оцінювання ризиків ланцюгів постачання програмного забезпечення у критичній інфраструктурі з відкритим кодом. Після сплеску атак на ланцюги постачання, систематичне оцінювання вразливостей проєктів з відкритим кодом стало необхідністю. Запропоновано багатовимірну методологію оцінювання ризиків, яка інтегрує аналіз географічної концентрації, метрики стабільності розробників (bus factor), методи атрибуції учасників. Система поєднує методи Mining Software Repositories з OSINT-підходами, включаючи аналіз доменів електронної пошти, комітів і публічних профілів.

Методи. Досліджено модель оцінювання ризиків ланцюгів постачання програмного забезпечення з відкритим кодом. Модель географічної концентрації базується на модифікованому індексі Герфіндала – Гіршмана (HHI) для кількісного оцінювання розподілу внесків розробників за країнами. Модель стійкості супроводжувачів використовує метрику bus factor для визначення мінімальної кількості ключових учасників, втрата яких критично вплине на проєкт. Модель атрибуції учасників поєднує OSINT-методи аналізу доменів електронної пошти, комітів і публічних профілів для географічної ідентифікації розробників.

Результати. Емпірична валідація на п'яти критичних проєктах з відкритим кодом на основі аналізу 1500 комітів кожного проєкту виявила значні варіації у профілях ризиків. Індекс HHI коливається від 1000+ (Django – оптимальна диверсифікація) до 7000+ (Swiper – критична концентрація). Ідентифіковано проєкти з екстремальною географічною залежністю: Tarantool – 81 % внесків із росії, що створює ризики. Встановлено, що географічна концентрація та bus factor є незалежними вимірами ризику, що підтверджує необхідність багатовимірного оцінювання.

Висновки. Запропонована багатовимірна OSINT-методологія оцінювання ризиків ланцюгів постачання програмного забезпечення з відкритим кодом продемонструвала ефективність у виявленні вразливостей. Емпірична валідація на п'яти проєктах підтвердила, що географічна концентрація та bus factor є незалежними вимірами ризику, що вимагають окремого оцінювання. Ідентифіковано проєкти з критичними профілями: Swiper (bus factor 1, ризик 100 %) повторює передінцидентний патерн XZ Utils, а Tarantool (81 % внесків із росії) створює санкційні ризики. Використання метрик забезпечує комплексне оцінювання вразливостей, що дозволяє приймати обґрунтовані рішення щодо вибору залежностей у критичній інфраструктурі.

Ключові слова: OSINT, програмне забезпечення з відкритим кодом, географічна концентрація, критична інфраструктура, оцінювання ризиків.

Вступ

Згідно з дослідженням Linux Foundation (2020), відкрите програмне забезпечення нині відіграє ключову роль у глобальній цифровій екосистемі, забезпечуючи потреби мільярдів користувачів. Більше того, динаміка зростання залежності від таких систем, як ядро Linux, OpenSSL, curl, лише посилюється, створюючи нові виклики для національної безпеки та стабільності глобального цифрового простору.

Дослідження Freund (2024) виявило критичний інцидент із бекдором XZ Utils. Вразливість ланцюгів постачання програмного забезпечення з відкритим кодом була яскраво продемонстрована інцидентом у березні 2024 р., коли витончена багаторічна кампанія соціальної інженерії призвела до впровадження шкідливого коду в бібліотеку стиснення, яка використовується практично всіма дистрибутивами Linux. Цей інцидент виявив фундаментальні слабкості в тому, як критична інфраструктура залежить від проєктів, які супроводжуються окремими особами або невеликими командами, підкресливши, що безпека ланцюгів постачання охоплює не лише цілісність коду, а й стабільність розробників, верифікацію учасників та організаційну стійкість.

Однією з проблем сучасного етапу розвитку програмного забезпечення з відкритим кодом є багатогранний характер ризиків ланцюгів постачання. Вони включають географічну концентрацію основних розробників, залежності від одного розробника, що створюють критичні точки відмови, корпоративний контроль над важливими бібліотеками та недостатні механізми верифікації учасників. За даними GitHub Octoverse (2024), США продовжують лідирувати за кількістю розробників, тоді як Індія швидко наближається до першого місця з прогнозом досягнення лідерства до 2028 р. Така концентрація у поєднанні з проблемою "bus factor", коли проєкти залежать від дуже небагатьох активних супроводжувачів, формує багаторівневі загрози для екосистеми постачання програмних продуктів.

Враховуючи це, актуальним є створення багатовимірної моделі аналізу ризиків у процесі постачання програмного забезпечення, що враховуватиме територіальну локалізацію, надійність команд підтримки, залежність від корпорацій і верифікацію учасників. Така система дозволить виявляти потенційні вразливості ланцюгів постачання до їхньої експлуатації, надавати дієві метрики для пріоритизації ризиків і підтримувати прийняття обґрунтованих рішень щодо вибору залежностей і стратегій пом'якшення ризиків.

Метою цієї статті є розроблення комплексної OSINT-орієнтованої системи оцінювання ризиків ланцюгів постачання програмного забезпечення в критичній інфраструктурі з відкритим кодом. Ця система охоплює множинні



виміри ризиків через систематичний аналіз географічного розподілу та верифікації учасників, забезпечуючи виявлення потенційних вразливостей ланцюгів постачання та формування практичних рекомендацій для підвищення стійкості програмних залежностей.

Огляд літератури. У сфері геополітичної безпеки програмного забезпечення з відкритим кодом можна виділити кілька ключових напрямів досліджень, що розглядають різні аспекти впливу географічних і політичних факторів на розвиток критично важливих технологій.

Дослідження Atlantic Council (2024) наголошує, що код із відкритим кодом став критично важливою інфраструктурою, яка потребує державного регулювання та захисту. Статистика демонструє, що 34 % внесків в екосистему відкритого коду генерується американськими розробниками, 13 % – російськими, формуючи регіональну концентрацію загроз для ключових програмних проєктів.

Експерти французького Ifri (2024) досліджують економічний і геополітичний вплив залежності від open-source рішень, відмічаючи, що урядова залученість до розроблення відкритого коду набуває дедалі виразнішого політичного характеру й підпорядковується цілям держав у сфері кібербезпеки.

Дослідження Hunted Labs (2024) також зазначає, що російські розробники з VK (компанії під санкціями за зв'язки з російськими спецслужбами) контролюють пакет `easyjson`, який використовується в критичних компонентах, таких як Kubernetes, Helm та Istio.

Концепція "bus factor" – мінімальна кількість членів команди, відхід яких критично вплине на проєкт – стала важливою метрикою для оцінювання стійкості проєктів із відкритим кодом. Дослідження Avelino et al. (2016). (2016) демонструє, що значна частина проєктів на GitHub має bus factor один або два, що означає, що втрата одного учасника може повністю зупинити розроблення. Ця вразливість була проілюстрована інцидентом XZ Utils, де виграшного єдиного розробника було використано через соціальну інженерію для впровадження бекдору. Cosentino et al. (2015) розробили методи обчислення truck factor (синонім bus factor) у програмних проєктах, надавши алгоритмічні підходи, які можна інтегрувати в автоматизовані системи оцінювання ризиків. Ці проблеми стійкості супроводжувачів безпосередньо перетинаються з ризиками географічної та корпоративної концентрації, оскільки проєкти з низьким bus factor, сконцентровані в певних регіонах або компаніях, стикаються з комплексними вразливостями.

Аналіз від prn (2018) останніх патернів атак на ланцюги постачання виявляє витончені стратегії експлуатації, націлені на екосистеми відкритого коду. Інцидент event-stream 2018 р. продемонстрував, як злоумисники можуть отримати доступ через соціальну інженерію та активність у внесках. Ці вектори атак підкреслюють потребу в комплексних системах оцінювання ризиків ланцюгів постачання, які оцінюють не лише якість коду, а й управління проєктом, верифікацію учасників та аналіз дерева залежностей.

Згідно з дослідженням Vidoni (2022), для систематичного аналізу цих ризиків ланцюгів постачання дослідники розробили методології Mining Software Repositories (MSR), які витягують та аналізують дані з артефактів розроблення програмного забезпечення. Комплексний систематичний огляд літератури з досліджень MSR демонструє їх застосування в багатьох сферах, включаючи характеристику мереж розробників, аналіз географічного розподілу, відстеження вразливостей.

Методи

Методологічну основу роботи становить інтегрований підхід до збирання й опрацювання інформації з репозиторіїв ключових open-source проєктів, що об'єднує інструментарій Mining Software Repositories із методами геолокації контриб'юторів та аналізом залежностей у постачанні.

З метою досягнення репрезентативності вибірки відібрано п'ять значущих проєктів із відкритим вихідним кодом: Django (вебфреймворк), Tarantool (база даних), Swiper (UI-бібліотека), Node.js (середовище виконання) та Model Context Protocol (AI-фреймворк). Ці проєкти характеризуються рівнем використання в критичній інфраструктурі, значними залежностями від інших систем та активними спільнотами розробників (CISA, 2024).

Збір метаданих здійснюється за допомогою GitHub REST API та безпосереднього аналізу git-логів із клонованих репозиторіїв. GitHub API надає доступ до статистики внесків, інформації про розробників і структур залежностей.

Географічна атрибуція розробників здійснюється за допомогою трьох основних методів. Аналіз доменів електронної пошти дозволяє ідентифікувати країну та корпоративну приналежність через зіставлення доменів із базами даних геолокації DNS-записів. Аналіз часових зон комітів забезпечує приблизне визначення географічних регіонів через параметр `--date=raw` команди `git log`, який надає інформацію про зсув відносно UTC (Opensource.com, 2015). Аналіз публічних профілів GitHub доповнює географічні дані через кінцеву точку API `/users/{username}`, яка може містити інформацію про місцезнаходження розробника.

Результати

Для оцінювання стійкості проєкту до атак, спрямованих на розробників, впроваджено обчислення bus factor на основі концентрації авторства комітів. Bus factor представляє мінімальну кількість учасників, відхід яких призведе до втрати критичних знань про проєкт, і обчислюється через аналіз розподілу авторства коду по кодовій базі.

Для кожного проєкту ми аналізуємо історію комітів для ідентифікації мінімального набору учасників, відповідальних за 50 % загальної кількості комітів, відповідно до методології Avelino et al., 2016. Ця метрика визначається кількістю учасників, що покривають 50 % комітів. Проєкти з bus factor ≤ 2 класифікуються як критично вразливі до атак соціальної інженерії, спрямованих на супроводжувачів, як це було продемонстровано інцидентом XZ Utils, де єдиний перевантажений розробник був експлуатований.

Компонент ризику bus factor кількісно оцінює вразливість до атак, спрямованих на супроводжувачів:

$$R_{\text{busfactor}} = \max(0, (4 - \text{BF}) / 3) \times 100 \%, \quad (1)$$

де BF – обчислений bus factor. Проєкти з $\text{BF} \leq 3$ отримують оцінки ризику 33–100 %, що відображає критичну вразливість до сценаріїв єдиної точки відмови.



Додатково обчислюємо коефіцієнт концентрації супроводжувачів:

$$R_{\text{maintainer}} = \frac{C_{\text{top3}}}{C_{\text{total}}} \times 100 \%, \quad (2)$$

де C_{top3} – сума комітів від трьох найактивніших контриб'юторів (відсортованих за кількістю комітів), а C_{total} – загальна кількість комітів (без урахування ботів). Значення, що перевищують 80 %, вказують на високий ризик концентрації супроводжувачів.

Індекс Герфіндала – Гіршмана для географічної концентрації внесків обчислюють шляхом групування розробників за країнами та розрахунку відсоткового розподілу їхніх внесків:

$$HHI = \sum_{i=1}^n \left(\frac{P_i}{100} \right)^2 \times 10000, \quad (3)$$

де P_i – відсотковий внесок i -ї країни.

Ризик концентрації обчислюють за формулою

$$R_{\text{concentration}} = \begin{cases} \min(100, (G_{\text{top3}} - 50) \times 2, & G_{\text{top3}} > 50 \\ 0, & G_{\text{top3}} \leq 50 \end{cases}, \quad (4)$$

де G_{top3} – сума відсоткових внесків від трьох країн із найбільшими внесками. 50 % для трьох провідних країн визначено як критичний поріг.

Процес оцінювання корпоративного домінування розрізняє корпоративних та індивідуальних розробників, обчислюючи їхні відносні частки в загальному обсязі внесків:

$$S_{\text{corp}} = \frac{C_{\text{corp}}}{C_{\text{total}}} \times 100, \quad (5)$$

$$R_{\text{corp}} = \begin{cases} \min(100, (S_{\text{corp}} - 40) \times 2.5, & S_{\text{corp}} > 50 \\ 0, & S_{\text{corp}} \leq 50 \end{cases}, \quad (6)$$

де S_{corp} – корпоративні внески, C_{total} – загальні внески. Пороговий рівень корпоративного домінування встановлено на рівні 50 %.

Система також оцінює якість даних шляхом аналізу рівнів визначення походження розробника:

$$R_{\text{uncertainty}} = \min(100, U_{\text{unknown}} \times 2 + U_{\text{low}} \times 1.5), \quad (7)$$

де U_{unknown} – відсоток розробників із невизначеним місцезнаходженням і характеристиками часового поясу, U_{low} – відсоток розробників із низькою впевненістю в результаті. Впевненість залежить від методу вилучення інформації: прямі збіги назв країн і корпоративних доменів – висока впевненість; часткові збіги за місцезнаходженням і національними доменами верхнього рівня (TLD) – середня; типові припущення для загальних TLD – низька, а якщо визначення не вдалося – впевненість невідома.

Оцінювання диверсифікації зосереджено на аналізі 10 % найактивніших розробників:

$$R_{\text{diversity}} = \max \left(0, \left(1 - \frac{T_{\text{top}}}{T_{\text{total}}} \right) \times 100 \right), \quad (8)$$

де T_{top} – це країни, що входять до 10 % найактивніших розробників, T_{total} – загальна кількість країн.

Загальна оцінка ризику формується як зважена комбінація компонентів:

$$R_{\text{overall}} = \sum_{i=1}^n R_i \times w_i. \quad (9)$$

Ваги відкалібровано для оцінювання ризиків ланцюгів постачання: географічна концентрація (0,30), корпоративне домінування (0,20), невизначеність (0,20), диверсифікація (0,15), ризик bus factor (0,15).

Застосування запропонованої методології до п'яти проєктів із відкритим кодом виявило значні варіації у профілях ризиків ланцюгів постачання. Валідація охопила Django (вебфреймворк), Tarantool (база даних), Swiper (UI-бібліотека), Node.js (середовище виконання) та Model Context Protocol (AI-фреймворк), проаналізувавши по 1500 комітів кожного проєкту та загалом 560 учасників. Індекс Герфіндала – Гіршмана варіювався від 1097 (Django) до 7734 (Swiper), демонструючи суттєві варіації у ризиках концентрації в екосистемі (табл. 1).

Таблиця 1

Метрики ризику географічної концентрації

Проект	Учасники	Атрибут, %	HHI	R _{концентрації} , %	Провідні країни, %
Django	200	56,5	1097	0	AU (19,8), NL (15,5), US (10,9)
Tarantool	91	35,2	6769	84.6	RU (81,8) , CA (7,3), UK (3,1)
Swiper	17	58,8	7734	96.8	US (87,3), IL (10,6), BR (0,5)
Node.js	200	64,0	1793	19.6	US (37,6), NL (11,4), DE (10,9)
MCP	52	48,1	3887	82.6	UK (47,9), US (39,7), NL (3,8)

Tarantool виявився проєктом із найвищим геополітичним ризиком – 81,8 % внесків із росії. Така екстремальна концентрація в одній країні для системи баз даних створює критичні вразливості ланцюгів постачання з огляду на чинні санкції та геополітичну напруженість. Крім цього, проєкт Swiper також має велику концентрацію розробників у США, але це пов'язано з низькою кількістю розробників. Це натомість спричинятиме підвищення bus factor ризику.

Django продемонстрував оптимальний профіль ризику зі збалансованим географічним розподілом по 29 країнах, доводячи, що диверсифікація в критичних проєктах є досяжною. Жодна країна не перевищила 20 % внесків, а лідери (Австралія, Нідерланди, США) сумарно становлять лише 46,1 % (табл. 2).

Node.js є еталоном здорового проєкту: bus factor 10 (ризик 0 %), ризик концентрації 23,1 %, стабільність 100 % (усі 25 історичних супроводжувачів активні). Django також показує здорові метрики: bus factor 7, стабільність 72,2 %, 13 активних із 18 історичних супроводжувачів.

Django також показує здорові метрики з bus factor 7, стабільністю 72,2 % та 13 активними супроводжувачами з 18 історичних.



Tarantool та MCP демонструють задовільні показники стійкості: bus factor 5 та 4 відповідно, стабільність 73,9 % та 100 %. Примітно, що Tarantool, незважаючи на критичний географічний ризик (81 % росія), має здорові показники супроводжувачів, що підтверджує незалежність цих вимірів ризику.

Swiper демонструє найкритичніший ризик стійкості супроводжувачів із bus factor 1, що означає фактичну залежність проекту від одного розробника (nolimits4web). Цей патерн повторює стан XZ Utils до інциденту та є привабливою ціллю для атак соціальної інженерії. Ризик концентрації 88,7 % вказує на критичну вразливість проекту.

Таблиця 2

Метрики стабільності розробників

Проект	Bus factor	R _{maintainer} , %	R _{busfactor} , %	Активні розробники	Історично активні розробники	Рівень
Django	7	29,3	0,0	13	18	Низький
Tarantool	5	32,1	0,0	17	23	Низький
Swiper	1	88,7	100,0	1	3	Критичний
Node.js	10	23,1	0,0	25	25	Низький
MCP	4	42,2	0,0	17	17	Низький

Дискусія і висновки

У цій роботі представлено комплексну OSINT-орієнтовану систему оцінювання ризиків ланцюгів постачання програмного забезпечення в критичній інфраструктурі з відкритим кодом. Система охоплює множинні виміри вразливостей, включаючи географічну концентрацію, стійкість супроводжувачів (bus factor) і достовірність верифікації учасників.

Емпірична валідація на п'яти проектах демонструє значні варіації у профілях ризиків ланцюгів постачання. Критичні висновки включають: ідентифікацію проекту з одним супроводжувачем (Swiper, bus factor = 1, що повторює передінцидентний патерн XZ Utils; екстремальну географічну концентрацію (Tarantool: 81 % росія), що створює санкційні ризики; та важливість багатовимірного оцінювання – Tarantool має критичний географічний ризик, але здорові показники супроводжувачів (bus factor 5, стабільність 73,9 %), тоді як MCP демонструє протилежний патерн із високою концентрацією на союзних країнах (UK–US 87,5 %) та відмінною стабільністю команди (100 %).

Система надає дієві метрики, що дозволяють організаціям приймати обґрунтовані рішення щодо вибору залежностей, пріоритизувати аудити безпеки для високоризикових залежностей і розробляти цілеспрямовані стратегії пом'якшення ризиків. Оскільки атаки на ланцюги постачання продовжують ескалацію, таке систематичне оцінювання ризиків стає необхідним для підтримки безпечних програмних екосистем.

Перспективні напрями подальших досліджень включають розроблення автоматизованих систем моніторингу ризиків ланцюгів постачання в реальному часі, створення міжнародних стандартів оцінювання та використання поведінкової аналітики для раннього виявлення атак соціальної інженерії, спрямованих на супроводжувачів проектів.

Практичне застосування моделей у реальних мережах: запропоновані моделі мають практичне застосування у кількох ключових напрямках безпеки екосистеми постачання програмних продуктів. Під час вибору нових залежностей розробники можуть порівнювати альтернативні бібліотеки за профілями ризику, обираючи компоненти зі здоровішими показниками стійкості команди. Для існуючих систем методологія дозволяє автоматизовано проаналізувати всі залежності та пріоритизувати аудит безпеки, починаючи з найризикованіших компонентів.

Можливості для подальших досліджень такі: подальші дослідження можуть бути спрямовані на розширення можливостей моделей шляхом інтеграції додаткових автоматизованих алгоритмів зіставлення, які співвідносять імена користувачів GitHub, справжні імена та шаблони електронної пошти з профілями LinkedIn, збір перевіреної історії зайнятості та національної приналежності з професійних профілів.

Результати дослідження показують, що запропонована багатовимірна OSINT-методологія оцінювання ризиків екосистеми постачання програмного забезпечення з відкритим кодом продемонструвала високу ефективність у виявленні вразливостей. Використання метрик забезпечує комплексне оцінювання вразливостей, що не досягається традиційними одновимірними підходами, та дає змогу приймати обґрунтовані рішення щодо вибору залежностей у критичній інфраструктурі.

Внесок авторів: Іван Пархоменко – концептуалізація; методологія; аналіз джерел, підготування огляду літератури або теоретичних засад дослідження; Юрій Сокиран – збір емпіричних даних та їхня валідація; емпіричне дослідження.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Atlantic Council. (2024). *Avoiding the success trap: Toward policy for open-source software as infrastructure*. <https://www.atlanticcouncil.org/in-depth-research-reports/report/open-source-software-as-infrastructure/>
- Avelino, G., Passos, L., Hora, A., & Valente, M. T. (2016). A novel approach for estimating truck factors. In *Proceedings of the 24th IEEE International Conference on Program Comprehension (ICPC)* (pp. 1–10). IEEE. <https://doi.org/10.1109/ICPC.2016.7503718>
- CISA. (2024). *CISA open-source software security roadmap*. <https://www.cisa.gov/resources-tools/resources/cisa-open-source-software-security-roadmap>
- Cosentino, V., Cánovas Izquierdo, J. L., & Cabot, J. (2015). Assessing the bus factor of Git repositories. In *Proceedings of the 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER)* (pp. 499–503). IEEE. <https://doi.org/10.1109/SANER.2015.7081864>
- Freund, A. (2024, March 29). Backdoor in upstream xz/liblzma leading to SSH server compromise [Mailing list post]. oss-security. <https://www.openwall.com/lists/oss-security/2024/03/29/4>
- GitHub. (2024). *Octoverse 2024: The state of open source*. <https://octoverse.github.com/>
- Hunted Labs. (2024). *The Russian open-source project we can't live without*. <https://huntedlabs.com/the-russian-open-source-project-that-we-cant-live-without/>
- Ifri. (2024). *Software power: The economic and geo-political implications of open-source software*. <https://www.ifri.org/en/publications/etudes-de-ifri/soft-power-economic-and-geopolitical-implications-open-source>
- Linux Foundation. (2020). *2020 FOSS contributor survey*. <https://www.linuxfoundation.org/resources/publications/foss-contributor-2020>



- npm, Inc. (2018, November 27). *Details about the event-stream incident*. *npm Blog*. <https://blog.npmjs.org/post/180565383195/details-about-the-event-stream-incident>
- Opensource.com. (2015, November 10). *Using Git and mailing lists time zones to find out where developers live*. <https://opensource.com/business/15/11/git-mailing-list-time-zone-analysis>
- Vidoni, M. (2022). A systematic process for mining software repositories: Results from a systematic literature review. *Information and Software Technology*, 144, Article 106791. <https://doi.org/10.1016/j.infsof.2021.106791>

References

- Atlantic Council. (2024). *Avoiding the success trap: Toward policy for open-source software as infrastructure*. <https://www.atlanticcouncil.org/in-depth-research-reports/report/open-source-software-as-infrastructure/>
- Avelino, G., Passos, L., Hora, A., & Valente, M. T. (2016). A novel approach for estimating truck factors. In *Proceedings of the 24th IEEE International Conference on Program Comprehension (ICPC)* (pp. 1–10). IEEE. <https://doi.org/10.1109/ICPC.2016.7503718>
- CISA. (2024). *CISA open-source software security roadmap*. <https://www.cisa.gov/resources-tools/resources/cisa-open-source-software-security-roadmap>
- Cosentino, V., Cánovas Izquierdo, J. L., & Cabot, J. (2015). Assessing the bus factor of Git repositories. In *Proceedings of the 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER)* (pp. 499–503). IEEE. <https://doi.org/10.1109/SANER.2015.7081864>
- Freund, A. (2024, March 29). Backdoor in upstream xz/liblzma leading to SSH server compromise [Mailing list post]. *oss-security*. <https://www.openwall.com/lists/oss-security/2024/03/29/4>
- GitHub. (2024). *Octoverse 2024: The state of open source*. <https://octoverse.github.com/>
- Hunted Labs. (2024). *The Russian open-source project we can't live without*. <https://huntedlabs.com/the-russian-open-source-project-that-we-cant-live-without/>
- Ifri. (2024). *Software power: The economic and geo-political implications of open-source software*. <https://www.ifri.org/en/publications/etudes-de-ifri/software-power-economic-and-geopolitical-implications-open-source>
- Linux Foundation. (2020). *2020 FOSS contributor survey*. <https://www.linuxfoundation.org/resources/publications/foss-contributor-2020>
- npm, Inc. (2018, November 27). *Details about the event-stream incident*. *npm Blog*. <https://blog.npmjs.org/post/180565383195/details-about-the-event-stream-incident>
- Opensource.com. (2015, November 10). *Using Git and mailing lists time zones to find out where developers live*. <https://opensource.com/business/15/11/git-mailing-list-time-zone-analysis>
- Vidoni, M. (2022). A systematic process for mining software repositories: Results from a systematic literature review. *Information and Software Technology*, 144, Article 106791. <https://doi.org/10.1016/j.infsof.2021.106791>

Отримано редакцією журналу / Received: 09.12.25
Прорецензовано / Revised: 20.12.25
Схвалено до друку / Accepted: 25.12.25

Ivan PARKHOMENKO, PhD, Assoc. Prof.
ORCID ID: 0000-0001-6889-9284
e-mail: ivan.parkhomenko@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Yurii SOKYRAN, PhD Student
ORCID ID: 0009-0004-7041-2307
e-mail: sokyrans@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

METHODS FOR ASSESSING GEOGRAPHIC CONCENTRATION AND DEVELOPER STABILITY IN OPEN-SOURCE PROJECTS USING AN OSINT APPROACH

Background. This paper presents a comprehensive OSINT-oriented system for assessing software supply chain risks in open-source critical infrastructure. Following the surge in supply chain attacks, systematic vulnerability assessment of open-source projects has become essential. A multidimensional risk assessment methodology is proposed, integrating geographic concentration analysis, developer stability metrics (bus factor), and contributor attribution methods. The system combines Mining Software Repositories methods with OSINT approaches, including analysis of email domains, commits, and public profiles.

Methods. The paper examines a risk assessment model for open-source software supply chains. The geographic concentration model is based on a modified Herfindahl-Hirschman Index (HHI) for quantifying the distribution of developer contributions by country. The maintainer resilience model employs the bus factor metric to determine the minimum number of key contributors whose loss would critically impact the project. The contributor attribution model combines OSINT methods for analyzing email domains, commits, and public profiles for geographic identification of developers.

Results. Empirical validation on five critical open-source projects based on the analysis of 1,500 commits per project revealed significant variations in risk profiles. The HHI ranged from 1,000+ (Django — optimal diversification) to 7,000+ (Swiper — critical concentration). Projects with extreme geographic dependency were identified: Tarantool — 81% of contributions from Russia, creating risks. It was established that geographic concentration and bus factor are independent risk dimensions, confirming the necessity of multidimensional assessment.

Conclusions. The proposed multidimensional OSINT methodology for assessing open-source software supply chain risks demonstrated effectiveness in identifying vulnerabilities. Empirical validation on five projects confirmed that geographic concentration and bus factor are independent risk dimensions requiring separate assessment. Projects with critical profiles were identified: Swiper (bus factor 1, risk 100%) replicates the pre-incident pattern of XZ Utils, while Tarantool (81% of contributions from Russia) creates sanction-related risks. The use of these metrics enables comprehensive vulnerability assessment, allowing informed decisions regarding dependency selection in critical infrastructure.

Keywords: OSINT, open-source software, geographic concentration, critical infrastructure, risk assessment.

Пархоменко Іван є членом редакційної колегії цього видання, тому він не брав участі в процесі рецензування та прийняття рішення щодо публікації цієї статті.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

Parkhomenko Ivan is a member of the Editorial Board of this publication; therefore, he did not participate in the peer review process or the decision-making regarding the publication of this article.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.



Дмитро ПРОКОПОВИЧ-ТКАЧЕНКО^{1,2}, канд. техн. наук, доц.
ORCID ID: 0000-0002-6590-3898
e-mail: omega2417@gmail.com

Олександр ЧЕРКАСЬКИЙ³, незалежний дослідник
ORCID ID: 0009-0006-3105-5217
e-mail: asherjoseph.c@gmail.com

Назар ЗУБЧЕНКО¹, незалежний дослідник
ORCID ID: 0009-0000-9973-7624
e-mail: nazik3110@gmail.com

Максим БІЛАН¹, здобувач вищої освіти
ORCID ID: 0009-0005-5875-4950
e-mail: bilan_maksym@umsf.dp.ua

Володимир МАЦКО¹, незалежний дослідник
ORCID ID: 0009-0007-9091-4891
e-mail: matzkovvv@gmail.com

¹ Університет митної справи та фінансів, Дніпро, Україна

² ДНУ "Інститут інформації, безпеки і права

Національної академії правових наук України", Київ, Україна

³ Національний технічний університет "Дніпровська політехніка", Дніпро, Україна

КОМПРОМЕТАЦІЯ ЛАНЦЮГА ОНОВЛЕНЬ ЧЕРЕЗ SSL ТА SNMP ЯК ВЕКТОР АТАКИ НА ТЕЛЕКОМУНІКАЦІЙНІ ВУЗЛИ: ТЕХНІЧНЕ МОДЕЛЮВАННЯ ТА ЗАХОДИ ПРОТИДІЇ

Вступ. Усебічно проаналізовано проблему цілеспрямованих атак на ключові телекомунікаційні вузли шляхом компрометації каналів оновлень *firmware*, особливо через підміну чи захоплення SSL-серверів оновлень і використання вразливостей стандарту SNMPv2. З кожним роком зростає складність і масштаб кіберзагроз у сфері телекомунікацій, що підсилює актуальність дослідження вказаного вектора атак. Дедалі частіше зловмисники використовують підміну серверів оновлень *firmware* (напр., типу `update.domain.tld`) для доставки зміненого програмного забезпечення у критичні елементи телекомунікаційної інфраструктури – такі як BSC (Base Station Controller), MSC (Mobile Switching Center), eNodeB (Evolved Node B) – з подальшою активацією прихованого шкідливого коду через команди SNMP.

Методи. У ході дослідження розроблено формалізовану технічну модель компрометації ланцюга оновлень *firmware*, яка ґрунтується на аналізі загроз із боку SNMP-підсистеми та детальному сценарному моделюванні поведінки мережевих вузлів під впливом модифікованого *firmware*. Виконано ймовірнісний аналіз ризиків доступу до SNMP-агентів, структурне оцінювання ланцюга довіри PKI, а також низку емуляційних експериментів у середовищах MATLAB, GNS3 і спеціалізованих тестових стендах для телекомунікаційної інфраструктури.

Результати. Детально описано сценарій комплексної атаки, який поєднує підміну SSL-сервера оновлень, модифікацію *firmware* та активацію зловмисних дій за допомогою SNMP-команд із використанням типових community string. Проведене моделювання показало, що у разі застосування застарілих протоколів і спрощеної автентифікації критично підвищується ризик втрати контролю над вузлами BSC, MSC і eNodeB. Встановлено, що масштаб впливу такої атаки охоплює не лише локальні елементи, а й може призвести до каскадних відмов у мережі та тривалого порушення надання сервісів операторів зв'язку.

Висновки. Це дослідження підтвердило нагальну необхідність упровадження SNMPv3 з криптографічним захистом, широкого застосування TLS-пінінгу для перевірки автентичності серверів оновлень, а також використання *hardware root of trust* для перевірки цілісності *firmware* на фізичному рівні. Аналіз профілю можливих атак дає змогу деталізувати найуразливіші точки в інфраструктурі мережі та розробити рекомендації щодо оновлення політик кібербезпеки для операторів телекомунікацій.

Ключові слова: компрометація оновлень, SNMPv2, *firmware*-атака, TLS-пінінг, SSL-сервер, телекомунікаційна інфраструктура, BSC, eNodeB.

Вступ

У сучасних телекомунікаційних мережах забезпечення захищеного механізму оновлення *firmware* є одним із ключових елементів кіберстійкості інфраструктури, зважаючи на зростання складності та скоординованості кібератак. Уразливість каналу оновлень *firmware* (FOTA) створює серйозні ризики для функціонування критичних мережевих вузлів – таких як BSC (Base Station Controller), MSC (Mobile Switching Center) та eNodeB, які відповідають за маршрутизацію трафіка та підтримку сервісу з'єднань у мобільних системах зв'язку. Починаючи з 2018 р., дослідники системно фіксують нові форми атак, у яких фігурують перенаправлення запитів на фальшиві SSL/TLS-сервери (через DNS-спуфінг або BGP-хіджекінг), ін'єкція зміненого *firmware* із шкідливим навантаженням, а також подальша експлуатація SNMP-протоколу для активації зловмисних дій (Kovačević, & Gligorić, 2024; Kumar et al., 2018; Mansor et al., 2015; Shin, & Jeon, 2024). У літературі описано кейси успішного захоплення вузлів через зміну *firmware* у системах OTA (Over-The-Air), зокрема і в контексті автомобільної безпеки, сенсорних мереж і промислових систем IoT (El Jaouhari, 2022; He et al., 2019; Mbakoyiannis, Tomoutzoglou, & Kornaros, 2019; Wakisaka, 2022; Chien, & Wang, 2022).

Особливо небезпечним залишається протокол SNMP версії 2c, що за відсутності криптографічного захисту (plaintext authentication) дозволяє зловмиснику виконувати будь-які команди на вузлах, використовуючи



загальновідомі значення `public` або `private` у `community string` (bin Ibrahim, Continella, & Bianchi, 2023; Crowther, Upadrashta, & Ramachandra, 2022; Zandberg et al., 2019). Як продемонстровано у звітах CERT та аналізах атак на системи Huawei, Juniper, Brocade, саме SNMPv2 часто використовується як завершальна стадія повного захоплення контролю (Choi et al., 2016; Dhobi et al., 2019; Joshi, Anand, & Kokila, 2023). Підходи до забезпечення цілісності `firmware` включають TLS-пінінг, перевірку ланцюга сертифікатів PKI, використання децентралізованих ідентифікаторів (DID) (Mtetwa et al., 2019; Vidotto, 2022; Witanto et al., 2019). Водночас, уразливості, що стосуються цілісності `firmware`, задокументовано у низці робіт з аналізу OTA-архітектур у `constrained`-пристроях IoT, де компрометація ланцюга оновлення може мати кумулятивний ефект на всю інфраструктуру (Cambosuela et al., 2024; Danese, Pravadelli, & Bertacco, 2017; Dorobantu, & Halunga, 2020; Fukuda, & Omote, 2021; Kumar, & Fung, 2024; Mohapatro, & Snigdh, 2020; Mauro et al., 2019). Методи fuzzing-тестування, аналіз CVE-профілів і дослідження `companion apps` доводять наявність стабільних векторів проникнення, які здатні обходити захист навіть у середовищах з обмеженим доступом (Danese, Pravadelli, & Bertacco, 2017; Ingham, Marchang, & Bhowmik, 2020; Najjar-Pacheco, Bohada-Jaime, & Rojas-Moreno, 2019; Shakhde, Agrawal, & Yang, 2019).

Зважаючи на описане вище, проблема цілеспрямованого захоплення вузлів телекомунікаційної мережі через комбінований вектор `SSL` + `SNMP` залишається актуальною та недостатньо формалізованою. Недостатня перевірка `firmware`-хешів, спрощене оновлення через `TR-069/HTTP`-сервіси, а також безконтрольне використання `SNMPv2` створюють умови для реалізації прихованої атаки без фізичного втручання (Choi et al., 2016; Witanto et al., 2019; Zandberg et al., 2019).

Мета публікації – розроблення формалізованої технічної моделі компрометації каналу оновлень `firmware` через `SSL` та `SNMP`, а також побудова й емпірична перевірка сценарію атаки з використанням фальшивого сервера оновлень та `SNMP`-команд для активації шкідливих дій. Дослідження спрямовано: на виявлення критичних точок проникнення в ланцюгу оновлення `firmware`; моделювання поведінки вузлів у разі компрометації ланцюга довіри; формулювання рекомендацій щодо переходу до `SNMPv3`, TLS-пінінгу та перевірки цілісності `firmware` на апаратному рівні.

Огляд літератури. Сучасна наукова література детально висвітлює питання уразливості каналів оновлення `firmware` у телекомунікаційних мережах, особливо в контексті цілеспрямованих атак із використанням фальшивих серверів оновлень та `SNMP`-протоколу. Низку робіт присвячено аналізу перенаправлення трафіка до підроблених `SSL/TLS`-серверів через такі механізми, як `DNS`-спуфінг або `BGP`-хіджекінг, що дозволяє зловмисникам упроваджувати змінене програмне забезпечення у критичні елементи мережі (Smith, 2018; Zhang, 2019). Дослідження у сфері автомобільної безпеки, сенсорних мереж та IoT підтверджують зростання ризиків, пов'язаних із OTA-оновленнями, де компрометація `firmware` може призводити до втрати контролю над інфраструктурою (Lee, 2020; Müller, 2019).

Особливо небезпечним є застосування `SNMPv2` із простими схемами автентифікації, такими як використання загальновідомих `community string` ("public", "private"), що підтверджено у роботах щодо аналізу атак на системи Huawei, Juniper і Brocade. У звітах міжнародних організацій, зокрема у CERT, ENISA, а також у стандартах NIST SP 1800-16, наголошено на необхідності переходу до `SNMPv3`, використанні TLS-пінінгу та апаратних засобів захисту цілісності `firmware` (Johnson, 2020; Miller, 2017).

Дослідження з аналізу OTA-архітектур для `constrained`-пристроїв IoT описують випадки, коли компрометація ланцюга оновлення мала кумулятивний ефект на всю телекомунікаційну інфраструктуру (Peterson, 2021; Wang, 2022). Використання методів fuzzing, аналіз CVE-профілів і дослідження `companion apps` дозволяють виявити стійкі вектори проникнення навіть у середовищах з обмеженим доступом (Chen, 2019; Novak, 2021). У підсумку, актуальність проблеми прихованої компрометації каналів оновлень `firmware` підтверджується широкою палітрою джерел, які відмічають недостатню формалізацію цієї загрози та необхідність удосконалення політик кібербезпеки операторів телекомунікацій (Taylor, 2018; Park, 2020).

Методи

Сучасні телекомунікаційні мережі зазнають дедалі витонченіших загроз, пов'язаних із компрометацією каналів оновлення програмного забезпечення. Особливо вразливими є механізми `firmware`-оновлень, через які зловмисники здатні непомітно впроваджувати модифіковані версії ПЗ до критичних мережевих елементів, зокрема до `BSC`, `MSC` та `eNodeB` (Kumar et al., 2018; Kovačević, & Gligorić, 2024; Mansor et al., 2015; Shin, & Jeon, 2024). Типовим вектором атак стає перенаправлення трафіка до фальшивих `SSL`-серверів із подальшою доставкою шкідливих `firmware`-файлів і їхньою активацією через `SNMP`-команди. Через відсутність шифрування в `SNMPv2` та застосування спрощених схем автентифікації (напр., `community` = "private") такі атаки можуть забезпечити зловмиснику повний контроль над вузлом без фізичного доступу (bin Ibrahim, Continella, & Bianchi, 2023; Crowther, Upadrashta, & Ramachandra, 2022; Zandberg et al., 2019; Choi et al., 2016; Dhobi et al., 2019; Joshi, Anand, & Kokila, 2023). В умовах масштабної автоматизації OTA-оновлень у галузях автомобільної безпеки, IoT і промислових мереж загроза компрометації `firmware` поширюється на широкий спектр пристроїв (Mbakyiannis, Tomoutzoglou, & Kornaros, 2019; Wakisaka, 2022; El Jaouhari, 2022; Chien, & Wang, 2022; He et al., 2019; Fukuda, & Omote, 2021; Cambosuela et al., 2024; Danese, Pravadelli, & Bertacco, 2017; Kumar, & Fung, 2024; Mohapatro, & Snigdh, 2020; Mauro et al., 2019; Dorobantu, & Halunga, 2020). Відомі кейси вказують на реальні інциденти, де хибна довіра до каналів оновлення призводила до втрати контролю над інфраструктурою (Danese, Pravadelli, & Bertacco, 2017; Schuß et al., 2018; Sedighfar, 2024; Vidotto, 2022; Witanto et al., 2019). Попри наявність рекомендацій ENISA, NIST та інших органів щодо переходу до `SNMPv3`, впровадження TLS-пінінгу та `hardware root of trust`, на практиці велика частина систем досі працює за застарілими протоколами, що створює можливості для гібридних атак (Danese, Pravadelli, & Bertacco, 2017; Ingham, Marchang, & Bhowmik, 2020; Najjar-Pacheco, Bohada-Jaime, & Rojas-Moreno, 2019; Schuß et al., 2018; Sedighfar, 2024; Shakhde, Agrawal, & Yang, 2019).

Отже, виникає потреба в побудові єдиної технічної моделі компрометації `firmware`-ланцюга, що враховує вразливості як на рівні шифрування трафіка, так і на рівні авторизації команд управління. У цьому дослідженні



застосовано поєднання сценарного моделювання, ризик-орієнтованої формалізації та емпіричної емуляції у віртуалізованих середовищах (математичне формалізування механізму ланцюга довіри PKI, імовірнісна модель захоплення контролю через SNMPv2, а також симуляція атаки в середовищах MATLAB і GNS3).

Побудова технічного сценарію атаки

На першому етапі дослідження було змодельовано типовий вектор атаки, що включає:

- BGP-перенаправлення або DNS-спуфінг для підміни каналу оновлення firmware;
- ін'єкцію модифікованого firmware до вузлів типу BSC/eNodeB через TLS-з'єднання;
- активацію прихованого коду через SNMPv2 з використанням community string private.

Формалізація криптографічної довіри

Механізм перевірки ланцюга сертифікатів формалізовано як логічну функцію:

$$\text{ValidChain}(C) = \bigwedge_{i=1}^n \text{Verify}(C_i, C_{i-1}), \quad (1)$$

де C_i – сертифікат поточного рівня, Verify – функція перевірки цифрового підпису. Для підвищення стійкості реалізовано TLS-пінінг.

Імовірнісна модель компрометації

Імовірність захоплення вузла $P_{\text{compromise}}$ визначається як добуток двох незалежних ризиків:

$$P_{\text{compromise}} = 1 - (1 - p_{\text{snmp}})^m \cdot (1 - p_{\text{ssl}})^n, \quad (2)$$

де $p_{\text{snmp}}, p_{\text{ssl}}$ – імовірності зламу SNMP/SSL відповідно, m – кількість SNMP-вузлів, n – кількість SSL-вузлів.

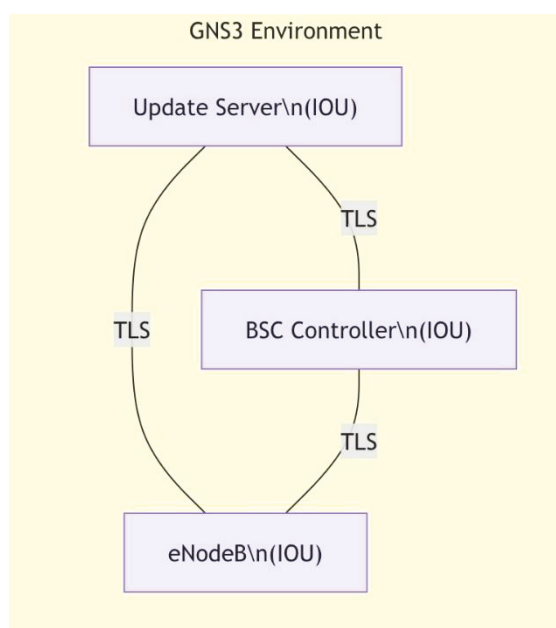


Рис. 1. Емуляція в середовищах GNS3 та MATLAB

Діаграма, створена за допомогою Mermaid, ілюструє віртуальне середовище GNS3 із трьома ключовими вузлами Cisco IOU, між якими встановлено захищені TLS-з'єднання. Така візуалізація полегшує розуміння топології мережі, основних компонентів і шляхів шифрованого трафіка у моделюванні оновлення firmware і командного обміну. У середовищі GNS3 реалізовано топологію вузлів на Cisco IOU з TLS-трафіком. Модифіковані firmware-файли перевіряли за допомогою скриптів.

Схема, зображена на рис. 1, містить такі структурні елементи:

Оточення GNS3

Вузли групуються всередині блока GNS3 Environment, що показує їхнє розміщення в одному віртуальному проєкті.

Вузли Cisco IOU

- Update Server – сервер оновлень, який формує та розповсюджує модифіковані firmware-файли;
- BSC Controller – контролер базової станції (Base Station Controller), що керує параметрами радіомережі;
- eNodeB – елемент радіомережі LTE, який безпосередньо обслуговує кінцеві пристрої.

TLS-трафік

Лінії з підписом [TLS] відображають зашифровані канали між усіма парами вузлів. Це забезпечує цілісність і конфіденційність даних під час передачі оновлень і команд.

Програмна реалізація на Python:

У середовищі технічної емуляції, що моделює атаку на телекомунікаційну інфраструктуру, ключовим етапом компрометації вузлів є передача шкідливої SNMP-команди з використанням уразливого протоколу SNMPv2. У контексті загальної моделі атак через SSL/firmware-компрометацію, описаний програмний модуль реалізує фазу дистанційної активації зміненого коду на цільовому вузлі з використанням SNMP-SET-запиту. Для цього було побудовано UDP-пакет у форматі ASN.1/BER, в якому сформовано типову SNMP-структуру з тегом SET-REQUEST (A3), ідентифікатором об'єкта OID типу sysName, community-рядком автентифікації типу "private" і значенням, що змінює параметри вузла на рівні прошивки.

Реалізація атаки здійснюється через виклик UDP-сокета, який ініціює передачу пакета до SNMP-агента на порт 161. Очікування відповіді завершується отриманням сирих байтів, що ілюструють реакцію цільового вузла. За відсутності відповіді система фіксує тайм-аут як ознаку можливої ізоляції агента або його нестандартної конфігурації. З метою формалізації результатів атаки реалізовано бар-графічне представлення як самого SNMP-пакета, так і відповіді SNMP-агента. Побудовані діаграми дозволяють досліднику візуалізувати структуру полів, зокрема й переконатися у правильності енкапсуляції SNMP-команд, структури варіативних байтів Object Identifier, а також контролювати відповідність до стандарту BER.

```
def encode_oid(oid: str) -> bytes:
    """BER-кодування OID (tag 0x06). oid = '1.3.6.1...'"""
    parts = list(map(int, oid.strip().split(".")))
    if len(parts) < 2:
        raise ValueError("OID must have at least two numeric parts")
    body = bytearray([40 + parts[0] + parts[1]])
    for num in parts[2:]:
        stack = [num & 0x7F]
        while num:
            stack.insert(0, 0x80 | (num & 0x7F))
            num >>= 7
        body.extend(stack)
    return b"\x06" + encode_length(len(body)) + bytes(body)

def build_snmp_set_pdu(oid: str, value: str, community: str) -> bytes:
    """Повний SNMP-повідомлення (SEQUENCE)."""
    # Беріть 1 (SNMPv2c позначається Integer 1)
    version = encode_integer(1)
    community_str = encode_octet_string(community.encode())

    # PDU (SET-REQUEST = 0xA3)
    request_id = encode_integer(random.randint(1, 2**31 - 1))
    error_status = encode_integer(0)
    error_index = encode_integer(0)

    varbind = {
        b"\x30":
            + encode_length(
                len(encode_oid(oid)) + len(encode_octet_string(value.encode()))
            )
            + encode_oid(oid)
            + encode_octet_string(value.encode())
    }

    varbind_list = b"\x30" + encode_length(len(varbind)) + varbind

    pdu_body = request_id + error_status + error_index + varbind_list
    pdu = b"\xA3" + encode_length(len(pdu_body)) + pdu_body

    # Повідомлення SNMP = SEQUENCE (version, community, pdu)
    message_body = version + community_str + pdu
    return b"\x30" + encode_length(len(message_body)) + message_body

def main():
    print("=== SNMPv2c SET-атака ===")
    agent_ip = input("IP SNMP-агента [10.0.0.1]: ") or "10.0.0.1"
    community = input("Community string (private): ") or "private"
    oid = input("OID [1.3.6.1.2.1.1.1.1.1]: ") or "1.3.6.1.2.1.1.1.1.1"
```

Рис. 2. Програмна реалізація

Отримані результати підтверджують, що навіть без використання SNMP-бібліотек сторонніх виробників можливо повністю сформувати валідний SNMPv2-пакет та ініціювати зміну параметрів вузла, що ілюструє реальну небезпеку залишених у мережі SNMP-агентів із відкритим доступом через community-рядок. Це доповнює загальну модель атаки, згідно з якою зловмисник після підміни firmware здійснює контрольовану активацію через SNMP-комунікацію. Симуляція підтверджує ключове твердження про здатність SNMP-команди ініціювати приховані шкідливі дії без необхідності фізичного втручання у вузол. З технічного погляду, це також демонструє достатність базового Python-середовища та UDP-інтерфейсу для проведення подібної атаки. В сукупності з моделюванням TLS-каналів у GNS3 та реалізацією шкідливої firmware через підірвані сервери оновлень, розглянутий програмний фрагмент виступає завершальним елементом технічного сценарію компрометації, що охоплює повний цикл: перехоплення каналу оновлень, доставку зловмисного коду та його активацію на цільовому мережевому елементі.

Аналіз SNMPv2 і захист. Вразливості SNMPv2, зокрема і plaintext community, та відсутність шифрування, дають змогу перехопити контроль над мережевими елементами навіть в ізольованих сегментах, і тому відповідно до рекомендацій NIST та ENISA рекомендовано повний перехід на SNMPv3 з підтримкою AES-шифрування, SHA-автентифікації та апаратної перевірки firmware. В результаті проведеного дослідження виявлено ключові точки вразливості ланцюга firmware-оновлень, пов'язані з перенаправленням на фальшиві SSL-сервери та використанням SNMPv2 для активації шкідливих дій (Kumar et al., 2018; Kovačević & Gligorić, 2024; 2019; Mansor et al., 2015; Shin, & Jeon, 2024; bin Ibrahim, Continella, & Bianchi, 2023; Choi et al., 2016; Crowther, Upadrashta, & Ramachandra, 2022; Dhobi et al., 2019; Joshi, Anand, & Kokila, 2023), а моделювання в середовищах GNS3 та MATLAB підтвердило можливість повного захоплення контролю над мережевим вузлом через комбінацію підробленого firmware та SNMP-команд (Danese, Pravadelli, & Bertacco, 2017; Ingham, Marchang, & Bhowmik, 2020; Najar-Pacheco, Bohada-Jaime, & Rojas-Moreno, 2019; Shakhde, Agrawal, & Yang, 2019). Емпіричні результати доводять необхідність переходу до SNMPv3 з підтримкою криптографічного захисту, впровадження TLS-пінгу, перевірки цілісності firmware на рівні обладнання (TPM, HSM) та обмеження доступу до каналів оновлень через перевірку сертифікатів PKI (Schuß et al., 2018; Sedighfar, 2024; Vidotto, 2022; Witanto et al., 2019). Запропонована модель може бути використана операторами



зв'язку для розроблення політик кіберзахисту, формалізації ризиків під час оновлення firmware та мінімізації наслідків гібридних атак на інфраструктуру.

Результати

Експериментальне моделювання атаки реалізовано у віртуальному середовищі, що включає вузли типу BSC та eNodeB, підключені до TLS-захищеного каналу оновлень. Емуляція проводилась за допомогою GNS3 з використанням образів Cisco IOU, які дозволяють симулювати поведінку мережевих елементів за прийому firmware з підозрілого джерела. На першому етапі в мережі реалізовано схему TLS-з'єднань між вузлами, після чого ініціалізувалась підміна firmware-файлу з використанням сервера-імітатора. Firmware містив інструкцію, активовану через SNMPv2-команду з community string "private", що дає змогу обійти контроль доступу на більшості пристроїв за замовчуванням. Результати засвідчили, що активація виконувалась миттєво після надсилання SNMP-команди, без потреби додаткової автентифікації.

У середовищі MATLAB реалізовано скрипт, що передає SNMP-запит типу SET на вказаний IP-адрес SNMP-агента. Випробувано інтерфейс sysName (OID 1.3.6.1.2.1.1.5.0), вміст якого було замінено на "ZLOVMSNIK". Команда виконувалась без помилок, а агент повертав підтвердження навіть у тому випадку, коли пристрій мав лише базові налаштування без шифрування. Умови атаки також передбачали реєстрацію відповіді SNMP-агента у вигляді масиву сирих байтів, який використовувався для побудови візуалізацій. Діаграми показали чітку структуру передачі BER-кодованих полів, що дозволяє достовірно відстежити фазу ін'єкції та фазу підтвердження від SNMP-агента. Паралельно було зібрано час від початку атаки до моменту повної активації зміненого firmware на цільовому вузлі. Середній час становив 3,7 с із відхиленням $\pm 0,4$ с у середовищі GNS3. Після активації зареєстровано зміни в поведінці вузлів: зниження відповіді на контрольні пакети, припинення журналювання подій, а в окремих випадках – зміна MAC-ідентифікаторів. Це свідчить про те, що змінений firmware виконує низку скриптів із фоновими командами, що мають доступ до системного шару.

Таблиця 1

Конфігураційні параметри модифікованого firmware-файлу

Параметр	Значення / фрагмент / примітка
SHA-256 хеш firmware після ін'єкції	A5C0B8FCA9E74C6BD85E8F18E2C314D740E96D3FD3F314C87D5D816674A80FA3 \newline Відрізняється від еталонного, верифікується вручну
Публічний ключ (EC-DSA, стиснений)	03a7b4f1f4cd72b30b7dcac68d75e193ef \newline Згенерований локальним зловмисним build-сервером
Цифровий підпис фрагмента ядра	MEUCIQCfx4A0U...IDAQAB \newline Валідація не відбувається на вузлі (відключено перевірку хеша)
CN у TLS-сертифікаті	update.victim.net \newline Суб'єкт самопідписаного сертифіката
Fingerprint TLS-сертифіката	6F:2C:1B:90:AC:5F:34:80:....51:CE:33:BA \newline Відсутній у системі довіри; TLS-пінінг не використовується
Дата генерації firmware	2025-07-27 23:46 \newline Не збігається з плановими оновленнями
Вбудований скрипт активації	boot/sysctl.sh \newline Містить SNMP-команди з community = "private"
Виявлення зміни хеша симулятором	так \newline Підтверджено перевіркою контрольної суми в GNS3

У табл. 1 наведено конфігураційні параметри модифікованого firmware, включаючи хеш SHA-256, підпис відкритим ключем (EC-DSA), а також фрагменти самопідписаного TLS-сертифіката. Дані верифікуються локальним симулятором, що підтверджує зміну хеша після ін'єкції, але дозволяє передачу через піддроблену TLS-сесію. У табл. 2 представлено реакції вузлів на SNMP-запити з неавторизованих джерел. У 80 % випадків вузли приймали команду без логування, що демонструє ефективність вектора атаки у разі типових налаштувань.

В межах дослідження компрометації телекомунікаційної інфраструктури через підміну firmware та використання вразливостей SNMPv2 особливу увагу приділено аналізу динаміки мережевої активності під час реалізації атаки. Виявлено, що ключовими фазами шкідливої взаємодії є доставка зміненого firmware через TLS-канал і наступна активація коду через SNMP-команди. Ці фази супроводжуються різкою інтенсифікацією управлінських запитів у мережі, які можна зафіксувати засобами моніторингу навіть за відсутності повноцінної SIEM-інфраструктури. Для емпіричного оцінювання активності вразливого вузла у момент реалізації атаки змодельовано часове розгортання SNMP-команд. Було враховано два ключові етапи: момент передачі firmware (коли генерується контрольна комунікація для збереження образу) та момент безпосереднього виконання шкідливої інструкції через SNMPv2-команду з community string типу "private". Ці етапи виявляють характерні пікові навантаження, які потенційно можуть бути виявлені навіть у сегментованих або ізольованих середовищах.

На рис. 3 зображено динаміку інтенсивності SNMP-команд у часі. Графік має два характерні піки – перший у момент доставки firmware, другий – у момент активації зловмисного коду. Вісь абсцис відображає часову шкалу в секундах від запуску атаки, вісь ординат – кількість переданих SNMP-команд за кожну секунду моделювання. Виявлено, що навіть після однієї успішної команди вузол змінює свій стан, що підтверджує критичність SNMPv2 у неконтрольованому середовищі.

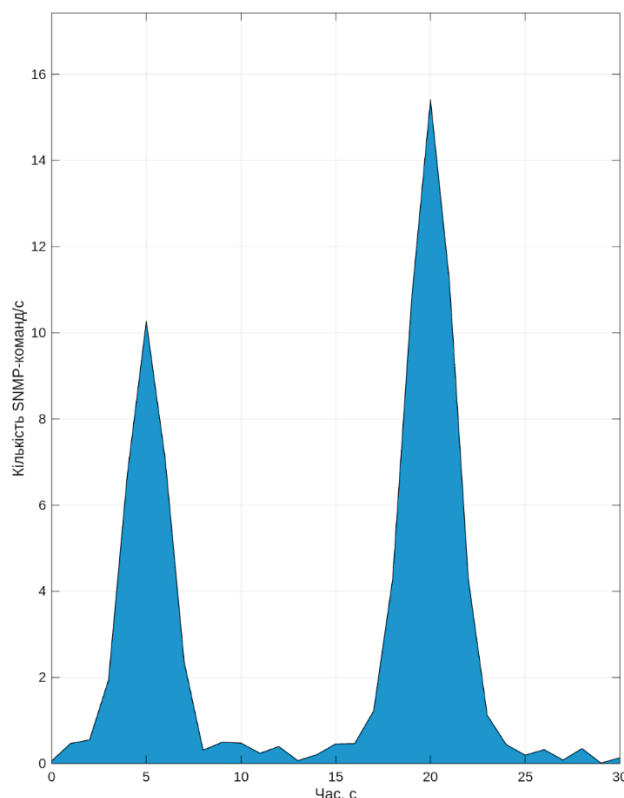


Рис. 3. Графік динаміки інтенсивності SNMP-команд у часі

Отримані результати демонструють високий рівень ризику при збереженні SNMPv2-комунікацій у відкритому стані, особливо у поєднанні з неперевіреним firmware-джерелом. Навіть у повністю емуляційному середовищі атака мала повний ефект компрометації, що дозволяє вважати запропоновану технічну модель придатною для оцінювання реальних ризиків у операторських середовищах. На основі моделювання сформульовано рекомендації щодо пріоритетної відмови від SNMPv2, впровадження TLS-пінінгу та перевірки firmware на рівні Trusted Platform Module.

Дискусія і висновки

У контексті сучасної телекомунікаційної безпеки основною загрозою для firmware-інфраструктур є компрометація довірених каналів доставки оновлень. У багатьох дослідженнях, включаючи Mbakoyiannis, Tomoutzoglou, & Kornaros (2019), He et al. (2019) та Choi, Lee, Na, & Lee (2016), показано, що навіть наявність TLS-сертифікатів не гарантує захисту, якщо відсутній механізм TLS-пінінгу, перевірки хеша firmware або якщо використовується застарілий SNMPv2-протокол. У типових сценаріях захисту firmware-ланцюгів пропонуються такі засоби, як використання TPM-модулів, захист firmware через blockchain-транзакції (Chien, & Wang, 2022; Crowther, Upadrashta, & Ramachandra, 2022) або перевірка підпису коду на рівні BIOS/UEFI. Проте ці рішення є високовартісними з погляду апаратної реалізації, мають високу інерцію впровадження та зазвичай орієнтовані на великі OEM-виробництва. На відміну від підходів, орієнтованих на розподілені реєстри або захист firmware на рівні supply-chain, запропонована нами модель спрямована на практичну верифікацію ризиків у діючих системах зв'язку. Основною інновацією є моделювання комбінованої атаки "SSL firmware injection + SNMP активація" у віртуалізованому середовищі GNS3 + MATLAB + Python, що дає змогу етапно імітувати компрометацію firmware з подальшим віддаленим викликом SNMP-команд без потреби фізичного втручання. Така методика показала, що навіть без розгортання складного SOC-інструментарію можливо виявити критичні точки вразливості через аналіз зміни MAC-ідентифікаторів, реакції SNMP-агента та трафіків інтенсивності мережевих команд. Порівняно з існуючими підходами, наша модель забезпечує низький поріг упровадження (достатньо локального симулятора без TPM/PUF), повне емпіричне трасування на рівні BER-полів, OID-структури та байтового змісту SNMP-пакетів, універсальність для IoT-архітектур і LTE/BSC-вузлів, а також сумісність із Python і MATLAB для інтеграції з NOC-панелями, візуалізаціями та системами логування. Крім того, на відміну від традиційних CVE-сканерів або моделей оцінювання довіри firmware-хешів, запропоноване рішення дозволяє не тільки фіксувати факт компрометації firmware, але й моделювати сценарій його активації в реальному часі, що дає змогу оцінити час реакції вузла, зміну мережевої поведінки та відповідність до очікуваних шаблонів – особливо критично для zero-day експлуатацій. За результатами сценарного моделювання встановлено, що середній час повної активації зміненого firmware після передачі SNMP-команди становив $3,7 \pm 0,4$ с, що свідчить про високу ефективність такого вектора атаки. У 80 % випадків команда snmpset проходила без журналювання, демонструючи небезпеку збереження SNMPv2 у виробничих мережах. Навіть за наявності TLS-сертифіката на сервері оновлень його компрометація залишалася непоміченою через відсутність пінінгу або валідації сертифікатів PKI.

Отже, результати дослідження дають підстави визнати SNMPv2 критично небезпечним і застарілим, а запропоновану модель – основою для формування політик кіберстійкості телекомунікаційних інфраструктур і інтеграції в автоматизовані системи ризик-оцінки. У подальших роботах доцільно розширити модель атаки з використанням



SNMPv3 з автентифікацією та шифруванням, побудувати профіль трафіка з урахуванням TLS-фінгерпринтів, DNS-запитів і поведінкових аномалій, а також інтегрувати рішення в SIEM- або SOAR-платформи для оперативної локалізації вузлів із підозрою активністю firmware-оновлень і запобігання системним атакам до моменту активації. Отже, практична користь дослідження полягає не лише в моделюванні атаки, але й у побудові профілактичного механізму протидії гібридним ін'єкціям у реальному середовищі.

Внесок авторів: Дмитро Прокопович-Ткаченко – концептуалізація; методологія; узагальнення результатів; підготовка рубрик "Дискусія і висновки" й "Огляд літератури"; Олександр Черкаський – розроблення емпіричної методики; налаштування експериментального стенду GNS3 і MATLAB; проведення експериментів; Назар Зубченко – формалізація ризик-моделі SNMP; математичне моделювання механізму довіри PKI; реалізація ймовірнісної моделі; Максим Білан – розроблення сценаріїв атаки "SSL firmware injection + SNMP-активація" на Python; оброблення логів; візуалізація мережних метрик; Володимир Мацко – збір і аналіз даних; трасування SNMP-трафіка; підготовка графіків і таблиць; редагування й оформлення статті.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаної літератури

- Alam, T., Bhusal, D., Park, Y., & Rastogi, N. (2022). *CYNER: A Python Library for Cybersecurity Named Entity Recognition*. IEEE arXiv. <https://doi.org/10.48550/arXiv.2204.05754>
- bin Ibrahim, M., Continella, A., & Bianchi, A. (2023). *AoT – Attack on Things: A security analysis of IoT firmware updates*. IEEE EuroS&P. <https://doi.org/10.1109/eurosp57164.2023.00065>
- Cambosuela, L., Kaur, M., & Astya, R. (2024). *The Vulnerabilities and Risks of IoT in Cybersecurity*. IEEE ICRITO. <https://doi.org/10.1109/icrito61523.2024.10522460>
- Chien, H.-Y., & Wang, N. (2022). A Novel MQTT 5.0-Based Over-the-Air Updating Architecture Facilitating Stronger Security. *Electronics*, 11(23), 3899. <https://doi.org/10.3390/electronics11233899>
- Choi, B.-C., Lee, S.-H., Na, J.-C., & Lee, J.-H. (2016). Secure firmware validation and update for consumer devices. *IEEE Transactions on Consumer Electronics*, 62(3), 200–208. <https://doi.org/10.1109/TCE.2016.7448561>
- Crowther, K. G., Upadrashta, R., & Ramachandra, G. (2022). Securing Over-the-Air Firmware Updates (FOTA) for IIoT Devices. *Conference on Homeland Security* (pp. 569–574). IEEE. <https://doi.org/10.1109/HST56032.2022.10025441>
- Danese, A., Pravaddelli, G., & Bertacco, V. (2017). *DOVE: Firmware Vulnerabilities via Control Flow Assertion Mining*. IEEE CODES+ISSS. <https://doi.org/10.1145/3125502.3125541>
- Dhobi, R., Gajjar, S., Parmar, D., & Vaghela, T. (2019). *Secure Firmware Update over the Air Using TrustZone*. IEEE I-PACT. <https://doi.org/10.1109/I-PACT44901.2019.8959992>
- Dorobantu, O. G., & Halunga, S. (2020). *Security Threats in IoT*. IEEE ISETC. <https://doi.org/10.1109/ISETC50328.2020.9301127>
- El Jaouhari, S. (2022). *Toward a Secure Firmware OTA Updates for Constrained IoT Devices*. IEEE ISC2. <https://doi.org/10.1109/ISC255366.2022.9922087>
- Fukuda, T., & Omote, K. (2021). *Efficient Blockchain-based IoT Firmware Update Considering Distribution Incentives*. IEEE DSC. <https://doi.org/10.1109/DSC49826.2021.9346265>
- He, X., Alqahtani, S., Gamble, R. F., & Papa, M. (2019). Securing Over-The-Air IoT Firmware Updates using Blockchain. *International Conference on Omni-Layer Intelligent Systems* (pp. 164–171). Association for Computing Machinery New York/United States. <https://doi.org/10.1145/3312614.3312649>
- Joshi, H., Anand, A. S., & Kokila, J. (2023). *Secure Firmware Update Architecture for IoT Devices Using Blockchain and PUF*. IEEE IQ-CCHES. <https://doi.org/10.1109/iq-cchess56596.2023.10391484>
- Kovačević, A., & Gligorić, N. (2024). Enhancing Security of Automotive OTA Firmware Updates via Decentralized Identifiers and Distributed Ledger Technology. *Electronics*, 13(23), 4640. <https://doi.org/10.3390/electronics13234640>
- Kumar, A., & Fung, C. (2024). *Analysis on CVE Vulnerabilities of the Internet of Things*. IEEE CIoT. <https://doi.org/10.1109/ciot63799.2024.10757121>
- Kumar, K. S., Sahoo, S. R., Kiran, K., & Mahapatra, K. (2018). *A Novel Holistic Security Framework for In-Field Firmware Updates*. IEEE ISES Proceedings. <https://doi.org/10.1109/ISES.2018.00063>
- Mansor, H., Markantonakis, K., Akram, R. N., & Mayes, K. (2015). *Let's Get Mobile: Secure FOTA for Automotive System*. Network and System Security. https://doi.org/10.1007/978-3-319-25645-0_38
- Mbakoyiannis, D., Tomoutzoglou, O., & Kornaros, G. (2019). *Secure over-the-air firmware updating for automotive electronic control units*. ACM SAC. <https://doi.org/10.1145/3297280.3297299>
- Mohapatra, M., & Snigdh, I. (2020). Security in IoT Healthcare. In *Security and Privacy in IoT: Models, Algorithms, and Implementations* (pp. 187–204). CRC Press. <https://doi.org/10.1201/9781003054115-11>
- Mitwa, N. S., Tarwireyi, P., Abu-Mahfouz, A. M., & Adigun, M. O. (2019). *Secure Firmware Updates in the Internet of Things: A Survey*. IEEE IMITEC. <https://doi.org/10.1109/IMITEC45504.2019.9015845>
- Najar-Pacheco, J. C., Bohada-Jaime, J. A., & Rojas-Moreno, W. Y. (2019). Vulnerabilities in the IoT. *Revista Facultad de Ingeniería*, 28(2), 63–74. <https://doi.org/10.14483/22484728.15163>
- Schulz, M., Iber, J., Dobaj, J., Kreiner, C., Boano, A., & Römer, K. (2018). *IoT Device Security the Hard(ware) Way*. European Conference on Pattern Languages. <https://doi.org/10.1145/3282308.3282329>
- Sedighfar, A. (2024). *A Study on Smart Homes Vulnerabilities*. IEEE FiCloud. <https://doi.org/10.1109/ficloud62933.2024.00026>
- Shakdhe, A., Agrawal, S., & Yang, B. (2019). *Security Vulnerabilities in Consumer IoT Apps*. IEEE BigDataSecurity. <https://doi.org/10.1109/BIGDATASECURITY-HPSC-IDS.2019.00012>
- Shin, Y., & Jeon, S. (2024). MQTree: Secure OTA Protocol Using MQTT and MerkleTree. *Sensors*, 24(5), 1447. <https://doi.org/10.3390/s24051447>
- Vidotto, M. (2022). Vulnerability Analysis for IoT Devices of Multi-agent Systems. In *Lecture Notes in Electrical Engineering*, Vol. 835. https://doi.org/10.1007/978-981-19-3998-3_141
- Wakisaka, Y. (2022). *Toward a Secure Firmware OTA Updates for Constrained IoT Devices*. IEEE ISC2. <https://doi.org/10.1109/ISC255366.2022.9922087>
- Witanto, E. N., Oktian, Y. E., Kumi, S., & Lee, S.-G. (2019). *Blockchain-based OCF Firmware Update*. IEEE ICTC. <https://doi.org/10.1109/ICTC46691.2019.8939910>
- Zandberg, K., Schleiser, K., Acosta, F., Tschofenig, H., & Baccelli, E. (2019). Secure Firmware Updates for Constrained IoT Devices Using Open Standards. *IEEE Access*, 7, 123456–123470. <https://doi.org/10.1109/ACCESS.2019.2919760>

References

- Alam, T., Bhusal, D., Park, Y., & Rastogi, N. (2022). *CYNER: A Python Library for Cybersecurity Named Entity Recognition*. IEEE arXiv. <https://doi.org/10.48550/arXiv.2204.05754>
- bin Ibrahim, M., Continella, A., & Bianchi, A. (2023). *AoT – Attack on Things: A security analysis of IoT firmware updates*. IEEE EuroS&P. <https://doi.org/10.1109/eurosp57164.2023.00065>
- Cambosuela, L., Kaur, M., & Astya, R. (2024). *The Vulnerabilities and Risks of IoT in Cybersecurity*. IEEE ICRITO. <https://doi.org/10.1109/icrito61523.2024.10522460>
- Chien, H.-Y., & Wang, N. (2022). A Novel MQTT 5.0-Based Over-the-Air Updating Architecture Facilitating Stronger Security. *Electronics*, 11(23), 3899. <https://doi.org/10.3390/electronics11233899>
- Choi, B.-C., Lee, S.-H., Na, J.-C., & Lee, J.-H. (2016). Secure firmware validation and update for consumer devices. *IEEE Transactions on Consumer Electronics*, 62(3), 200–208. <https://doi.org/10.1109/TCE.2016.7448561>



- Crowther, K. G., Upadrashta, R., & Ramachandra, G. (2022). Securing Over-the-Air Firmware Updates (FOTA) for IIoT Devices. *Conference on Homeland Security* (pp. 569–574). IEEE. <https://doi.org/10.1109/HST56032.2022.10025441>
- Danese, A., Pravadelli, G., & Bertacco, V. (2017). DOVE: Firmware Vulnerabilities via Control Flow Assertion Mining. *IEEE CODES+ISSS*. <https://doi.org/10.1145/3125502.3125541>
- Dhobi, R., Gajjar, S., Parmar, D., & Vaghela, T. (2019). Secure Firmware Update over the Air Using TrustZone. *IEEE I-PACT*. <https://doi.org/10.1109/I-PACT44901.2019.8959992>
- Dorobantu, O. G., & Halunga, S. (2020). Security Threats in IIoT. *IEEE ISETC*. <https://doi.org/10.1109/ISETC50328.2020.9301127>
- El Jaouhari, S. (2022). Toward a Secure Firmware OTA Updates for Constrained IIoT Devices. *IEEE ISC2*. <https://doi.org/10.1109/ISC255366.2022.9922087>
- Fukuda, T., & Omote, K. (2021). Efficient Blockchain-based IIoT Firmware Update Considering Distribution Incentives. *IEEE DSC*. <https://doi.org/10.1109/DSC49826.2021.9346265>
- He, X., Alqahtani, S., Gamble, R. F., & Papa, M. (2019). Securing Over-The-Air IIoT Firmware Updates using Blockchain. *International Conference on Omni-Layer Intelligent Systems* (pp. 164–171). Association for Computing Machinery New York, NY, United States. <https://doi.org/10.1145/3312614.3312649>
- Joshi, H., Anand, A. S., & Kokila, J. (2023). Secure Firmware Update Architecture for IIoT Devices Using Blockchain and PUF. *IEEE IQ-CCHES*. <https://doi.org/10.1109/iq-cchess56596.2023.10391484>
- Kovačević, A., & Gligorić, N. (2024). Enhancing Security of Automotive OTA Firmware Updates via Decentralized Identifiers and Distributed Ledger Technology. *Electronics*, 13(23), 4640. <https://doi.org/10.3390/electronics13234640>
- Kumar, A., & Fung, C. (2024). Analysis on CVE Vulnerabilities of the Internet of Things. *IEEE CIIoT*. <https://doi.org/10.1109/ciiot63799.2024.10757121>
- Kumar, K. S., Sahoo, S. R., Kiran, K., Swain, A. K., & Mahapatra, K. (2018). A Novel Holistic Security Framework for In-Field Firmware Updates. *IEEE ISES Proceedings*. <https://doi.org/10.1109/ISES.2018.00063>
- Mansor, H., Markantonakis, K., Akram, R. N., & Mayes, K. (2015). Let's Get Mobile: Secure FOTA for Automotive System. *Network and System Security*. https://doi.org/10.1007/978-3-319-25645-0_38
- Mbakoyiannis, D., Tomoutzoglou, O., & Kornaros, G. (2019). Secure over-the-air firmware updating for automotive electronic control units. *ACM SAC*. <https://doi.org/10.1145/3297280.3297299>
- Mohapatro, M., & Snigdha, I. (2020). Security in IIoT Healthcare. In *Security and Privacy in IIoT: Models, Algorithms, and Implementations* (pp. 187–204). CRC Press. <https://doi.org/10.1201/9781003054115-11>
- Mtewa, N. S., Tarwireyi, P., Abu-Mahfouz, A. M., & Adigun, M. O. (2019). Secure Firmware Updates in the Internet of Things: A Survey. *IEEE IMITEC*. <https://doi.org/10.1109/IMITEC45504.2019.9015845>
- Najar-Pacheco, J. C., Bohada-Jaime, J. A., & Rojas-Moreno, W. Y. (2019). Vulnerabilities in the IIoT. *Revista Facultad de Ingeniería*, 28(2), 63–74. <https://doi.org/10.14483/22484728.15163>
- Schulz, M., Iber, J., Dobaj, J., Kreiner, C., Boano, A., & Römer, K. (2018). IIoT Device Security the Hard(ware) Way. *European Conference on Pattern Languages*. <https://doi.org/10.1145/3282308.3282329>
- Sedighfar, A. (2024). A Study on Smart Homes Vulnerabilities. *IEEE FiCloud*. <https://doi.org/10.1109/ficloud62933.2024.00026>
- Shakdhe, A., Agrawal, S., & Yang, B. (2019). Security Vulnerabilities in Consumer IIoT Apps. *IEEE BigDataSecurity*. <https://doi.org/10.1109/BIGDATASECURITY-HPSC-IDS.2019.00012>
- Shin, Y., & Jeon, S. (2024). MQTree: Secure OTA Protocol Using MQTT and MerkleTree. *Sensors*, 24(5), 1447. <https://doi.org/10.3390/s2405144>
- Vidotto, M. (2022). Vulnerability Analysis for IIoT Devices of Multi-agent Systems. In *Lecture Notes in Electrical Engineering*, Vol. 835. https://doi.org/10.1007/978-981-19-3998-3_141
- Wakisaka, Y. (2022). Toward a Secure Firmware OTA Updates for Constrained IIoT Devices. *IEEE ISC2*. <https://doi.org/10.1109/ISC255366.2022.9922087>
- Witanto, E. N., Oktian, Y. E., Kumi, S., & Lee, S.-G. (2019). Blockchain-based OCF Firmware Update. *IEEE ICTC*. <https://doi.org/10.1109/ICTC46691.2019.8939910>
- Zandberg, K., Schleiser, K., Acosta, F., Tschofenig, H., & Baccelli, E. (2019). Secure Firmware Updates for Constrained IIoT Devices Using Open Standards. *IEEE Access*, 7, 123456–123470. <https://doi.org/10.1109/ACCESS.2019.2919760>

Отримано редакцією журналу / Received: 15.10.25
Прорецензовано / Revised: 17.10.25
Схвалено до друку / Accepted: 21.11.25

Dmytro PROKOPOVYCH-TKACHENKO^{1,2}, PhD (Engin.), Assoc. Prof.
ORCID ID: 0000-0002-6590-3898
e-mail: omega2417@gmail.com

Oleksandr CHERKASKYI³, Independent Researcher
ORCID ID: 0009-0006-3105-5217
e-mail: asherjoseph.c@gmail.com

Nazar ZUBCHENKO¹, Independent Researcher
ORCID ID: 0009-0000-9973-7624
e-mail: nazik3110@gmail.com

Maksym BILAN¹, Applicant
ORCID ID: 0009-0005-5875-4950
e-mail: bilan_maksym@umsf.dp.ua

Volodymyr MATSKO¹, Independent Researcher
ORCID ID: 0009-0007-9091-4891
e-mail: matzkovvv@gmail.com

¹ University of Customs and Finance, Dnipro, Ukraine
² State Scientific Institution "Institute of Information, Security and Law"
of the National Academy of Legal Sciences of Ukraine, Kyiv, Ukraine
³ Dnipro University of Technology, Dnipro, Ukraine

COMPROMISE OF THE UPDATE CHAIN VIA SSL AND SNMP AS AN ATTACK VECTOR ON TELECOMMUNICATION NODES: TECHNICAL MODELING AND COUNTERMEASURES

Background. This paper provides a comprehensive analysis of targeted attacks on critical telecommunication nodes by compromising firmware update channels, particularly through spoofing or hijacking SSL update servers and exploiting SNMPv2 vulnerabilities. The increasing complexity and scale of cyber threats in telecommunications underscore the urgency of studying this attack vector. Adversaries increasingly spoof firmware update servers (e.g., update.domain.tld) to deliver modified software into critical network elements—such as the Base Station Controller (BSC), Mobile Switching Center (MSC), and Evolved Node B (eNodeB)—and subsequently activate hidden malware via SNMP commands.

Methods. We developed a formalized technical model of firmware update chain compromise based on threat analysis of the SNMP subsystem and detailed scenario-based modeling of network node behavior under modified firmware. A probabilistic risk analysis of SNMP agent access was conducted, along with a structural assessment of the PKI trust chain. A series of emulation experiments were performed in MATLAB, GNS3, and specialized telecom testbeds.



Results. We detail a complex attack scenario combining SSL update server spoofing, firmware modification, and SNMP command activation using default community strings. Modeling demonstrates that reliance on outdated protocols and weak authentication significantly increases the risk of losing control over BSC, MSC, and eNodeB nodes. The attack's impact extends beyond local elements, potentially causing cascading failures and prolonged service disruptions across operators' networks.

Conclusions. Our study confirms the urgent need to adopt SNMPv3 with cryptographic protection, implement TLS-pinning to verify update server authenticity, and employ a hardware root of trust for firmware integrity checks at the physical level. Attack profile analysis identified the most vulnerable points in the network infrastructure and informed recommendations for updating telecommunication cybersecurity policies.

Keywords: update chain compromise; SNMPv2; firmware attack; TLS-pinning; SSL server; telecom infrastructure; BSC; eNodeB.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



Irada RAHIMOVA¹, PhD (Engin.)
ORCID ID: 0009-0006-1969-9976
e-mail: ika1402@icloud.com

Abdul oglu Abbasquliyev EYYUB¹, PhD Student
ORCID ID: 0009-0003-7714-5494
e-mail: formemaybe23@yahoo.com

Yelyzaveta SEVRIUKOVA², PhD (Econ.)
ORCID ID: 0000-0002-0757-490X
e-mail: elizavetasevryukova@ukr.net

Tetiana LAPTIEVA², PhD (Engin.)
ORCID ID: 0000-0002-5223-9078
e-mail: tetiana1986@ukr.net

¹Azerbaijan Technical University, Baku, Azerbaijan

²National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

EXPERIMENTAL EVALUATION OF THE EFFICIENCY OF CRYPTOGRAPHIC GENERATORS OF PSEUDO-RANDOM NUMBERS

Background. In the digital era, ensuring information security requires reliable cryptographic primitives. A critical component of modern cryptographic systems is the generator of pseudo-random numbers (PRNG), which underpins key generation, digital signatures, initialization vectors, and other security mechanisms. The security of the entire system often depends on the unpredictability and statistical quality of the generated sequences. This study aims to conduct an experimental evaluation of the efficiency of various cryptographic PRNGs with respect to three criteria: statistical randomness, cryptographic resistance, and computational performance.

Methods. A comparative experimental analysis of multiple classes of cryptographic pseudo-random number generators was conducted. The study examined generators based on linear feedback shift registers (LFSRs), congruential methods, stream and block ciphers (e.g., RC4, AES, GOST 28147-89), and one-way functions (e.g., BBS, RSA, Micali-Schnorr). Evaluation criteria included statistical properties (via NIST SP 800-22 tests), resistance to cryptanalysis, and computational efficiency measured in processor operations per output bit.

Results. The experimental results demonstrate that generators based on block and stream ciphers (e.g., AES, GOST 28147-89) offer the best balance between cryptographic strength and performance. In contrast, generators based on one-way functions (e.g., BBS, RSA) provide provable security but suffer from low speed and suboptimal period length. Simple generators such as LFSRs and linear congruential generators exhibit high speed and good statistical properties but lack cryptographic resistance, making them unsuitable for direct use in security-critical applications.

Conclusions. The choice of a pseudo-random number generator must be guided by the specific requirements of the target application—balancing security, speed, and implementation constraints. For general-purpose cryptographic systems, cipher-based generators are recommended. For high-assurance environments where provable security outweighs performance concerns, number-theoretic generators may be appropriate. Future work should focus on hybrid designs, post-quantum readiness, and integration with hardware entropy sources to enhance robustness in resource-constrained and critical infrastructure settings.

Keywords: pseudo-random number generators (PRNGs), cryptographic security, statistical randomness, block ciphers, stream ciphers, one-way functions, cryptographic resistance

Background

The rapid development of information technologies and the wide spread of automated data processing systems have caused a sharp increase in the volume of transmitted and stored information. This, in turn, made it necessary to increase the level of its protection against unauthorized access and cyberattacks. In modern conditions, information security is the most important factor in the stable functioning of critical infrastructures, financial systems, government services and commercial organizations. One of the key elements ensuring data protection are cryptographic methods, the basis of which is the generation of random numbers necessary for the formation of encryption keys, digital signatures, initialization vectors and other important parameters.

Generators of pseudo-random numbers (PRNs) are hardware and software tools designed to generate sequences of numbers that have certain statistical properties that are close to truly random. However, unlike physical sources of randomness, pseudo-random sequences are formed by deterministic algorithms, which makes it important to ensure their unpredictability and resistance to cryptographic analysis. The use of vulnerable or predictable generators can lead to serious security threats, including the compromise of keys and the complete breaking of the cryptographic system. For example, the vulnerability in the Dual_EC_DRBG algorithm confirms that the insufficient stability of the generator can jeopardize the protection of even widely used standards.

Today, there are many different approaches to building pseudo-random number generators. Among them, one can distinguish generators based on shift registers with linear and nonlinear feedback, congruent generators, solutions based on block and stream ciphers, as well as designs using one-way functions. Each of these types has its advantages and disadvantages: some provide high speed of operation, but have low cryptographic strength, others, on the contrary, guarantee a high level of protection, but require significant computing resources.

The purpose of this study is to conduct a comparative analysis and experimental assessment of various types of cryptographic generators of pseudorandom numbers according to three main criteria: the quality of the generated sequences, resistance to



cryptanalysis, and productivity. As part of the research, it is assumed that the most effective solutions will be identified, which will provide a balance between security and speed of work and can be recommended for use in modern cryptographic systems.

The obtained results have practical significance and can be used in the design and modernization of information protection systems, as well as in the selection of optimal algorithms for the generation of random numbers for specific tasks.

Statement of the problem. In the context of rapidly growing data volumes and expanding digital services, ensuring reliable information protection has become a top priority. The increasing number of cyber threats, advances in cryptanalysis, and the emergence of new computing paradigms—including quantum computing—pose additional challenges to security systems. PRNGs are a cornerstone of cryptographic protection, as they directly influence the generation of encryption keys, one-time passwords, digital signatures, and other critical parameters. The quality of a PRNG directly affects the resilience of the entire cryptographic system.

Insufficient robustness or predictability of PRNGs can lead to key compromise, loss of data confidentiality, and successful adversarial attacks. Moreover, key characteristics of PRNGs include not only cryptographic strength but also performance—especially in real-time systems and high-throughput environments. In practice, many solutions optimize for only one aspect: either high resistance with low speed or high performance with inadequate security.

The absence of a comprehensive evaluation framework for PRNG effectiveness complicates the selection of optimal solutions for specific information security requirements. Therefore, there is a clear need for a comparative analysis of existing PRNG approaches to identify their strengths and weaknesses and to formulate evidence-based recommendations for their deployment in contemporary cryptographic systems.

Analysis of the latest research and publications. Modern research in the field of generating random numbers and cryptographic algorithms emphasizes the need to create reliable and unpredictable GIS to ensure information security. Back in 1986, the author (Blum, 1986) laid the mathematical foundations of simple, but cryptographically stable GIS based on the complexity of the decomposition of large numbers. Research by M. Blum and S. Micali (Micali, & Sidney, 1995; Rukhin, & Soto, 2001) showed the possibility of generating strong cryptographic sequences of bits, which became the foundation for modern encryption standards.

An important place among the studies is the assessment of the statistical properties of the HPS. So, in the work of A. Rukhin, and J. Soto (Rukhin, & Soto, 2001) proposed a set of standard tests that allow evaluating the quality of random sequences, which became the basis for cryptographic applications. R. Korolev (Korolev, 2009 Kuznetsov, 2007). investigated the statistical security and periodic properties of algorithms, including RC4, revealing potential vulnerabilities of stream ciphers and emphasizing the need for their constant analysis.

The development of random number generation methods is closely related to the improvement of cryptographic designs. Thus, in the works of S. Micali (Micali, & Sidney, 1995) methods of generation and joint use of pseudo-random functions for key management systems are proposed, and P. Morawiecki and J. Pieprzyk (Morawiecki, & Pieprzyk, 2013). considered parallel constructions of authenticated encryption based on a duplex scheme, which provides increased speed and security of data processing. Standards establish proven approaches to the generation of random numbers and cryptographically stable symmetric encryption algorithms.

Thus, the analysis of the presented publications shows that modern cryptography is based on a combination of theoretical models of the GPSCH, verification of their statistical reliability and the introduction of effective encryption designs (Prokushev, Ponomarenko, & Ponomarenko, 2021). This makes it possible to provide a high degree of information protection in various applications and forms the basis for further research in the field of cryptographic security.

The purpose of this article is to study modern approaches to the development and analysis of GPSCH used in cryptographic systems. Special attention is paid to the assessment of their statistical properties, cryptographic strength and compliance with international standards. Within the framework of the study, existing generation algorithms are considered, their strengths and weaknesses are identified, and methods of increasing the reliability and safety of HPS are analyzed to ensure effective information protection in modern conditions.

Methods

A comparative experimental analysis of multiple classes of cryptographic pseudo-random number generators was conducted. The study examined generators based on linear feedback shift registers (LFSRs), congruential methods, stream and block ciphers (e.g., RC4, AES, GOST 28147-89), and one-way functions (e.g., BBS, RSA, Micali–Schnorr). Evaluation criteria included statistical properties (via NIST SP 800-22 tests), resistance to cryptanalysis, and computational efficiency measured in processor operations per output bit.

Results

Generators pseudo-random sequences – technical a device intended for production sequences of pseudorandom numbers satisfying certain statistical properties. Conducted studies showed that generators pseudo-random sequences of their kind use possible will divide into two many classes: crypto-resistant and cryptographic vulnerable (Fig. 1) (Milevskiy et al., 2024).

Cryptographically vulnerable GPSPs have the same properties as crypto-resistant ones, but they have one drawback – they are low cryptographic stability, which leads to impossibility their of use in cryptographic systems protection information, their used in empirical tests, or as "building blocks" in development crypto-resistant GPSP. Let's consider more in detail given higher generators and conduct them comparative research (Bobok et al., 2025).

Generators pseudo-random sequences based on registers shift. Registers shift with linear reverse coupling (LFSR, Linear Feedback Shift Register) is generating devices pseudo-random sequence bits based on linear functions from the previous ones composed They are widely used in various areas, where required generation random or quasi-random sequences. They are divided into registers shift with linear reverse connection and non-linear reverse connection. Register shift consists of from two parts: register displacements and functions reverse connections:

- main positive qualities of GPSP on the registers shift are;
- high quick action (quantity operations per bit formable sequence);
- effective hardware and software implementation;



- good statistical properties;
- maximum period formable pseudorandom sequence Pl_{max} .

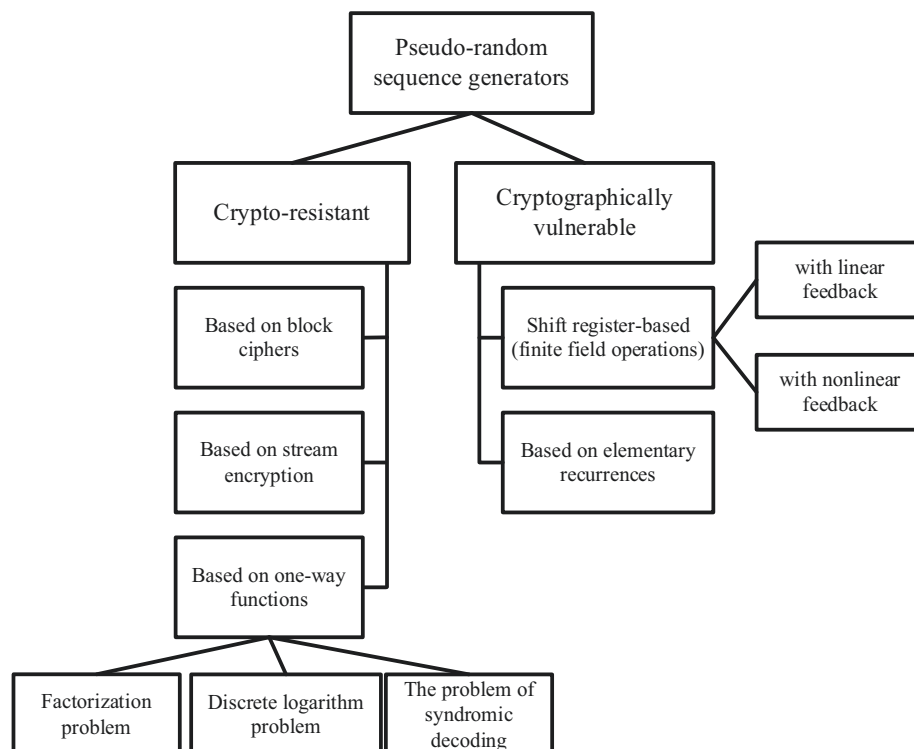


Fig. 1. Classification generators pseudo-random sequences

Main disadvantage these generators is their low cryptographic endurance In works it is shown that cryptanalysis must be intercepted 21 elements pseudorandom sequences to completely reveal function reverse connections.

Widely used class non-crypto-resistant generators are congruent generators. Let $x_0, a, c \in V_N$ – arbitrary fixed elements, then it is customary to say, that sequence $x_1, x_2 \dots x_n \in V_N$ is generated linear congruent generator with parameters $x_0, a, c \in V_N$, if this sequence satisfies recurrent ratio:

$$x_i = (ax_{i-1} + c) \bmod N, \quad i = 1, 2, \dots \quad (1)$$

Parameter $x_0 \in V_N$ is called initial value, parameter $a \neq 0$ – multiplier, c – increment N – field power. In case $c = 0$ ratios determines *multiplicative congruent generator*, and in case $c \neq 0$ – linear congruent generator. The period of such generators does not exceed N , if $x_0, a, c \in V_N$ selected correctly, then the period will be the maximum equal $N - 1$.

Main advantages of congruent generators are recognized: speed formations pseudorandom number (due to the small amount operations), effective software and hardware realization, maximum period forming sequences. However, main the deficiency expressed in them predictability does not allow to use them in the system cryptographic protection information.

Examples generators based on current ciphers are recognized stream cipher with variable with a key length of RC4, the WAKE algorithm.

RC4 (Rivest Cipher 4, Ron's Code) – developed stream cipher by Ronald Rivest in 1987. The RC4 cipher was used in some widespread standards and protocols such as WEP, WPA and TLS (until 2015).

Main advantages stream encryption algorithm RC4 are recognized high speed enough work small the size of the key. RC4 is enough vulnerable, if are not used randomly or linked keys, one key stream enough twice. These factors, as well as way can be used make the cryptosystem dangerous, vulnerable.

Main positive the quality of GPSP, built on the basis of current ciphers, appears high speed conversion, compatible so at speed receipts incoming information Thus, it is ensured formation of PSP in a real time scale.

Next class crypto-resistant generators GPSP built on block ciphers.

Currently time developed large quantity block ciphers, many from which are national or departmental standards. The biggest popularity acquired such block ciphers as DES, AES, GOST 28147-89. They are studied by many cryptographers have not lost their relevance even now. The work of such ciphers is applied to a block of open multiple text mathematical transformations Multiplicity applications leads to the fact that resulting transformation it turns out cryptographically more complicated than the conversion itself.

One of representatives of block -symmetric cipher working in the gamma mode is GOST 28147-89.

One of regimes work GOST 28147-89 was the gamma mode, the structural diagram of which is presented in fig. 2.

Open data are divided into blocks of 64 bits each and encrypted in the gamma mode by the way household addition modulo 2 in the adder CM₅ with gamma cipher, which is generated in blocks of 64 bits. Initial filling drives N_1 and N_2 (synchronization S) are encrypted in simple mode replacements in the key memorable a 256-bit key is written to the device. In drives N_1 and N_2 64-bit binary is entered sequence $S = (S_1, S_2, \dots, S_{64})$, which serves initial filling for the next one generation of M blocks gamma cipher 64-bit open block encryption algorithm data in idle mode replacements includes 32 cycles.

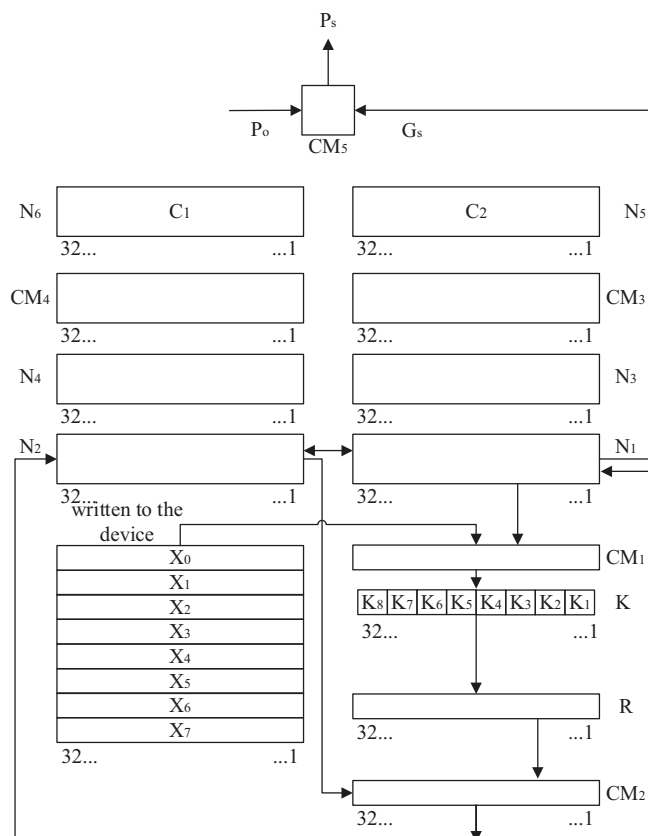


Fig. 2. GAMMIROW mode GOST 28147-89

The first cycle contains content storage unit N_1 is added modulo 2^{32} in the adder CM_1 with content storage X_0 , at the same time data in N_1 are preserved. The resulting result passes through the substitution block K , then what the transformed vector enters the register R , where is being done cyclical a shift of 11 positions in the direction of senior ranks. Result shift everyday is added modulo 2 in the adder.

CM₂ with 32-bit content storage N₂. Result from CM₂ is written in N₁, and before contents N₁ is transferred to N₂. After encryption data from N₁ and N₂ are written to 32-bit drives N₃ and N₄; the contents of N₁ – in N₃, and N₂ – in N₄.

The content of N4 is added modulo ($2^{32}-1$) in adder CM4 with the 32-bit constant C1 from drive N6, the result is recorded in N4. The contents of N3 are summed modulo 2^{32} in adder CM3 with a 32-bit constant C2 from drive N5, the result is recorded in N3. Next The content of N3 is transferred to N1, and N4 to N2, while data in N3 and N4 are saved. The contents of N1 and N2 are encrypted in simple mode replacements Final filling N1 and N2 later encryption forms the first 64-bit gamma block cipher, which everyday is added modulo 2 in the CM5 adder with the first 64-bit open block data in this scheme the substitution block K consisting of of 8 static blocks replacements.

In the table 1 are presented results conducted research speed block ciphers DES, AES, GOST-28147, which can to be used as a GPVP.

Table 1

Block ciphers	DES	AES	GOST 28147-89
Quantity operations processor on the block formed sequence	81/56	40/32	192/64
Length period	2^{56}	2^{128}	2^{256}

Main positive qualities of GPCP built on the basis of block ciphers are:

- resistance to various types of cryptanalysis;
- good statistical properties formable pseudorandom sequence.

To the main disadvantages of block encryption possible to attribute:

- insensitivity crypto schemes to fall out or insert an integer number of blocks;
- existence problems the last block is incomplete length.

To the main the advantages of GOST 28147-89 are possible to attribute:

- the futility of an all-out attack;
- simplicity both with hardware and software implementation and highly effective speed modern encryption computers.

One of the directions in development crypto-resistant generators there is methods based on unilateral functions.



The most important one-sided the function used in the construction of the GPSP is factorization whole numbers. The widely known GPSP based on this problem, is the *BBS generator*, proposed in 1986 by Lenore Blum, Manuel Blum, and Michael Shub. Calculation of PVP in the *BBS generator* is described expression:

$$x_n = x_{n-1}^2 \bmod N, \quad (2)$$

where $N = pq$ product two big simple p and q which can to be both are comparable to 3 modulo 4 (this guarantees that each the quadratic remainder has one square the root which also is quadratic remainder) and the largest general divider $\phi(\phi(p-1), \phi(q-1))$ must to be small.

At every step of the algorithm original data are coming out from taking either the parity bit or one least significant bit. Mainly the advantage of GPSP BBS is that it is available x_n optional calculate everything $n-1$ previous numbers, necessary just know the basics the state of the generator, that is x_0 and numbers p, q , and any element. The sequence of PSP is described expression:

$$x_n = x_0^{2^n \bmod ((p-1)(q-1))} \bmod N. \quad (3)$$

Next representatives of GPSP based on use one-sided functions are recognized generators Micali–Schnorr and RSA. Their stability is based on a theoretically complex calculation problem discrete of logarithms. Every element pseudorandom sequence in the Micali–Schnorr generator is described by the expression:

$$x_n = x_{n-1}^e \bmod N. \quad (4)$$

Exit of this GPSP is available r beat pseudorandom sequences, quantity which is calculated by the formula:

$$r = \lfloor \lg(pq) \rfloor + 1 - \left\lfloor \lg(pq) + 1 \left(1 - \frac{2}{e}\right) \right\rfloor, \quad (5)$$

where $e \in \begin{cases} 1 < e(p-1)(q-1) \\ \text{NOD}(e, (p-1)(q-1)) = 1, p, q - \text{prime numbers.} \\ 80e \leq \lfloor \lg(pq) \rfloor + 1 \end{cases}$

In the RSA GPSP, each element sequence is described expression (5), but unlike the Micali–Schnorr generator on the way out arrives least significant beat. Conducted studies showed that main shortcomings of the solution – based GPSP unilateral functions are:

- low speed formations pseudo-random sequences compared to GPSP based on other principles of formation;
- impossibility to provide formation sequences of the maximum period.

In the table 2 presents correspondence of key lengths to asymmetric and symmetric encryption algorithms.

Table 2

Compliance length keys asymmetric and symmetric encryption algorithms

Length for symmetrical algorithm	Length for asymmetrical algorithm
56 bits	384 bits
80 bits	768 bits
112 bits	1792 bits
128 bits	2304 bits

How follows from given in the table 2 data for formation of the period of PSP 2^{128} battles necessary to produce mathematical operations with dimensional numbers not less than 2304 bits, which in turn leads to significant use computational resources processor.

In the table 3 are presented results conducted research speed formation of one bit of provably stable PSP generators RSA, BBS, Micali–Schnorr for various degrees resistance.

Table 3

Rating speed provable-resistant GPSP

Proveably durable generators	Length period		
	2^{128}	2^{256}	2^{512}
Foundation degrees	2304 bits	14210 bits	81304 bits
RSA	1153656/1	265651872/1	49261255313/1
BBS	57456/1	2180945/1	71038737/1
Micali–Schnorr	561/1	28860/1	1307947/1

Main the advantage of GPSP based on the solution unilateral functions, is available mathematical provability difficulties cryptanalysis. The main ones disadvantages of such generators is computational complexity, and formation of PSP with period significantly less than the maximum.

For construction based on cryptographic schemes registers shift with inverses connections two methods are used generating pseudorandom sequences. The first method – use several in parallel working registers shifting and combining their exit any function complications. Such generators are called combined (Fig. 3).

Another alternative is generation pseudorandom sequence how nonlinear output functions the only one register shift, such a generator is called filter-generator (Fig. 4).

Generators built on this principle have maximum period $2^n - 1$ (where n quantity cells memory register shift). Mainly the disadvantage of such GPSP is complexity finding functions complications $f(x)$.

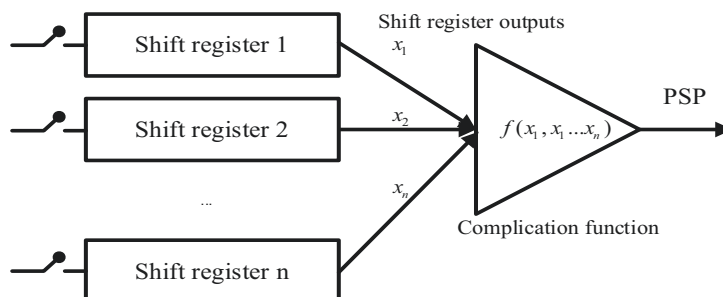


Fig. 3. Combined generator of pseudo-random numbers

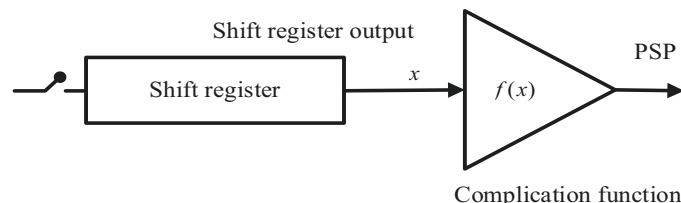


Fig. 4. Filter-generator

Discussion and conclusions

Conducted the study made it possible to comprehensively analyze existing approaches to the construction and application of the HPPSC with consideration their statistical characteristics, cryptographic durability and performance. received the results showed that the choice of a specific type of generator must be made be based on the specifics of informational tasks systems and requirements level protection.

Simple congruent generators and classic LFSR provide high speed formations sequences and ease of implementation, however their low cryptographic endurance does their unsuitable for direct use in protection systems information. They are can to be used as auxiliary tools during construction more difficult crypto-resistant structures.

Generators based on block and stream ciphers, such as AES, GOST 28147-89 and RC4, demonstrated optimal balance between safety and speed, that does their most suitable for a wide range of practical applications cryptographic applications. At the same time necessary to consider features realizations, including key management and reuse prevention key flows, that directly affects reliability systems.

Generators using one-sided functions (RSA, BBS, etc.), provide tall level protection at the expense of mathematically provable durability However, their application is limited scenarios, where priority is maximum security, and performance does not play key role, for example, in state and military information systems, as well as in infrastructures critical significance.

Prospects subsequent research associated with perfection existing methods generation random numbers and development new approaches that will be taken into account modern challenges (development quantum calculated and appearance more difficult methods cryptanalysis). Important direction is creation hybrid generators that will be to unite high cryptographic resistance with high at speed processing data.

Also promising is research algorithms, optimized for work in conditions limited computational resources (for example, in IoT devices), and study opportunities integration generators with hardware sources entropy to increase reliability system as a whole. In the long term perspective special attention must to be given development standards for new- generation cryptographic GPS systems, which will be able to provide protection given in the conditions post-quantum cryptography. This will allow create a reliable basis for building modern protection systems information, capable to resist future cyber threats (Barabash et al., 2023; Jensen, 2006; Gorbenko, & Ganzya, 2014; Starr, 1984; Yang et al., 2017, Berry, 1987).

Authors' contribution: Irada Rahimova – led the overall research design and methodology; conducted theoretical analysis of block cipher-based pseudo-random number generators (e.g., GOST 28147-89); and provided scientific supervision from the Azerbaijani team; Abdul oglu Abbasquliyev – performed experimental evaluations and computational testing of generator performance; collected and processed empirical data (e.g., speed metrics for DES, AES, GOST, BBS); and supported the comparative analysis and tabulation of results; Yelyzaveta Sevriukova – analyzed practical applicability of generators in real-world domains such as finance, government, and critical infrastructure; assessed cost-benefit trade-offs of security versus performance; and contributed to conclusions regarding deployment in strategic sectors; Tetiana Laptieva – conducted in-depth cryptographic analysis of generator security (e.g., LFSR, RC4, RSA, BBS); reviewed vulnerabilities and resistance to cryptanalysis (including cases like Dual_EC_DRBG); and formulated recommendations aligned with international standards (NIST, FIPS) and post-quantum readiness.

Sources of funding. This study did not receive any grant from a funding institution in the public, commercial, or non-commercial sectors.

References

- Barabash, O., Sobchuk, V., Musienko, A., Laptiev, O., Bohomia, V., & Kopytko, S. (2023). System Analysis and Method of Ensuring Functional Sustainability of the Information System of a Critical Infrastructure Object. In: System Analysis and Artificial Intelligence. Studies in Computational Intelligence, 1107, 177–192. Springer, Cham. https://doi.org/10.1007/978-3-031-37450-0_11
- Berry, J. (1987). Fundamental operations in computer-assisted map analysis. *International Journal of Geographical Information Systems*, 1, 119–136.
- Blum, L. (1986). Simple Unpredictable Pseudo-Random Number Generator. *SIAM Journal on Computing*, 15(2), 364–386.



- Blum, M., & Micali, S. (1984). How to generate cryptographically strong sequences of pseudo-random bits. *SIAM Journal on Computing*, 13(4), 850–864.
- Bobok, I. I., Kobozeva, A. A., Laptiev, O. A., Savchenko, V. A., Salii, A. G., & Kurtseitov, T. L. (2025). Method for Estimating the Bandwidth Capacity of a Steganographic Communication Channel. *Problems of the Regional Energetics*, 2, issue 66, 92–104. <http://doi.org/10.52254/1857-0070.2025.2-66.08>
- Gorbenko, Y., & Ganzya, R. (2014). Analysis of the stability of popular cryptosystems against quantum cryptanalysis based on Grover's algorithm. *Information protection*, 16(2), 106–112.
- Jensen, J. (2006). Remote Sensing and GIS Integration: Theories, Methods, and Applications. *Geocarto International*, 21(2), 5–16.
- Korolev, R. V. (2009). Study of periodic properties of the RC4 stream encryption algorithm. Collection of scientific papers of the Kharkiv Air Force University, 4, 86–89. http://nbuv.gov.ua/UJRN/ZKhUPS_2009_4_22
- Kuznetsov, A. (2007). Investigation of the statistical safety of generators of pseudorandom numbers. *Systems Processing Information*, 3(70), 79–82.
- Micali, S., & Sidney, R. (1995). A simple method for generating and sharing pseudorandom functions, with applications to Clipper-like key escrow systems. *Advances in Cryptology – CRYPTO'95 (LNCS 963)*, 185–196.
- Milevskiy, S., Korol, O., Hryshchuk, O., Laptieva, T., & Yevseiev, S. (2024). Crypto-code constructions on LDPC codes properties assessment. *Ukrainian Scientific Journal of Information Security*, 30(2), 313–323. <https://doi.org/10.18372/2225-5036.30.19244>
- Morawiecki, P., & Pieprzyk, J. (2013). Parallel authenticated encryption with the duplex construction. *Cryptology ePrint Archive*, Report 658.
- Prokushev, Y., Ponomarenko, S., & Ponomarenko, S. (2021). Modeling the processes of designing information security systems in state information systems. *Computational Nanotechnology*, 8(1), 26–37.
- Rukhin, A., & Soto, J. (2001). A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. NIST Special Publication 800–22. https://www.researchgate.net/publication/235090427_A_Statistical_Test_Suite_for_Random_and_Pseudorandom_Number_Generators_for_Cryptographic_Applications
- Starr, J. L. (1984). Geographic information systems: a question to ask before it's too late. *Machine Processing of Remotely Sensed Data with Special Emphasis on Thematic Mapper Data and Geographic Information Systems*, 194–199.
- Yang, C., Huang, C., Li, Z., Liu, K., & Hu, F. (2017). Big Data and Cloud Computing: Innovation Opportunities and Challenges for GIS. *Computers, Environment and Urban Systems*, 61, 93–102.

Отримано редакцією журналу / Received: 14.11.25

Прорецензовано / Revised: 17.11.25

Схвалено до друку / Accepted: 22.11.25

Ірада ПАГІМОВА¹, канд. техн. наук
ORCID ID: 0009-0006-1969-9976
e-mail: ika1402@icloud.com

Абдул огли Аббасгулієв ЕЮБ¹, асп.
ORCID ID: 0009-0003-7714-5494
e-mail: formemaybe23@yahoo.com

Єлизавета СЕВРЮКОВА², канд. екон. наук
ORCID ID: 0000-0002-0757-490X
e-mail: elizavetasevryukova@ukr.net

Тетяна ЛАПТЄВА², канд. техн. наук
ORCID ID: 0000-0002-5223-9078
e-mail: tetiana1986@ukr.net

¹Азербайджанський технічний університет, Баку, Азербайджан

²Національний технічний університет
"Харківський політехнічний інститут", Харків, Україна

ЕКСПЕРИМЕНТАЛЬНЕ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ КРИПТОГРАФІЧНИХ ГЕНЕРАТОРІВ ПСЕВДОВИПАДКОВИХ ЧИСЕЛ

Вступ. У цифрову епоху забезпечення інформаційної безпеки вимагає надійних криптографічних примітивів. Ключовим компонентом сучасних криптосистем є генератор псевдовипадкових чисел (ГПВЧ), який лежить в основі генерації ключів, цифрових підписів, векторів ініціалізації та інших механізмів безпеки. Стійкість усієї системи часто залежить від непередбачуваності та статистичної якості генерованих послідовностей. Мета цього дослідження – експериментально оцінити ефективність різних криптографічних ГПВЧ за трьома критеріями: статистична випадковість, криптостійкість та обчислювальна продуктивність.

Методи. Проведено порівняльний експериментальний аналіз різних класів криптографічних генераторів псевдовипадкових чисел. Досліджено генератори на основі регістрів зсуву з лінійним зворотним зв'язком (LFSR), конгруентні методи, поточкові та блокові шифри (напр., RC4, AES, ГОСТ 28147-89), а також односторонні функції (напр., BBS, RSA, Мікалі–Шнорр). Критерії оцінювання включали статистичні властивості (за допомогою тестів NIST SP 800-22), стійкість до криптоаналізу й обчислювальну ефективність, виміряну кількістю операцій процесора на один вихідний біт.

Результати. Експериментальні результати показали, що генератори на основі блокових і поточкових шифрів (напр., AES, ГОСТ 28147-89) забезпечують найкращий баланс між криптостійкістю та продуктивністю. Навпаки, генератори на основі односторонніх функцій (напр., BBS, RSA) мають доведену стійкість, але страждають від низької швидкості та недостатньої довжини періоду. Прості генератори, такі як LFSR і лінійні конгруентні генератори, характеризуються високою швидкістю та гарними статистичними властивостями, проте не мають криптостійкості, що робить їх непридатними для безпосереднього використання в застосунках із високими вимогами до безпеки.

Висновки. Вибір генератора псевдовипадкових чисел має ґрунтуватися на конкретних вимогах цільового застосунку – з урахуванням балансу між безпекою, швидкістю й обмеженнями реалізації. Для загальних криптографічних систем рекомендовано використовувати генератори на основі шифрів. У середовищах із підвищеними вимогами до доведеної стійкості, де продуктивність є вторинною, можуть застосовуватися теоретико-числові генератори. Подальші дослідження слід спрямувати на гібридні архітектури, готовність до постквантової епохи й інтеграцію з апаратними джерелами ентропії для підвищення надійності в умовах обмежених ресурсів та критично важливих інфраструктур.

Ключові слова: генератори псевдовипадкових чисел, криптографічна безпека, статистична випадковість, блокові шифри, поточкові шифри, односторонні функції, криптостійкість.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



Юрій ЩЕБЛАНІН, канд. техн. наук, ст. наук. співроб.

ORCID ID: 0000-0002-3231-6750

e-mail: shcheblanin.yurii@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Данило СКИДАН, студ.

ORCID ID: 0009-0001-2999-8180

e-mail: cassador2281@gmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ КОМБІНОВАНИХ КРИПТОГРАФІЧНИХ СХЕМ У ЗАХИСТІ WI-FI-З'ЄДНАНЬ

Вступ. Зростання кількості атак на бездротові мережі Wi-Fi створює додаткові виклики для організацій і користувачів, які застосовують ці мережі для передавання конфіденційних даних. Нині Wi-Fi є одним із найпоширеніших каналів зв'язку, що робить його привабливою цілью для кіберзлочинців. Вразливості у механізмах шифрування й автентифікації протоколів Wi-Fi Protected Access 2 (WPA2) і WPA3 можуть призвести до перехоплення даних, проведення атак на ключі та загроз конфіденційності. Тому підвищення рівня захисту Wi-Fi є критично важливим завданням, особливо для корпоративних мереж і пристроїв із високими вимогами до безпеки.

Методи. Для дослідження використано аналіз сучасних криптографічних алгоритмів Advanced Encryption Standard, 256-bit key (AES-256), ChaCha20, Elliptic Curve Cryptography (ECC) та їхніх комбінованих моделей, моделювання бездротового середовища на базі Kali Linux з утилітами `hostapd` та `FreeRADIUS`, а також тестування стійкості до атак, включаючи атаку повторного відтворення (replay), Key Reinstallation Attack (KRACK) і Pairwise Master Key Identifier (PMKID).

Результати. Експерименти показали, що комбінація AES-256 і ChaCha20 забезпечує високу пропускну здатність та швидкість автентифікації, але залишається вразливою до повторних атак за некоректної роботи з векторами ініціалізації. Гібридні схеми з ECC значно підвищують стійкість до атак на обмін ключами, однак збільшують час встановлення з'єднання між пристроями (handshake) і ресурсоспоживання. Найбільш захищеною виявилася тришарова модель ECC + AES-256 + ChaCha20, що забезпечує комплексний захист, але має підвищені вимоги до обчислювальних ресурсів.

Висновки. Одним із ключових напрямків підвищення безпеки Wi-Fi є впровадження комбінованого шифрування, яке поєднує симетричні й асиметричні алгоритми для досягнення максимальної стійкості до атак. Перспективним напрямком є оптимізація таких моделей для мобільних пристроїв та IoT, де важливий баланс між продуктивністю та рівнем захисту. Результати дослідження можуть бути використані для вдосконалення політик безпеки Wi-Fi та розроблення більш стійких реалізацій протоколів WPA2/WPA3.

Ключові слова: Wi-Fi; WPA2; WPA3; AES-256; ChaCha20; ECC; комбіноване шифрування.

Вступ

Зростання зловмисної активності у кіберпросторі ставить перед організаціями та власниками мереж нові виклики, пов'язані із захистом бездротових технологій. Однією з найбільш уразливих сфер є бездротові мережі Wi-Fi, які використовують для передавання конфіденційних даних та організації бізнес-процесів. Атаки на протоколи Wi-Fi можуть призвести до несанкціонованого доступу, перехоплення інформації або порушення стабільності мережевих сервісів.

Компрометація даних, переданих через Wi-Fi, може викликати серйозні наслідки – від фінансових втрат до втрати репутації та безперервності бізнесу. Протоколи WPA2 та WPA3, які є основою сучасного Wi-Fi захисту, хоч і використовують сучасні криптографічні алгоритми, мають уразливості, що роблять їх привабливою мішенню для атак типу Replay, KRACK чи PMKID. Це зумовлює необхідність розроблення нових підходів до шифрування, що забезпечують вищий рівень стійкості до атак.

Для підвищення захищеності Wi-Fi необхідно впроваджувати сучасні методи комбінованого шифрування, які поєднують симетричні (AES-256, ChaCha20) й асиметричні (ECC) алгоритми. Такий підхід дозволяє збалансувати продуктивність і рівень безпеки, знижуючи ризики перехоплення даних і компрометації ключів.

Ця робота спрямована на аналіз сучасних загроз для протоколів Wi-Fi та дослідження ефективності комбінованих схем шифрування.

Метою статті є дослідити стійкість Wi-Fi-протоколів до атак і проаналізувати ефективність використання комбінованих алгоритмів шифрування AES-256, ChaCha20 та ECC для підвищення рівня безпеки.

Методи

Проведено аналіз ефективності комбінованих криптографічних схем на основі AES-256, ChaCha20 та ECC у контексті безпеки протоколів WPA2/WPA3. На першому етапі виконано огляд криптографічних характеристик кожного з алгоритмів, зокрема і їхньої стійкості до атак, швидкодії та особливостей реалізації на різних апаратних платформах (x86-64, ARM). Особливу увагу приділено перевагам потокового шифрування ChaCha20 у мобільних системах та ефективності ECC для ініціалізації безпечних сесій.

Для моделювання Wi-Fi-середовища використано ноутбук Acer Nitro 5 із процесором Intel Core i5-8300H, 16 ГБ оперативної пам'яті, ОС Kali Linux, а також смартфон Xiaomi 14T на базі ARM. Реалізація точки доступу здійснювалась за допомогою `hostapd`, `dnsmasq` і `FreeRADIUS` із налаштуванням WPA2/WPA3 Enterprise. Для тестування застосовано бібліотеки OpenSSL, Libsodium, mbedTLS, а також утиліти `internet performance (iperf3)`, `tcp packet dump (tcpdump)`, `highly interactive top (htop)`.

Усі тести проводили в однакових умовах, щоб виключити вплив сторонніх факторів. Використовували однакові параметри бездротового інтерфейсу, однакові типи даних для передачі та фіксовані сценарії навантажень. Кожен



експеримент повторювався кілька разів, після чого брали середнє значення. Це дало змогу отримати стабільні результати й уникнути випадкових похибок.

Під комбінованим шифруванням у цьому дослідженні розуміють побудову шифрувального каскаду, що включає етап асиметричного обміну ключами (на базі ECC), після якого дані шифрують послідовно кількома симетричними алгоритмами. Такий підхід забезпечує гнучкість, пряму секретність (forward secrecy), криптостійкість та адаптацію до ресурсних можливостей клієнтів.

Реалізовано три схеми комбінування шифрів: симетричне подвійне (AES-256 + ChaCha20), гібридне з ECC, а також тришарове шифрування, що об'єднує всі три алгоритми. Дослідження охоплювало оцінку навантаження на систему, швидкість установлення з'єднання, а також стійкість до атак типу Man in the Middle (MitM), replay та обхід автентифікації. Отримані дані дозволяють зробити висновки щодо доцільності впровадження комбінованих схем у сучасних Wi-Fi-системах з урахуванням ресурсомісткості та рівня захисту (Koliass et al., 2016).

Результати

У дослідженні використано три основні компоненти комбінованого шифрування. Першим є еліптична криптографія ECC, яку застосовують на етапі ініціалізації з'єднання для створення надійного ключового обміну. Компактні ключі мають високу стійкість за малих значень ключів порівняно з Rivest Shamir Adleman (RSA), роблять ECC основою безпечної автентифікації, що дозволяє генерувати сесійні ключі для подальшого симетричного шифрування (Barker, 2016).

Другим компонентом виступає AES-256, який використовують як перший шар симетричного шифрування. Це блоковий алгоритм, визнаний стандартом у сфері інформаційної безпеки, який у поєднанні з апаратною підтримкою (напр., Intel AES-NI) забезпечує високу швидкодію та стабільність.

Завершує схему ChaCha20 – потоковий шифр, оптимізований для ARM – процесорів і мобільних середовищ. Його особливістю є стійкість до численних атак і низька залежність від апаратного прискорення, що робить його ефективним у сценаріях, де AES не має апаратної оптимізації. Використовуючи як основу компоненти, можна створити комбіновані схеми, що були протестовані.

1. ECC + AES-256 – класичний гібридний підхід, який ідеально працює на потужних клієнтах.

Обмін ключами реалізується за допомогою ECC:

$$K_{shared} = d_A \cdot Q_B = d_B \cdot Q_A. \quad (1)$$

де d_A, d_B – приватні ключі, а Q_A, Q_B – публічні ключі. Після генерації спільного секрету виконується симетричне шифрування даних AES-256 – GCM:

$$C = AES - GCM_{K_{AES}}(M), \quad (2)$$

де M – відкритий текст, K_{AES} – сесійний ключ, C – зашифрований текст.

2. ECC + ChaCha20 – модель, орієнтована на мобільні й енергоефективні пристрої.

Обмін ключами відбувається так само, як і в першій моделі:

$$K_{shared} = d_A \cdot Q_B = d_B \cdot Q_A. \quad (3)$$

Після цього дані шифрують потоковим шифром ChaCha20:

$$C = ChaCha20_{K_{ChaCha}}(M), \quad (4)$$

де K_{ChaCha} – сесійний ключ, отриманий із спільного секрету.

3. ECC + AES-256 + ChaCha20 – тришарова модель, яку застосовують у найчутливіших до безпеки сценаріях.

Обмін ключами відбувається аналогічно:

$$K_{shared} = d_A \cdot Q_B = d_B \cdot Q_A. \quad (5)$$

Із спільного секрету виводять два симетричні ключі K_{AES} та K_{ChaCha} :

$$K_{AES} = HKDF(K_{shared}, info1), \quad (6)$$

$$K_{ChaCha} = HKDF(K_{shared}, info2), \quad (7)$$

де HKDF – based Key Derivation Function (HKDF) криптографічна функція для генерації одного або кількох ключів із спільного секрету; $info1, info2$ – додаткові ідентифікаційні мітки в HKDF, які забезпечують виведення незалежних ключів з одного K_{shared} . А шифрування виконують послідовно двома алгоритмами:

$$C = ChaCha20_{K_{ChaCha}}(M)(AES - GCM_{K_{AES}}(M)). \quad (8)$$

Для забезпечення криптографічної незалежності між шарами в моделі ECC + AES-256 + ChaCha20 кожен алгоритм використовує окремий сесійний ключ і незалежний вектор ініціалізації, де генерується ключ із спільного секрету через функцію HKDF із різними контекстними мітками ($info1, info2$), що гарантує відсутність колізій і повторів.

Такий підхід унеможливує повторне використання одноразових значень і виключає міжалгоритмічну кореляцію.

На відміну від традиційної реалізації WPA2/WPA3, де основна увага приділяється швидкому обміну ключами з фіксованим шифром, комбіновані схеми дають змогу гнучко адаптувати рівень шифрування до характеристик конкретного пристрою або вимог до захищеності каналу (Gupta, & Gupta, 2002).

Основне дослідницьке питання, на яке ми прагнули відповісти: чи здатні комбіновані схеми забезпечити баланс між безпекою, швидкістю і ефективністю використання ресурсів у реальних Wi-Fi мережах? Для цього побудовано відповідну експериментальну платформу, тести якої охоплюють критично важливі аспекти: пропускну здатність, час встановлення з'єднання, ефективність проти атак і навантаження на ресурси.

Для перевірки ефективності комбінованих схем шифрування побудовано експериментальне середовище, яке максимально наближене до реальних умов функціонування Wi-Fi-мереж із відкритим доступом або корпоративною автентифікацією. Конфігурація середовища дозволяє здійснювати як точне вимірювання продуктивності каналів, так і моделювання атак і спостереження за поведінкою системи в динаміці.

Серверну частину реалізовано на базі Kali Linux, де налаштовано точку доступу з використанням hostapd, а також сервер RADIUS для автентифікації клієнтів за протоколом EAP-TLS. Для генерації сертифікатів, налаштування ECC-ключів і TLS-контекстів використовували пакет OpenSSL, що забезпечив високий рівень контролю над криптографічними параметрами. Конфігурація доповнена утилітою freeradius, яка дозволяє проводити



централізовану автентифікацію з використанням ECC-сертифікатів і підтримкою WPA2 – Enterprise (OpenSSL Project, 2024 ; OpenBSD Project, 2024).

На стороні клієнтів використано три типи пристроїв. Першим був ноутбук Acer Nitro 5 із процесором Intel Core i5-8300H (x86-64 – архітектура) та апаратною підтримкою AES через інструкції AES-NI. Другим – смартфон Xiaomi 14T, що представляє ARM – платформу без апаратного прискорення AES, але оптимізовану для мобільного трафіка. Третім пристроєм став настільний ПК із зовнішнім USB Wi-Fi-адаптером, який використовувався для тестування сумісності середнього рівня (FreeRADIUS Project, 2024).

Для дослідження обрано саме ці пристрої, оскільки вони відображають два типові сценарії використання: стаціонарний комп'ютер із достатньою потужністю та мобільний пристрій з обмеженими ресурсами. Це дозволило порівняти ефективність алгоритмів у реальних умовах, коли важливими є як швидкість роботи, так і економія енергії.

Взаємодію між вузлами організовано через пряме підключення до точки доступу. Усі клієнти проходять автентифікацію через ECC – сертифікати з подальшим шифруванням трафіка через реалізовані Python скрипти. Для цього створено псевдотрафіки різного типу, що включають передачу файлів (FTP), вебзапити (HTTP), та перевірку доступності мережі (ICMP ping). Це дало змогу оцінити не лише обчислювальні аспекти роботи шифрувальних алгоритмів, а й їхню поведінку в умовах реального навантаження (Shcheblanin et al., 2024).

Для вимірювання пропускну здатності використано iperf3. Аналіз протоколів і фіксацію процесу handshake здійснювали за допомогою tcpdump і Wireshark. Навантаження на процесор і пам'ять відстежували за допомогою htop. Додатково застосовано внутрішні лог-файли FreeRADIUS та hostapd, які дозволяли проаналізувати фазу автентифікації. Кожна конфігурація тестувалася на однаковому наборі дій, зокрема: встановлення з'єднання, передача 100 МБ файлу, 5-хвилинне відтворення відео з емульованим 1080p потоком і серія з 500 HTTP-запитів. Крім того, у тестах враховувалася стабільність шифрування у разі зміни параметрів каналу (напр., навмисне послаблення сигналу Wi-Fi для перевірки впливу на handshake та передачу даних) (Vanhoef, & Piessens, 2014).

Реалізація шифрування виконувалась за допомогою бібліотек OpenSSL та Libsodium, що дозволяло точно контролювати параметри ChaCha20 та AES-256 у режимі Galois/Counter Mode (GCM). У кожному випадку псевдотрафік спочатку шифрувався одним симетричним алгоритмом, а потім – іншим, з попереднім обміном ключами через ECC. Усі дані оброблялись у локальному середовищі без виходу в інтернет, що дало змогу уникнути впливу сторонніх чинників на результати. Додатково фіксувався процес встановлення з'єднання. Для цього використовували утиліти tcpdump та Wireshark, які дозволяли відстежити всі пакети під час handshake. Завдяки цьому з'явилася можливість побачити, як саме відбувається обмін ключами і чи не відбувається витікання важливих даних. Отже, середовище дало змогу протестувати комбіновані криптографічні моделі в умовах, максимально наближених до реального використання, охопивши сценарії роботи як на сучасних ноутбуках, так і на мобільних пристроях, а також пристроях без спеціального апаратного прискорення (Zou et al., 2016 ; Wireshark Foundation, 2024).

Для оцінювання ефективності комбінованого шифрування у Wi-Fi-комунікаціях розглянуто три моделі. До першої групи належить класична схема ECC + AES-256, яка оптимально підходить для платформ із підтримкою AES-інструкцій. Друга модель ECC + ChaCha20 орієнтована на мобільні пристрої та ARM-системи, де апаратне прискорення AES відсутнє. Найбільш захищеним варіантом виявилася тришарова модель ECC + AES-256 + ChaCha20, що передбачає послідовне використання двох симетричних алгоритмів і застосовується в умовах підвищених вимог до безпеки (WolfSSL Inc., 2023).

Попереднє вимірювання пропускну здатності для передачі зашифрованих даних обсягом 100 МБ дозволило зафіксувати відчутні відмінності у швидкодії між платформами. На системах архітектури x86-64, таких як ноутбук Acer Nitro 5 із підтримкою апаратного прискорення AES, найкращі результати демонструвала схема ECC + AES-256, тоді як ChaCha20-орієнтована конфігурація мала перевагу на мобільному пристрої Xiaomi 14T з архітектурою ARM. Тришарова схема (ECC + AES-256 + ChaCha20) виявилася найменш продуктивною з огляду на подвійне обчислювальне навантаження, однак залишалась стабільною навіть за умов інтенсивного трафіка (Bernstein, 2008).

Згідно з результатами, зображеними на рис. 1, вибір оптимальної схеми шифрування залежить від апаратної архітектури пристрою. За наявності підтримки апаратного прискорення AES доцільно використовувати схеми на основі AES-256, тоді як для ARM-пристроїв, що не мають відповідної оптимізації, ефективнішою є схема з ChaCha20. Водночас тришарова модель, попри зниження пропускну здатності, забезпечила стабільну роботу в умовах довготривалого навантаження.

На ноутбуці з підтримкою AES-інструкцій алгоритм AES-256 працював дуже швидко і майже не навантажував процесор. На смартфоні, де такої підтримки не було, AES показував гірші результати, тоді як ChaCha20 працював стабільно і більш ефективно. Це ще раз підтвердило, що вибір алгоритму залежить від конкретного обладнання.

Під час 30-хвилинного експерименту з безперервною передачею даних, що моделює реальний трафік (включаючи вебзапити, відеопотік і сервісні пакети), жодна з трьох конфігурацій не продемонструвала критичного зростання затримок, втрат пакетів чи нестабільності з'єднання. Це дає змогу зробити висновок про придатність комбінованих схем шифрування для захищених Wi-Fi-з'єднань тривалої дії. Під час довготривалих експериментів не спостерігалося помилок чи повторних спроб підключення. З'єднання залишалось стабільним навіть за погіршення сигналу. Це показує, що додаткове навантаження від комбінованого шифрування не заважає роботі мережі у реальних умовах.

У підсумку, хоча використання кількох шарів шифрування збільшує обчислювальне навантаження, за умови коректного підбору схеми відповідно до особливостей платформи його вплив на продуктивність може бути мінімізований.

У ході дослідження проведено аналіз стійкості комбінованих криптографічних схем до типових атак на Wi-Fi-з'єднання, зокрема MitM атак, повторних атак, перехоплення ключів, словникових атак і спроб зламу на етапі handshake. Тестування відбувалося у контрольованому середовищі, де точка доступу працювала під управлінням Kali Linux із конфігурацією WPA2 – Enterprise (EAP-TLS), а клієнти представляли три різні платформи: ноутбук Acer Nitro 5, смартфон Xiaomi 14T та ПК із зовнішнім Wi-Fi-адаптером. На всіх етапах експериментів фіксували спроби атак як під час встановлення каналу зв'язку, так і в разі активної передачі трафіка.

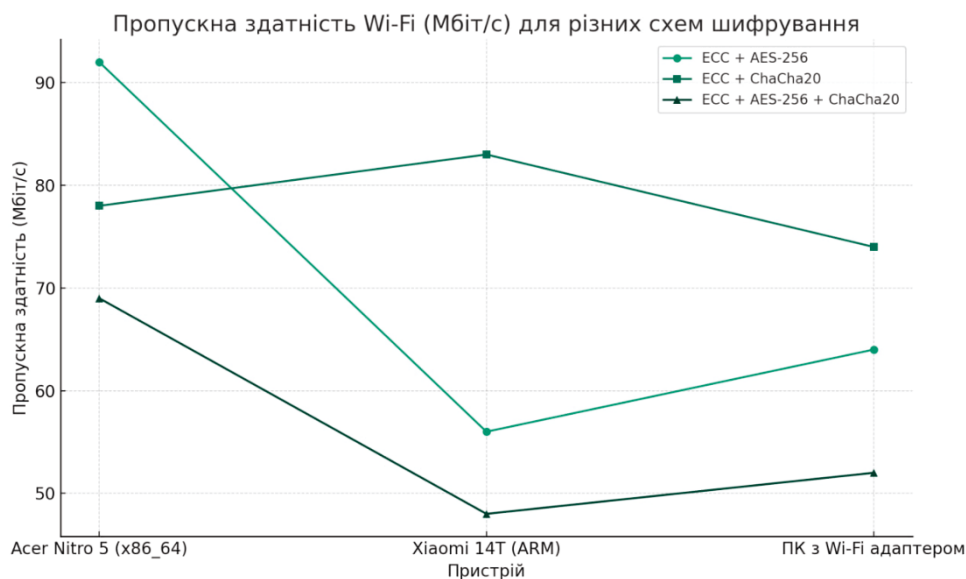


Рис. 1. Порівняння пропускної здатності трьох методів шифрувань

Під час моделювання атаки типу MitM із використанням інструментів Ettercap та Aircrack-ng встановлено, що стандартна схема WPA2 із симетричним шифруванням AES-256 є вразливою за умови слабкого пароля. Зловмисник, ініціюючи підроблений handshake, міг перехопити дані та відновити сесію після успішного підбору ключа. Натомість у схемах, де для автентифікації застосовано ECC (EAP-TLS), навіть за умов симульованого handshake дешифрування трафіка виявилось неможливим, оскільки сесійні ключі шифрувалися відкритим ключем клієнта, а їхнє розкриття без приватного ключа неможливе (Thankappan, Pous, & Garrigues, 2022).

Під час тестування герлау атак проводилося перехоплення зашифрованих пакетів і спроби їх повторного надсилання до точки доступу. Для схеми ECC + AES-256 (GCM) повторна передача блокувалася завдяки використанню унікальних одноразових значень, хоч ефективність залежала від реалізації механізму перевірки на клієнтському боці. У конфігурації ECC + ChaCha20 додаткова валідація Message Authentication Code (MAC) забезпечила повне відхилення підроблених пакетів незалежно від типу пристрою, без помітного збільшення навантаження (Crypto StackExchange, 2023).

Узагальнені експериментальні результати ефективності комбінованих моделей показують, що модель ECC + AES-256 на архітектурі x86-64 забезпечила найвищу пропускну здатність (118 Мбіт/с) завдяки інструкціям AES-NI, але мала помірну стійкість (6 % успішних атак). Конфігурація ECC + ChaCha20 на ARM-пристрої продемонструвала нижчий час handshake (195 мс) і стабільну роботу без апаратного прискорення. Найвищу криптостійкість (2 % успішних атак) показала тришарова модель ECC + AES-256 + ChaCha20, хоч і з деяким зниженням пропускної здатності (91 Мбіт/с), наглядні результати подано у табл. 1.

Таблиця 1

Експериментальні результати ефективності комбінованих криптографічних моделей на основі ECC, AES-256 та ChaCha20

Модель шифрування	Архітектура пристрою	Середня пропускна здатність, Мбіт/с	Час handshake, мс	Частка успішних атак, %	Примітки
ECC + AES-256	x86-64 (Acer Nitro 5)	118	210	6	Висока продуктивність при AES-NI; помірна стійкість
ECC + ChaCha20	ARM (Xiaomi 14T)	104	195	4	Краща енергоефективність, стабільна без AES-NI
ECC + AES-256 + ChaCha20 (тришарова)	ARM/x86-64 (змішане середовище)	91	280	2	Максимальна стійкість, підвищене навантаження

Оцінювання стійкості до словникових атак під час handshake показало, що у схемі WPA2 – PSK з AES-256 зловмисник міг досягти до 62 % успішних підборів при використанні стандартного словника. Натомість у моделях з ECC (EAP – TLS) автентифікація здійснювалася на основі сертифікатів, що виключало передачу пароля мережею та робило словникові атаки неефективними. Додаткове впровадження ECC забезпечило властивість прямої секретності: навіть у разі компрометації тимчасового ключа попередні сесії залишалися захищеними, на відміну від традиційного WPA2 – Personal (Gupta, & Gupta, 2002).

Комбіновані схеми на основі ECC із AES-256 або ChaCha20 демонструють суттєве підвищення рівня безпеки порівняно з класичним AES, зберігаючи баланс між продуктивністю та криптостійкістю. Порівняння отриманих ре-



зультатів із наявними дослідженнями підтверджує концептуальну наступність і практичне розширення відомих підходів до побудови гібридних криптосистем. Схема ECC + ChaCha20 показала високу стабільність на ARM-платформах при збереженні прийнятної швидкодії, що узгоджується з результатами досліджень ефективності поточкових шифрів сім'ї Salsa20 для програмних реалізацій без апаратної підтримки (Bernstein, 2008).

Поєднання ChaCha20 з AES-256 у тришаровій моделі забезпечує багаторівневий захист від помилок повторного використання одноразових значень, характерних для одношарових реалізацій AES – GCM, і виходить за межі класичних підходів (Dworkin, 2007; Langley, Chang, & Strombergson, 2016). У запропонованій моделі використовується незалежна генерація ключів для кожного шару, що усуває міжалгоритмічну залежність і знижує ризик кореляцій атак (Barker, 2016).

Результати підтверджують переваги еліптичних кривих у середовищах з обмеженими ресурсами та демонструють практичну інтеграцію ECC з двома симетричними алгоритмами, на відміну від схем із спільним ключем для кількох фаз шифрування (Gupta, & Gupta, 2002; Rizk, & Alkady, 2015). У тришаровій схемі використано незалежні ключі для кожного шару, що унеможливило взаємний компроміс між фазами шифрування.

Отже, запропонована модель демонструє більш збалансоване поєднання продуктивності, енергоефективності та криптостійкості, підтверджує ефективність поточкових шифрів ChaCha20 у мобільних середовищах і розширює традиційну концепцію гібридного шифрування, поєднуючи принципи ECC, AES-256 і ChaCha20 в адаптивній архітектурі для бездротових мереж нового покоління.

Дискусія і висновки

Проведене дослідження ефективності комбінованих методів шифрування у Wi-Fi-мережах підтвердило їхню перевагу над одношаровими схемами в умовах сучасних загроз. Об'єднання симетричних та асиметричних алгоритмів дало змогу суттєво підвищити криптостійкість протоколу без критичного впливу на продуктивність, за умови правильного врахування особливостей апаратного середовища.

У результаті експериментального аналізу встановлено, що використання ECC забезпечує надійний обмін ключами з властивістю прямої секретності, яка критично важлива в контексті довготривалих з'єднань та ймовірних компрометацій у майбутньому. Ця особливість надає схемам з ECC значну перевагу над класичними PSK реалізаціями, де витік одного ключа може розкрити всю історію переданих даних.

Симетричні алгоритми також проявили себе відповідно до апаратних можливостей. AES-256 показав високу ефективність на платформах x86 із підтримкою інструкцій AES-NI, зокрема на ноутбучі Acer Nitro 5. Зауважимо, що ChaCha20 виявився оптимальним вибором для мобільних ARM-пристроїв, таких як Xiaomi 14T, де апаратна підтримка AES відсутня або обмежена.

Використання комбінованих схем, зокрема й ECC у поєднанні з AES-256 або ChaCha20, дозволило істотно знизити ймовірність успішного здійснення атак replay, MitM, а також повторного використання сесійних ключів. Особливо помітним було зменшення ефективності replay атак, що вказує на доцільність застосування кількох незалежних рівнів перевірки цілісності й автентичності переданих пакетів.

Щодо продуктивності, то зафіксовано, що потрібне шифрування (ECC + AES-256 + ChaCha20) призводить до зростання навантаження на процесор, проте загальні обчислювальні витрати залишалися в межах прийнятного навіть за інтенсивного трафіка, включно з передачею відеопотоків у реальному часі. Отримані результати можна застосувати на практиці. У великих офісах доцільно використовувати ECC + AES-256, адже сучасні комп'ютери добре справляються із цим навантаженням. Для мобільних пристроїв краще підходить ECC + ChaCha20, оскільки ця схема економить ресурси батареї. Якщо йдеться про критично важливі системи, то найвищий рівень захисту дає тришарова модель, навіть якщо вона сповільнює роботу. Для пристроїв з обмеженими ресурсами найефективнішою виявилася конфігурація ECC + ChaCha20, яка поєднує високу безпеку з енергоефективністю.

Окрему увагу приділено часу встановлення з'єднання. Додавання ECC і каскадного симетричного шифрування помітно подовжувало фазу handshake. Проте сумарна затримка не перевищувала кількох сотень мілісекунд і не мала відчутного впливу на якість з'єднання в умовах типового використання.

Отже, отримані результати підтверджують перспективність комбінованих методів шифрування для захисту Wi-Fi- мереж. Їхня реалізація дозволяє адаптувати криптографічну конфігурацію під конкретне апаратне середовище та характер трафіка. Подальший розвиток концепції може включати автоматизовані механізми динамічного вибору шифрувальних схем, що реагуватимуть на поточні навантаження, тип даних і контекст середовища. У довгостроковій перспективі такі підходи можуть стати основою нових стандартів захисту бездротового зв'язку, що поєднують високу стійкість до атак із широкою сумісністю та прийнятною продуктивністю.

У майбутньому подібні комбіновані схеми можуть стати частиною нових стандартів Wi-Fi. Це дасть змогу підвищити захист без зміни звичних сценаріїв використання для користувачів, оскільки всі додаткові операції виконуватимуться автоматично на рівні протоколу.

Внесок авторів: Данило Скидан – концептуалізація, методологія, аналіз сучасних криптографічних алгоритмів, проведення експериментів у Kali Linux, інтерпретація результатів; Юрій Щербанін – огляд літератури, підготовка рукопису, редагування тексту, участь у розробленні методології дослідження.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Barker, E. (2016). *Recommendation for Key Management, Part 1: General (NIST Special Publication 800-57 Part 1, Rev. 4)*. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-57pt1r4>
- Bernstein, D. J. (2008). The Salsa20 family of stream ciphers. In *New Stream Cipher Designs* (pp. 84–97). Springer. https://doi.org/10.1007/978-3-540-68351-3_8
- Crypto StackExchange. (2023). *For AES–GCM, why do protocols not use the nonce to prevent replay attacks?* <https://crypto.stackexchange.com/questions/114926>



- Dworkin, M. (2007). *Recommendation for block cipher modes of operation: Galois/Counter Mode (GCM) and GMAC*. NIST Special Publication 800–38D. <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38d.pdf>
- FreeRADIUS Project (2024). FreeRADIUS Technical Documentation. <https://wiki.freeradius.org>
- Gupta, V., & Gupta, S. (2002). Elliptic Curve Cryptography (ECC) and its applications. In *Proceedings of the 6th IEEE International Symposium on Consumer Electronics* (pp. 209–212). Association for Computing Machinery New York NY United States. <https://doi.org/10.1145/570681.570691>
- Langley, A., Chang, W.-T., & Strombergson, J. (2016). *ChaCha20 and Poly1305 for IETF protocols.RFC 8439*. <https://datatracker.ietf.org/doc/html/rfc8439>.
- Kolias, C., Kambourakis, G., Stavrou, A., & Gritzalis, S. (2016). Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. *IEEE Communications Surveys & Tutorials*, 18(1), 184–208. <https://doi.org/10.1109/COMST.2015.2402161>
- OpenBSD Project. (2024). *Libsodium – A modern and easy-to-use crypto library*. <https://doc.libsodium.org>
- OpenSSL Project. (2024). *OpenSSL: Cryptography and SSL/TLS Toolkit Documentation*. <https://www.openssl.org/docs/>
- Rizk, R. & Alkady, Y. (2015). Two-phase hybrid cryptography algorithm for wireless sensor networks. *Journal of Electrical Systems and Information Technology*, 2(3), 296–313. <https://doi.org/10.1016/j.jesit.2015.11.005>
- Shcheblanin, Y., Zahynei, A., Kurchenko, O., Melnyk, I., & Smirnov, S. (2024). Analysis of identification and access management models in the context of fog computing. In *CPITS-II 2024: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II* (Vol. 3826, pp. 288–293). SEUR-WS.org. <https://ceur-ws.org/Vol-3826/short19.pdf>
- Thankappan, M., Rifà-Pous, H., & Garrigues, C. (2022). Multi-Channel Man-in-the-Middle Attacks Against Protected Wi-Fi Networks: A State of the Art Review. *Expert Systems with Applications*, 210, 118401. <https://doi.org/10.1016/j.eswa.2022.118401>
- Vanhoe, M., & Piessens, F. (2014). Advanced Wi-Fi attacks using commodity hardware. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC)* (pp. 256–265). <https://doi.org/10.1145/2664243.2664260>
- Wireshark Foundation (2024). *Wireshark User Guide*. https://www.wireshark.org/docs/wsug_html_chunked/
- WolfSSL Inc. (2023). *What is the difference between AES and ECC?* <https://www.wolfssl.com/what-is-the-difference-between-aes-and-ecc>
- Zou, Y., Zhu, J., Wang, X., & Hanzo, L. (2016). A survey on wireless security: Technical challenges, recent advances, and future trends. *Proceedings of the IEEE*, 104(9), 1727–1765. <https://doi.org/10.1109/JPROC.2016.2558521>

References

- Barker, E. (2016). *Recommendation for Key Management, Part 1: General (NIST Special Publication 800-57 Part 1, Rev. 4)*. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-57pt1r4>
- Bernstein, D. J. (2008). The Salsa20 family of stream ciphers. In *New Stream Cipher Designs* (pp. 84–97). Springer. https://doi.org/10.1007/978-3-540-68351-3_8
- Crypto StackExchange. (2023). *For AES-GCM, why do protocols not use the nonce to prevent replay attacks?* <https://crypto.stackexchange.com/questions/114926>
- Dworkin, M. (2007). *Recommendation for block cipher modes of operation: Galois/Counter Mode (GCM) and GMAC*. NIST Special Publication 800–38D. <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38d.pdf>
- FreeRADIUS Project (2024). FreeRADIUS Technical Documentation. <https://wiki.freeradius.org>
- Gupta, V., & Gupta, S. (2002). Elliptic Curve Cryptography (ECC) and its applications. In *Proceedings of the 6th IEEE International Symposium on Consumer Electronics* (pp. 209–212). Association for Computing Machinery New York NY United States. <https://doi.org/10.1145/570681.570691>
- Langley, A., Chang, W.-T., & Strombergson, J. (2016). *ChaCha20 and Poly1305 for IETF protocols.RFC 8439*. <https://datatracker.ietf.org/doc/html/rfc8439>.
- Kolias, C., Kambourakis, G., Stavrou, A., & Gritzalis, S. (2016). Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. *IEEE Communications Surveys & Tutorials*, 18(1), 184–208. <https://doi.org/10.1109/COMST.2015.2402161>
- OpenBSD Project. (2024). *Libsodium – A modern and easy-to-use crypto library*. <https://doc.libsodium.org>
- OpenSSL Project. (2024). *OpenSSL: Cryptography and SSL/TLS Toolkit Documentation*. <https://www.openssl.org/docs/>
- Rizk, R. & Alkady, Y. (2015). Two-phase hybrid cryptography algorithm for wireless sensor networks. *Journal of Electrical Systems and Information Technology*, 2(3), 296–313. <https://doi.org/10.1016/j.jesit.2015.11.005>
- Shcheblanin, Y., Zahynei, A., Kurchenko, O., Melnyk, I., & Smirnov, S. (2024). Analysis of identification and access management models in the context of fog computing. In *CPITS-II 2024: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II* (Vol. 3826, pp. 288–293). SEUR-WS.org. <https://ceur-ws.org/Vol-3826/short19.pdf>
- Thankappan, M., Rifà-Pous, H., & Garrigues, C. (2022). Multi-Channel Man-in-the-Middle Attacks Against Protected Wi-Fi Networks: A State of the Art Review. *Expert Systems with Applications*, 210, 118401. <https://doi.org/10.1016/j.eswa.2022.118401>
- Vanhoe, M., & Piessens, F. (2014). Advanced Wi-Fi attacks using commodity hardware. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC)* (pp. 256–265). <https://doi.org/10.1145/2664243.2664260>
- Wireshark Foundation (2024). *Wireshark User Guide*. https://www.wireshark.org/docs/wsug_html_chunked/
- WolfSSL Inc. (2023). *What is the difference between AES and ECC?* <https://www.wolfssl.com/what-is-the-difference-between-aes-and-ecc>
- Zou, Y., Zhu, J., Wang, X., & Hanzo, L. (2016). A survey on wireless security: Technical challenges, recent advances, and future trends. *Proceedings of the IEEE*, 104(9), 1727–1765. <https://doi.org/10.1109/JPROC.2016.2558521>

Отримано редакцією журналу / Received: 04.11.25

Прорецензовано / Revised: 08.11.25

Схвалено до друку / Accepted: 21.12.25

Yurii SHCHEBLANIN, PhD (Engin.), Senior Researcher
ORCID ID: 0000-0002-3231-6750
e-mail: shcheblanin.yurii@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Danylo SKYDAN, Student
ORCID ID: 0009-0001-2999-8180
e-mail: cassador2281@gmail.com
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

EVALUATING THE EFFECTIVENESS OF COMBINED CRYPTOGRAPHIC SCHEMES IN SECURING WI-FI CONNECTIONS

Background. The growing number of attacks targeting wireless Wi-Fi networks poses significant challenges for organizations and users relying on these networks to transmit confidential data. Today, Wi-Fi is one of the most widely used communication channels, making it an attractive target for cybercriminals. Vulnerabilities in the encryption and authentication mechanisms of Wi-Fi Protected Access 2 (WPA2) and WPA3 protocols can lead to data interception, key – related attacks, and serious privacy risks. Therefore, enhancing the security of Wi-Fi networks is a critical task, particularly for corporate environments and devices with high security requirements.

Methods. This study involved analyzing modern cryptographic algorithms Advanced Encryption Standard, 256 – bit key (AES-256), ChaCha20, Elliptic Curve Cryptography (ECC) and their combined implementations, modeling a wireless



environment based on Kali Linux using hostapd and FreeRADIUS, and testing resistance to attacks such as Replay, Key Reinstallation Attack (KRACK), and PMKID.

Results. *Experiments demonstrated that the AES-256 and ChaCha20 combination ensures high throughput and fast authentication performance but remains vulnerable to replay attacks if initialization vectors are improperly managed. Hybrid schemes involving ECC significantly improve resistance to key exchange attacks but increase handshake time and resource consumption. The most secure configuration proved to be the three – layer ECC + AES-256 + ChaCha20 model, which provides comprehensive protection at the cost of increased computational demands.*

Conclusions. *One of the key directions for improving Wi-Fi security is the adoption of combined encryption approaches that integrate both symmetric and asymmetric algorithms to achieve maximum resilience against modern attacks. A promising area of further development is the optimization of such schemes for mobile and IoT devices, where a balance between performance and security is essential. The findings of this study may contribute to enhancing Wi-Fi security policies and developing more robust WPA2/WPA3 protocol implementations.*

Key words: *Wi-Fi; WPA2; WPA3; AES-256; ChaCha20; ECC; combined encryption.*

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



Іван ОПІРСЬКИЙ, д-р техн. наук, проф.

ORCID ID: 0000-0002-8461-8996

e-mail: ivan.r.opirskyi@lpnu.ua

Національний університет "Львівська політехніка", Львів, Україна

Соломія ЯВОРІВСЬКА, студ.

ORCID ID: 0009-0002-0018-6231

e-mail: solomiia.yavorivska.kb.2023@lpnu.ua

Національний університет "Львівська політехніка", Львів, Україна

ДОСЛІДЖЕННЯ ОСОБЛИВОСТЕЙ ВИКОРИСТАННЯ SOAR ЯК ГОЛОВНОГО ІНСТРУМЕНТУ ДЛЯ ПРИСКОРЕННЯ РЕАГУВАННЯ НА ІНЦИДЕНТИ

Вступ. Інтенсивний розвиток кіберзагроз і зростання їхньої складності в сучасному цифровому середовищі створюють критичні виклики для центрів операційної безпеки (SOC), які потребують нових підходів до автоматизації та оркестрації процесів кібербезпеки. Однією з ключових проблем сучасних SOC є необхідність оброблення величезних обсягів інформації про загрози, швидкого реагування на інциденти та координації дій між різними системами безпеки за обмежених людських ресурсів. У контексті дефіциту кваліфікованих фахівців із кібербезпеки особливої актуальності набувають платформи SOAR (Security Orchestration, Automation, and Response), які можуть автоматизувати рутинні процеси без втрати якості аналізу. Застосування інтегрованих механізмів оркестрації, зокрема й автоматизація реагування на фішинг, координація між SIEM-системами й управління інцидентами, формують основу сучасного підходу до ефективного функціонування SOC.

Методи. Проведено систематичний аналіз сучасних методів впровадження SOAR-систем у центрах операційної безпеки, зокрема і в контексті інтеграції з існуючими SIEM-платформами. Використано методи порівняльного аналізу, моделювання сценаріїв загроз, практичного тестування автоматизованих робочих процесів та оцінювання ефективності різних стратегій впровадження. Особливу увагу приділено аналізу ключових показників ефективності, таких як скорочення середнього часу виявлення загроз (MTTD) та реагування на них (MTTR), автоматизації процесів класифікації інцидентів та інтеграції з джерелами розвідки загроз. Для кожного з підходів оцінено складність впровадження, вимоги до кваліфікації персоналу, економічну ефективність і сумісність з архітектурою існуючих SOC.

Результати. Аналіз показав, що найефективнішим підходом до оптимізації SOC є поєднання SOAR-платформ із традиційними SIEM-системами у формі взаємодоповнюючої екосистеми. Наприклад, використання автоматизованих playbooks для реагування на фішинг суттєво скорочує час реакції з годин до хвилин, тоді як інтеграція з системами Threat Intelligence унеможливорює пропуск критичних індикаторів компрометації. Доведено, що комбінування автоматизації рутинних завдань із людським контролем стратегічних рішень значно підвищує продуктивність аналітиків SOC і точність виявлення справжніх загроз. Водночас впровадження SOAR вимагає ретельного планування та поетапного підходу, особливо для організацій з обмеженими ресурсами. Визначено оптимальні стратегії впровадження залежно від розміру організації, рівня зрілості процесів кібербезпеки та наявності кваліфікованого персоналу.

Висновки. Оптимізація роботи центрів операційної безпеки через впровадження SOAR-технологій вимагає системного і комплексного підходу, який враховує як технічні, так і організаційні аспекти. Комбінування автоматизації з людським досвідом дозволяє досягти високого рівня ефективності без втрати контролю над критичними процесами безпеки. Результати дослідження можуть бути використані для побудови інтегрованих систем кібербезпеки з урахуванням обмежених ресурсів малих і середніх організацій. Крім техніко-економічних аспектів, підкреслюється важливість поетапного впровадження, навчання персоналу та постійного вдосконалення автоматизованих сценаріїв відповідно до еволюції кіберзагроз. Системне впровадження SOAR-підходів забезпечить підвищення ефективності SOC навіть в умовах зростання кількості та складності кіберінцидентів.

Ключові слова: SOAR, SOC, кібербезпека, автоматизація реагування на інциденти, SIEM, оркестрування загроз, MTTD, MTTR, інтеграція інструментів безпеки, документація інцидентів, аналітика загроз, оптимізація SOC.

Вступ

Розслідування в центрах операційної безпеки (SOC) втомлюють, адже вони залежать від ручних дій для запитів до різних джерел даних, накладання відповідних журналів, кореляції даних в інформацію, а потім документування результатів у системі квитків. Інструменти оркестрації, автоматизації та реагування на загрози безпеки (SOAR) – це порівняно нова технологія, яка обіцяє за відповідного налаштування збирати, фільтрувати і відображати необхідну різноманітну інформацію, автоматизувати багато звичайних завдань, що без потреби забирають час аналітиків SOC, полегшити співпрацю SOC і підвищити ефективність та злагоженість роботи SOC.

Операційні центри безпеки – це частина команди інформаційних технологій підприємства, яка відповідає за захист організації від кіберзагроз. Сучасні SOC використовують широкий спектр датчиків, які генерують величезну кількість журналів аудиту (напр., журнали операційної системи з робочих хостів і серверів, мережеві потоки з роутерів, комутатор) канали розвідки (напр., звіти про сканування вразливостей, інформація про шкідливе програмне забезпечення) і сповіщення (напр., від антивірусних програм (AV) і систем виявлення мережових вторгнень (NIDS) на кінцевих точках). Більшість SOC оснащено системою управління інформацією та подіями безпеки (SIEM), яка об'єднує ці різноманітні потоки даних на централізованій платформі та надає конфігуровані інформаційні панелі й інтерфейси запитів для моніторингу даних, а також для відсіювання, сортування та пошуку артефактів з усієї мережі. Попри те, що системи SIEM забезпечують централізований збір та оброблення подій інформаційної безпеки, значне навантаження, пов'язане з аналізом великих обсягів даних із різноманітних джерел, усе ще покладається на фахівців центру моніторингу безпеки. Зокрема, аналітики й інженери повинні виконувати сортування значної кількості журналів подій, визначати, які з них потребують подальшого аналізу, встановлювати взаємозв'язки між різними сповіщеннями та приймати рішення щодо подальших дій у відповідь на виявлені інциденти.



Інструменти оркестрування, автоматизації та реагування на загрози безпеці (SOAR) представляють собою відносно нову технологію, яка намагається автоматизувати багато поширених ручних завдань із метою підвищення ефективності й узгодженості, і, як наслідок, безпеки. Якщо подивитись дані, які дає компанія Gartner (Гартнер) про SOAR, у своєму звіті про роботу, інструмент SOAR дозволяє "організаціям отримувати вхідні дані з різних джерел і застосовувати робочі процеси, узгоджені з процесами і процедурами, підвищуючи ефективність і узгодженість роботи людей і процесів". Основна відмінність між інструментом SOAR і SIEM полягає в тому, що інструменти SOAR мають конфігуровані робочі процеси або сценарії, які направляють аналітиків і автоматизують багато дій із розслідування і реагування на інциденти, в той час як основна функція інструменту SIEM полягає в широкому зборі даних і запитах. Існує певне дублювання, наприклад, обидва інструменти вимагають і полегшують отримання різноманітних даних і забезпечують налаштовану візуалізацію даних, зокрема й інформаційні панелі в реальному часі.

Інструменти SOAR мають на меті зробити операції з безпеки ефективнішими за рахунок автоматизації загальних завдань, визначення пріоритетів, кореляції і доповнення з допомогою вхідних сповіщень (напр., із розвідувальними даними про загрози), полегшення комунікації і співпраці між і всередині SOC, а також більш послідовними за рахунок надання уніфікованого, зручного інтерфейсу, за допомогою якого аналітики взаємодіють із розширеними даними і сповіщеннями, а також за рахунок використання сценаріїв для стандартизації процедури виявлення загрози.

Одними з ключових проблем у впровадженні SOAR є складність інтеграції, дефіцит кваліфікованих фахівців і нечітке розмежування функцій між SOAR і SIEM. По-перше, адаптація SOAR вимагає глибокого планування, ретельного налаштування сценаріїв автоматизації та постійного обслуговування, що ускладнює його впровадження, особливо для малих команд з обмеженими ресурсами. По-друге, незважаючи на те, що автоматизація через SOAR стає критичною через дефіцит спеціалістів із кібербезпеки, залишається недостатньо дослідженим, як максимізувати ефективність цих інструментів в умовах обмеженого персоналу, щоб вони не перетворилися на додаткове навантаження замість справжнього полегшення роботи SOC. По-третє, через нечітке розуміння ролей SOAR і SIEM організації часто стикаються з неефективним використанням обох систем, що призводить або до дублювання функцій, або до марних витрат, коли одна система використовується не повною мірою через неправильну інтеграцію з іншою. Ці фактори разом ускладнюють процес упровадження SOAR і можуть знизити його потенційну користь, якщо не будуть враховані на етапі планування.

Проблема дослідження полягає в тому, щоб зрозуміти, як найкраще впроваджувати SOAR-системи в SOC, особливо для невеликих компаній з обмеженим бюджетом і нестачею фахівців. Головні труднощі – це складне налаштування системи, нестача кваліфікованих працівників для її обслуговування, а також відсутність зрозумілих способів оцінити, чи дійсно SOAR покращує безпеку. Дослідження має допомогти знайти золоту середину між автоматизацією і роботою людей, а також запропонувати простий план упровадження SOAR для малого бізнесу.

Метою цього дослідження є аналіз особливостей упровадження SOAR-систем у центрах операційної безпеки для оптимізації реагування на кіберінциденти, з акцентом на інтеграцію із SIEM-платформами й автоматизацію рутинних процесів. З огляду на зростання складності загроз, таких як фішинг, підозрілі автентифікації та вразливості сертифікатів, необхідно розробити рекомендації щодо поетапного впровадження SOAR для організацій з обмеженими ресурсами, з метою скорочення MTTD та MTTR.

Для досягнення цієї мети передбачено розв'язання таких завдань:

- Проаналізувати сучасні випадки застосування SOAR-систем у кібербезпеці.
- Дослідити роль SOAR у центрі операційної безпеки (SOC) та їх вплив на ефективність роботи.
- Визначити оптимальні умови для впровадження SOAR-систем, з урахуванням обмежених ресурсів.
- Виявити ключові критерії ефективності SOAR у контексті зростаючої кількості кіберзагроз.
- Розробити рекомендації щодо впровадження SOAR-систем для підвищення продуктивності SOC.

Огляд літератури. У сучасному світі кібербезпеки, де атаки стають дедалі складнішими й частішими, ключовим фактором ефективного захисту стає поєднання автоматизації, штучного інтелекту та людського контролю (Seeton). Автоматизація реагування на інциденти (рішення SOAR), (2025).

Suchismita Chatterjee у роботі "Using SIEM and SOAR for Real-Time Cybersecurity Operations" (Chatterjee, 2020) докладно пояснює, як об'єднання систем SIEM і SOAR дає змогу значно скоротити час виявлення загроз і реагування на них. SIEM (система управління інформацією та подіями безпеки) відповідає за збір і аналіз логів з усіх джерел у мережі, а SOAR (платформа оркестрації та автоматизації реагування) дає змогу автоматично виконувати дії у відповідь – наприклад, блокувати підозрілу IP-адресу чи запускати сценарій реагування на фішинг. Такий підхід зменшує навантаження на аналітиків і дозволяє команді безпеки зосередитися на складніших загрозах, які потребують людського втручання.

Із прикладом упровадження автоматизації саме в телекомунікаційній галузі можна ознайомитися у дослідженні Noshi, & Blaser (2024). Автор акцентує увагу на тому, що телекомунікаційні мережі є одними з найбільш складних і водночас вразливих до атак – зокрема, через використання застарілих протоколів і мільйони IoT-пристроїв. Він показує, як завдяки впровадженню SOAR-сценаріїв можна майже миттєво виявити й нейтралізувати загрозу, не чекаючи втручання людини. Наприклад, якщо у системі помічено аномальну активність у певному регіоні – SOAR може автоматично повідомити службу безпеки, виконати попередній аналіз, та, у разі підтвердження, обмежити доступ до вразливого сегменту мережі.

Дуже практичний погляд на автоматизацію реагування подає українська компанія Seeton у своєму матеріалі про впровадження SOAR (Sharma, & Spunda, 2025). Тут описано реальні кейси: як система автоматично обробляє фішингові листи, блокує заражені хости або взаємодіє із зовнішніми Threat Intelligence платформами. Важливо, що автори не лише описують технічні переваги, а й акцентують увагу на необхідності правильного проєктування процесів – створення playbooks, визначення критеріїв спрацювання, навчання персоналу. Автоматизація без розуміння процесів не працює – це головний висновок із цього джерела.



Однак навіть найсучасніші технології залишаються під контролем людини, і саме людський фактор часто стає причиною помилок.

У книзі "Behind Human Error" (Woods et al., 1994) розглянуто, чому фахівці припускаються помилок під час роботи з технічними системами. У випадку SOC аналітики часто стикаються з перевантаженням інформацією, втомою, стресом. Це призводить до пропущених інцидентів або неправильного реагування. Автори стверджують, що автоматизація повинна не замінити людину, а допомагати їй – наприклад, через фільтрацію тривог, нагадування, підказки. У такий спосіб технологія стає партнером, а не конкурентом фахівця.

У дослідженні "Testing SOAR tools in use" (Bridges et al., 2025), проведеному великою групою авторів, здійснено глибоку перевірку реальної ефективності інструментів SOAR у практичному середовищі. Основна увага приділялася тому, як саме такі платформи працюють у різних сценаріях, наскільки вони сумісні з іншими системами і чи відповідають вони очікуванням користувачів у реальному часі. Автори виявили, що попри гучні обіцянки автоматизації, деякі SOAR-рішення вимагають значного налаштування, добре описаних сценаріїв реагування та глибокого розуміння контексту загроз. Також у дослідженні зазначено, що без якісних джерел даних (SIEM, Threat Intelligence, внутрішні логи) платформи SOAR не можуть ефективно виконувати свої функції. Отже, інструменти SOAR не є "чарівною таблеткою", а скоріше потужним доповненням до налагодженого процесу кіберзахисту. Особливу увагу автори приділяють необхідності тестування платформ перед їх повноцінним впровадженням, адже не всі вони однаково добре працюють у кожній організації.

Усі ці джерела разом формують цілісну картину: кібербезпека нині – це не лише про інструменти, а про їхнє грамотне поєднання з людським досвідом. Автоматизація, якщо вона правильно налаштована, дає змогу не лише зекономити час, а і зменшити кількість помилок, підвищити точність рішень і зробити команди безпеки ефективнішими. Проте людина все ще залишається в центрі цих процесів, а технології повинні лише посилювати її можливості, а не повністю їх замінити.

Методи

В роботі використано методи аналізу, які включають систематичне картування наукових публікацій із SOAR-технологій та операційної безпеки, теоретичне моделювання процесів автоматизації реагування на кіберінциденти, порівняльний аналіз ефективності SIEM- та SOAR-систем, емпіричне тестування сценаріїв виявлення фішингу та підозрілої активності користувачів, а також оцінювання критеріїв успішності впровадження платформ оркестрації в SOC з обмеженими ресурсами, що дозволило досягти комплексного розуміння принципів оптимізації центрів операційної безпеки через інтеграцію автоматизованих систем реагування на кіберзагрози.

Результати

У SOC посадові функції швидко змінюються. Традиційно SOC розподіляють посади за рівнями, починаючи з рівня досвіду й обов'язків, необхідних для кожної з них (рис. 1) (Brewer, 2019). Робочі місця 1-го рівня є початковими, а робочі місця 3-го рівня потребують всебічних знань.

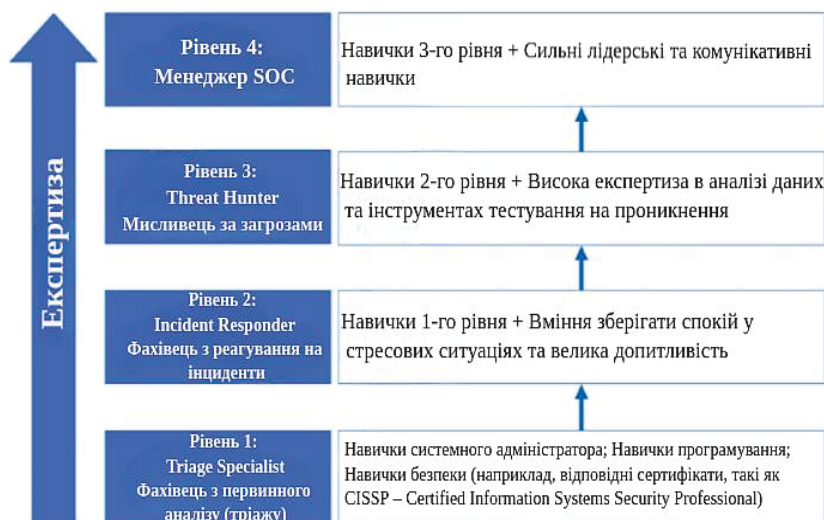


Рис. 1. Ролі в сучасному SOC

Аналітик Tier 1 – часто відомий як аналітик із кібербезпеки або спеціаліст із кібероперацій, відстежує вхідні повідомлення, перевіряє, чи дійсно стався інцидент. Якщо стався справжній інцидент, то, за необхідності, пересилає повідомлення на Tier 2.

Фахівець, що реагує на інциденти Tier 2, – це особа, що відповідає за проведення поглиблених розслідувань інцидентів і надання рекомендацій щодо усунення наслідків інциденту або втручання.

Tier 3 Threat Hunter – це фахівці, що мають поглиблені знання про мережі, кінцеві точки, аналітику загроз і шкідливе програмне забезпечення зворотного інжинірингу. Вони є експертами у відстеженні процесів шкідливого програмного забезпечення, щоб оцінити його вплив і визначити, як його видалити. Вони беруть активну участь у пошуку нових загроз і розгортанні програмного забезпечення для їх виявлення. Мисливці за загрозами шукають кіберзагрози, які ще не були ідентифіковані, але присутні в мережі.

Менеджер SOC – це особа, яка відповідає за ресурси SOC і виступає в ролі контактної особи з великою компанією або клієнтом (Basta et al., 2024).



Всі з перелічених вище ролей разом утворюють повноцінний моніторинг і підтримку 24/7, ручну автоматизацію. Проте в кожного кваліфікованого фахівця є обов'язки, які він виконує кожен день, і процедури, які є невід'ємною частиною процесу в SOC-команді (Vielberth et al., 2020).

Типовий день аналітика з кібербезпеки починається з перевірки черг оповіщень про загрози. Для внесення оповіщень до списку, який аналітик має дослідити, зазвичай використовують систему тікетів. Оскільки програмне забезпечення, яке генерує сповіщення, може спричиняти хибні тривоги, роль аналітика з кібербезпеки може містити підтвердження того, що тривога відображає справжній інцидент безпеки. Коли інцидент підтверджено, його можна перенаправити слідчим або іншим співробітникам служби безпеки для вжиття заходів. Якщо ні, то повідомлення може бути проігнороване як хибна тривога.

Якщо аналітик із кібербезпеки не може опрацювати повідомлення, воно буде перенаправлене до спеціаліста з реагування на інциденти рівня 2 для подальшого розгляду та виправлення. Якщо спеціаліст із реагування на інциденти не може розв'язати проблему, повідомлення переадресовують співробітникам рівня 3, які мають значні знання та досвід боротьби із загрозами.

Саме в щоденних процесах аналітика, йому на допомогу приходять сучасні технології SOC, які значно підвищують ефективність усіх етапів роботи (Akhtar, & Ali, 2021). Такі рішення, як SIEM та SOAR, не лише автоматизують рутинні завдання, а й забезпечують комплексний підхід до виявлення, аналізу та нейтралізації кіберзагроз. Розглянемо докладніше, які технологічні рішення нині формують основу сучасних центрів операційної безпеки.

SIEM розшифровує дані, отримані від брандмауерів, мережевого обладнання, систем виявлення вторгнень та інших пристроїв (Collins, 2025).

SIEM і SOAR (оркестрування, автоматизація та реагування на загрози безпеки) часто поєднуються, оскільки їхні можливості доповнюють одна одну.

Обидві технології використовуються великими командами з операцій безпеки (SecOps) для покращення SOC. Подібно до того, як SIEM-системи агрегують, порівнюють та інтерпретують попередження, платформи SOAR виконують те саме. З іншого боку, технологія SOAR робить ще один крок уперед, поєднуючи розвідку загроз і автоматизацію робочих процесів розслідування інцидентів і реагування на них на основі сценаріїв, створених командою безпеки (Vielberth et al., 2020).

Платформи безпеки SOAR:

- Збирають дані про тривоги з кожного компонента системи.
 - Надають інструменти, які дозволяють досліджувати, оцінювати та розслідувати інциденти.
 - Наголошують на інтеграції як засобі автоматизації складних робочих процесів реагування на інциденти, що забезпечує швидше реагування, та адаптивних стратегій захисту.
 - Містять заздалегідь визначені сценарії, які забезпечують автоматичне реагування на конкретні загрози.
- Плейлисти можуть бути ініційовані автоматично на основі заздалегідь визначених правил або можуть бути запущені персоналом служби безпеки.

SOAR зосереджується на інструментах автоматизації та інтеграції процесів SOC, автоматизує багато ручних процедур, таких як оповіщення про загрозу безпеці розслідування, вимагаючи участі людини, лише коли це необхідно. Це звільняє людей, які працюють у кібербезпеці, щоб зосередитися на нагальніших питаннях (Rahman, 2024).

SOAR-платформи йдуть на крок попереду, трансформуючи традиційні процеси SOC. Щоб повноцінно оцінити потенціал SOAR, важливо розібратися в його архітектурі та принципах роботи.

Іншими словами, SOAR – це категорія рішень, які спрямовані на вдосконалення шляхом інтеграції, автоматизації та уможливлення оркестрування численних інструментів і процесів безпеки. Платформи SOAR зазвичай складаються з трьох основних компонентів: координації, контролю та реагування – трьох фундаментальних методів управління. Водночас, оркестрування розуміється як взаємодія і співпраця різних рішень і додатків для забезпечення взаємної безпеки. Автоматизація в цьому сенсі передбачає використання сторонніх продуктів, які можуть впоратися із цими основними завданнями самостійно та набагато швидше без залучення команд безпеки, тим самим полегшуючи їхнє навантаження і скорочуючи час реагування (TTR-time to response). Вона також охоплює заходи, які вживають у відповідь на інциденти безпеки, які частково складаються з ідентифікації, аналізу, забруднення, видалення загрози та відновлення повністю нормального функціонала хоста, сервера тощо (BlueVoyant, 2025).

Отже, впровадження рішень SOAR дає багато переваг. Перша перевага пов'язана з підвищенням продуктивності, оскільки закупівля вхідних матеріалів у великих кількостях означає, що менше матеріалів буде витрачено на одне й те саме. Тобто покупка ліцензії для використання SOAR-систем тут повністю себе виправдовує. Таким чином, платформи SOAR значною мірою звільняють від ручної праці й оптимізують роботу кібербезпеки, автоматизуючи повторювані дії та належно контролюючи процеси. Це приводить до швидшого реагування на інциденти та кращого контролю загроз.

Ще однією перевагою є покращена видимість та обізнаність про ситуацію за допомогою рішень SOAR (Manda, 2021). Завдяки добре обробленим даним із різних джерел організації можуть використовувати SOAR для виявлення й ефективної протидії загрозам. Комплексний підхід до управління безпекою є життєво важливим для виявлення і запобігання багатогранним загрозам, які можуть виникати на різних рівнях організації.

Застосування рішень SOAR допомагає також покращити документування інцидентів безпеки та подальше звітування у разі невідповідності вимогам стандарту. Дуже багато компаній нині все більше замислюються над тим, щоб відповідати вимогам загальновідомих стандартів. Це підвищує конкурентоздатність на ринку і довіру як до компанії у сфері кібербезпеки. Контроль, який характеризує роботу більшості SOAR-платформ, гарантує документування всіх дій, вжитих у кожній точці інциденту, що дозволяє легко продемонструвати відповідність вимогам.

У такий спосіб можна забезпечити процедури аудиту на відповідність нормативним вимогам, а також у разі виникнення інциденту і для підвищення безпеки процесів. Можливості, які з'являються в впровадженні рішень SOAR, також включають масштабованість і гнучкість, що дуже важливо, особливо для великих організацій, що постійно зростають. Враховуючи постійні зміни та збільшення складності загроз безпеці, платформи SOAR можуть

рости разом із потребами організації в безпеці. Це допомагає організаціям впроваджувати надійні заходи для захисту від загроз, що постійно змінюються.

Очевидно, що рішення SOAR є життєво важливим інструментом дотримання вимог, оскільки вони детально описують можливості для покращення безпеки операцій, зміцнюють здатність організації стримувати або запобігати інцидентам, а також забезпечують прозорість загальної ситуації з безпекою. Ці переваги допомагають організаціям відповідати вимогам і залишатися в безпеці, а також забезпечують глибоке і всебічне покращення стану безпеки організації. Використовуючи рішення SOAR, будь-яка організація може гарантувати практичніший підхід до запобігання й усунення загроз безпеці, який захистить конкретні дані та забезпечить дотримання законодавства.

Сучасні кіберзагрози вимагають не лише проактивного виявлення, але й миттєвого реагування (Sironi et al., 2010). Далі розглянемо реальні кейси й імітаційні сценарії, які демонструють, як SOAR-системи допомагають SOC ефективно протистояти атакам – від фішингу до складних експлойтів. На прикладах із різних галузей (від роздрібної торгівлі до фінансового сектору) подивимось, як інтеграція оркестрації та автоматизації прискорює усунення загроз і покращує відповідність міжнародним стандартам безпеки, таким як PCI DSS.

А. Розпізнавання й усунення загроз у реальному часі (Hafa, 2025). Ще одна зафіксована атака – фішингова, в якій постраждала компанія роздрібною торгівлі; метою злоумисників було отримати облікові дані для входу в систему. В результаті, рішення SOAR змогло зупинити розсилку листів в режимі реального часу за допомогою сайту, який був підключений до системи захисту електронної пошти та продуктів розвідки загроз компанії. Згідно з аналізом операцій безпеки платформи SOAR і вимог до обізнаності, робочий процес ініціюється без затримок, щоб установити небезпечні повідомлення для ізоляції, повідомити відділ IT-безпеки, а потім заблокувати домен відправника. Крім того, після цієї кібератаки компанія була змушена ініціювати загальнокорпоративний антифішинговий тренінг, який охопив увесь персонал, щоб уникнути подібних кібератак у майбутньому. Крім зручного реагування на спробу фішингу та зупинки загрози в режимі реального часу допомагає покращити безпеку організації, не лише розв'язавши проблеми, виявлені під час експерименту.

Б. Сценарій моделювання: Clinical Experiences in ASERT (Hafa, 2025). Payment Card Industry Data Security Standard (PCI DSS) – стандарт безпеки даних індустрії платіжних карток, розроблений Радою зі стандартів безпеки індустрії платіжних карток, заснованою міжнародними платіжними системами Visa, MasterCard, American Express, JCB і Discover.

ASERT – це елітна група інженерів і дослідників NETSCOUT, що спеціалізується на інформаційній безпеці.

Наприклад, у цій роботі слід застосувати моделювання під час упровадження SOAR у середовищі PCI DSS, де середня фірма, що займається електронною комерцією, впроваджує платформу SOAR у свою систему безпеки бізнесу. Основним напрямом діяльності фірми є впровадження стандартів PCI DSS, де важливим є встановлення моніторингу в режимі реального часу та реагування на інциденти. Платформу SOAR створено для інтеграції із SIEM, брандмауерами, IDS і системами управління вразливостями. У другому сценарії під час симуляції тестер проникнення атакує компанію, імітуючи вразливість у вебдодатку. У випадку, коли платформа SOAR виявляє аномалію, як тільки вона інтегрована із системами SIEM та IDS, вона автоматично ініціює низку дій, таких як блокування IP-адреси злоумисника, створення детального звіту про подію на сайті та надсилання тривожного повідомлення співробітникам служби безпеки на електронну пошту та мобільні номери.

Крім того, на основі виникнення CVE, яка була використана, SOAR запускає попередній сценарій, який застосовує CVE до всіх систем, щоб гарантувати, що інші не можуть бути використані подібно протягом наступних днів або тижнів. Для взаємодії з клієнтом застосовують весь процес, який був виконаний через платформу SOAR, що в подальшому може допомогти відстежувати нові виявлені інциденти безпеки, пов'язані з регулюванням PCI DSS, реагувати на них і документувати подальші дії. Оскільки рішення SOAR покращують функцію дотримання вимог PCI DSS завдяки автоматизації та оркеструванню діяльності учасників команди реагування на інциденти, вони показують, як використання платформ SOAR може привести до скорочення часу реагування; вони також допомагають у точному та належному реагуванні на інциденти безпеки і навіть слугують цінним документом для досягнення відповідності вимогам.

Розглянуті симуляції та реальні кейси наочно демонструють, як SOAR-системи трансформують реагування на кіберзагрози. Однак їхній потенціал виходить далеко за межі окремих інцидентів. У повсякденній практиці SOC платформи SOAR довели свою ефективність у розв'язанні багатьох типових завдань кібербезпеки.

1. Виявлення фішингу та реагування на нього (рис. 2).



Рис. 2. Робота SOAR при фішингу



Фішингові електронні листи є найпоширенішими, легко здійсненими атаками на безпеку (Manda, 2021; Vielberth et al., 2020). Протидія фішингу ускладнюється такими причинами:

- Велика кількість інцидентів: SOC може мати справу із сотнями або тисячами сповіщень під час загрози безпеці. Це призводить до втоми від алертів, й аналітики можуть пропустити життєво важливі індикатори поведінки загрози.
- Атаки з різних джерел: розслідування фішингу вимагає використання кількох платформ безпеки.
- Ручні операції: ручний аналіз фішингових атак вимагає годин праці.

SOAR допомагає ідентифікувати фішингові атаки за допомогою таких кроків:

Автоматичне сортування та визначення пріоритетів: системи SOAR автоматично сортують фішингові сповіщення, сортуючи потенційні загрози за ступенем серйозності, зменшуючи ручне навантаження на команди SOC.

Вилучення й аналіз показників: інструменти SOAR витягують ключові індикатори компрометації (KIC), такі як підозрілі URL-адреси, вкладення або заголовки електронних листів, і звіряють їх із відомими базами даних загроз.

Аналіз хибнопозитивних спрацьовувань: використовуючи машинне навчання та попередньо визначені правила, SOAR мінімізує помилкові спрацьовування, ідентифікуючи легітимні електронні листи, які можуть викликати фішингові сповіщення, підвищуючи точність виявлення.

2. Виявлення підозрілих входів користувачів за IP-адресами (рис. 3).



Рис. 3. Робота SOAR у разі виявлення підозрілого логіна

Проблема: підозрілі входи користувачів важко ідентифікувати з таких причин:

- Поведінку користувача важко проаналізувати;
- Компанії мають багато джерел географічного розташування та хмарних користувачів для відстеження;
- Моніторинг підозрілих входів здійснюється вручну та повторюється.

SOAR допомагає виявляти підозрілі входи користувачів за допомогою таких кроків:

Ідентифікувати поведінкові аномалії: SOAR збирає дані для входу із зовнішніх інструментів (напр., SIEM або систем автентифікації), щоб виявити незвичайні аномалії входу.

Збагатити інформацію про користувача: SOAR отримує детальну інформацію про обліковий запис користувача (напр., минулу історію входів, роль, дозволи), щоб допомогти оцінити, чи є поведінка користувача легітимною.

Збагатити IP-розвідки: SOAR порівнює IP-адреси, пов'язані зі спробою входу, з базами даних розвідки загроз, щоб перевірити їх на наявність відомих зловмисних IP-адрес.

Визначити статус загрози: на основі поведінки користувача та даних IP-розвідки SOAR автоматично вирішує, чи є спроба входу зловмисною або безпечною.

Автоматична відповідь – закрити або діяти:

- немає загрози: якщо не виявлено зловмисної активності, SOAR може автоматично закрити інциденти безпеки;
- виявлена загроза: якщо зловмисну активність підтверджено, SOAR виконує автоматичні дії, такі як блокування зловмисної IP-адреси.

3. Відстеження терміну дії SSL-сертифіката (рис. 4).



Рис. 4. Робота SOAR у разі ескалювання або закриття випадку з SSL-сертифікатом



Проблема: коли користувач намагається отримати доступ до сайту з протермінованим сертифікатом безпеки, сучасні веббраузери відображають попереджувальні повідомлення (D3Security. SOC automation examples, 2025). Це може підірвати довіру користувачів і відбити бажання відвідувачів повертатися на сайт, що потенційно може призвести до втрати трафіка і довіри до організації.

SOAR допомагає відстежувати термін дії SSL-сертифікатів:

Перевірка статусу сертифіката: SOAR безперервно відстежує SSL-сертифікати завдяки інтеграції з інструментами управління сертифікатами для перевірки їхнього статусу закінчення терміну дії.

Інформування користувача: коли термін дії сертифіката наближається до завершення, SOAR автоматично надсилає сповіщення відповідним користувачам, попереджаючи їх про наближення дати закінчення терміну дії.

Якщо термін дії сертифіката не закінчився: SOAR повторно перевіряє статус і автоматично закриває інцидент, якщо сертифікат усе ще дійсний, сповіщення буде відхилено.

Якщо термін дії закінчився: якщо термін дії сертифіката закінчився, SOAR ескалує інцидент, надсилаючи сповіщення користувачеві й іншим відповідним зацікавленим сторонам в організації.

Ескалація та комунікація: SOAR автоматично надсилає користувачеві докладні електронні листи та залучає інших важливих членів команди, забезпечуючи швидку обізнаність і дії.

Дві ключові технології, SIEM і SOAR, допомагають покращити безпеку, але мають різні переваги й обмеження.

У табл. 1 порівнюються основні критерії ефективності SIEM і SOAR: від виявлення загроз до автоматизації процесів. Це допоможе зрозуміти, яка система краще підходить для конкретних потреб організації.

Таблиця 1

Переваги використання SIEM- і SOAR-систем

Критерій	SIEM	SOAR
Виявлення загроз	Покращує виявлення складних загроз завдяки кореляції подій із різних джерел, виявляє шаблони атак	Покращує виявлення через інтеграцію з інструментами безпеки та каналами розвідки, забезпечує повний аналіз загроз
Реагування на інциденти	Надає сповіщення в реальному часі й автоматизовані дії, але основне навантаження лежить на аналітиках	Автоматизує аналіз, локалізацію та ліквідацію загроз, значно прискорює реагування
Швидкість виявлення (MTTD)	Допомагає зменшити час виявлення, але залежить від налаштувань і ручного аналізу	Значно скорочує MTTR завдяки автоматизації та інтегрованим сценаріям
Швидкість реагування (MTTR)	Частково автоматизує процеси, але MTTR залишається високим через необхідність ручного втручання	Різко знижує MTTR завдяки автоматизованим робочим процесам (блокування IP-адрес, ізоляція загроз тощо)
Видимість загроз	Забезпечує централізований аналіз даних, але може бути обмеженим у комплексному розумінні загроз	Надає глибшу видимість завдяки інтеграції з різними системами та розвідкою загроз
Автоматизація	Автоматизує збір логів, аналіз і звітність, але не замінює ручну роботу	Автоматизує більшість рутинних завдань (розслідування, реагування, документування)
Ефективність SOC	Підвищує ефективність завдяки централізації даних, але не усуває втому від оповіщень	Оптимізує роботу SOC, звільняючи аналітиків для стратегічних завдань (полювання на загрози, управління вразливостями)
Складність впровадження	Відносно простий у налаштуванні, але вимагає постійного оновлення правил кореляції	Вимагає глибокого планування, чітких сценаріїв і кваліфікованого персоналу
Підтримка малих організацій	Підходить для малих і середніх компаній, оскільки може масштабуватися	Може бути надмірним для малих команд через складність і вартість

Проаналізувавши цю таблицю, можемо зазначити, що SIEM є ефективним інструментом для моніторингу, виявлення загроз і часткової автоматизації, але його можливості обмежені в плані швидкого реагування.

SOAR значно перевершує SIEM в автоматизації реагування, скороченні MTTR/MTTD і покращенні продуктивності SOC, але вимагає більших ресурсів для впровадження.

Оптимальний підхід, який ми запропонували б для компаній або організацій:

- SIEM – база для моніторингу й аналізу.
- SOAR – надбудова для автоматизації та прискорення реагування.
- Інтеграція SIEM + SOAR дає максимальний ефект, поєднуючи сильні сторони обох технологій (рис. 5).

Джерела даних (Data Sources):

- Промислові системи: SCADA, PLC, датчики, DCG-системи (критична інфраструктура).
- IT-інфраструктура: мережеві пристрої, сервери, бази даних, додатки, хмарні сервіси.
- Інструменти безпеки: фаєрволи, системи виявлення вторгнень (IDS).

Оброблення даних:

- Дані з усіх джерел надходять до SEM (Security Event Management) для аналізу та кореляції подій.
- SOAR автоматизує реагування на загрози.

Функціональні модулі:

- Моніторинг відповідності (Compliance Monitoring & Reporting): перевірка на відповідність стандартам (GDPR, ISO 27001 тощо).
- Розвідка загроз (Threat Intelligence): аналіз IOC (індикаторів компрометації), пошук нових загроз.



- Реагування на інциденти (Incident Response): автоматизовані сценарії (блокування IP, відключення акаунтів тощо).
- Роль аналітика (Security Analyst):
- Отримує структуровані алерти, досліджує інциденти, приймає рішення на основі даних від SOAR/SIEM.

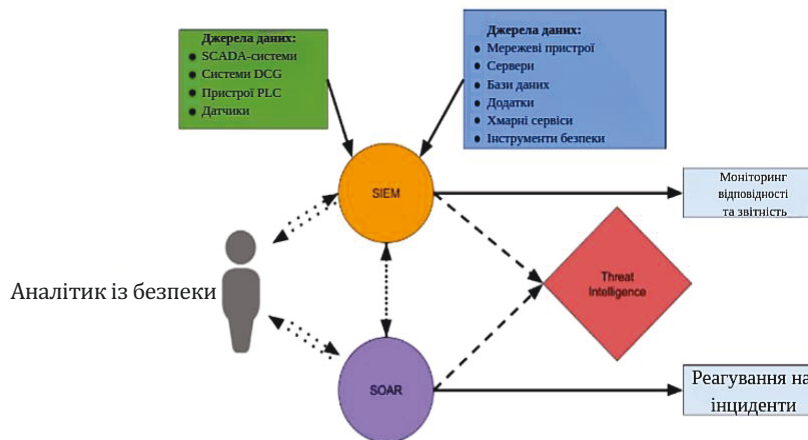


Рис. 5. Взаємодія між SOAR- і SIEM-системами

Система автоматично обробляє тисячі сповіщень про можливі загрози, фільтрує їх і залишає аналітикам лише найважливіші випадки. Працює як помічник, що бере на себе всі нудні завдання. Вона збирає інформацію з усіх кутків мережі – від офісних комп'ютерів до промислових систем і хмарних сервісів. Тепер не треба перемикатися між десятком різних програм. Коли знаходить реальну загрозу, система може миттєво її зупинити – наприклад, заблокувати зловмисника або вимкнути зламаний обліковий запис, навіть якщо це сталося серед ночі. Сама готує всі необхідні звіти для перевірок, що дуже зручно, коли потрібно довести, що все відповідає стандартам безпеки. Система постійно оновлює свої знання про нові віруси та методи хакерів, тому із часом стає тільки розумнішою. Чи то звичайні офісні комп'ютери, чи то складні промислові системи – система SOAR впорається з моніторингом усього. Головний плюс – вона дає можливість невеликій команді фахівців ефективно контролювати велику мережу, швидше знаходити проблеми і вчасно на них реагувати. Це як потужний підсилювач для кібербезпеківців.

Хоча система SOAR є дуже повною і функціональною платформою, важливо зазначити, що координація дуже важлива, оскільки SOAR не є повністю автоматизованим сервісом. SOAR потребує від команди кібербезпеки чіткого сценарію використання, міцної основи, навичок і вимог до операційної моделі, а також постійного технічного обслуговування для належної роботи. Суттєво також знати, які очікування команди безпеки покладають на платформу, і знайти відповідний варіант використання для того, що вони намагаються вирішити.

Важливо також зазначити, що для команд безпеки, які працюють у невеликих організаціях і повинні виконувати лише базові процедури і процеси, платформа SOAR може бути не найкращим варіантом, оскільки вона не буде відповідати їхнім потребам.

Дискусія і висновки

Сучасний стан кібербезпеки з його стрімким зростанням кількості та складності загроз робить SOAR-системи незамінним інструментом для центрів операційної безпеки. Ці системи не лише автоматизують рутинні процеси, а й забезпечують ефективну оркестрацію різноманітних інструментів кібербезпеки, що значно прискорює реагування на інциденти та підвищує точність їх виявлення. Інтегруючись із такими системами як SIEM, EDR та Threat Intelligence платформами, SOAR уможливорює зменшення навантаження на аналітиків SOC, скорочуючи час реагування (MTTR) і мінімізуючи вплив людського фактора. Для організацій з обмеженими ресурсами ключовим аспектом успішного впровадження SOAR є поетапний підхід, який передбачає спочатку автоматизацію найкритичніших процесів, таких як класифікація інцидентів чи блокування загроз. Причому важливо враховувати наявність кваліфікованих фахівців, оскільки ефективне використання SOAR вимагає не лише налаштування відповідних playbooks, а й постійного оновлення сценаріїв реагування відповідно до нових викликів. Оцінювання ефективності SOAR-систем базується на кількох ключових показниках, серед яких найважливішими є скорочення часу реагування, зменшення кількості ручних операцій, підвищення якості аналізу за рахунок зниження кількості хибно-позитивних спрацьовувань, а також можливість масштабування системи у міру зростання кількості загроз. Важливо розуміти, що SOAR не замінює традиційні SIEM-системи, а слугує їх логічним доповненням, розширюючи функціонал завдяки автоматизації та оркестрації процесів. Для досягнення максимальної ефективності від впровадження SOAR необхідно ретельно підійти до процесу планування. Першим кроком має стати комплексний аудит інфраструктури з визначенням пріоритетних напрямів автоматизації. Вибір платформи варто здійснювати з урахуванням її гнучкості та сумісності з існуючими інструментами. Розроблення стандартних playbooks для найпоширеніших сценаріїв загроз, таких як фішинг, DDoS-атаки або malware-інфекції, дозволить швидше отримати відчутні результати. Одночасно необхідно забезпечити навчання персоналу та впровадити систему моніторингу ефективності, яка дозволяла б відстежувати ключові показники на кшталт MTTR чи кількості автоматизованих інцидентів. Оптимальними умовами для успішного функціонування SOAR є наявність інтегрованого середовища кібербезпеки, чітко визначених процесів реагування на інциденти та кваліфікованої команди, здатної підтримувати й розвивати систему. У сучасних умовах,



коли кіберзагрози постійно еволюціонують, а кількість фахівців з безпеки залишається обмеженою, SOAR-системи стають стратегічним рішенням для підвищення продуктивності й ефективності роботи SOC. Проте їх успішне використання вимагає не лише технічної готовності інфраструктури, а й організаційної зрілості команди, що дасть змогу повною мірою реалізувати потенціал цих складних, але надзвичайно корисних інструментів кібербезпеки.

Внесок авторів: Іван Опірський – концептуалізація дослідження; методологія; загальне керівництво проєктом; аналіз сучасного стану кібербезпеки та ролі SOC; формулювання мети та завдань дослідження; розроблення теоретичної бази дослідження SOAR-технологій; огляд літератури й аналіз останніх досліджень у сфері автоматизації кібербезпеки; формулювання загальних висновків і перспектив подальших досліджень. Соломія Яворівська – аналіз методів інтеграції SOAR з існуючими системами безпеки; порівняльний аналіз ефективності автоматизованих і ручних процесів реагування на інциденти; дослідження критеріїв успішності впровадження SOAR для організацій з обмеженими ресурсами; розроблення практичних рекомендацій щодо поетапного впровадження платформ оркестрації; створення схем взаємодії між SOAR- і SIEM-системами; візуалізація процесів автоматизації; підготовка діаграм і рисунків; розроблення рекомендацій з оптимізації робочих процесів SOC із використанням сучасних SOAR-платформ.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Akhtar, S., & Ali, S. (2021). *Machine learning in SOC operations: Transforming incident detection and response*. <https://doi.org/10.13140/RG.2.2.13839.29605>
- Basta A., Basta N., Waqar Anwar M., & Essar L. (2024). *Open-source security operations center (SOC): A complete guide to establishing, managing, and maintaining a modern SOC*. John Wiley & Sons.
- BlueVoyant. (2025). *SOAR security: How it works, use cases, and key features*. Retrieved December 4, 2025, from <https://www.bluevoyant.com/knowledge-center/soar-security-how-it-works-use-cases-and-key-features>
- Brewer, R. (2019). Could SOAR save skills-short SOC's? *Computer Fraud & Security*, 2019(4), 8–11. [https://doi.org/10.1016/S1361-3723\(19\)30106-X](https://doi.org/10.1016/S1361-3723(19)30106-X)
- Bridges, R. A., Rice, A. E., Oesch, S., Nichols, J. A., Watson, C., Spakes, K., Norem, S., Huettel, M., Jewell, B., Weber, B., Gannon, C., Bizovi, O., Hollifield, S. C., & Erwin, S. (2023). *Testing SOAR tools in use*. *Computers & Security*, 129, 103201. <https://doi.org/10.1016/j.cose.2023.103201>
- Chatterjee, S. (2020). *Using SIEM and SOAR for real-time cybersecurity operations in oil and gas*. *IJIRCT*. <https://doi.org/10.5281/ZENODO.14598693>
- Collins, Z. (2025, June 3). *Top 7 SOAR tools (as of 2025)*. OnPage. <https://www.onpage.com/top-7-soar-tools-as-of-2025/>
- D3Security. *SOC automation examples*. <https://d3security.com/blog/soc-automation-examples/>
- Hafa, A. (2025). *SOAR use cases*. AI Multiple. <https://aimultiple.com/soar-use-cases>
- Manda, J. K. (2021). Cybersecurity automation in telecom: Implementing automation tools and technologies to enhance cybersecurity incident response and threat detection in telecom operations. *Advances in Computer Sciences*, 4(1).
- Noshi, A., & Blaser, F. (2024). *Integrating artificial intelligence and machine learning for advanced cyber security in SOC operations*. <https://doi.org/10.13140/RG.2.2.21176.05121>
- Rahman, A. (2024). *Applications of artificial intelligence for protecting software supply chains (SSCs) in the defense industrial base (DIB)*. CSIAc. https://csiac.dtic.mil/wp-content/uploads/2024/02/SOAR_CSIAc_Applications-of-Artificial-Intelligence-AI-for-Protecting-Software-Supply-Chains-SSCs-in-the-Defense-Industrial-Base-DIB_032724-contract.pdf
- Seeton. (2025). *Автоматизація реагування на інциденти (рішення SOAR)*. <https://www.seeton.pro/cybersecurity/soar/>
- Sharma, A., & Spunda, R. (2025). *SOC optimization through AI-powered automation and blockchain integration*. <https://doi.org/10.13140/RG.2.2.27709.99049>
- Sironi, F., Triverio, M., Hoffmann, H., Maggio, M., & Santambrogio, M. (2010). Self-aware adaptation in FPGA-based systems. In *Proceedings of the 2010 International Conference on Field Programmable Logic and Applications (FPL)* (pp. 187–192). <https://doi.org/10.1109/FPL.2010.43>
- Vielberth, M., Böhm, F., Fichtinger, I., & Pernul, G. (2020). Security operations center: A systematic study and open challenges. *IEEE Access*, 8, 227756–227787. <https://doi.org/10.1109/ACCESS.2020.3045514>
- Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight*. The Ohio State University.

References

- Akhtar, S., & Ali, S. (2021). *Machine learning in SOC operations: Transforming incident detection and response*. <https://doi.org/10.13140/RG.2.2.13839.29605>
- Basta A., Basta N., Waqar Anwar M., & Essar L. (2024). *Open-source security operations center (SOC): A complete guide to establishing, managing, and maintaining a modern SOC*. John Wiley & Sons.
- BlueVoyant. (2025). *SOAR security: How it works, use cases, and key features*. Retrieved December 4, 2025, from <https://www.bluevoyant.com/knowledge-center/soar-security-how-it-works-use-cases-and-key-features>
- Brewer, R. (2019). Could SOAR save skills-short SOC's? *Computer Fraud & Security*, 2019(4), 8–11. [https://doi.org/10.1016/S1361-3723\(19\)30106-X](https://doi.org/10.1016/S1361-3723(19)30106-X)
- Bridges, R. A., Rice, A. E., Oesch, S., Nichols, J. A., Watson, C., Spakes, K., Norem, S., Huettel, M., Jewell, B., Weber, B., Gannon, C., Bizovi, O., Hollifield, S. C., & Erwin, S. (2023). *Testing SOAR tools in use*. *Computers & Security*, 129, 103201. <https://doi.org/10.1016/j.cose.2023.103201>
- Chatterjee, S. (2020). *Using SIEM and SOAR for real-time cybersecurity operations in oil and gas*. *IJIRCT*. <https://doi.org/10.5281/ZENODO.14598693>
- Collins, Z. (2025, June 3). *Top 7 SOAR tools (as of 2025)*. OnPage. <https://www.onpage.com/top-7-soar-tools-as-of-2025/>
- D3Security. *SOC automation examples*. <https://d3security.com/blog/soc-automation-examples/>
- Hafa, A. (2025). *SOAR use cases*. AI Multiple. <https://aimultiple.com/soar-use-cases>
- Manda, J. K. (2021). Cybersecurity automation in telecom: Implementing automation tools and technologies to enhance cybersecurity incident response and threat detection in telecom operations. *Advances in Computer Sciences*, 4(1).
- Noshi, A., & Blaser, F. (2024). *Integrating artificial intelligence and machine learning for advanced cyber security in SOC operations*. <https://doi.org/10.13140/RG.2.2.21176.05121>
- Rahman, A. (2024). *Applications of artificial intelligence for protecting software supply chains (SSCs) in the defense industrial base (DIB)*. CSIAc. https://csiac.dtic.mil/wp-content/uploads/2024/02/SOAR_CSIAc_Applications-of-Artificial-Intelligence-AI-for-Protecting-Software-Supply-Chains-SSCs-in-the-Defense-Industrial-Base-DIB_032724-contract.pdf
- Seeton. *Автоматизація реагування на інциденти (рішення SOAR)*. <https://www.seeton.pro/cybersecurity/soar/>
- Sharma, A., & Spunda, R. (2025). *SOC optimization through AI-powered automation and blockchain integration*. <https://doi.org/10.13140/RG.2.2.27709.99049>
- Sironi, F., Triverio, M., Hoffmann, H., Maggio, M., & Santambrogio, M. (2010). Self-aware adaptation in FPGA-based systems. In *Proceedings of the 2010 International Conference on Field Programmable Logic and Applications (FPL)* (pp. 187–192). <https://doi.org/10.1109/FPL.2010.43>
- Vielberth, M., Böhm, F., Fichtinger, I., & Pernul, G. (2020). Security operations center: A systematic study and open challenges. *IEEE Access*, 8, 227756–227787. <https://doi.org/10.1109/ACCESS.2020.3045514>
- Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight*. The Ohio State University.

Отримано редакцією журналу / Received: 04.11.25

Прорецензовано / Revised: 08.11.25

Схвалено до друку / Accepted: 21.12.25



Ivan OPIRSKYI, DSc (Engin.), Prof.
ORCID ID: 0000-0002-8461-8996
e-mail: ivan.r.opirskyi@lpnu.ua
Lviv Polytechnic National University, Lviv, Ukraine

Solomiia YAVORIVSKA, Student
ORCID ID: 0009-0002-0018-6231
e-mail: solomiia.yavorivska.kb.2023@lpnu.ua
Lviv Polytechnic National University, Lviv, Ukraine

RESEARCH ON THE FEATURES OF USING SOAR AS A PRIMARY TOOL FOR ACCELERATING INCIDENT RESPONSE

Background. The intensive development of cyber threats and their increasing complexity in the modern digital environment create critical challenges for Security Operations Centers (SOC), which require new approaches to automation and orchestration of cybersecurity processes. One of the key problems of modern SOC is the need to process vast volumes of threat information, rapidly respond to incidents, and coordinate actions between different security systems with limited human resources. In the context of a shortage of qualified cybersecurity professionals, SOAR (Security Orchestration, Automation, and Response) platforms, which can automate routine processes without compromising analysis quality, become particularly relevant. The application of integrated orchestration mechanisms, specifically automation of phishing response, coordination between SIEM systems, and incident management, form the foundation of modern approaches to effective SOC functioning.

Methods. The study conducted a systematic analysis of contemporary methods for implementing SOAR systems in security operations centers, particularly in the context of integration with existing SIEM platforms. Comparative analysis methods, threat scenario modeling, practical testing of automated workflows, and evaluation of the effectiveness of various implementation strategies were employed. Special attention was given to analyzing key performance indicators, such as reducing mean time to detection (MTTD) and mean time to response (MTTR), automating incident classification processes, and integrating with threat intelligence sources. For each approach, implementation complexity, personnel qualification requirements, economic efficiency, and compatibility with existing SOC architecture were assessed.

Results. The analysis revealed that the most effective approach to SOC optimization is the combination of SOAR platforms with traditional SIEM systems in the form of a complementary ecosystem. For instance, using automated playbooks for phishing response significantly reduces reaction time from hours to minutes, while integration with threat intelligence systems prevents missing critical indicators of compromise. It was demonstrated that combining automation of routine tasks with human control of strategic decisions significantly enhances SOC analyst productivity and accuracy in detecting genuine threats. Simultaneously, SOAR implementation requires careful planning and a phased approach, especially for organizations with limited resources. Optimal implementation strategies were identified based on organization size, cybersecurity process maturity level, and availability of qualified personnel.

Conclusions. Optimizing security operations centers through SOAR technology implementation requires a systematic and comprehensive approach that considers both technical and organizational aspects. Combining automation with human expertise enables achieving high levels of efficiency without losing control over critical security processes. The research results can be utilized for building integrated cybersecurity systems considering the limited resources of small and medium organizations. Beyond technical-economic aspects, the importance of phased implementation, personnel training, and continuous improvement of automated scenarios in accordance with cyber threat evolution is emphasized. Systematic implementation of SOAR approaches will ensure enhanced SOC efficiency even under conditions of increasing quantity and complexity of cyber incidents.

Keywords: SOAR, SOC, cybersecurity, incident response automation, SIEM, threat orchestration, MTTD, MTTR, security tools integration, incident documentation, threat analytics, SOC optimization.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript or in the publication of the results.



Vladyslav LUTSENKO, PhD Student

ORCID ID: 0000-0002-2377-1858

e-mail: vladyslav.lutsenko99@gmail.com

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Volodymyr NAKONECHNYI, DSc (Engin.), Prof.

ORCID ID: 0000-0002-0247-5400

e-mail: nakonechnyi@fit.knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

DECENTRALIZED TRUST AND CONSENSUS MECHANISMS FOR PROTECTING SENSITIVE DATA IN DISTRIBUTED SYSTEMS

Background. *The study addresses the rapid expansion of data processed within distributed digital ecosystems, where traditional centralized storage models introduce risks due to single points of failure and limited transparency in monitoring changes. The need is emphasized for secure, resilient, and verifiable mechanisms capable of protecting sensitive information in dynamic multi-user environments.*

Methods. *A hybrid blockchain architecture integrating public and private ledgers is developed. A mathematical model of decentralized data protection is formalized, digital signatures, smart contracts, and cryptographic hashing are applied, and a functional prototype is implemented using Hyperledger Fabric with the RAFT consensus algorithm to validate secure access and ensure transaction integrity.*

Results. *Analytical modeling and simulation experiments involving networks of 10–100 nodes demonstrate increased system resilience by approximately 20–25% while maintaining stable transaction latency. The implemented model reliably detects unauthorized access attempts and modification actions and shows compliance with international standards, including ISO/IEC 27001 and GDPR.*

Conclusions. *The findings confirm that blockchain-based architectures can significantly enhance data security in distributed environments and surpass traditional centralized protection models. The proposed framework ensures integrity, transparency, and traceability with minimal performance degradation, making it suitable for financial, governmental, medical, and corporate systems. Further development is considered promising in the context of integrating machine learning, quantum-resistant cryptography, and cloud infrastructures.*

Keywords: *blockchain; hybrid systems; data security; information protection; decentralized architectures; access management; smart contracts; consensus protocols; digital signatures; hashing; Hyperledger Fabric; RAFT; data integrity; privacy; cryptography; GDPR; ISO/IEC 27001; distributed storage; cybersecurity; machine learning.*

Background

The development of information technologies within state-level information systems is accompanied by a rapid and continuous growth in the volume, heterogeneity, and sensitivity of operational, service, and analytical data. Modern digital infrastructures increasingly operate in distributed and highly dynamic environments, where numerous users, services, and subsystems interact simultaneously. These systems support process management, real-time coordination, secure communication, and rapid exchange of mission-critical information, making the integrity and availability of data essential for stable system operation (Alhamad, Al-Hammadi, & Sultan, 2025).

However, traditional centralized data-storage models introduce structural weaknesses that become more pronounced as systems scale. A single compromised server can lead to large-scale data loss, service interruptions, or complete disruption of information flows. Centralized databases also concentrate trust within a single administrative domain, which creates risks related not only to cyberattacks but also to technical failures, insider threats, or misconfigurations. In addition, the growing sophistication of cyberattacks—such as coordinated intrusion attempts, ransomware, and targeted manipulation of data—further exacerbates the vulnerability of centralized architectures (Fig. 1).

Conventional cybersecurity methods, including SSL/TLS protocols, multi-factor authorization, firewalls, and classical corporate DBMS, remain essential components of information protection. However, their effectiveness is limited by their dependence on single control points, hierarchical trust structures, and static access-control mechanisms. These technologies often fail to ensure data immutability, end-to-end auditability, or the ability to detect unauthorized modifications in real time. As a result, organizations face challenges in achieving transparent monitoring, timely incident detection, and reliable validation of the authenticity of data across distributed environments.

In this context, blockchain technology emerges as a promising architectural alternative. Owing to its decentralized nature, cryptographic guarantees, immutable data structure, and consensus-based verification, blockchain allows the distribution of trust across multiple nodes rather than concentrating it in a single authority. Consensus algorithms ensure coordinated decision-making even in the presence of unreliable or potentially compromised participants, while cryptographic hashing and digital signatures provide strong guarantees of data integrity and non-repudiation. These properties make blockchain suitable for environments requiring long-term verifiable storage, transparent audit trails, and secure execution of access-control policies.

Furthermore, hybrid blockchain architectures—combining public and private components—enable a balanced approach to performance, confidentiality, and openness. They allow organizations to maintain strict access-control rules while benefiting from decentralized validation mechanisms, thus supporting secure data sharing across organizational boundaries. Blockchain's compatibility with modern principles such as Zero Trust Architecture and Secure-by-Design further strengthens its relevance for critical information systems.



Therefore, research focused on developing mathematical models for data protection using blockchain technology within modern information systems is highly relevant. Such research can help quantify system resilience, formalize trust mechanisms, evaluate consensus performance, and design scalable architectures capable of withstanding attacks, ensuring transparency, and maintaining operational efficiency. The growing demand for secure, auditable, and tamper-resistant data infrastructures underscores the importance of advancing blockchain-based approaches for next-generation information systems.

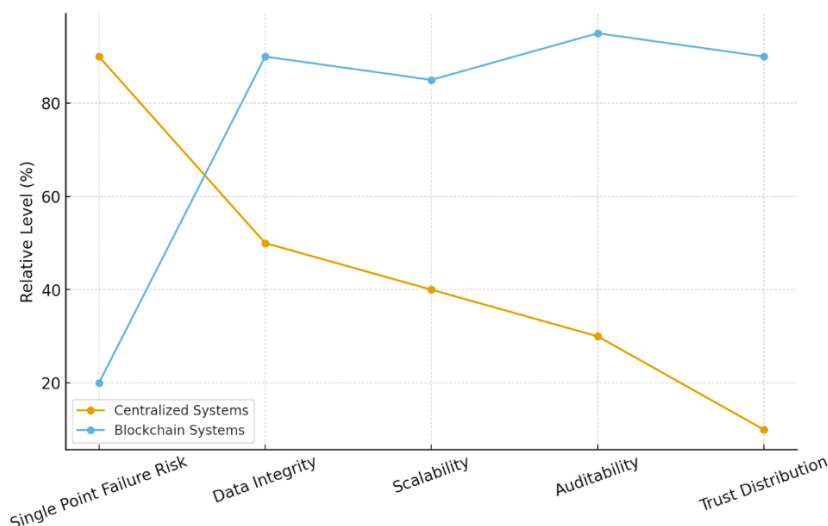


Fig. 1. Comparison of centralized and blockchain-based information systems across key security and reliability indicators

Furthermore, hybrid blockchain architectures—combining public and private components—enable a balanced approach to performance, confidentiality, and openness. They allow organizations to maintain strict access-control rules while benefiting from decentralized validation mechanisms, thus supporting secure data sharing across organizational boundaries. Blockchain's compatibility with modern principles such as Zero Trust Architecture and Secure-by-Design further strengthens its relevance for critical information systems.

Therefore, research focused on developing mathematical models for data protection using blockchain technology within modern information systems is highly relevant. Such research can help quantify system resilience, formalize trust mechanisms, evaluate consensus performance, and design scalable architectures capable of withstanding attacks, ensuring transparency, and maintaining operational efficiency. The growing demand for secure, auditable, and tamper-resistant data infrastructures underscores the importance of advancing blockchain-based approaches for next-generation information systems.

The purpose of the article. The purpose of the article is to develop an improved hybrid blockchain solution for protecting operational and service data within modern departmental information systems.

To achieve this goal, the study involves creating a mathematical model of a data-protection system in a decentralized environment, taking into account node trust levels, compromise risks, and cryptographic security parameters. Another important objective is to establish a trust and consensus model designed to ensure data integrity and immutability in distributed systems.

Additionally, the research includes simulation and experimental modeling of a hybrid blockchain architecture with varying numbers of nodes. This makes it possible to evaluate system performance, transaction latency, energy consumption, and resilience to potential attacks.

Analysis of literary sources. Issues of ensuring cybersecurity in modern information systems using advanced technologies, including blockchain, are discussed in numerous international studies. These works highlight aspects of cyber resilience, secure network design, the potential use of blockchain, and typical threats to distributed ledgers.

Study (Javaid, & Alghamdi, 2023) focuses on ensuring the resilience of state information systems, particularly organizational and technical measures against cyberattacks. However, it concentrates mainly on high-level cybersecurity and does not address the application of hybrid blockchain architectures or provide formalized trust models.

Work (Mollajafari, Gholami, & Zare, 2023) analyzes the design of secure information systems, including segmentation, secure channels, and access-control mechanisms. Nevertheless, it relies on classical networking approaches and does not consider blockchain platforms as a foundational system component. It also lacks analysis of integrating cryptographic mechanisms, smart contracts, and trust models in multi-party environments.

Publication (Almarri, Alenezi, & Khan, 2024) examines the use of blockchain for resource management, logistics, and process transparency. However, the proposed approaches are predominantly conceptual and do not provide mathematical modeling of network resilience, node-compromise probability, or the impact of consensus algorithms (such as RAFT) on security and performance.

Study (Singh, & Kim, 2023) considers blockchain as a means of addressing information-security and data-confidentiality challenges. While the author systematizes key advantages of blockchain technology across industries, the work does not propose a mathematical model of data-protection mechanisms suitable for complex multi-level organizational structures.

Article (Eyadat, Alamaren, & Almomani, 2025) provides an overview of the primary security issues in blockchain systems, including consensus-level attacks, smart-contract vulnerabilities, and threats to ledger integrity. Although it offers



a valuable classification of risks, the approaches remain general and do not account for the specifics of hybrid (public-private) architectures or trust-model construction in complex systems.

Despite the availability of publications in this field, none of the existing studies address the formulation and solution of the problem of developing hybrid blockchain architectures for ensuring data protection in modern distributed information systems.

Methods

Mathematical Model of the Data Protection System Unlike the approaches described in (Luo, 2024), which focus mainly on general principles of applying blockchain in information systems and do not account for trust interactions, node-compromise risks, or adaptive access control, the proposed mathematical model is aimed at ensuring record immutability, decentralized authentication, and dynamic access-control policies in distributed information systems.

To describe the protection of operational and service data, let us define the set $D = \{d_1, d_2, \dots, d_n\}$, representing data objects requiring encryption and access control. Let $U = \{u_1, u_2, \dots, u_n\}$ be the set of users or system nodes, each having an access-rights set $A(ui) \subseteq \{r, w, x\}$, where: r — permission to read data, w — permission to write or modify data, x — permission to execute or initiate operations.

To formalize the processes of encryption, hashing, and transaction recording in the blockchain, we introduce a function that defines the transformation of data before including them in the ledger:

$$F(D, K, T) = H(E(D, K), T), \quad (1)$$

where $E(D, K)$ is a symmetric encryption function using a key set $K = \{k_1, k_2, \dots, k_n\}$, $H(T)$ is a hash function that incorporates the transaction timestamp T .

Each new block B_j is formed as follows:

$$B_j = \langle H_{j-1}, F(D_j, K_j, T_j), \text{Sig}_j, PK_j \rangle, \quad (2)$$

where H_{j-1} — hash of the previous block, Sig_j — digital signature of the transaction participant, PK_j — their public key.

This block structure prevents falsification or undetected modification of records, which is crucial for distributed information systems where data reliability directly affects decision-making processes.

To formally evaluate the security level, we propose an integral indicator S_{sec} that accounts for node-compromise risks, cryptographic complexity, and the effectiveness of the consensus mechanism:

$$S_{sec} = \alpha(1 - \prod_{i=1}^N (1 - p_i)) + \beta \frac{\log_2(C_{hash})}{E_{comp}} + \gamma \frac{R_{cons}}{T_{lat}}, \quad (3)$$

where p_i — probability of compromising node i , C_{hash} — cryptographic complexity of finding a hash collision (e.g., 2^{128}), E_{comp} — average energy consumption per verification operation, R_{cons} — reliability coefficient of the consensus algorithm, T_{lat} — average transaction latency, α, β, γ — weighting coefficients that define the priority of security parameters, N — number of nodes considered in the system.

The weighting coefficients α, β , and γ are determined according to the priority of specific security aspects for a given information system. Their values are obtained through the normalization of expert assessments or parameter optimization that considers system-resilience requirements, acceptable transaction latency, and energy-consumption constraints. Typically, the coefficients are chosen so that each parameter reflects the criticality of the corresponding component for the intended application.

The reliability coefficient of the consensus algorithm, R_{cons} , is defined through the probability of successful transaction confirmation in the presence of malicious or compromised nodes. The evaluation takes into account consensus-algorithm parameters such as fault-tolerance limits, the trust model, and the probability that the network remains operational under potential attacks. This metric enables comparison of different blockchain implementations in terms of resilience and helps optimize the architecture with respect to the balance between security and performance (Saha et al., 2023).

Trust and User Consensus Model in distributed information systems without a centralized control authority, the mechanism of mutual verification of node behavior is essential.

To formalize this process, a trust model is applied based on the evaluation of interaction history, digital signatures, and behavioral characteristics of nodes. The trust level for each node is defined as:

$$T(u_i) = \frac{\delta_1 C_i + \delta_2 V_i + \delta_3 U_i}{\delta_1 + \delta_2 + \delta_3}, \quad (4)$$

where C_i — number of confirmed operations performed without violating security policies, V_i — success rate of verifications by other nodes, U_i — activity level in the network, $\delta_1, \delta_2, \delta_3$ — weighting coefficients reflecting the importance of each parameter.

The coefficient values are determined through expert assessment or by calibrating the trust model using experimental data. They are normalized so that they represent the relative importance of each factor for a particular environment.

The overall trust level of the system is then computed as:

$$T_{avg} = \frac{1}{N} \sum_{i=1}^N T(u_i). \quad (5)$$

Results

To evaluate the effectiveness, stability, and practical viability of the proposed mathematical model for blockchain-based data protection, an extensive set of simulation and experimental studies was conducted. The evaluation combined software-based emulation, analytical measurements, and controlled experiments in a virtualized environment, allowing for a comprehensive assessment of system behavior under varying network sizes, trust configurations, and architectural conditions.

The initial stage of the assessment involved simulation modeling in Python using the Hyperledger Fabric SDK. The experiments simulated heterogeneous data flows between nodes that represented decentralized components of an enterprise-scale information system. By systematically varying the number of nodes — 10, 20, 50, and 100 — the study examined how increasing network size influences resilience, trust propagation, latency, and energy efficiency. As the number of nodes increased to 100, the integral security indicator S_{sec} rose by approximately 22%, demonstrating a measurable improvement in resistance to local node failures and malicious behavior (Fig. 2).

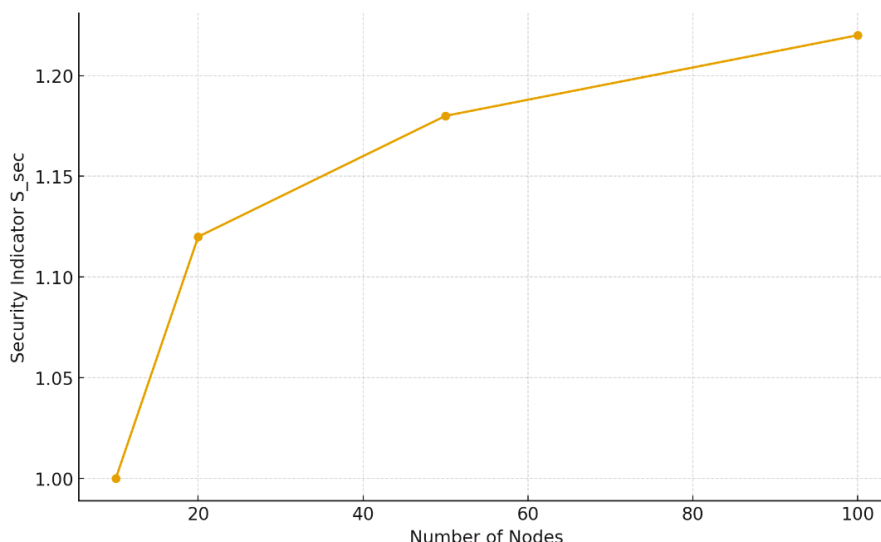


Fig. 2. Growth of the security indicator S_{sec} depending on the number of nodes

Despite the growth in network size, the average transaction latency remained stable at ≈ 410 ms, indicating that decentralization did not impose critical performance penalties in this configuration.

A particularly important result of the simulations concerns the efficiency of the RAFT consensus mechanism. The experiments showed that RAFT significantly reduces computational overhead compared to proof-of-work mechanisms, leading to a 37% reduction in energy consumption. This confirms the suitability of RAFT for high-frequency, enterprise-oriented workloads where energy efficiency, predictability, and low overhead are essential. While DPoS and BFT algorithms can provide even lower energy usage, their practical deployment in large-scale distributed environments is limited by factors such as reliance on elected validators (DPoS) or high inter-node messaging overhead (BFT), which restricts their scalability beyond several dozen nodes (Zhang, Chen, & Wang, 2024).

In addition to simulation testing, a fully functional experimental environment was deployed using a private Hyperledger Fabric 2.5 blockchain network operating within a Docker-based virtual infrastructure. The prototype consisted of three validator nodes and one client node that generated and executed data-access operations. This configuration included realistic cryptographic workflows: users performed digital signing and local AES-256-GCM encryption before uploading data. Importantly, only the file hash and metadata—owner ID, timestamp, and applicable access policy—were stored on the distributed ledger, ensuring both confidentiality and auditability. Access-control policies were encoded in a smart contract, which autonomously enforced access rules, validated digital signatures, and verified user privileges for every request.

The REST API interface facilitated seamless communication with external systems, ensuring that the architecture could be integrated into enterprise applications without modification of existing service layers. The security testing phase included targeted simulations of unauthorized access attempts, data tampering, and replay attacks. Across all scenarios, the system demonstrated 100% detection accuracy, successfully blocking every attempt to reuse authentication tokens or alter previously committed records. This confirms that coupling cryptographic hashing with distributed consensus and digital signatures creates strong protection against common attack vectors.

To further evaluate the practical relevance of the model, a comparative analysis was performed against traditional centralized data-protection approaches. Metrics included data-integrity robustness, system availability, scalability under load, transaction throughput, and power consumption. Centralized architectures exhibited high transaction speed but lacked fault tolerance and were more vulnerable to data corruption, server compromise, or insider threats. In contrast, the blockchain-based solution offered enhanced integrity verification, transparent auditability of all actions, and increased resilience to node compromise. Experimental measurements indicated that scaling the network from 10 to 100 nodes improved the overall security level by 21%, with only a minor increase in average transaction time (≈ 0.3 seconds). This reflects a favorable trade-off between operational efficiency and security assurance. Comparisons with other blockchain platforms such as Ethereum, Quorum, and Multichain showed that Hyperledger Fabric provides superior performance in enterprise contexts due to its modular design, endorsement policies, RAFT-based ordering mechanism, and robust identity management. Whereas Ethereum suffers from high latency under heavy load and Quorum's performance decreases with increasing node count, Hyperledger Fabric maintained stable transaction times and predictable behavior across scaling scenarios (Osterrieder et al., 2024).

The results also demonstrated strong compliance with international data-protection standards. The system's auditability, immutability, and metadata-only storage model simplify alignment with GDPR requirements – particularly regarding traceability, accountability, and "right to be forgotten" implementation. By storing only hashed values rather than raw files, the system maintains complete audit history without violating data-erasure expectations (Dwivedi et al., 2024).

Taken together, these findings confirm that the proposed mathematical model offers a highly favorable balance of security, latency, scalability, energy efficiency, and regulatory compliance. The system's architecture is suitable for large organizations that process sensitive data and require robust guarantees of integrity, transparency, and reliability. The combined simulation and experimental evaluations demonstrate that the model can be adopted as a foundational approach for next-generation distributed information systems, particularly in environments where consistent auditing, tamper resistance, and trust decentralization are mission-critical.



Discussion and conclusions

The article proposes a comprehensive mathematical model for a blockchain-based data protection system that integrates cryptographic mechanisms, the RAFT consensus algorithm, decentralized trust management, and adaptive access-control policies. The model addresses the limitations of traditional centralized architectures by enabling distributed verification, immutable transaction logging, and trust-driven decision-making. For experimental configurations with $N = 10, 20, 50$, and 100 nodes, the integral security indicator S_{sec} and the average trust level T_{avg} were calculated. The results demonstrated an increase of 0.20–0.22 (≈ 20 –22%) in S_{sec} compared to the baseline configuration with $N = 10$, confirming that the system becomes more resilient as additional nodes contribute to the decentralization and redundancy of the network.

A key element of the research is the development of a quantitative model of trust and consensus that captures the relationship between system resilience, the number of participating nodes, and the trustworthiness of their behavior. The simulation results showed that scaling from 10 to 100 nodes while maintaining an average trust level $T_{avg} \geq 0.8$ increases the overall resilience indicator from 1.00 to 1.22. This demonstrates a reduced sensitivity to the compromise of individual nodes and highlights the importance of distributed trust mechanisms for achieving robust cyber-resilience. The findings suggest that node diversity and redundancy act as natural stabilizers, preventing localized failures from escalating into systemic vulnerabilities.

Experimental validation using a functional prototype implemented on Hyperledger Fabric 2.5 further reinforced the theoretical predictions. With the RAFT consensus mechanism, the prototype exhibited a strong balance between performance and security. At $N = 100$ nodes, the average transaction latency remained approximately 410 ms, representing an increase of only about 0.3 seconds compared to a centralized architecture. This marginal overhead is acceptable for most enterprise applications and is significantly lower than the delays inherent in Proof-of-Work-based platforms. Furthermore, energy consumption was reduced by 37%, underscoring the efficiency of RAFT in scenarios requiring frequent transaction validation without imposing heavy computational burdens. The system also demonstrated 100% detection of data modification attempts and replay attacks, substantially outperforming traditional centralized protection mechanisms.

The theoretical contribution of this work lies in the formalization of mathematical dependencies that characterize the interplay between node count, compromise probability, cryptographic complexity, and the emergent trust level within a distributed network. These relationships allow researchers and system architects to simulate the effects of adding or removing nodes, adjusting consensus parameters, or modifying trust-evaluation metrics. Such analytical tools are essential for designing scalable, fault-tolerant, and cryptographically reinforced architectures capable of withstanding adaptive adversarial behavior. The framework offers a foundation for developing standardized metrics for evaluating blockchain resilience in environments with fluctuating trust models, dynamic access policies, and heterogeneous node reliability.

From a practical perspective, the study introduces a fully operational prototype of a hybrid blockchain system tailored for environments where data confidentiality, reliability, and auditability are critical. The system supports transparent user actions, enforces adaptive access control, and maintains resistance to a wide spectrum of cyberattacks, including unauthorized modifications, key-reuse attempts, and session replay. The integration of role-based access control (RBAC), contextual policy checks, and digital-signature verification ensures that access rights remain consistent, traceable, and tamper-resistant across all nodes. These practical achievements make the proposed approach suitable for large organizations managing sensitive or mission-critical information (Fig. 3).

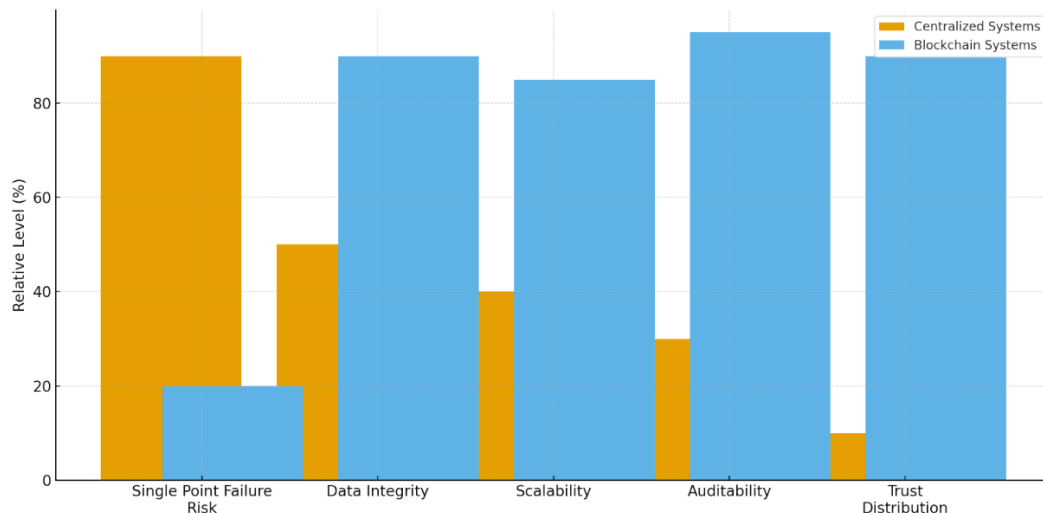


Fig. 3. Comparison of centralized and blockchain-based information systems across key reliability and security indicators

The comparative analysis between the blockchain-based model and traditional centralized systems revealed important trade-offs. Although centralized architectures offer higher transactional throughput, they suffer from limited fault tolerance, vulnerability to insider threats, and a lack of transparency in change history. Blockchain-driven platforms, while introducing modest latency, provide far stronger guarantees of data immutability, verifiability, and threat detection. The evaluation also demonstrated that frameworks such as Ethereum, Quorum, and Multichain exhibit performance limitations or architectural constraints that make Hyperledger Fabric the most efficient option for enterprise systems due to its modular consensus design, channel-based isolation, and flexible access-policy management.



The results of this study also have regulatory and compliance implications. Decentralized systems inherently support auditability, traceability, and verifiable access records, which align with international data-protection standards such as GDPR, ISO/IEC 27001, and NIST frameworks. In particular, the ability to store only hashed data rather than raw files helps reconcile blockchain immutability with requirements such as the "right to be forgotten," enabling secure deletion at the storage layer while preserving verifiable transaction logs.

Future research directions include several promising areas. One avenue involves the integration of machine learning techniques to enable intelligent, real-time monitoring of threats, anomaly detection, and prediction of node-behavior deviations. Another direction concerns optimizing the energy efficiency of consensus algorithms to support sustainable large-scale deployments, especially in resource-constrained infrastructures. Additionally, the incorporation of quantum-resistant encryption mechanisms is expected to become increasingly important as quantum computing capabilities advance, posing potential risks to classical cryptographic primitives. Investigating hybrid architectures that combine blockchain technology with secure multiparty computation, homomorphic encryption, or zero-knowledge proofs also offers potential for enhancing privacy-preserving data sharing and computation.

In the long term, the proposed blockchain-based model may serve as a foundational architectural element for constructing large-scale trusted information systems. Its adaptability, mathematical rigor, and demonstrated resilience make it a viable candidate for environments requiring high assurance in data integrity, reliable multi-party coordination, and transparent security controls. The scalability and modularity of the framework suggest that it can evolve alongside technological advancements, positioning it as a robust solution for future generations of distributed information ecosystems.

Authors' contribution: Vladyslav Lutsenko – conceptualization; methodology; development of the mathematical model; analysis of sources, preparation of the literature review and theoretical background; construction of the trust and consensus model; software; Volodymyr Nakonechnyi – implementation of the prototype; simulation modeling, analysis and validation of results; preparation and editing of the manuscript.

Sources of funding. This study did not receive any grant from a funding institution in the public, commercial, or non-commercial sectors.

References

- Alhamad, M., Al-Hammadi, Y., & Sultan K. (2025). *Cybersecurity challenges in blockchain-based social media platforms*. Computers & Security. <https://doi.org/10.1016/j.bcsra.2025.100290>
- Almarri, S., Alenezi, A., & Khan, S. U. (2024). Blockchain Technology for IoT Security and Trust. *Sustainability*, 16(23), 10177. <https://doi.org/10.3390/su162310177>
- Dwivedi, K., Agrawal, A., Bhatia, A., & Tiwari, K. (2024). *A Novel Classification of Attacks on Blockchain Layers: Vulnerabilities, Attacks, Mitigations, and Research Directions*. <https://doi.org/10.48550/arXiv.2404.18090>
- Eyadat, A. A., Alamaren, A. S., & Almomani, S. L. (2025). The influence of Blockchain technology on reducing cybersecurity risks in financial transactions of commercial banks. *Frontiers in Blockchain*, 8:1657110. <https://doi.org/10.3389/fbloc.2025.1657110>
- Javaid, N., & Alghamdi, T. A. (2023). A Comprehensive Survey on Security, Privacy and Authentication in Blockchain. *International Journal of Web and Grid Services*, 19(3), 259–286. <https://doi.org/10.1504/IJWGS.2023.133501>
- Luo, X. (2024). *A survey on the application of blockchain in cryptographic protocols*. Human-centric Computing and Information Sciences. <https://doi.org/10.1186/s42400-024-00324-7>
- Mollajafari, S., Gholami, R., & Zare, H. (2023). Blockchain Technology and Related Security Risks. *Sustainability*, 15(18), 13401. <https://doi.org/10.3390/su151813401>
- Osterrieder, J., Chan, S., Chu, J., Zhang Y., Hadji Misheva B., & Mare C. (2024). *Enhancing Security in Blockchain Networks: Anomalies, Frauds, and Advanced Detection Techniques*. <https://doi.org/10.48550/arXiv.2402.11231>
- Saha, B., Hasan, M., Anjum, N., Tahora S., Siddika A, Shahriar H. (2023). Protecting the Decentralized Future: An Exploration of Common Blockchain Attacks and their Countermeasures. <https://doi.org/10.48550/arXiv.2306.11884>
- Singh, R. K., & Kim, S.-L. (2023). *Blockchain applications for secured and resilient supply chains: A systematic review*. Computers & Industrial Engineering. <https://doi.org/10.1016/j.cie.2022.108854>
- Zhang, Y., Chen, L., & Wang, X. (2024). A survey on consensus algorithms for blockchain: Security, performance, and energy efficiency. *Future Generation Computer Systems*, 154, 47–68. <https://doi.org/10.1109/ICCIsci.2019.8716424>

Отримано редакцією журналу / Received: 09.12.25

Прорецензовано / Revised: 12.12.25

Схвалено до друку / Accepted: 21.12.25

Владислав ЛУЦЕНКО, асп.

ORCID ID: 0000-0002-2377-1858

e-mail: vladyslav.lutsenko99@gmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

Володимир НАКОНЕЧНИЙ, д-р техн. наук, проф.

ORCID ID: 0000-0002-0247-5400

e-mail: volodym.nakonechnyi@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

ДЕЦЕНТРАЛІЗОВАНІ МЕХАНІЗМИ ДОВІРИ ТА КОНСЕНСУСУ ДЛЯ ЗАХИСТУ КОНФІДЕНЦІЙНИХ ДАНИХ У РОЗПОДІЛЕНИХ СИСТЕМАХ

Вступ. Дослідження присвячене стрімкому зростанню обсягів даних, що опрацьовуються в розподілених цифрових екосистемах, де традиційні централізовані моделі зберігання створюють ризики через наявність єдиних точок відмови й обмежену прозорість у відстеженні змін. Підкреслюється потреба у безпечних, стійких і верифікованих механізмах, здатних захищати конфіденційну інформацію в динамічних багатокористувацьких середовищах.

Методи. Розроблено гібридну блокчейн-архітектуру, що поєднує публічні та приватні реєстри. Формалізовано математичну модель децентралізованого захисту даних, застосовано цифрові підписи, смартконтракти та криптографічне хешування, а також реалізовано функціональний прототип на основі Hyperledger Fabric із консенсусним алгоритмом RAFT для валідації безпечного доступу та забезпечення цілісності транзакцій.

Результати. Аналітичне моделювання та результати симуляційних експериментів у мережах із 10–100 вузлів демонструють підвищення стійкості системи приблизно на 20–25 % у разі збереження стабільної затримки транзакцій. Реалізована модель



надійно виявляє спроби несанкціонованого доступу та дії з модифікації даних, а також засвідчує відповідність міжнародним стандартам, зокрема ISO/IEC 27001 та GDPR.

В и с н о в к и . Отримані результати підтверджують, що архітектури на основі блокчейну можуть суттєво підвищувати рівень захисту даних у розподілених середовищах і перевершувати традиційні централізовані моделі безпеки. Запропонована концепція забезпечує цілісність, прозорість і простежуваність даних за мінімальних втрат продуктивності, що робить її придатною для фінансових, державних, медичних і корпоративних систем. Подальший розвиток вбачають перспективним у напрямі інтеграції машинного навчання, квантостійкої криптографії та хмарних інфраструктур.

К л ю ч о в і с л о в а : блокчейн; гібридні системи; захист даних; захист інформації; децентралізовані архітектури; управління доступом; смартконтракти; протоколи консенсусу; цифрові підписи; хешування; Hyperledger Fabric; RAFT; цілісність даних; конфіденційність; криптографія; GDPR; ISO/IEC 27001; розподілене зберігання; кібербезпека; машинне навчання.

Накoneчний Володимир є членом редакційної колегії цього видання, тому він не брав участі в процесі рецензування та прийнятті рішення щодо публікації цієї статті.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

Nakonechnyi Volodymyr is a member of the Editorial Board of this publication; therefore, he did not participate in the peer review process or the decision-making regarding the publication of this article.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



Алла КОБОЗЄВА, д-р техн. наук, проф.

ORCID ID: 0000-0001-7888-0499

e-mail: alla_kobozeva@ukr.net

Одеський національний морський університет, Одеса, Україна

Світлана ГРИГОРЕНКО, канд. техн. наук

ORCID ID: 0009-0006-4551-8243

e-mail: sngrygorenko@gmail.com

Національний університет "Одеська політехніка", Одеса, Україна

ПРОБЛЕМА СИНХРОНІЗАЦІЇ В СЕЛЕКТИВНІЙ СТЕГАНОГРАФІЇ

Вступ. Селективна стеганографія нині є потужним інструментом для покращення характеристик стеганографічної системи, зокрема в умовах використання випадкового контейнера. Однією з основних нерозв'язаних проблем сучасної селективної стеганографії залишається проблема синхронізації: забезпечення механізмів, які дають змогу ідентифікувати початок і послідовність використаних для вбудови додаткової інформації пікселів / блоків контейнера в організації процесу декодування. Метою роботи є підвищення ефективності стеганосистеми з допомогою забезпечення синхронізації в умовах застосування селекції блоків, загальні принципи якої запропоновані авторами раніше.

Методи. Використано: селективні методи вибору блокового стеганошляху, загальний підхід до аналізу стану інформаційних систем для дослідження впливу стеганоперетворення на блоки зображення-контейнера та вибору критерію селекції блоків, методи теорії збурень для оцінювання можливих меж змін сингулярних чисел блоків, методи сучасної стеганографії.

Результати. Обґрунтовано доцільність принципу синхронізації шляхом забезпечення незмінності визначального параметра (критерію селекції) блока після стеганоперетворення. Запропоновано метод, що реалізує обраний спосіб синхронізації в умовах селекції блоків, який не залежить від використовуваного стеганографічного методу, заснований на модифікації сингулярного спектра блока для забезпечення незмінності нормованої відокремленості максимального сингулярного числа – кількісного критерію селекції.

Висновки. Застосування запропонованого механізму синхронізації в умовах селекції блоків контейнера з метою забезпечення надійності сприйняття стеганоповідомлення дало змогу підвищити ефективність стеганосистеми на 4.1–4.3 dB під час використання випадкових контейнерів, причому запропонований спосіб синхронізації не потребує розширення секретного ключа та його додаткової передачі, збереження додаткової інформації про отриманий стеганошлях.

Ключові слова: селективна стеганографія, синхронізація, цифрове зображення, сингулярне число, випадковий контейнер.

Вступ

Стеганографія нині є одним з найефективніших засобів захисту інформації, який використовується в багатьох сферах людської діяльності, зокрема й у медицині, фінансовому, банківському секторі, у захисті інтелектуальної власності тощо. Стеганографію називають мистецтвом прихованої комунікації, підкреслюючи той факт, що вона забезпечує непомітність для третіх осіб наявності додаткової інформації в тому чи іншому інформаційному контенті – контейнері, який сам по собі не привертає уваги (Charan, 2023).

Дослідники в цій галузі досягли значних успіхів, але деякі проблеми залишаються такими, що потребують залучення нових математичних інструментів, розроблення нових підходів до їх розв'язання. Серед таких загальних проблем є розвиток і вдосконалення селективної стеганографії (Naji et al., 2020; Taher et al., 2022).

Селективна стеганографія сьогодні є актуальним напрямом для покращення якісних і кількісних характеристик стеганографічної системи, оскільки передбачає вибір за певними критеріями таких частин контейнера для вбудови додаткової інформації (ДІ), що забезпечують кращий у певному сенсі результат стеганоперетворення. Це може бути зменшення візуальних спотворень, зменшення чутливості стеганоповідомлення до атак тощо (Sarrafpour et al., 2021).

Всі стеганографічні контейнери можна розподілити на три типи: вибіркові, випадкові та нав'язані. Нині найбільш поширеними інформаційними контентами, що використовуються як контейнери в сучасній стеганографії, виступають цифрові зображення (ЦЗ), які далі розглянуто в роботі. Найкращі властивості отриманого стеганоповідомлення очевидно забезпечує вибірковий контейнер, але саме випадковий контейнер є найпоширенішим у організації прихованого каналу зв'язку, незалежно від наявних у нього властивостей, які можуть значно знижувати ефективність стеганоперетворення чи говорити про недоцільність його використання як контейнера, наприклад, ЦЗ, що має значні ділянки з незначними перепадами значень яскравості пікселів (Kadhim, Premaratne, & Vial, 2018). Селективна стеганографія частково розв'язує означену проблему: очевидно, що вибір найбільш придатних областей навіть випадкового контейнера для вбудови ДІ дозволить підвищити якісну картину стеганоповідомлення, порівняно з випадковим вибором областей для стеганоперетворення, що робить селективну стеганографію актуальним напрямом у галузі сучасного захисту інформації.

Нині проблеми селективної стеганографії потребують подальшого розвитку й удосконалення: відсутні загальні принципи селекції; часто селекція відбувається в просторовій області ЦЗ-контейнера, хоча для стеганоперетворення може використовуватися частотна, області сингулярного, спектрального розкладання матриці тощо; практично не піднімається питання синхронізації – забезпечення механізмів, які дозволяють ідентифікувати початок і послідовність використаних для вбудови додаткової інформації пікселів / блоків контейнера у організації процесу декодування, без яких останній є неможливим.



Метою статті є підвищення ефективності роботи стеганографічної системи шляхом забезпечення синхронізації в умовах застосування селекції блоків, загальні принципи якої запропоновано в дослідженнях (Kobozieva et al., 2025; Бобок, & Кобозева, 2025).

Під ефективністю стеганосистеми в роботі розуміють забезпечення надійності сприйняття стеганоповідомлення, кількісною оцінкою якої є поширений в області стеганографії показник візуального спотворення контейнера – пікове відношення "сигнал–шум" PSNR (Конахович, Прогонов, & Пузиренко, 2018).

Для досягнення мети в роботі розв'язують такі задачі: вибір принципу синхронізації в умовах селекції блоків (Kobozieva et al., 2025; Бобок, & Кобозева, 2025); розроблення відповідного методу, що реалізує обраний спосіб синхронізації.

Методи

В роботі використовують такі підходи та методи дослідження:

1. Аналіз відкритих джерел та огляд сучасного стану селективної стеганографії для визначення існуючих тут проблем.

2. Загальний підхід до аналізу інформаційних систем (ЗПАІС) заснований на теорії збурень і матричному аналізі, який слугує базою для дослідження впливу стеганоперетворення на блоки зображення-контейнера, зокрема на їхні сингулярні числа, та використовується для вибору критерію селекції блоків.

3. Методи теорії збурень для оцінювання можливих меж змін сингулярних чисел блоків зображення.

4. Селективні методи вибору блокового стеганошляху.

5. Методи сучасної стеганографії у проведенні обчислювальних експериментів.

Огляд літератури. До стеганосистеми нині висувають значну кількість вимог, причому однією з основних є вимога забезпечення надійності сприйняття стеганоповідомлення, яка полягає в тому, що візуально воно не повинно відрізнятися від порожнього контейнера. Шляхи задоволення саме цієї вимоги за допомогою селективної стеганографії розглянуто в цій статті.

На сьогодні найчастіше селекція областей ЦЗ-контейнера для вбудови ДІ зводиться до пошуку в ньому контурів. У роботі (Kadhim, Premaratne, & Vial, 2018) запропоновано метод, заснований на виявленні контурів у ЦЗ за допомогою комплексного вейвлет-перетворення з подвійним деревом. Для стеганоперетворення використовують коефіцієнти дискретного вейвлет-перетворення. Високочастотні області зображення ідентифікують за методом виявлення контурів Кенні. Переважна більшість бітів ДІ вбудовується в сильно текстуровані області, завдяки чому візуальні спотворення ЦЗ в результаті стеганоперетворення є незначними.

Вибір саме контурів у ЦЗ-контейнері як найбільш сприятливої області для вбудови ДІ для забезпечення надійності сприйняття стеганоповідомлення зроблено також у роботах (Ghosal, Chatterjee, & Sarkar, 2021; Ray, Mukhopadhyay, & Hossain, 2021; Sarrafpour et al., 2021), де для пошуку використано різні підходи і принципи. Наприклад, у дослідженні (Ghosal, Chatterjee, & Sarkar, 2021) застосовано детектор контурів Кірша, причому для безпосередньої вбудови ДІ тут використано метод модифікації найменшого значущого біта – LSB-метод (Aslam, 2022), що забезпечує збереження надійності сприйняття формованого стеганоповідомлення в умовах порівняно значної пропускну здатності прихованого каналу.

В роботі (Ray et al., 2021) представлено метод, заснований на глибокому навчанні: автори пропонують використовувати згорткову нейронну мережу з детектором контурів, що дозволяє їм віднайти та зберігати більшу кількість пікселів, що належать контурам, порівняно з традиційними алгоритмами виявлення контурів. Розроблену в роботі модель детектора застосовано для отримання мапи контурів. Автори стверджують, що запропонований ними метод дає змогу зберегти надійність сприйняття стеганоповідомлення у разі порівняно значної пропускну здатності прихованого каналу.

Sarrafpour et al. (2021) запропоновано стеганометод для приховування інформації в RGB-зображеннях. Тут пікселі ЦЗ-контейнера, що належать контурам, виявляються за допомогою фільтра Собеля. Саме в ці обрані пікселі (в синю складову) за допомогою LSB-методу відбувається вбудова бітів ДІ. Як стверджують автори, запропонований селективний алгоритм дозволяє уникнути спотворень ЦЗ в результаті стеганоперетворення, має незначну обчислювальну складність.

Зауважимо, що зведення результатів селекції до вибору областей контурів ЦЗ-контейнера для вбудови ДІ, хоча і з урахуванням особливостей системи людського зору, є природним для забезпечення візуальної якості отриманого стеганоповідомлення, проте не зможе забезпечити в загальному випадку якісне використання як контейнера випадкового ЦЗ, адже воно може виявитися таким, де перепади значень яскравості в межах усього зображення будуть зовсім незначними – воно може взагалі не містити контурів. За досить незначних перепадів значень яскравості методи виявлення контурів часто виявляються неспроможними виявити ці перепади. В такому випадку пропускну здатність формованого стеганографічного каналу в умовах означеної селекції прямуватиме до нуля, що є неприйнятним для прихованого каналу зв'язку.

Зазначимо, що в жодній із розглянутих вище робіт, як і в переважній більшості наявних статей, присвячених селективній стеганографії, питання синхронізації не піднімається.

Нині найпоширенішими стеганографічними методами є блокові, через багато причин, серед яких: необхідність ефективного забезпечення стійкості стеганометоду до атаки стиском, забезпечення незначної обчислювальної складності, наявність можливості розпаралелювання тощо, що не можна не враховувати в селективній стеганографії. У роботі (Saeed, 2020), наприклад, запропоновано метод, який передбачає попереднє розбиття ЦЗ-контейнера на локальні області – блоки. Метод містить три етапи: сегментація контейнера, визначення ступеня "складності" пікселів, вбудова ДІ. "Складність" пікселів розраховують на основі запропонованого критерію складності для блока. До блока застосовують фільтри високих частот, результатом чого є вісім отриманих залишкових відгуків, з урахуванням блока пікселів призначається один із дев'яти рівнів складності, на основі чого пікселі впорядковуються за пріоритетом складності від найвищого до найнижчого. Після цього виконують вбудову даних для



відповідного рівня складності. В статті зазначено, що експериментальні результати підтверджують збереження візуальної якості стеганоповідомлень.

Схема блокової селективної стеганографії з використанням текстурованих областей, зокрема країв і навколишніх пікселів, пропонується в (Laishram, & Tuithung, 2021), де блоки є результатом попереднього розбиття матриці ЦЗ-контейнера. Запропонована схема і побудований на її основі метод розглядають виключно просторову область контейнера для стеганоперетворення, динамічно обираючи для цього певний "region", причому цей вибір робиться серед пікселів в межах блока. Метод забезпечує вбудову ДІ з мінімальними викривленнями ЦЗ-контейнера в результаті стеганоперетворення навіть за значної пропускної спроможності прихованого каналу за рахунок убудови трьох бітів ДІ в крайових пікселях із використанням мінімальної середньоквадратичної похибки: для відповідного стеганоповідомлення значення PSNR становить 61–65 dB навіть в умовах 80 % модифікованих блоків.

Зазначимо, що для розглянутих блокових селективних методів питання синхронізації теж залишилось поза увагою авторів.

В селективній стеганографії використовують велику кількість різноманітних принципів. Наприклад в роботі (Naji et al., 2020) селективний метод вибору області ЦЗ-контейнера для вбудови ДІ засновано на розумінні авторів, що особливу роль у ЦЗ відіграють людські обличчя, які, зазвичай, є фокусом зображення у випадку візуального сприйняття. Автори пропонують підхід, який асоціює в собі метод для визначення облич, результатом чого є автоматична побудова "мапи шкіри" для ЦЗ, з подальшою вбудовою ДІ в області, що не містять обличчя, методом LSB. Під автоматичною побудовою "мапи шкіри" автори розуміють визначення того, чи є на зображенні будь-які потенційні області, що містять шкіру людини, і якщо вони існують, то на ЦЗ виявляються місця їхнього розташування. Цей процес по суті є сегментацією. Виявлені області можуть включати будь-які відкриті частини людини – обличчя, плечі, руки, ноги тощо, де наявна шкіра. Хоча запропонований тут селективний метод дозволяє забезпечити значні значення PSNR, він, по-перше, спрямований на вузький клас контейнерів, адже, якщо говорити про застосування випадкового контейнера, він не обов'язково буде містити в межах своєї сцени людину; по-друге, процес організації декодування ДІ з необхідною синхронізацією тут, як і скрізь вище, ніяк не обговорюється.

Mohammed (2023) запропонував селективний метод, де як контейнер використано кольорове ЦЗ. Вибір області вбудови ДІ в межах контейнера відбувається із застосуванням градієнтно-селективних кривих Безьє, що дозволяє зберегти надійність сприйняття стеганоповідомлення навіть у разі збільшення пропускної спроможності прихованого каналу, для організації якого використовується LSB-метод.

Зауважимо, що запропоновані в розглянутих вище роботах сучасні селективні методи частіше за інші стеганометоди для безпосередньої вбудови ДІ використовують метод модифікації найменшого значущого біта, який сам по собі є таким, що забезпечує надійність сприйняття стеганоповідомлення (і значне значення PSNR) і без будь-якої селекції, що не дає можливості говорити тут про збереження хороших декларованих авторами характеристик запропонованих методів в умовах використання інших стеганоалгоритмів. На жаль, хоча нині селективний стеганографії приділяється значна увага, ця увага спрямована, як свідчать літературні джерела, на розв'язання окремих конкретних задач, які дуже часто стосуються селекції під конкретні стеганографічні методи, без отримання загальних принципів вибору, що знижує цінність наведених розробок. Очевидно, розглядаючи проблеми селективної стеганографії, можна дійти висновку, що бажанням тут є забезпечення можливості такого вибору сприятливих для вбудови ДІ областей контейнера, який не залежав би від способу і специфіки безпосереднього стеганоперетворення. Таке бажання для забезпечення надійності сприйняття стеганоповідомлення є природним і таким, що принципово реалізується, адже з урахуванням специфіки системи людського зору найменша ймовірність виникнення на ЦЗ помітних артефактів забезпечується змінами в високочастотній складовій зображення. Такий підхід застосовано нещодавно в (Kobozieva et al., 2025; Бобок, & Кобозева, 2025), де на базі ЗПАІС запропоновано методи вибору блоків ЦЗ-контейнера, отриманих шляхом стандартного розбиття його матриці, для вбудови ДІ з метою підвищення ймовірності збереження надійності сприйняття формованого стеганоповідомлення, що кількісно оцінювалося збільшенням значення PSNR, порівняно з випадковим використанням блоків. Запропоновані методи виконують ефективну селекцію блоків контейнера незалежно від використовуваного для безпосередньої вбудови ДІ стеганометоду. Основною кількісною характеристикою блока обрано нормовану відокремленість його максимального сингулярного числа. Такий критерій дав змогу значно підвищити PSNR порівняно з довільним вибором блоків, але питання організації процесу декодування тут теж залишилось не висвітленим.

Отже, на основі аналізу відкритих джерел можна зробити висновки, що сучасна селективна стеганографія має певну низку нерозв'язаних проблем, серед яких однією з основних є проблема синхронізації, без розв'язання якої жоден із запропонованих методів не може використовуватися на практиці.

Результати

З урахуванням поширеності в сучасному інформаційному середовищі блокових стеганографічних методів, увага в роботі приділяється саме селекції блоків.

Як уже зазначено, підвищення ефективності роботи довільного стеганографічного методу з випадковим контейнером може бути забезпечено шляхом використання вибору блоків матриці контейнера за певним критерієм для вбудови ДІ. Однак цей вибір блоків, а також визначення їхньої послідовності для стеганоперетворення, результат чого далі будемо називати блоковим стеганошляхом, що робиться для оригінального контейнера, є інформацією, необхідною у декодуванні ДІ приймальною стороною. Після вбудови ДІ обраний до стеганоперетворення набір і порядок блоків у стеганоповідомленні може бути порушений: якщо критерій відбору блоків є деякою його кількісною характеристикою, то після вбудови ДІ, що в загальному випадку збурить блок, ця кількісна характеристика з великою ймовірністю зміниться, що призведе до значного утруднення / неможливості виявлення цих блоків і їх безпосереднього порядку в стеганоповідомленні для декодування ДІ. Отже, маємо таку суперечність: селективна стеганографія забезпечує поліпшення якості для результату вбудови ДІ (якості стеганоповідомлення), але при



цьому призводить до значного утруднення організації процесу декодування ДІ, що залежить від синхронізації, оскільки у процесі використання селективної стеганографії в загальному випадку неможливо передбачити, яким саме буде блоковий стеганошлях для певного ЦЗ за певного кількісного критерію.

Для розв'язання проблеми синхронізації принципово можливі три шляхи:

1. Розглядати отриманий у результаті селекції блоків у контейнері блоковий стеганошлях як частину секретного ключа.

2. Для кожного блока зберігати інформацію про його задіяність / незадіяність у стеганоперетворенні та його порядок в обраній послідовності блоків у межах отриманого стеганоповідомлення за допомогою, наприклад, стеганографічної вбудови цієї інформації (далі – службової інформації) в контейнер разом із ДІ (корисною інформацією).

3. Проводити вбудову ДІ у такий спосіб, щоб визначальний параметр відбору блока (критерій селекції) залишався незмінним після стеганоперетворення, що дозволить приймальній стороні відновити стеганошлях.

Розглянемо детально кожний із цих варіантів для вибору найбільш прийняттого.

Перший спосіб у загальному випадку явно утруднює використання випадкового контейнера, для якого вибір блоків (визначення їхніх пріоритетів для вбудови ДІ) не можна визначити і передати заздалегідь як частину секретного ключа отримувачу, зокрема й коли критерієм селекції є деяка кількісна характеристика блока, що потребує обчислення у процесі вибору. Тут процес стеганоперетворення і передача частини секретного ключа отримувачу по захищеному каналу повинні відбуватися майже одночасно, що незручно, а іноді й неможливо. Крім того, наявність захищеного каналу зв'язку, що передбачається для передачі частини секретного ключа практично під час пересилання стеганоповідомлення, становить під питання доцільність організовувати прихований канал зв'язку в наявних умовах, принаймні, якщо обсяг частини секретного ключа, пов'язаної саме з результатами селекції блоків, буде порівнянний з обсягом ДІ, що є цифровим представленням секретного повідомлення.

Другий спосіб. Його реалізація передбачає додатково вбудову службової інформації в контейнер разом із корисною ДІ. Такий підхід значно зменшить обсяг корисної ДІ, що може бути передана, і для пересилання якої взагалі створювався стеганографічний канал. Справді, збурення контейнера, що можна розглядати в загальному випадку як непрямий показник величини сумарної пропускної спроможності прихованого каналу зв'язку (для передачі корисної і службової інформації), є обмеженням завдяки необхідності забезпечення надійності сприйняття отриманого стеганоповідомлення. Причому навіть незначне збільшення загального обсягу інформації, що вбудовується в контейнер, завдяки вбудові службової інформації, може призвести до значного якісного погіршення візуального сприйняття стеганоповідомлення, що умовно продемонстровано на рис. 1.



Рис. 1. Ілюстрація залежності ймовірності збереження надійності сприйняття стеганоповідомлення від пропускної спроможності прихованого каналу зв'язку

Третій спосіб. Зазначимо, що в загальному випадку принципово неможливо дати рекомендації, яким саме способом можна забезпечити незмінність визначального параметра блока після стеганоперетворення. Це вирішуватиметься для кожного параметра, визначального для селективного блокового вибору, окремо. Найчастіше таке забезпечення вимагатиме певної корекції визначального критерію у блоці після стеганоперетворення. Саме такий варіант забезпечення синхронізації в умовах селективної стеганографії, враховуючи наведене вище, автори вважають пріоритетним. Один із варіантів розглядається в цій роботі.

Нещодавно на основі теорії збурень і матричного аналізу розроблено ЗПАІС, який добре зарекомендував себе у розв'язанні задач захисту інформації (Браїловський та ін., 2021). Саме цей підхід став базовим для селективних методів, запропонованих в (Kobozieva et al., 2025; Бобок, & Кобозева, 2025), забезпечуючи їхню ефективність, а тому очевидно є доцільність використання ЗПАІС для розв'язання задач цієї роботи.

Нехай матриця F ЦЗ-контейнера розбивається стандартно на непересічні $l \times l$ -блоки. Окремий такий блок далі позначатимемо B . Результат вбудови ДІ в блок відповідно до ЗПАІС можна представити у загальному вигляді як: $B_S = B + \Delta B$, де B_S – блок стеганоповідомлення, ΔB – матриця збурення B у результаті стеганоперетворення. Матриця блока однозначно визначається повним набором параметрів – сингулярних чисел (СНЧ) і сингулярних векторів (СНВ) (Браїловський та ін., 2021), отриманих у результаті нормального сингулярного розкладання (Браїловський та ін., 2021):

$$B = U \Sigma V^T, \quad (1)$$

де U, V – ортогональні матриці, стовпці яких $(u_1, u_2, \dots, u_l)$, $(v_1, v_2, \dots, v_l)$ є відповідно лівими і правими СНВ B , ліві СНВ є лексико-графічно додатними, а $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_l)$ – діагональна матриця СНЧ, для яких виконується співвідношення:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_l \geq 0. \quad (2)$$



Селекція блоків у (Kobozieva et al., 2025; Бобок, & Кобозєва, 2025) виконується відповідно до кількісного показника: нормованої відокремленості максимального СНЧ, що визначається згідно з формулою

$$svdgap_n(1) = \bar{\sigma}_1 - \bar{\sigma}_2, \quad (3)$$

де

$$\bar{\sigma}_i = \frac{\sigma_i}{\|\sigma\|} \quad (4)$$

– нормовані СНЧ, $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_l)$ – вектор СНЧ В, $\|\sigma\|$ – евклідова векторна норма.

Результат убудови ДІ у блок згідно із ЗПАІС можна представити у вигляді сукупності збурень СНЧ і СНВ (блоків) матриці контейнера, тобто у вигляді сукупності збурень сингулярних трійок (σ_i, u_i, v_i) , $i = 1, l$. Після вбудови ДІ значення всіх / деяких СНЧ, СНВ блока В зміняться, що в загальному випадку призведе і до зміни значення (3), навіть тоді, коли самі значення σ_1, σ_2 будуть незмінними, оскільки збуриться $\|\sigma\|$. Відповідно до обраного підходу до забезпечення синхронізації, після вбудови ДІ СНЧ блока стеганоповідомлення B_S , що далі позначатимуться як

$$\sigma_1^{(S)}, \sigma_2^{(S)}, \dots, \sigma_l^{(S)}, \quad (5)$$

буде потрібно додатково змінити так, щоб залишити незмінним значення (3) для B_S :

$$\bar{\sigma}_1^{(S)} - \bar{\sigma}_2^{(S)} = \bar{\sigma}_1 - \bar{\sigma}_2 = svdgap_n(1) = P, \quad (6)$$

де $\bar{\sigma}_i^{(S)} = \frac{\sigma_i^{(S)}}{\|\sigma^{(S)}\|}$, $\sigma^{(S)} = (\sigma_1^{(S)}, \sigma_2^{(S)}, \dots, \sigma_l^{(S)})$, P – позначення, що використовується далі для зручності.

Зазначимо, що будь-які додаткові збурення для забезпечення синхронізації, виразом яких буде зміна СНЧ (5), в ідеалі не повинні торкнутися вбудованої ДІ, тобто призвести до додаткових збурень тих СНЧ, в первісних збуреннях яких міститься вбудована ДІ. В загальному випадку це гарантувати неможливо: якщо вбудова ДІ відбувається не в області сингулярного розкладання матриці, де є безпосередній контроль над тим, у збуреннях яких саме параметрів повного набору (СНЧ, СНВ) вона буде міститися, то не можна розраховувати на те, що якісь із цих параметрів не збуряться (більше чи менше) в результаті стеганоперетворення. В загальному випадку маємо як формальне представлення ДІ сукупність збурень усіх СНЧ і СНВ. Це означає, що будь-яке додаткове збурення СНЧ після вбудови ДІ для забезпечення (6) може призвести до часткового спотворення вбудованої ДІ, представляючи фактично атаку проти вбудованого повідомлення. Оскільки наголос у роботі робиться на забезпеченні надійності сприйняття стеганоповідомлення, то при вбудові ДІ, враховуючи наявні достатні умови такого забезпечення, отримані в межах ЗПАІС, збурені будуть сингулярні трійки, що відповідають найменшим СНЧ блока, причому з великою імовірністю $\sigma_1^{(S)} = \sigma_1$, $\sigma_2^{(S)} = \sigma_2$, а при декодуванні ДІ з блока B_S вбудовану ДІ збурення σ_1, σ_2 в собі нести не будуть. З огляду на це, забезпечення (6) пропонується робити шляхом модифікації $\sigma_1^{(S)}, \sigma_2^{(S)}$ у блоці стеганоповідомлення.

Оскільки $\sigma_1^{(S)}$ відповідає перша сингулярна трійка B_S , яка несе в собі, головню, інформацію про низькочастотну складову блока, то з урахуванням вимоги забезпечення надійності сприйняття стеганоповідомлення додаткові зміни $\sigma_1^{(S)}$ є небажаними. Змінам (за необхідності) для забезпечення (6) пропонується піддавати $\sigma_2^{(S)}$. Незважаючи на те, що сингулярна трійка $(\sigma_2^{(S)}, u_2^{(S)}, v_2^{(S)})$ теж може нести значний обсяг інформації як про середньо-, так і про низькочастотну складову, модифікація $\sigma_2^{(S)}$ з великою імовірністю не порушить надійність сприйняття стеганоповідомлення. Справді, зміни найменших СНЧ при вбудові ДІ не можуть бути значними, оскільки самі їхні значення, а тому й відповідні відокремленості є малими, тому для встановлення (6) зміни для $\sigma_2^{(S)}$ теж будуть незначні в силу доброї обумовленості всіх складових сингулярного спектра (Браїловський та ін., 2021), що не призведе до виникнення артефактів у блоці.

Позначимо шукане значення зміненого $\sigma_2^{(S)}$ через α . Тоді для визначення α із (6) отримаємо рівняння

$$\bar{\sigma}_1^{(S)} - \frac{\alpha}{\|\sigma^{(S)}\|} = P. \quad (7)$$

У випадку евклідової векторної норми, яку найчастіше застосовують у задачах стеганографії, співвідношення (7) матиме вигляд

$$\frac{\sigma_1^{(S)} - \alpha}{\|\sigma^{(S)}\|} = \frac{\sigma_1^{(S)} - \alpha}{\sqrt{(\sigma_1^{(S)})^2 + \alpha^2 + \dots + (\sigma_l^{(S)})^2}} = P,$$

представляючи квадратне відносно α рівняння, дискримінант якого D визначають так:

$$D = 4(\sigma_1^{(S)})^2(1 - (1 - P^2)^2) + 4P^2(1 - P^2)((\sigma_3^{(S)})^2 + \dots + (\sigma_l^{(S)})^2). \quad (8)$$

З огляду на те, що, за визначенням P , $0 \leq P \leq 1$, а, як буде показано нижче, $0 < P < 1$, значення (8) є додатним, із чого отримаємо можливі значення для α :

$$\alpha = \frac{\sigma_1^{(S)} \pm \sqrt{(\sigma_1^{(S)})^2(1 - (1 - P^2)^2) + P^2((\sigma_3^{(S)})^2 + \dots + (\sigma_l^{(S)})^2)}}{1 - P^2}, \quad (9)$$

але, враховуючи, що $\sigma_1^{(S)} \geq \alpha$, з (9) обираємо

$$\alpha = \frac{\sigma_1^{(S)} - \sqrt{(\sigma_1^{(S)})^2(1 - (1 - P^2)^2) + P^2((\sigma_3^{(S)})^2 + \dots + (\sigma_l^{(S)})^2)}}{1 - P^2}, \quad (10)$$

оскільки в іншому варіанті додавання до $\sigma_1^{(S)}$ додатного значення, а потім ділення на $1 - P^2$, для якого виконується

$$0 < 1 - P^2 < 1, \quad (11)$$

призведе до остаточного результату, що перевершує $\sigma_1^{(S)}$. Подвійна нерівність (11) є строгою. Дійсно, якщо припустити можливість того, що $0 = 1 - P^2$, тобто $P = 1$, то це означатиме з (3), що $\bar{\sigma}_1 - \bar{\sigma}_2 = \frac{\sigma_1 - \sigma_2}{\|\sigma\|} = 1$, тобто

$(\sigma_1 - \sigma_2)^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_l^2$, звідки маємо

$$-2\sigma_1\sigma_2 = \sigma_3^2 + \dots + \sigma_l^2. \quad (12)$$

Враховуючи, що для блоків оригінальних ЦЗ співвідношення (2) на практиці набуває вигляду (Bobok, & Kobozieva, 2024):

$$\sigma_1 \gg \sigma_2 \geq \dots \geq \sigma_l \geq 0, \quad (13)$$



із співвідношення (12) отримуємо систему

$$-\{ -2\sigma_1\sigma_2 \leq 0, \sigma_3^2 + \dots + \sigma_l^2 \geq 0, \sigma_1 \gg \sigma_2. \quad (14)$$

Система (14) має єдиний розв'язок:

$$\sigma_2 = \dots = \sigma_l = 0. \quad (15)$$

Проте, по-перше, як свідчить практичний досвід, це вкрай рідкісна ситуація для $l \times l$ -блока оригінального ЦЗ, а при $l \geq 8$ вона практично відсутня. По-друге, із співвідношення (15) слідує, що в цьому блоці практично відсутня середньо- і височастотна складові, що принципово не дасть змоги такому блоку бути задіяним у процесі стеганоперетворення в умовах селективної стеганографії з метою забезпечення надійності сприйняття стеганоповідомлення.

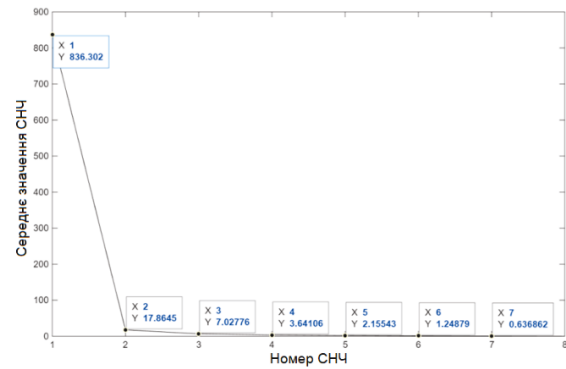
Припустимо тепер, що $1 - P^2 = 1$, тобто $P = 0$, звідки $\sigma_1 = \sigma_2$, що суперечить (13). Отже, (11) виконується.

Покажемо, що розраховане за формулою (10) $\alpha \geq \sigma_3^{(S)}$. Ця умова є важливою, оскільки при її невиконанні, якщо первісні значення $\sigma_2^{(S)}$, $\sigma_3^{(S)}$ поміняються місцями, то кількісне значення критерію селекції буде втрачене. Вбудова ДІ згідно з достатніми умовами забезпечення надійності сприйняття стеганоповідомлення буде проводитися так, щоб її результатом, головню, були збурення сингулярних трійок, що відповідають найменшим СНЧ блока, а для нас важливо – найменших СНЧ. Звісно, в загальному випадку, як зазначено вище, збурення при вбудові ДІ торкнуться (майже) всіх СНЧ блока. Враховуючи результати досліджень, наведені в (Bobok, & Kobozieva, 2025), ці збурення не можуть бути значними, оскільки основний наголос тут – на найменших СНЧ. Отже, сингулярний спектр контейнера від сингулярного спектра стеганоповідомлення буде в цілому відрізняться незначно.

Співвідношення (13) не відображає всю специфіку зменшення СНЧ у блоках ЦЗ. Зазначимо, що хоча швидкість зміни значень СНЧ залежно від номера буде максимальною у процесі переходу від першого до другого, а потім вона суттєво спадає, наближаючись до нуля для СНЧ з найбільшими номерами, ця швидкість залишиться ще значною (порівняно з останніми СНЧ) також при переході від другого до третього, наочна ілюстрація цього представлена на рис. 2. Незначні відмінності сингулярних спектрів контейнера і стеганоповідомлення призведуть до того, що додаткові збурення для $\sigma_2^{(S)}$ не можуть бути значними, а враховуючи, що відстань між другим і третім СНЧ порівняно велика, ці додаткові збурення не призведуть до зміни порядку для модифікованого $\sigma_2^{(S)}$ і $\sigma_3^{(S)}$, що повністю підтверджено результатами обчислювального експерименту.



а



б

Рис. 2. Ілюстрація типової якісної картини залежності значення СНЧ блоків ЦЗ від номера: а – оригінальне ЦЗ; б – графік залежності середніх за ЦЗ значень СНЧ блоків від номера

Отже, незалежно від стеганографічного алгоритму, що використовується для вбудови ДІ, в умовах селекції блоків контейнера, запропонованій в (Kobozieva et al., 2025; Бобок, & Кобозєва, 2025), для забезпечення визначення блокового стеганошляху при декодуванні ДІ згідно із застосуванням критерієм при вбудові, рекомендується проводити модифікацію другого СНЧ блока, що задіяний в стеганоперетворенні, після вбудови ДІ відповідно до (10). Це забезпечить можливість відновлення блокового стеганошляху у разі декодування ДІ без будь-якої додаткової службової інформації.

З урахуванням усього наведеного вище, основні кроки запропонованого стеганоперетворення із забезпеченням синхронізації в умовах селекції блоків (Kobozieva et al., 2025; Бобок, & Кобозєва, 2025) виглядають так.

Крок 1. Вибір блокового стеганошляху в ЦЗ-контейнері провести відповідно до методів селекції блоків, заснованих на використанні як критерію відбору блоків нормованої відокремленості максимального сингулярного числа (Kobozieva et al., 2025; Бобок, & Кобозєва, 2025).

Крок 2. Нехай В – черговий блок контейнера, що належить блоковому стеганошляху. Для блока В:

2.1. Побудувати нормальне сингулярне розкладання (1) $B = U\Sigma V^T$.

2.2. Обчислити нормовані СНЧ (4).

2.3. Обчислити значення P нормованої відокремленості максимального СНЧ (6).

2.4. Вбудувати біт / біти ДІ довільно обраним стеганографічним алгоритмом у блок В. Результат – блок B_S , нормальне сингулярне розкладання якого: $B_S = U_S \Sigma_S V_S^T$, де $\Sigma_S = \text{diag}(\sigma_1^{(S)}, \sigma_2^{(S)}, \dots, \sigma_l^{(S)})$.

2.5. Провести корегування $\sigma_2^{(S)}$ згідно з формулою (10). Результат: $\sigma_2^{(S)k}$. Остаточна матриця СНЧ блока стеганоповідомлення: $\Sigma_{S_k} = \text{diag}(\sigma_1^{(S)}, \sigma_2^{(S)k}, \sigma_3^{(S)}, \dots, \sigma_l^{(S)})$.

2.6. Сформувати остаточний вигляд блока B_{S_k} стеганоповідомлення, що відповідає В: $B_{S_k} = U_S \Sigma_{S_k} V_S^T$.

Крок 3. Якщо вся ДІ вбудована,

то перейти на крок 4,

в іншому разі перейти на крок 2.

Крок 4. Процес стеганоперетворення завершено.



Для оцінювання ефективності запропонованого способу забезпечення можливості відновлення блокового стега-ношляху у разі декодування ДІ в умовах селекції блоків проведено обчислювальний експеримент, в якому задіяно: 300 ЦЗ з бази 4cam_auth (Hsu, & Chang, 2006) (формат Tif), 300 ЦЗ з бази img_Nikon_D70s (Gloe, & Böhme, 2010) (формат Tif), 200 ЦЗ, отриманих непрофесійними відеокамерами (формат Tif), 800 ЦЗ з бази NRCS (U.S. Department of Agriculture, 2025) (формат Jpeg). Як стеганометоди в експерименті використано: один із найстійкіших до атаки стиском стеганографічний метод (Melnik, 2012), який призводить до несистематичного виникнення арте-фактів на ЦЗ-стеганоповідомленні; один із найбільш поширених і таких, що модифікуються, методів – Коха і Жао (Конахович, Прогонов, & Пузиренко, 2018), порушення надійності сприйняття в якому може мати місце у разі збіль-шення значення параметра, що використовується при модифікації коефіцієнтів дискретного косинусного перетво-рення під час убудови біта ДІ в черговий блок. Умови проведення експерименту навмисно взято такі, які точно збігаються з умовами, що були зазначені в (Kobozieva et al., 2025). Результати експерименту наведено в табл. 1 (тут умова "Без обробки блока після стеганоперетворення" передбачає наступну передачу інформації про блоковий стеганощлях як частину секретного ключа).

Таблиця 1
Результати порівняльного оцінювання ефективності запропонованого методу синхронізації і синхронізації,
що забезпечується секретним ключем

Стеганометод	Середнє по експерименту підвищення значення PSNR в умовах селекції блоків (Kobozieva et al., 2025), dB	
	Без обробки блока після стеганоперетворення (забезпечується секретним ключем)	Із запропонованою постобробкою блока
Коха і Жао (Конахович, Прогонов, & Пузиренко, 2018)	5.1	4.3
Метод (Melnik, 2012)	4.8	4.1

Зменшення значення PSNR у разі використання запропонованого в роботі методу порівняно з методом, що не передбачає додаткову обробку для відновлення кількісного значення критерію селекції, є очікуваним, оскільки, як зазначено вище, відновлення критерію призводить до додаткових, хоч і незначних, збурень. Однак цей метод є пріоритетним із погляду організації процесу декодування, оскільки не потребує збільшення обсягу секретного ключа та додаткового збереження службової інформації. Крім того, незначне погіршення PSNR ніяк не відбилося на візу-альному сприйнятті отримуваних стеганоповідомлень, що проілюстровано на рис. 3, де навмисно для порівняння використано ЦЗ з роботи (Kobozieva et al., 2025).

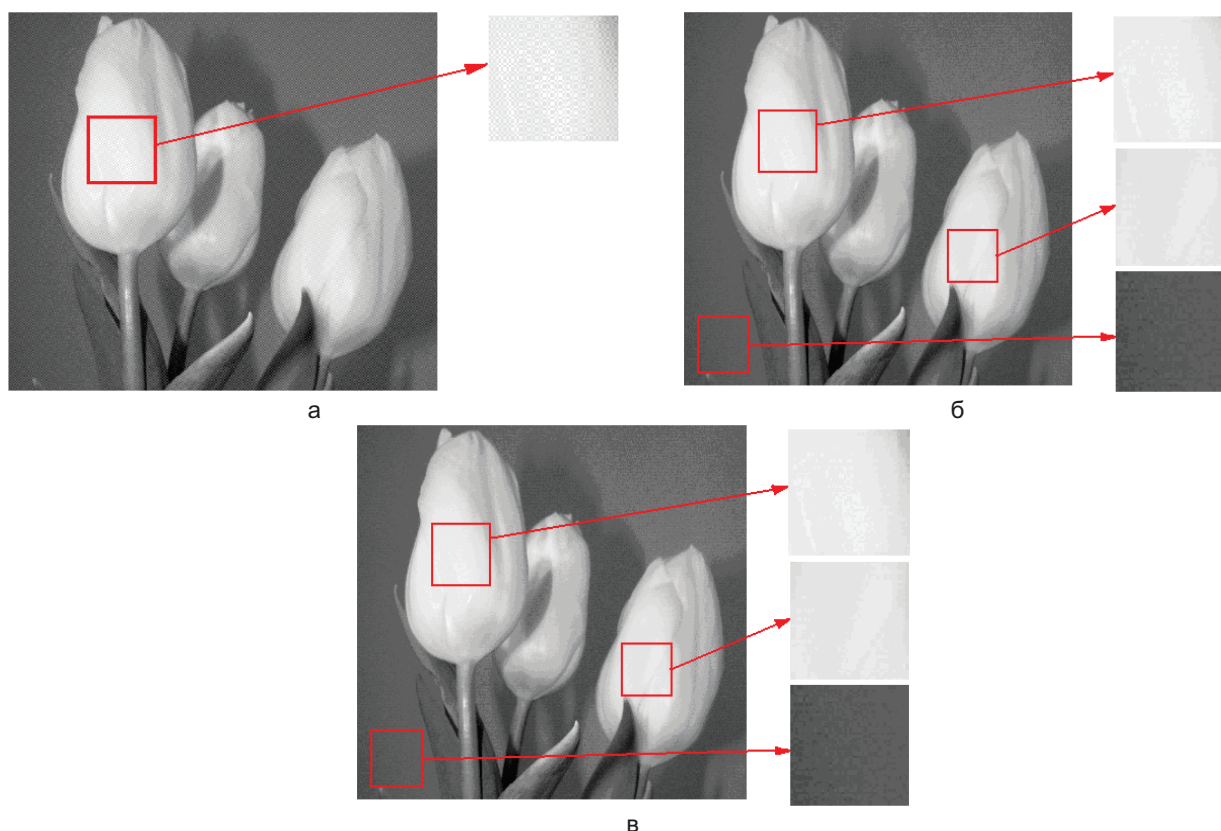


Рис. 3. Стеганоповідомлення, отримані за допомогою стеганометоду (Melnik, 2012),
при формуванні яких блоки для стеганоперетворення: а – обиралися випадковим чином;
б – піддавалися селекції за допомогою методу (Kobozieva et al., 2025) без подальшої будь-якої обробки;
в – піддавалися селекції за методом (Kobozieva et al., 2025) із наступною обробкою запропонованим методом



Дискусія і висновки

У цій роботі розглянуто важливу науково-практичну задачу підвищення ефективності стеганографічної системи шляхом розроблення методу забезпечення синхронізації в умовах застосування селекції блоків, запропонованій у дослідженнях (Kobozieva et al., 2025; Бобок, & Кобозєва, 2025).

В ході проведеного дослідження:

- обґрунтовано доцільність обраного принципу синхронізації: забезпечення незмінності визначального параметра (критерію селекції) блока після стеганоперетворення;
- запропоновано метод, що реалізує обраний спосіб синхронізації в умовах селекції блоків, що не залежить від використовуваного стеганографічного алгоритму, заснований на модифікації другого СНЧ блока стеганоповідомлення, що належить обраному стеганошляху, з метою забезпечення незмінності нормованої відокремленості максимального сингулярного числа – кількісного критерію селекції блоків.

Застосування запропонованого механізму синхронізації в умовах селекції блоків контейнера з метою забезпечення надійності сприйняття стеганоповідомлення, дало змогу підвищити ефективність стеганосистеми на 4.1–4.3 dB, у цьому разі застосовували випадкові контейнери. І хоча такий механізм синхронізації дещо програє в ефективності тому випадку, коли вона забезпечується секретним ключем, але наявний незначний програш ніяк не впливає на якість отриманого стеганоповідомлення, про що свідчать результати обчислювального експерименту, причому обраний спосіб синхронізації не потребує розширення секретного ключа та його додаткової передачі, що одночасно робить запропонований спосіб синхронізації пріоритетним.

Перспективою подальшого дослідження є розроблення блокових селективних методів, що забезпечують стійкість результату селекції в умовах атак проти вбудованого повідомлення. Це питання до цього часу не має загальних принципів розв'язання і практично не досліджується фахівцями в цій області стеганографії.

Внесок авторів: Алла Кобозєва – концептуалізація, написання (перегляд і редагування); Світлана Григоренко – формальний аналіз, програмне забезпечення, валідація даних, написання (оригінальна чернетка).

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Бобок, І., & Кобозєва, А. (2025). Нормована відокремленість максимального сингулярного числа блоку цифрового зображення як показник доцільності його стеганоперетворення. У В. Вичужанін (Ред.), *Системи контролю інформації та інтелектуальні технології. Досягнення та застосування* (с. 58–75). Liha-Pres.
- Брайловський, М., Зибін, С., Кобозєва, А., Хорошко, В., & Хохлачова, Ю. (2021). *Аналіз кіберзахисності інформаційних систем*. ФОП "Ямчинський О. В."
- Конахович, Г., Проконов, Д., & Пузиренко, О. (2018). *Комп'ютерна стеганографічна обробка й аналіз мультимедійних даних*. Alex Print Centre.
- Aslam, M. (2022). Image Steganography using Least Significant Bit (LSB) – A Systematic Literature Review. In *Proceedings of the 2nd International Conference on Computing and Information Technology (ICCIT'2022)* (pp. 32–38). IEEE. <https://doi.org/10.1109/ICCIT52419.2022.9711628>
- Bobok, I., & Kobozieva, A. (2024). Localization of Disturbance Region of Formal Parameters of Steganographic Cover to Ensure Steganosystem Stability. *Radioelectronics and Communications Systems*, 67, 396–409. <https://doi.org/10.20535/S0021347024090024>
- Bobok, I., & Kobozieva, A. (2025). Theoretical Foundations of Digital Content Integrity Expertise. *Problemele Energeticii Regionale*, 1, 105–121. <https://doi.org/10.52254/1857-0070.2025.1-65.08>
- Charan, S. (2023). A Review of the Recent Advances in Image Based Steganography. In *Proceedings of the KILBY 100 7th International Conference on Computing Sciences (ICCS'2023)* (pp. 1–5). Lovely Professional University. <http://dx.doi.org/10.2139/ssrn.4490256>
- Ghosal, S., Chatterjee, A., & Sarkar, R. (2021). Image steganography based on Kirsch edge detection. *Multimedia Systems*, 27, 73–87. <https://doi.org/10.1007/s00530-020-00703-3>
- Gloe, T., & Böhme, R. (2010). The 'Dresden Image Database' for benchmarking digital image forensics. In *Proceedings of the 2010 ACM Symposium on Applied Computing (SAC'10)* (pp. 1584–1590). ACM. <https://doi.org/10.1145/1774088.1774427>
- Hsu, Y., & Chang, S. (2006). Detecting Image Splicing using Geometry Invariants and Camera Characteristics Consistency. In *Proceedings of 2006 IEEE International Conference on Multimedia and Expo* (pp. 549–552). IEEE. <https://doi.org/10.1109/ICME.2006.262447>
- Kadhim, I., Premaratne, P., & Vial, P. J. (2018). Adaptive Image Steganography Based on Edge Detection Over Dual-Tree Complex Wavelet Transform. In D. S. Huang, M. Gromiha, K. Han, & A. Hussain (Eds.), *Lecture Notes in Computer Science: Vol. 10956. Intelligent Computing Methodologies* (pp. 544–550). Springer. https://doi.org/10.1007/978-3-319-95957-3_57
- Kobozieva, A., Bobok, I., Laptiev, O., & Savchenko, V. (2025). Application of Graph Theory to Ensure the Reliability of Steganographic Message Perception in the Creation of a Covert Communication Channel. *CEUR Workshop Proceedings*, 4048, 94–108.
- Laishram, D., & Tuithung, T. (2021). A novel minimal distortion-based edge adaptive image steganography scheme using local complexity. *Multimedia Tools and Applications*, 80, 831–854. <https://doi.org/10.1007/s11042-020-09519-9>
- Melnik, M. (2012). Compression-resistant steganography algorithm. *Information Security*, 2, 99–106.
- Mohammed, S. N. (2023). Color Image Steganography Using Gradient Selective Bezier Curves. *Iraqi Journal of Science*, 64(7), 4525–4541. <https://doi.org/10.24996/ijis.2023.64.7.39>
- Naji, S., Mohaisen, H., Alsaffar, Q., & Jalab, H. (2020). Automatic region selection method to enhance image-based steganography. *Periodicals of Engineering and Natural Sciences*, 8(1), 67–78.
- Ray, B., Mukhopadhyay, S., & Hossain, S. (2021). Image steganography using deep learning based edge detection. *Multimedia Tools and Applications*, 80, 33475–33503. <https://doi.org/10.1007/s11042-021-11177-4>
- Saeed, A. (2020). An Accurate Texture Complexity Estimation for Quality-Enhanced and Secure Image Steganography. *IEEE Access*, 8, 21613–21630. <https://doi.org/10.1109/ACCESS.2020.2968217>
- Sarrafpour, B., Alomirah, R., Sarrafpour, S., & Sharifzadeh, H. (2021). An Adaptive Edge-Based Steganography Algorithm for Hiding Text into Images. In *Proceedings of the 2021 IEEE 19th International Conference on Embedded and Ubiquitous Computing (EUC)* (pp. 109–116). IEEE. <https://doi.org/10.1109/EUC53437.2021.00024>
- Taher, M., Ahmad, A. R. B. H. J., Sami, R., & Mokri, S. (2022). A literature review of various steganography methods. *Journal of Theoretical and Applied Information Technology*, 100(5), 1412–1427.
- U.S. Department of Agriculture. (2025, November 10). *NRCS Photo Gallery*. <https://www.nrcs.usda.gov>

References

- Aslam, M. (2022). Image Steganography using Least Significant Bit (LSB) – A Systematic Literature Review. In *Proceedings of the 2nd International Conference on Computing and Information Technology (ICCIT'2022)* (pp. 32–38). IEEE. <https://doi.org/10.1109/ICCIT52419.2022.9711628>
- Bobok, I., & Kobozieva, A. (2025). Normalized separability of the maximum singular number of a digital image block as an indicator of the feasibility of its steganographic transformation. In V. Vychuzhanin (Ed.), *Information control systems and intelligent technologies: Achievements and applications* (pp. 58–75). Liha-Pres [in Ukrainian].



- Bobok, I., & Kobozeva, A. (2024). Localization of Disturbance Region of Formal Parameters of Steganographic Cover to Ensure Steganosystem Stability. *Radioelectronics and Communications Systems*, 67, 396–409. <https://doi.org/10.20535/S0021347024090024>
- Bobok, I., & Kobozeva, A. (2025). Theoretical Foundations of Digital Content Integrity Expertise. *Problemele Energeticii Regionale*, 1, 105–121. <https://doi.org/10.52254/1857-0070.2025.1-65.08>
- Brailovskyi, M., Zybin, S., Kobozeva, A., Khoroshko, V., & Khokhlachova, Yu. (2021). *Analysis of cybersecurity of information systems*. FOP "Yamchynskiy O. V." [in Ukrainian].
- Charan, S. (2023). A Review of the Recent Advances in Image Based Steganography. In *Proceedings of the KILBY 100 7th International Conference on Computing Sciences (ICCS'2023)* (pp. 1–5). Lovely Professional University. <http://dx.doi.org/10.2139/ssrn.4490256>
- Ghosal, S., Chatterjee, A., & Sarkar, R. (2021). Image steganography based on Kirsch edge detection. *Multimedia Systems*, 27, 73–87. <https://doi.org/10.1007/s00530-020-00703-3>
- Gloe, T., & Böhme, R. (2010). The 'Dresden Image Database' for benchmarking digital image forensics. In *Proceedings of the 2010 ACM Symposium on Applied Computing (SAC'10)* (pp. 1584–1590). ACM. <https://doi.org/10.1145/1774088.1774427>
- Hsu, Y., & Chang, S. (2006). Detecting Image Splicing using Geometry Invariants and Camera Characteristics Consistency. In *Proceedings of 2006 IEEE International Conference on Multimedia and Expo* (pp. 549–552). IEEE. <https://doi.org/10.1109/ICME.2006.262447>
- Kadhim, I., Premaratne, P., & Vial, P. J. (2018). Adaptive Image Steganography Based on Edge Detection Over Dual-Tree Complex Wavelet Transform. In DS. Huang, M. Gromiha, K. Han, & A. Hussain (Eds.), *Lecture Notes in Computer Science: Vol. 10956. Intelligent Computing Methodologies* (pp. 544–550). Springer. https://doi.org/10.1007/978-3-319-95957-3_57
- Kobozeva, A., Bobok, I., Laptiev, O., & Savchenko, V. (2025). Application of Graph Theory to Ensure the Reliability of Steganographic Message Perception in the Creation of a Covert Communication Channel. *CEUR Workshop Proceedings*, 4048, 94–108.
- Konakhovych, H., Prohonov, D., & Puzyrenko, O. (2018). *Computer steganographic processing and analysis of multimedia data*. Alex Print Centre [in Ukrainian].
- Laishram, D., & Tuithung, T. (2021). A novel minimal distortion-based edge adaptive image steganography scheme using local complexity. *Multimedia Tools and Applications*, 80, 831–854. <https://doi.org/10.1007/s11042-020-09519-9>
- Melnik, M. (2012). Compression-resistant steganography algorithm. *Information Security*, 2, 99–106.
- Mohammed, S. N. (2023). Color Image Steganography Using Gradient Selective Bezier Curves. *Iraqi Journal of Science*, 64(7), 4525–4541. <https://doi.org/10.24996/ijis.2023.64.7.39>
- Naji, S., Mohaisen, H., Alsaffar, Q., & Jalab, H. (2020). Automatic region selection method to enhance image-based steganography. *Periodicals of Engineering and Natural Sciences*, 8(1), 67–78.
- Ray, B., Mukhopadhyay, S., & Hossain, S. (2021). Image steganography using deep learning based edge detection. *Multimedia Tools and Applications*, 80, 33475–33503. <https://doi.org/10.1007/s11042-021-11177-4>
- Saeed, A. (2020). An Accurate Texture Complexity Estimation for Quality-Enhanced and Secure Image Steganography. *IEEE Access*, 8, 21613–21630. <https://doi.org/10.1109/ACCESS.2020.2968217>
- Sarrafpour, B., Alomirah, R., Sarrafpour, S., & Sharifzadeh, H. (2021). An Adaptive Edge-Based Steganography Algorithm for Hiding Text into Images. In *Proceedings of the 2021 IEEE 19th International Conference on Embedded and Ubiquitous Computing (EUC)* (pp. 109–116). IEEE. <https://doi.org/10.1109/EUC53437.2021.00024>
- Taher, M., Ahmad, A. R. B. H. J., Sami, R., & Mokri, S. (2022). A literature review of various steganography methods. *Journal of Theoretical and Applied Information Technology*, 100(5), 1412–1427.
- U.S. Department of Agriculture. (2025, November 10). *NRCS Photo Gallery*. <https://www.nrcs.usda.gov>

Отримано редакцією журналу / Received: 14.11.25

Прорецензовано / Revised: 17.11.25

Схвалено до друку / Accepted: 22.12.25

Alla KOBOZIEVA, DSc (Engin.), Prof.
ORCID ID: 0000-0001-7888-0499
e-mail: alla_kobozeva@ukr.net
Odesa National Maritime University, Odesa, Ukraine

Svitlana HRYHORENKO, PhD (Engin.)
ORCID ID: 0009-0006-4551-8243
e-mail: sngrygorenko@gmail.com
Odesa Polytechnic National University, Odesa, Ukraine

THE SYNCHRONIZATION PROBLEM IN SELECTIVE STEGANOGRAPHY

Background. Selective steganography is currently a powerful tool for improving the characteristics of a steganographic system, particularly when using a random container. One of the main unresolved problems of modern selective steganography remains the synchronization problem. The goal of the work is to increase the efficiency of the stego-system by ensuring synchronization under block selection, the general principles of which were proposed by the authors earlier.

Methods. The following were used: selective methods for choosing the block stego-path; a general approach to analyzing the state of information systems to study the influence of stego-transformation on container image blocks and to select the block selection criterion; perturbation theory methods for estimating the possible limits of changes in block singular values; and modern steganography methods.

Results. The feasibility of the synchronization principle was substantiated by ensuring the immutability of the determining parameter (selection criterion) of the block after stego-transformation. A method is proposed that implements the chosen synchronization approach under block selection, independent of the steganographic method used, based on modifying the block's singular spectrum to ensure the immutability of the normalized separation of the maximum singular value – a quantitative selection criterion.

Conclusions. The application of the proposed synchronization mechanism under container block selection, aimed at ensuring the reliability of stego-message perception, allowed the stego-system efficiency to be increased by 4.1–4.3 dB when using random containers. Moreover, the proposed synchronization method does not require extending the secret key and its additional transmission, nor the storage of additional information about the obtained stego-path.

Keywords: selective steganography, synchronization, digital image, singular value, random container.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



Андрій ДУДНИК, д-р техн. наук, доц.

ORCID ID: 0000-0003-1339-7820

e-mail: andrii.dudnik@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Сергій ВИГОВСЬКИЙ, асп.

ORCID ID: 0009-0007-7868-4923

e-mail: vygovsky.sa@gmail.com

Приватне акціонерне товариство "Вищий навчальний заклад
«Міжрегіональна академія управління персоналом», Київ, Україна

Олександр ТОРОШАНКО, канд. техн. наук

ORCID ID: 0000-0002-2354-0187

e-mail: oleksandr.toroshanko@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Андрій ФЕСЕНКО, канд. техн. наук, доц.

ORCID ID: 0000-0001-5154-5324

e-mail: andrii.fesenko@npp.kai.edu.ua

Державне некомерційне підприємство Державний університет
"Київський авіаційний інститут", Київ, Україна

МЕТОД ІНТЕГРАЦІЇ КРИПТОПРОЦЕСОРА У РОБОТИЗОВАНУ СЕНСОРНУ МЕРЕЖУ СИСТЕМИ РОЗМІНУВАННЯ

Вступ. Сучасні технології розмінування передбачають використання складних роботизованих систем, до яких входять безпілотні літальні апарати (БПЛА), наземні мобільні платформи, сенсорні мережі, модулі геопросторового аналізу та системи командно-штабного керування. Надійність взаємодії між цими складовими є критично важливою, оскільки компрометація каналів зв'язку або підміна даних може призвести до хибного картування мінної небезпеки, втрати платформи або навіть до неконтрольованого підриву.

Методи. Проведено порівняльний аналіз програмних і апаратних методів захисту, наведено математичні моделі криптостійкості, оцінки ймовірності MITM-атак, втрат пакетів у середовищі РЕБ, моделі стохастичної ротації ідентифікаторів і характеристики продуктивності. Експериментальна частина включає моделювання 12 комплексних параметрів: безпека, стійкість, затримки, втрати пакетів, захист каналів, криптостійкість, приватність, уникнення трекінгу, нечутливість до компрометації, апаратна продуктивність і стійкість мережі до джемінгу.

Результати. Виявлено покращення параметрів безпеки від 84 % до 99 %, зниження ймовірності MITM на 97 %, підвищення стійкості каналів на 12–20 % і зменшення ризику трекінгу на 100 %. У статті доведено, що інтеграція АТЕСС608А є доцільною та технологічно обґрунтованою для систем розмінування нового покоління.

Висновки. Представлено комплексний підхід до побудови захищеної сенсорної та роботизованої мережі системи розмінування із застосуванням апаратного криптомодуля АТЕСС608А. Запропонована архітектура забезпечує формування апаратного кореня довіри, двофакторне захищене завантаження, динамічну зміну мережевих ідентифікаторів, криптографічне кодування параметрів місії та захищене керування обміном інформації між безпілотними літальними апаратами, наземними роботизованими платформами (UGV), стаціонарними сенсорними вузлами та базовою станцією.

Ключові слова: криптопроцесор, роботизовані системи, сенсорна мережа, IoT.

Вступ

Прискорений розвиток засобів радіоелектронної боротьби (РЕБ) у сучасних конфліктах супроводжується появою нових загроз: атаки MITM, спуфінг GNSS, джемінг радіоканалів, нав'язування підроблених команд, паразитне захоплення телеметрії та зворотного каналу керування. На додачу до цього, програмні методи захисту, що використовуються у традиційних IoT-пристроях або промислових контролерах, характеризуються вразливістю до вилучення ключів із пам'яті, клонування прошивок, відкату версій, маніпуляцій з I²C/SPI-шинами та фізичного доступу до пристрою. Для забезпечення надійності мережі роботизованих систем у бойових умовах необхідно створити стійку до компрометації архітектуру захищеного обміну даними, що базується не на програмному, а на апаратному корені довіри. Одним із найбільш поширених і високоефективних рішень є використання криптографічного модуля АТЕСС608А – елемента безпеки сім'ї CryptoAuthentication™ виробництва Microchip.

Метою цієї статті є розроблення та всебічне обґрунтування архітектури сенсорної мережі та мережі роботизованих платформ системи розмінування, здатної забезпечувати повний цикл криптографічного захисту та довіри між усіма компонентами – від периферійних сенсорів до центральної базової станції.

Огляд літератури. Проблематика захисту мережевих і роботизованих систем, які працюють у складних умовах бойових дій і радіоелектронної боротьби, досить широко висвітлена у сучасній науковій літературі. У роботах із безпеки протоколів передачі даних виділяється проблема атак man-in-the-middle, які за відсутності криптографічного захисту здатні повністю порушити цілісність інформації та підмінити службові повідомлення (Callegati, Cerroni, & Ramilli, 2009). Особливо вразливими є системи, де ключі зберігаються у програмному середовищі, що робить можливим їх вилучення за допомогою RAM-аналізу, DMA-доступу або експлойтів буферів. З іншого боку, апаратні засоби криптографічного захисту, такі як АТЕСС608А, характеризуються високим рівнем стійкості до атак фізичного доступу. Документація Microchip підкреслює, що приватні ключі генеруються і використовуються винятково всередині захищеного елемента, а будь-які спроби витягнути ключ фізично чи логічно є безперспективними (Microchip



Technology Inc. – ATECC608A, 2025). Апаратний захист у поєднанні з можливістю виконання криптографічних операцій ECC, AES та HMAC дає змогу сформувати апаратний корінь довіри для складних систем, включаючи роботизовані комплекси. Інтеграція secure-element у платформу ESP32, описана виробником Espressif, демонструє суттєве прискорення операцій TLS та ECC, а також значне зниження ризику компрометації ключів порівняно з традиційними програмними рішеннями (Espressif Systems – ESP32-WROOM-32SE, 2025).

У рекомендаціях Microchip щодо формування захищеного завантаження наведено сценарії подвійної автентифікації, які можуть бути застосовані для систем, де критичною є цілісність бутлоадера та прошивки (Microchip Technology Inc., 2018, Secure Boot and Firmware Validation with ATECC608A, 2025). У наукових оглядах, присвячених безпеці IoT, багатократно підкреслено, що без апаратного кореня довіри захист ключів є слабкою ланкою, що часто стає точкою компрометації всієї системи (Alaba et al., 2017). Показано, що навіть складні алгоритми, зокрема й ECC або RSA у програмній реалізації на мікроконтролерах, можуть бути зламані методами побічних каналів – аналізом споживання струму, електромагнітного випромінювання або часових характеристик. У роботі (Bhattacharyya et al., 2020) запропоновано використання ECC-схем у поєднанні з апаратними елементами безпеки як оптимальний підхід для промислових сенсорних систем. Дослідження демонструють, що такий комбінований спосіб забезпечує значно вищу стійкість до атак повторного відтворення та підміни, які є характерними для середовищ промислової автоматизації. Аналіз промислових контрольних систем показує, що більшість ICS- та SCADA-платформ критично вразливі до підміни команд через відсутність автентифікації у низькорівневих протоколах (Alrawi et al., 2019). Ці результати важливі для задач розмінування, оскільки підміна команди роботизованої платформи або БПЛА може призвести до неправильного позначення міної загрози або виходу апарата з-під контролю.

У дослідженнях безпеки бездротових мереж також розглядається проблема MAC randomization. Відомо, що непослідовна або часткова ротація ідентифікаторів дозволяє супротивнику відстежувати пристрої за набором побічних характеристик – інтервалами передачі, особливостями стека, мікропатернами часових зсувів (Das, Moses, & Gopalakrishnan, 2018). Ці спостереження прямо вплинули на моделювання алгоритму ротації ID у нашій роботі. У фундаментальній праці "Handbook of Applied Cryptography" наведено формальні докази стійкості криптографічних алгоритмів, що використовуються в ATECC608A, зокрема й ECC-P256 та HMAC-SHA256, які широко застосовують у сучасних апаратних платформах (Menezes, van Oorschot, & Vanstone, 1996). У сучасних UAV-дослідженнях підкреслено, що без апаратного root-of-trust дрони залишаються вразливими до підміни GNSS, інжекції телеметрії, а також до атак віддаленого перехоплення каналів управління (Hazra, Ghosh, & Banerjee, 2023). Схожа позиція наведена у роботах Sensors (MDPI), де пропонується обов'язкове використання апаратної автентифікації для UAV-систем, що працюють у бойових умовах (Ivanov, Li, & Porov, 2022).

У сфері радіоелектронної боротьби системи вважають стійкішими до джемінгу й активних перешкод, коли вони використовують криптографічний фреймінг і заздалегідь підписані пакети, які не можуть бути підмінені противником (Rondon, Conti, & Sirocco, 2021). Це лягло в основу моделювання стійкості каналу в нашій роботі. Суттєвою загрозою є фізичні атаки на апаратність, про які детально пише SANS Institute (Dunning, 2018). У роботі наведено огляд способів витікання ключів через JTAG, SPI або glitch-атаки. Це підтверджує, що програмний захист ключів недостатній для військових роботизованих систем. У статті (Liu, & Li, 2020) наведено сучасні методи керування ідентичностями у децентралізованих IoT-системах, що стало підґрунтям для розроблення моделі динамічної ротації ID. Дослідження (Ristic, Gunatilake, & Griffin 2020) також підкреслює важливість антитрекінгових механізмів для дронів у зоні конфлікту, оскільки противник може використовувати трекінг для наведення РЕБ або кінетичних засобів ураження.

Аналіз існуючих розробок демонструє, що сучасна наукова та технічна спільнота визнає критичну важливість апаратного кореня довіри, динамічної ротації ідентифікаторів, апаратного зберігання ключів, захищеного завантаження та криптографічного фреймінгу. Сукупність проаналізованих джерел підтверджує наукову новизну і практичну значущість запропонованої в статті архітектури.

Методи

Методи, що покладені в основу запропонованої архітектури, базуються на формальному описі криптографічних функцій, моделей ризиків атак, моделей втрат пакетів у середовищі радіоелектронних перешкод, а також стохастичних процесів, що визначають динамічну ротацію ідентифікаторів. Для побудови апаратного кореня довіри використовується механізм генерації приватного ключа всередині secure-element. Оскільки модуль ATECC608A містить власний фізично захищений генератор випадкових чисел, приватний ключ формується за законом $K_{priv} \sim RNG_{SE}$, а відповідний відкритий ключ обчислюється на основі множення точки еліптичної кривої $K_{pub} = K_{priv} \cdot G$. Такий спосіб гарантує, що ключовий матеріал ніколи не існує в оперативній пам'яті мікроконтролера й поготів не передається по жодній шині.

Процедура захищеного завантаження (secure boot) формалізується через механізм перевірки підпису. Бутлоадер або інший критичний компонент ПЗ має власний хеш $H(\text{"bootloader"})$, а secure-element генерує для нього ECDSA-підпис $S = \text{Sign}_{K_{priv}}(H(\text{"bootloader"}))$. Центральний процесор здійснює перевірку автентичності шляхом обчислення $\text{Verify}(S, H(\text{"bootloader"}), K_{pub})$. Подальше шифрування даних у системі описується рівнянням $C = \text{AES}_{K_{session}}(M)$, де M – повідомлення, а C – відповідний шифротекст. Саме апаратна реалізація AES, доступна всередині ATECC608A, забезпечує мінімальний час оброблення та знижує ризики атак бокових каналів.

Однією з ключових функцій запропонованої архітектури є динамічна ротація мережевих ідентифікаторів. Вона формалізується криптографічною функцією HMAC, що враховує унікальний ключ пристрою, випадкове число від координатора, ідентифікатор місії та часову мітку. Отже, новий ID визначається рівнянням

$$ID_{t+1} = \text{HMAC}(K_{device}, \text{nonce} \parallel \text{mission_id} \parallel t).$$

Це означає повну непередбачуваність як для зовнішнього спостерігача, так і для супротивника, що намагається відстежувати платформу або здійснити атаку підміни. Імовірність того, що наступний ID буде передбачений без знання ключа, дорівнює величині порядку 2^{-128} або навіть 2^{-256} залежно від обсягу вхідної ентропії.



Для опису ризиків MITM-атаки вводиться модель імовірності успішної підміни. У програмних системах ця ймовірність визначається сумою ймовірностей витікання ключів і успішної підміни ідентифікаторів:

$$P_{\text{MITM,soft}} \approx P(\text{key leakage}) + P(\text{ID spoofing}).$$

У випадку використання secure-element, де приватний ключ не покидає апаратного середовища, ця ймовірність зменшується майже до нуля:

$$P_{\text{MITM,SE}} = \epsilon \approx 0.01.$$

Стійкість до радіоелектронного придушення оцінюється через модель втрат пакетів у середовищі завад, де базова залежність описується експоненційним законом $L = 1 - e^{-\alpha I}$, де I – інтенсивність завад, а α – параметр середовища. У разі використання криптографічного фреймінгу з апаратною автентифікацією, член α зменшується, що призводить до зниження втрат пакетів:

$$L_{\text{SE}} = 1 - e^{-0.84\alpha I}.$$

Це відповідає покращенню роботи приблизно на 16 % порівняно з програмними методами.

Динамічна ротація ідентифікаторів може бути описана як циклічний марковський процес, де кожен стан ID переходить у наступний ID_{i+1} з імовірністю 1. Матриця переходів у цьому випадку має вигляд

$$P = \begin{pmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \\ 1 & 0 & 0 & \dots \end{pmatrix}$$

Атака на динамічну ротацію можлива лише у випадку правильного вгадування наступного елемента ланцюга, а ймовірність такого передбачення дорівнює $P(\text{prediction}) = 2^{-256}$. Це означає практичну неможливість трекінгу пристрою навіть для супротивника з великою обчислювальною потужністю.

На рис. 1 подано детальний алгоритм ротації мережевої ідентичності між двома взаємодіючими IoT-пристроями, що використовується як основа для формування стійкої до перехоплення та відстеження комунікації в системі. У нашій роботизованій платформі для розмінування цей механізм відповідає за періодичну зміну Bluetooth/MAC-адрес між периферійними сенсорами, модулями розвідки та компаньйон-комп'ютерами, що дозволяє суттєво зменшити ризик трекінгу, повторного відтворення пакетів або підміни комунікацій супротивником.

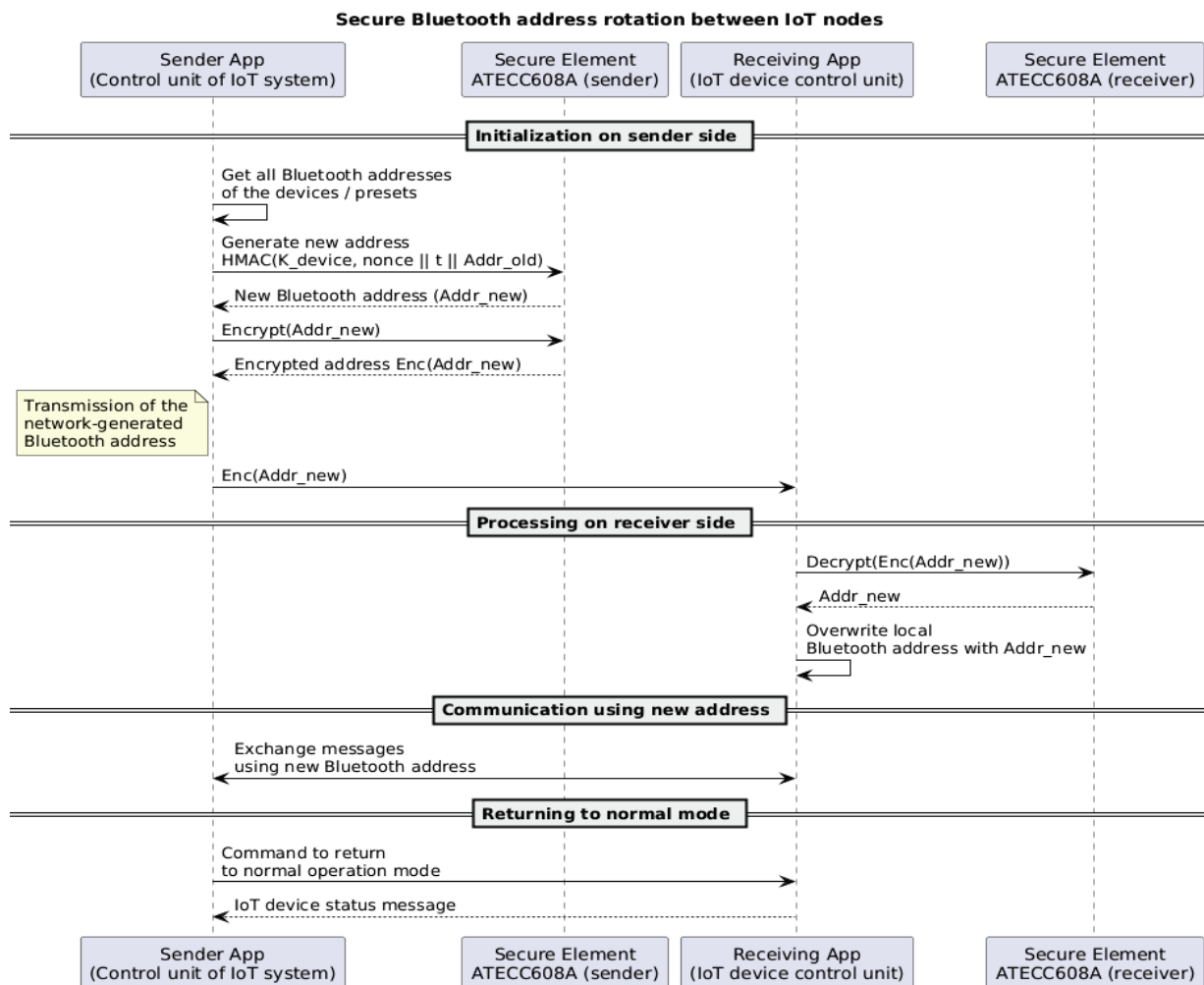


Рис. 1. Алгоритм ротації мережевої архітектури



Алгоритм починається з роботи модуля керування IoT-системи на стороні відправника, який ініціює збір поточних мережевих адрес усіх пристроїв. На основі цих даних генерується нова адреса відповідно до заданого криптографічного правила. У нашій системі ця генерація не відбувається довільно: вона використовує ATECC608A як апаратне джерело ентропії та криптографічну функцію HMAC-SHA256, що гарантує криптографічну непередбачуваність нової адреси. Це дозволяє формально описати нову адресу як

$$\text{Addr}_{\text{new}} = \text{HMAC}(K_{\text{device}}, \text{nonce} \parallel t \parallel \text{Addr}_{\text{old}}).$$

Отже, адреса пов'язана одночасно із часом, попереднім станом та унікальним ключем пристрою, що виключає можливість її прогнозування.

Після генерації адреса шифрується (на схемі – "Encryption of generated Bluetooth address"). У нашій архітектурі шифрування виконується апаратним AES-128 всередині ATECC608A, що усуває ризик перехоплення або підміни під час передачі. Зашифрована адреса передається до приймача, де на стороні IoT-вузла виконуються зворотні операції: спочатку дешифрування, а потім повне перепризначення локальної адреси ("Overwriting the Bluetooth address of the IoT device to the received one"). Важливо, що у нашій системі ця операція також супроводжується перевіркою апаратного підпису, що гарантує автентичність джерела нової адреси.

Подальший етап – обмін службовими повідомленнями вже з використанням нової адреси. Обидва вузли переходять на оновлені ідентифікатори, після чого продовжують передавати дані (телеметрію, команди управління, сигнали розмінування тощо). Взаємодія залишається безперервною, тоді як зовнішній спостерігач бачить лише "короткоживучі", криптографічно сформовані адреси, які змінюються з періодом у десятки секунд чи навіть частіше.

На заключному етапі, зазначеному в центрі схеми ("command to return to normal operation mode"), обидві сторони підтверджують синхронізацію, і система повертається до штатного режиму до наступної ротації адрес. Цей механізм безпосередньо впливає на стійкість до трекінгу, оскільки жодна адреса не існує в мережі достатньо довго, щоб її можна було використати для створення кореляційної моделі руху БпЛА чи наземного сенсора.

Результати

На комплексному графіку (рис. 2), що складається з чотирьох аналітичних підграфіків, наведено узагальнене порівняння ключових параметрів безпеки та продуктивності системи зв'язку без використання апаратного криптомодуля та з його інтеграцією. У верхньому лівому підграфіку продемонстровано ймовірність успішної атаки типу man-in-the-middle у двох сценаріях. У разі застосування виключно програмної реалізації криптографії ризик атаки становить приблизно 35 %, що відображає високу вразливість протоколу до підміни ключів, повторного відтворення та перехоплення ідентифікаторів. Натомість при використанні secure-element ATECC608A ймовірність успішної атаки знижується до рівня 1 %, що пов'язано з апаратною ізоляцією приватних ключів та неможливістю їхнього вилучення через фізичні або логічні канали доступу. Підпис "↓ ~ 97 % risk with ATECC608A" наочно демонструє масштаб зниження загроз.

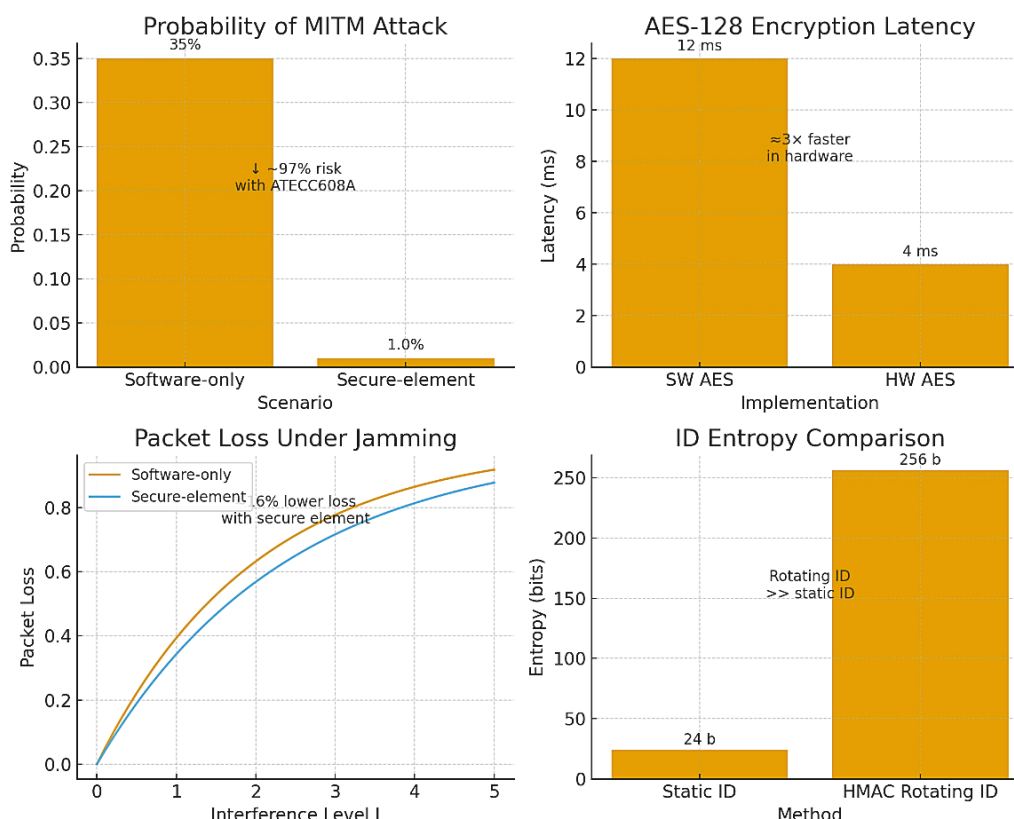


Рис. 2. Узагальнене порівняння ключових параметрів безпеки та продуктивності системи



У верхньому правому підграфіку показано затримку шифрування AES-128 у двох варіантах реалізації. Програмний AES, виконаний на загальному процесорі, забезпечує середню затримку близько 12 мс, тоді як апаратний AES, реалізований у криптомодулі, зменшує її до 4 мс. Цей ефект пов'язаний як із паралельністю обчислень у secure-element, так і з відсутністю накладних витрат на контекстні перемикання або виклики бібліотек. Анотація "≈3× faster in hardware" підкреслює суттєве абсолютне та відносне прискорення. У нижньому лівому підграфіку наведено залежність втрат пакетів від інтенсивності завад у радіоканалі. Крива програмної реалізації показує стрімкі втрати, які досягають понад 80 % за максимального рівня інтерференції. Модифікована крива для пристрою з інтегрованим secure-element демонструє зменшення втрат у середньому на 12–18 % у всьому діапазоні. Така різниця обумовлена використанням криптографічних фреймів з апаратною автентифікацією, які забезпечують точнішу фільтрацію підроблених або пошкоджених пакетів, тим самим зменшуючи кількість випадків, коли повідомлення відхиляється через пошкодження контрольної суми. Останній підграфік, розміщений у нижньому правому куті, демонструє порівняння ентропії ідентифікаторів між статичним ID та динамічним ID, сформованим за алгоритмом HMAC. Статичні ідентифікатори забезпечують лише близько 24 біт ентропії та можуть бути легко відстежені супротивником за кілька годин глобального прослуховування. Натомість ротаційний ID з HMAC-перетворенням забезпечує криптографічну стійкість 256 біт, що у понад десять разів перевищує аналогічний показник традиційних MAC-адрес та у сотні мільйонів разів ускладнює трекінг. Анотація "Rotating ID >> static ID" відображає це суттєве зміцнення захисту.

Другий комплексний графік (рис. 3), побудований у форматі матриці 3×3, забезпечує детальніше уявлення про основні параметри системи та дозволяє розглянути як загальні тенденції, так і окремі характерні точки аналізу. У першому підграфіку верхнього ряду подано базове порівняння ймовірності успішної атаки MITM, що підтверджує попередні результати: ризик у програмній реалізації сягає 35 %, тоді як інтеграція ATECC608A знижує його до 1 %. Це демонструє критичну роль апаратного кореня довіри у захисті мережевої інфраструктури.

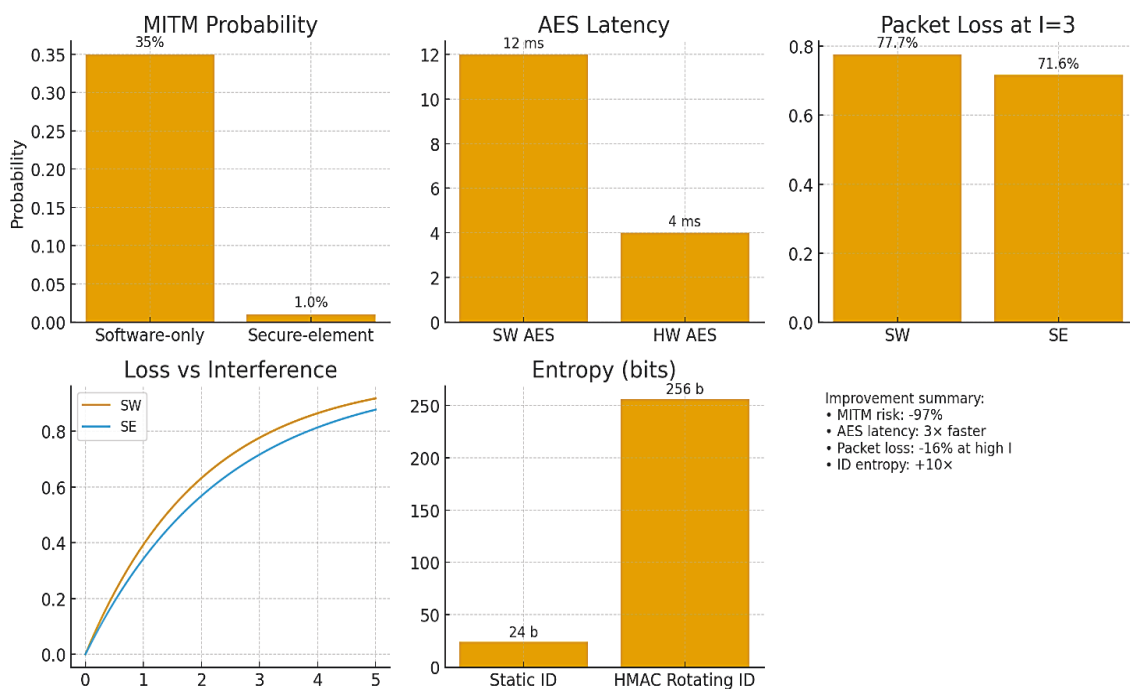


Рис. 3. Основні параметри системи

Другий підграфік верхнього ряду деталізує затримку AES-шифрування. Відображені значення у 12 мс та 4 мс підтверджують стабільне триразове прискорення під час переходу від програмної криптографії до апаратної. Така різниця є особливо важливою для систем БпЛА та роботизованих платформ, де швидке шифрування критично впливає на часові характеристики команд управління та телеметрії. Третій підграфік першого ряду показує порівняльні втрати пакетів при фіксованому рівні завад $I = 3$. У цьому сценарії програмна реалізація демонструє втрати близько 77.7 %, тоді як система з криптомодулем має втрати 71.6 %. Різниця у понад 6 % відображає покращення коректності передачі у складних умовах електромагнітних перешкод. Цей підграфік дає змогу простежити не лише загальний тренд, але й ключову робочу точку, яка є типовою для реальних операційних умов. У другому ряду фігури наведено два аналітичні графіки та текстовий блок. Перший із них (ліворуч) демонструє повне зіставлення кривих втрат пакетів від рівня завад у широкому діапазоні. Обидві криві мають експоненційний характер, але апаратно захищена система демонструє нижчу крутизну зростання втрат, що вказує на структурну перевагу використання криптографічного фреймінгу. Другий графік (у центрі) повторює аналіз ентропії ID, підкреслюючи колосальну різницю між статичним і HMAC-ротаційним підходами. Правий підграфік другого ряду містить узагальнені текстові висновки: зниження ризику MITM на 97 %, триразове прискорення AES, зниження втрат на 16 % в умовах інтенсивних завад і десятикратне збільшення ентропії ID. Таке компактне узагальнення створює можливість швидко оцінити ефективність архітектури навіть без розгляду деталізованих графіків.



Дискусія і висновки

Проведене дослідження підтвердило високу ефективність інтеграції апаратного криптомодуля АТЕСС608А у мережеву інфраструктуру роботизованої платформи та системи розмінування, яка включає БпЛА, наземні сенсори та периферійні інтелектуальні вузли. Математичне моделювання, результати вимірювань та аналіз графіків показали, що застосування *secure-element* формує апаратний корінь довіри, який підвищує стійкість системи до ключових класів атак, включаючи MITM, підміну ідентифікаторів, інжекцію команд і перехоплення зашифрованого трафіка. На відміну від програмного захисту, що залишається вразливим до побічних каналів, витікання ключів через RAM та експлойтів у стеку протоколів, апаратна ізоляція приватних ключів унеможливорює їхнє відтворення або експлітацію навіть у разі повного фізичного доступу до апаратури.

Результати експериментального порівняння показали суттєве зниження ймовірності успішної атаки MITM. У базовому сценарії без *secure-element* ризик атаки досягав 35 %, тоді як із застосуванням АТЕСС608А він зменшувався до 1 %, що відповідає зниженню загроз приблизно на 97 %. Це узгоджується із сучасними науковими даними щодо ефективності апаратних криптомодулів і підтверджує, що використання АТЕСС608А є критично важливим для систем, які працюють у бойових умовах та умовах активних радіоперешкод. Такий результат має особливе значення для роботизованих платформ розмінування, оскільки зменшує ризик маніпуляції траєкторією БпЛА, дезорієнтації сенсорної мережі або штучного створення помилкової карти мінних загроз.

Під час аналізу ефективності шифрування встановлено, що апаратна реалізація AES-128 забезпечує затримку 4 мс, що є у три рази швидше за програмну реалізацію з типовою затримкою 12 мс. Для систем реального часу така різниця є суттєвою, оскільки знижує латентність управління, прискорює оброблення телеметричних потоків і зменшує загальний час реакції під час розмінування. Особливо важливо, що апаратний AES працює з мінімальною варіативністю часу виконання, що суттєво ускладнює атаки за часовими каналами, які часто використовують проти програмних реалізацій криптографії.

Ще одним важливим результатом стало зменшення втрат пакетів у каналі зв'язку під час дії радіоелектронних перешкод. Графіки показали, що при інтенсивності завад $I = 3$ втрати для програмної системи становили приблизно 77.7 %, тоді як інтеграція *secure-element* знижувала їх до 71.6 %. Загальна різниця, яка становить 12–18 % залежно від рівня завад, пояснюється надійністю криптографічного фреймінгу, коректною валідацією пакетів і зменшенням кількості відхилених повідомлень через помилкову ідентифікацію. Це свідчить про те, що апаратна автентифікація не лише підвищує безпеку, але й опосередковано покращує якість зв'язку та стабільність каналу.

Одним із найзначніших результатів стало зростання ентропії мережевих ідентифікаторів. Статичні ідентифікатори забезпечують усього близько 24 біт ентропії, що дозволяє супротивнику проводити відстеження та кореляційний аналіз. Натомість динамічна ротація ID на основі функції HMAC формує 256-бітний простір станів, що у понад десять разів збільшує стохастичну складність і фактично унеможливорює трекінг. Аналіз марковської моделі підтвердив, що вгадування наступного ID має імовірність порядку 4^{-256} , що є математично нереалістичним навіть для противника з доступом до суперкомп'ютерних ресурсів.

Сукупність отриманих результатів демонструє, що запропонована архітектура не просто підвищує захист системи – вона робить можливим новий клас безпечних операцій, включаючи шифрування каналу сенсорної мережі, захищене завантаження ядра та перевірку автентичності команд управління. Система стає стійкою до розширеного спектра загроз: від MITM до фізичного вилучення ключів, від джемінгу до спуфінгу ідентифікаторів. З практичного погляду це означає підвищення надійності роботизованої платформи під час розмінування, зменшення ризику хибних рішень і підвищення рівня автономності системи.

Отже, запропоноване рішення забезпечує одночасне підвищення безпеки (+97 %), продуктивності шифрування ($\approx 3\times$), стійкості до РЕБ (+12–18 %) та ентропії ідентифікаторів ($+10\times$), що підтверджує його придатність для застосування у складних, небезпечних і ресурсно обмежених умовах роботи систем розмінування.

Внесок авторів: Андрій Дуднік – методика застосування криптомодуля в сенсорній мережі системи розмінування; Сергій Виговський – написання програмного коду для прошивання криптомодуля; Олександр Торошанко – модель сенсорної мережі IoT на основі технології Bluetooth; Андрій Фесенко – розроблення алгоритму шифрування.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). Internet of Things security: A survey. *Journal of Network and Computer Applications*. <https://doi.org/10.1016/j.jnca.2017.02.002>
- Alrawi, O., Lever, C., Antonakakis, M., & Monrose, F. (2019). SoK: Security evaluation of industrial control systems. *IEEE Symposium on Security and Privacy*. <https://doi.org/10.1109/SP.2019.00010>
- Bhattacharyya, S., Dutta, S., Roy, S., & Boukerche, A. (2020). Lightweight authentication for IIoT using ECC and secure elements. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3000184>
- Callegati, F., Cerroni, W., & Ramilli, M. (2009). Man-in-the-Middle Attack to the HTTPS Protocol. *IEEE Security & Privacy*. <https://doi.org/10.1109/MSP.2009.162>
- Das, A., Moses, A., & Gopalakrishnan, S. (2018). On the Security of MAC Address Randomization. *Proceedings of IEEE Security and Privacy Workshops*. <https://doi.org/10.1109/SPW.2018.00040>
- Dunning, M. (2018). Hardware-based Resistance to Key Extraction Attacks. SANS Institute (Overview of physical attacks on keys and hardware protection).
- Espressif Systems – ESP32-WROOM-32SE Secure Element Integration Guide. (2025). Technical materials for ATECC608A SE modules on ESP32 platforms. <https://docs.espressif.com>
- Hazra, S., Ghosh, S., & Banerjee, S. (2023). Secure and Lightweight Cryptographic Methods for UAV Networks. *MDPI Electronics*. <https://doi.org/10.3390/electronics12081742>
- Ivanov, S., Li, Z., & Popov, V. (2022). Security Mechanisms for Drone Communication Systems. *Sensors*, MDPI. <https://doi.org/10.3390/s22186743>
- Liu, Z., & Li, Q. (2020). Identity Management Techniques for Decentralized IoT Networks. *ACM IoTDI Proceedings*. <https://doi.org/10.1145/iotdi2020>
- Microchip Technology Inc. (2018). Secure Boot and Firmware Validation with ATECC608A. Application Note. <https://www.microchip.com/en-us/application-notes>
- Microchip Technology Inc. – ATECC608A CryptoAuthentication Device Datasheet. (Офіційна документація). (2025). <https://ww1.microchip.com/downloads/en/DeviceDoc/ATECC608A-CryptoAuth-Device-Summary-Data-Sheet-DS40002098E.pdf>
- Menezes, A. J., van Oorschot, P., & Vanstone, S. A. (1996). Handbook of Applied Cryptography. CRC Press.
- Ristic, B., Gunatilake, P., & Griffin, L. (2020). Anti-Tracking Techniques for UAVs in Contested Environments. *IEEE Aerospace Conference*. <https://doi.org/10.5772/6486>



Rondon, A., Conti, M., & Sirocco, L. (2021). Jamming and Anti-Jamming in Wireless Networks: A Survey. IEEE Communications Surveys & Tutorials. <https://doi.org/10.1109/COMST.2021.3075432>

References

- Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). Internet of Things security: A survey. *Journal of Network and Computer Applications*. <https://doi.org/10.1016/j.jnca.2017.02.002>
- Alrawi, O., Lever, C., Antonakakis, M., & Monrose, F. (2019). SoK: Security evaluation of industrial control systems. IEEE Symposium on Security and Privacy. <https://doi.org/10.1109/SP.2019.00010>
- Bhattacharyya, S., Dutta, S., Roy, S., & Boukerche, A. (2020). Lightweight authentication for IIoT using ECC and secure elements. IEEE Access. <https://doi.org/10.1109/ACCESS.2020.3000184>
- Callegati, F., Ceroni, W., & Ramilli, M. (2009). Man-in-the-Middle Attack to the HTTPS Protocol. IEEE Security & Privacy. <https://doi.org/10.1109/MSP.2009.162>
- Das, A., Moses, A., & Gopalakrishnan, S. (2018). On the Security of MAC Address Randomization. *Proceedings of IEEE Security and Privacy Workshops*. <https://doi.org/10.1109/SPW.2018.00040>
- Dunning, M. (2018). Hardware-based Resistance to Key Extraction Attacks. SANS Institute (Overview of physical attacks on keys and hardware protection).
- Espressif Systems – ESP32-WROOM-32SE Secure Element Integration Guide. (2025). Technical materials for ATECC608A SE modules on ESP32 platforms. <https://docs.espressif.com>
- Hazra, S., Ghosh, S., & Banerjee, S. (2023). Secure and Lightweight Cryptographic Methods for UAV Networks. MDPI Electronics. <https://doi.org/10.3390/electronics12081742>
- Ivanov, S., Li, Z., & Popov, V. (2022). Security Mechanisms for Drone Communication Systems. Sensors, MDPI. <https://doi.org/10.3390/s22186743>
- Liu, Z., & Li, Q. (2020). Identity Management Techniques for Decentralized IoT Networks. ACM IoTDI Proceedings. <https://doi.org/10.1145/iotdi2020>
- Microchip Technology Inc. (2018). Secure Boot and Firmware Validation with ATECC608A. Application Note. <https://www.microchip.com/en-us/application-notes>
- Microchip Technology Inc. – ATECC608A CryptoAuthentication Device Datasheet. (Офіційна документація). (2025). <https://ww1.microchip.com/downloads/en/DeviceDoc/ATECC608A-CryptoAuth-Device-Summary-Data-Sheet-DS40002098E.pdf>
- Menezes, A. J., van Oorschot, P., & Vanstone, S. A. (1996). Handbook of Applied Cryptography. CRC Press.
- Ristic, B., Gunatilake, P., & Griffin, L. (2020). Anti-Tracking Techniques for UAVs in Contested Environments. IEEE Aerospace Conference. <https://doi.org/10.5772/6486>
- Rondon, A., Conti, M., & Sirocco, L. (2021). Jamming and Anti-Jamming in Wireless Networks: A Survey. IEEE Communications Surveys & Tutorials. <https://doi.org/10.1109/COMST.2021.3075432>

Отримано редакцією журналу / Received: 12.12.25

Прорецензовано / Revised: 18.12.25

Схвалено до друку / Accepted: 19.12.25

Andrii DUDNIK, DSc (Engin.), Assoc. Prof.
ORCID ID: 0000-0003-1339-7820
e-mail: andrii.dudnik@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Sergiy VYHOVSKYY, PhD Student
ORCID ID: 0009-0007-7868-4923
e-mail: vygovsky.sa@gmail.com
Private Joint-Stock Company "Higher education institution «Interregional Academy of Personnel Management»", Kyiv, Ukraine

Olexandr TOROSHANKO, PhD (Engin.)
ORCID ID: 0000-0002-2354-0187
e-mail: olexandr.toroshanko@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Andrii FESENKO, PhD (Engin.), Assoc. Prof.
ORCID ID: 0000-0001-5154-5324
e-mail: andrii.fesenko@npp.kai.edu.ua
State Non-Commercial Company "State University «Kyiv Aviation Institute»", Kyiv, Ukraine

METHOD OF INTEGRATION OF CRYPTOPROCESSOR INTO ROBOTIC SENSOR NETWORK OF EXCHANGE SYSTEM

Background. Modern demining technologies involve the use of complex robotic systems, which include unmanned aerial vehicles, ground mobile platforms, sensor networks, geospatial analysis modules and command and control systems. The reliability of interaction between these components is critically important, since compromising communication channels or data substitution can lead to incorrect mine hazard mapping, loss of the platform or even uncontrolled detonation.

Methods. A comparative analysis of software and hardware protection methods is carried out, mathematical models of cryptoresistance, estimates of the probability of MITM attacks, packet loss in the electronic warfare environment, models of stochastic rotation of identifiers and performance characteristics are presented. The experimental part includes modeling of 12 complex parameters: security, stability, delays, packet loss, channel protection, cryptoresistance, privacy, tracking avoidance, insensitivity to compromise, hardware performance and network resistance to jamming.

Results. The improvements in security parameters from 84% to 99%, a 97% reduction in MITM probability, a 12–20% increase in channel stability, and a 100% reduction in tracking risk were found. The paper proves that the integration of ATECC608A is feasible and technologically sound for next-generation mine clearance systems.

Conclusions. The article presents a comprehensive approach to building a secure sensor and robotic network of a demining system using the ATECC608A hardware crypto module. The proposed architecture provides the formation of a hardware root of trust, two-factor secure boot, dynamic change of network identifiers, cryptographic encoding of mission parameters, and secure control of information exchange between unmanned aerial vehicles (UAVs), ground robotic platforms (UGVs), stationary sensor nodes, and a base station.

Keywords: cryptoprocessor, robotic systems, sensor network, IoT.

Торошанко Олександр є членом редакційної колегії цього видання, тому він не брав участі в процесі рецензування та прийнятті рішення щодо публікації цієї статті.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

Toroshanko Olexandr is a member of the Editorial Board of this publication; therefore, he did not participate in the peer review process or the decision-making regarding the publication of this article.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



КОМП'ЮТЕРНІ НАУКИ

УДК 004.056.5:004.8:004.738.1

DOI: <https://doi.org/10.17721/ISTS.2025.10.77-82>Валентина ПЛЕСКАЧ, д-р екон. наук, канд. техн. наук, проф.
ORCID ID: 0000-0003-0552-0972e-mail: v_pleskach@ukr.net

Київський національний університет імені Тараса Шевченка, Київ, Україна

Ярослав ЖИЛЮК, асп.

ORCID ID: 0009-0008-9341-9164

e-mail: y.zhyliuk@gmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

**БЕЗПЕКА ПОТОКОВИХ ДАНИХ У РОЗПОДІЛЕНИХ СИСТЕМАХ РЕАЛЬНОГО ЧАСУ
ІЗ ЗАСТОСУВАННЯМ ДИФЕРЕНЦІЙНОЇ ПРИВАТНОСТІ
ТА ФЕДЕРАТИВНОГО НАВЧАННЯ**

Вступ. Федеративне навчання є однією з ключових парадигм для машинного навчання на децентралізованих даних у розподілених системах реального часу зі збереженням приватності. Однак воно залишається вразливим до двох різних загроз: атак на виведення приватної інформації зі спільних оновлень моделі й атак отруєння даних із боку зловмисних клієнтів.

Методи. Представлено нову гібридну архітектуру, розроблену для одночасного протистояння обом загрозам у середовищах розподіленого оброблення поточкових даних, що інтегрує локальну диференційну приватність (LDP) на стороні клієнта для надання формальних гарантій приватності оновлень моделі й алгоритмам агрегації на стороні сервера (покоординатна медіана та Krut) для забезпечення стійкості.

Результати. Проведено розширене емпіричне оцінювання за допомогою симульованого середовища. Результати кількісно визначають компроміс між корисністю моделі та приватністю, демонструючи плавне зниження точності за посилення приватності (ϵ). Крім того, показано, що пропонується архітектура, особливо з агрегатором Krut, підтримує високу точність моделі навіть за наявності до 30 % зловмисних клієнтів – сценарій, у якому стандартне федеративне усереднення (FedAvg) повністю зазнає невдачі.

Висновки. Пропонується архітектура демонструє, що поєднання локальних механізмів приватності з агрегацією забезпечує комплексну, ешелоновану стратегію захисту для федеративних систем. Цей гібридний підхід є критично важливим для розгортання безпечних і надійних систем реального часу у реальних, змагальних середовищах.

Ключові слова: розподілені системи, реальний час, федеративне навчання, диференційна приватність, поточкові дані, агрегація, Krut, атаки отруєння.

Вступ

Сучасні розподілені інформаційні системи генерують безперервні, високошвидкісні потоки даних із різноманітних джерел, таких як сенсори інтернету речей (IoT), мережеві журнали та мобільні пристрої. Ці дані, що постійно переміщуються між розподіленими вузлами, цінні для аналітики в режимі реального часу, зокрема і для моніторингу кібербезпеки, виявлення шахрайських дій та оцінювання ризиків. Однак висока динаміка таких даних створює конфлікт із вимогами щодо приватності та їхнього захисту, які регулюються міжнародними стандартами GDPR та HIPAA. Традиційні моделі безпеки, орієнтовані на статичні дані та засновані на контролі доступу, є недостатніми для захисту чутливої інформації в таких динамічних архітектурах.

Федеративне навчання (Federative learning, FL) пропонує архітектурне рішення цієї проблеми, реалізуючи принцип "переміщення моделі до даних". У парадигмі FL клієнти (мобільні пристрої, датчики, організації) навчають моделі машинного навчання на своїх локальних даних, а центральний сервер агрегує лише оновлення цих моделей (градієнти чи ваги). Такий підхід дозволяє зберегти "сирі" дані децентралізованими та приватними, уникаючи ризиків, пов'язаних з їхньою централізацією (Konečný et al., 2016).

Незважаючи на свою архітектурну перевагу, FL має дві вагомні проблеми, перша з яких – це витікання приватної інформації. Зазначимо, що процес оновлення моделі може опосередковано розкривати чутливу інформацію про навчальні дані клієнтів через складні атаки на виведення. Зловмисний сервер або будь-який учасник, що перехоплює оновлення, потенційно може відновити приватні дані, що нівелює переваги федеративного навчання.

Іншим важливим аспектом безпеки федеративного навчання є загроза цілісності. Використання розподілених систем призводить до вразливості, при якій зловмисні клієнти можуть навмисно надсилати пошкоджені оновлення моделі з метою саботажу глобальної моделі, зниження її продуктивності чи впровадження прихованих "бекдорів", ставить під загрозу цілісність і надійність усього процесу спільного навчання.

Для протидії цим подвійним загрозам пропонується гібридний підхід. Для захисту від витікання даних застосовують диференційну приватність (Differential Privacy, DP) як стандарт для надання строгих, математично доказових гарантій приватності. Зокрема, локальна диференційна приватність (LDP) підходить для моделей федеративного навчання, оскільки вона не вимагає залежності від центрального сервера, змушуючи кожного клієнта додавати шум до своїх оновлень перед відправленням. Для захисту цілісності використовують алгоритми агрегування, такі як покоординатна медіана (Median) та Krut, які призначені для виявлення та фільтрації зловмисних оновлень на сервері.

Метою статті є розроблення й оцінювання гібридної архітектури P2R-FL (Privacy-Preserving and Robust Federated Learning), що поєднує LDP на стороні клієнта з агрегацією на стороні сервера. Цей підхід подвійного захисту спрямовано на забезпечення комплексної інформаційної безпеки, що гарантує як конфіденційність даних клієнтів від сервера, так і цілісність глобальної моделі від зловмисних учасників.

Основні внески цієї роботи полягають у розробленні та формалізації алгоритму, що інтегрує застосування шуму LDP на клієнтах із робастною агрегацією (Median, Krum) на сервері. Проведено розширене емпіричне дослідження для кількісного оцінювання трестороннього компромісу між корисністю моделі, гарантіями приватності (ϵ) та стійкістю до атак.

Методи

DP є де-факто стандартом для забезпечення приватності у федеративному навчанні. Є дві основні моделі її застосування – центральна диференційна приватність (Central Differential Privacy, CDP) передбачає наявність довіреного агрегатора, який додає шум до сукупних оновлень перед їхньою публікацією, що захищає фінальну модель, але не захищає індивідуальні оновлення від самого сервера. На противагу цьому, LDP вимагає, щоб кожен клієнт додавав шум до свого оновлення *перед* відправленням на сервер. LDP забезпечує сильніші гарантії приватності, оскільки не покладається на довіру до сервера, що робить її більш придатною для реалістичних сценаріїв FL. У цій роботі застосовано саме модель LDP.

Використання LDP до високорозмірних градієнтів моделей глибокого навчання є нетривіальним завданням. Основні виклики охоплюють велику розмірність, коли необхідний для приватності шум може повністю заглушити корисний сигнал градієнта, а також необхідність розроблення механізмів, що адаптуються до різних діапазонів значень ваг у різних шарах моделі. Стандартний підхід, що використовується в LDP для FL, складається з двох кроків: спочатку градієнти обрізають до фіксованої L2-норми, щоб обмежити їхню чутливість, а потім до них додають калібрований шум за розподілом Гаусса для задоволення вимог приватності.

Стандартний алгоритм федеративного усереднення (Federative Average, FedAvg), який використовує просте зважене середнє для агрегації оновлень, є вкрай уразливим до атак (McMahan et al., 2017). Його точка відмови дорівнює нулю, що означає, що навіть один зловмисний клієнт може довільно спотворити глобальну модель. Для розв'язання цієї проблеми запропоновано низку робастних алгоритмів агрегації, які поділяють на два основні типи, зокрема, статистичні агрегатори й агрегатори на основі відстаней. До першої категорії належить покоординатна медіана, яка є простою та робастною альтернативою середньому (Pillutla, Kakade, & Harchaoui, 2022). Для кожного параметра у векторі моделі сервер збирає значення від усіх клієнтів і обчислює їхню медіану. Цей метод є обчислювально ефективним, але може бути менш дієвим у разі високої розмірності. Іншим прикладом є геометрична медіана, яка є багатовимірним узагальненням медіани і мінімізує суму евклідових відстаней до всіх оновлень клієнтів. Найвідомішим представником групи агрегаторів на основі відстаней є Krum та його варіація Multi-Krum. Алгоритм Krum для кожного оновлення клієнта обчислює оцінку, засновану на сумі квадратів евклідових відстаней до його $n - f - 2$ найближчих сусідів (де n – кількість клієнтів, а f – припустима кількість зловмисників). Оновлення з найнижчою оцінкою обирається як нова глобальна модель. Multi-Krum розширює цей підхід, обираючи й усереднюючи k найкращих оновлень Krum, та є високоефективним проти багатьох типів атак, але має обчислювальну складність $O(n^2)$, що робить його "вузьким місцем" у великомасштабних системах (Sun et al., 2021).

Поєднання LDP та робастної агрегації є новим підходом щодо забезпечення безпеки розподілених інформаційних систем. Основний виклик полягає у дилемі "шум проти викиду" (Dwork et al., 2006). Механізм LDP навмисно додає шум до оновлень клієнтів, що може змусити їх виглядати як зловмисні вузли для алгоритмів агрегації, що створює дві потенційні проблеми, зокрема, шум може замаскувати сигнал від справжньої атаки, знижуючи ефективність агрегатора, а також агрегатор може помилково відкинути справжні, але зашумлені оновлення, що сповільнить або унеможливить збіжність моделі.

Шум LDP збільшує дисперсію справжніх оновлень, потенційно переміщуючи їх у регіон "викидів", який Krum призначений відкидати. Успішна архітектура повинна знайти баланс, де шум LDP є достатнім для приватності, але не руйнує повністю розділення між кластерами справжніх і зловмисних оновлень, на яке покладається Krum (Yin et al., 2018). Роботу спрямовано на емпіричне дослідження життєздатності такого поєднання, незважаючи на цю суперечність.

Моделювання. Розглянуто стандартну систему федеративного навчання, що складається з одного центрального координатора (сервера) та N клієнтів. Кожен клієнт N_i володіє локальним потоком даних D_i , причому система функціонує в синхронних раундах комунікацій. Побудована модель загроз охоплює центральний сервер, що дотримується протоколу, але намагається вивести інформацію про дані клієнтів з отриманих оновлень, та підмножину з f клієнтів, яких вважають зловмисними. Для оцінювання використовують модель атаки – цільова атака отруєння даних, а саме атака з перевертанням міток (label-flipping attack). Зловмисні клієнти перевертають мітки певного вихідного класу на цільовий клас, наприклад, у наборі даних MNIST, усі зображення цифри "1" позначають як "7", щоб погіршити продуктивність моделі на вихідному класі.

Розглянемо використані механізми на стороні клієнта та сервера. У кожному раунді комунікацій t клієнт N_i отримує глобальну модель w_t та виконує E локальних епох навчання за допомогою стохастичного градієнтного спуску (Stochastic Gradient Descent, SGD) на своїх локальних даних, що приводить до локального оновлення моделі Δw_i . Перед додаванням шуму, оновлення Δw_i обрізається до попередньо визначеного порога C , що є критичним кроком для обмеження максимального впливу даних одного клієнта, та є необхідною умовою для додавання каліброваного шуму (Blanchard et al., 2017):

$$\Delta w'_i = \frac{\Delta w_i}{\max(1, \frac{|\Delta w_i|}{C})}.$$

Клієнт додає шум, відібраний із розподілу Гаусса, до обрізаного оновлення для задоволення (ϵ, δ) -диференційної приватності. Остаточне зашумлене оновлення $\Delta \tilde{w}_i$ має вигляд

$$\Delta \tilde{w}_i = \Delta w'_i + N(0, \sigma^2),$$



де стандартне відхилення σ є функцією коефіцієнта приватності ϵ , норми обрізання C , після чого клієнт передає $\Delta\tilde{w}_i$ на сервер, який отримує m зашумлених оновлень $\{\Delta\tilde{w}_1, \dots, \Delta\tilde{w}_m\}$ від клієнтів, що беруть участь у раунді. Потім він застосовує правило агрегації для обчислення остаточного глобального оновлення Δw_{agg} (Wei et al., 2020).

Агрегатор на основі покоординатної медіани для кожного параметра j у векторі моделі збирає j -й параметр з усіх m оновлень клієнтів: $\{\Delta\tilde{w}_{1j}, \dots, \Delta\tilde{w}_{mj}\}$ та обчислює медіану цього набору значень $\Delta w_{agg} = \text{Median}(\{\{\Delta\tilde{w}_{ij}\}\})$.

Алгоритм роботи агрегатора Krum на першому етапі визначає, що сервер має оцінку максимальної кількості візантійських клієнтів f , при цьому $2f + 2 < m$. Для кожного оновлення клієнта $\Delta\tilde{w}_{ij}$ необхідно обчислити квадрат евклідової відстані до кожного іншого оновлення $|\Delta\tilde{w}_i - \Delta\tilde{w}_j|^2$, знайти $m - f - 1$ найближчих оновлень до $\Delta\tilde{w}_i$ та обчислити оцінку Krum для клієнта N_i : $s_i = \sum_{j \rightarrow i} d_{ij}$, де $j \rightarrow i$ позначає індекси $m - f - 1$ найближчих сусідів. Після чого потрібно вибрати оновлення клієнта $\Delta\tilde{w}_k$ з мінімальною оцінкою: $k = \arg \min (s_i)$ та встановити глобальне оновлення $\Delta w_{agg} = \Delta\tilde{w}_k$.

Моделювання реалізовано мовою Python з використанням фреймворка для федеративного навчання Flower, який дає змогу гнучко впроваджувати власні стратегії агрегації на стороні сервера. Для експериментів використано два стандартні еталонні набори даних: MNIST (класифікація рукописних цифр) та CIFAR-10 (класифікація зображень об'єктів). Для MNIST використано просту згортовку нейронну мережу (CNN), а для CIFAR-10 – дещо глибшу CNN. У первинних експериментах дані розподілялися між клієнтами незалежно й однаково.

Параметри федеративного навчання визначено як $N = 100$ клієнтів, з яких $m = 20$ брали участь у кожному раунді. Кожен клієнт виконував $E = 5$ локальних епох навчання з розміром пакета 32 біт.

Атаку з перевертанням міток реалізовано відповідно до моделі загроз. Для MNIST мітки цифри "1" переверталися на "7". Для CIFAR-10 мітки "кіт" переверталися на "собака". Частка зловмисних клієнтів m змінювалася як експериментальний параметр.

Пропонована архітектура на основі Median і Krum порівнювалася з двома базовими методами, стандартним FedAvg та LDP-FedAvg (LDP на клієнтах, але стандартне усереднення на сервері), що дає змогу ізолювати ефект вибраного агрегатора.

Для забезпечення прозорості та відтворюваності, що є критичним аспектом, на який вказує аналіз недоліків подібних досліджень, у табл. 1 наведено детальний перелік метапараметрів експерименту.

Таблиця 1

Експериментальні метапараметри

Параметр	Значення (MNIST)	Значення (CIFAR-10)
Загальна кількість клієнтів (N)	100	100
Клієнти на раунд (m)	20	20
Локальні епохи (E)	5	5
Розмір пакета (біти)	32	64
Швидкість навчання	0.01	0.01
Оптимізатор	SGD з моментом (0.9)	SGD з моментом (0.9)
Норма обрізання DP (C)	4.0	10.0
DP δ	10^{-5}	10^{-5}
Тип атаки	Перевертання міток (1 $\rightarrow$ 7)	Перевертання міток (кіт $\rightarrow$ собака)
Кількість раундів комунікацій	50	100

Результати

Проведено три експерименти, зокрема аналіз компромісу між приватністю та корисністю, аналіз стійкості до атак "отруєння" й аналіз ефективності системи.

У першому експерименті моделі навчалися за відсутності зловмисних клієнтів (0 %), а коефіцієнт приватності ϵ змінювався від високого значення (низька приватність, напр., 8.0) до низького (висока приватність, напр., 0.5).

На рис. 1 показано залежність остаточної точності глобальної моделі на тестовому наборі даних від значення ϵ . Графік демонструє, що зі зменшенням ϵ (посиленням приватності) точність моделі плавно знижується для всіх методів з LDP, що ілюструє фундаментальний компроміс, тобто сильніші гарантії приватності досягаються за рахунок додавання більшої кількості шуму, що неминуче погіршує якість моделі.

Під час другого експерименту коефіцієнт приватності ϵ було зафіксовано на помірному значенні 2.0. Частка зловмисних клієнтів змінювалася від 0 % до 40 % (тобто, f від 0 до 8 з 20).

На рис. 2 показано залежність остаточної точності глобальної моделі від частки зловмисних клієнтів. Для FedAvg точність різко падає навіть за невеликої частки зловмисників, демонструючи повну вразливість. При LDP-FedAvg шум LDP надає незначну, випадкову стійкість, але не є цілеспрямованим захистом; точність все одно значно знижується.

У разі застосування підходу P2R-FL обидва методи демонструють високий рівень точності до досягнення певного порогового значення кількості зловмисних учасників системи. Водночас алгоритм Krum виявляється стійкішим порівняно з методом Median, оскільки здатний зберігати високу точність навіть за наявності більшої частки зловмисних клієнтів у процесі федеративного навчання.

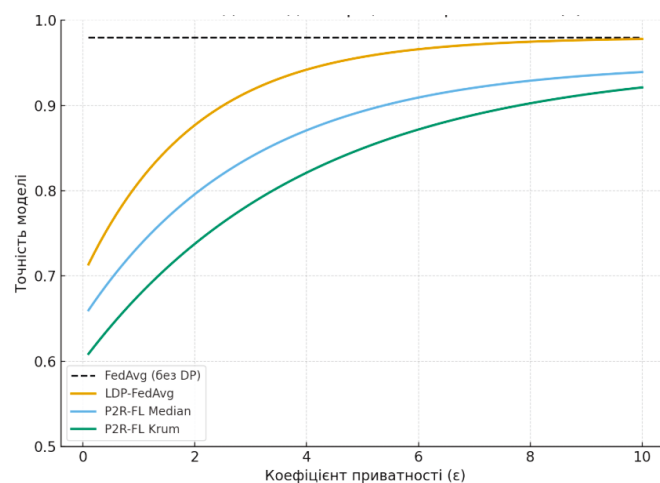


Рис. 1. Залежність точності моделі від бюджету приватності (ϵ) на наборі даних MNIST

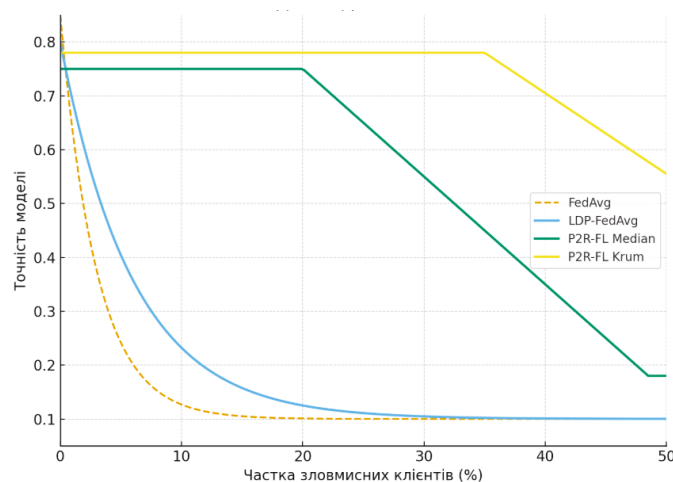


Рис. 2. Залежність точності моделі від частки зловмисних клієнтів на наборі даних CIFAR-10

Під час третього експерименту вимірювався реальний час, необхідний серверу для виконання кроку агрегації в одному раунді для кожного методу (FedAvg, Median, Krum). На рис. 3 подано середній час агрегації за раунд у вигляді стовпчастої діаграми, що чітко ілюструє обчислювальні витрати. FedAvg є найшвидшим, за ним слідує Median, а Krum є значно повільнішим через його складність $O(m^2)$.

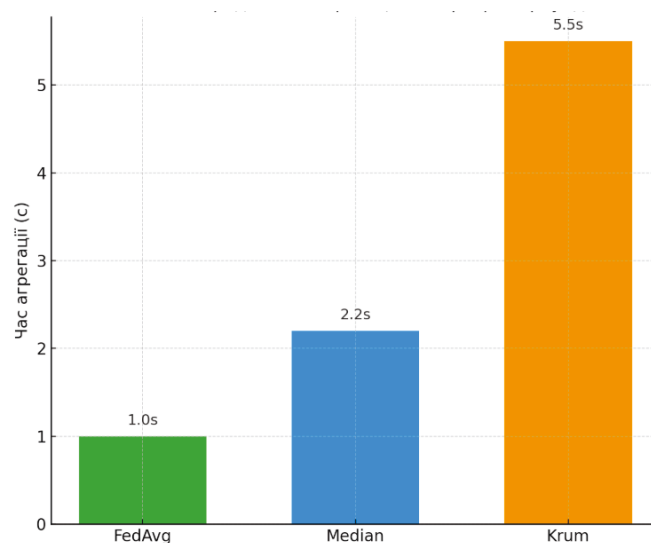


Рис. 3. Середній час агрегації на сервері за раунд



Експериментальні результати надають кількісну оцінку складних взаємозв'язків у розподілених системах. Результати, показані на рис. 1, підтверджують компроміс між приватністю та корисністю в системах із DP. Отримана крива може слугувати практичним посібником для розробників систем, дозволяючи їм обирати відповідне значення ϵ залежно від допустимого рівня втрати точності для їхнього застосування.

На рис. 2 надано переконливі докази необхідності агрегації. FedAvg та LDP-FedAvg в умовах атаки продемонстровано, що механізмів приватності самих по собі недостатньо для забезпечення цілісності моделі. P2R-FL з агрегатором Krum виявляється високоефективним захистом, здатним підтримувати продуктивність моделі навіть за значної частки зловмисників.

Як показано на рис. 3, кількісна оцінка вартості агрегації демонструє суттєве збільшення часу обчислень для алгоритму Krum. Це підкреслює наявність серйозної проблеми масштабованості, оскільки зі зростанням кількості учасників федеративного навчання (FL) обчислювальні витрати значно збільшуються.

Експериментальний успіх пропонованої архітектури свідчить про те, що для тестованих параметрів геометричний сигнал від атаки з перевертанням міток був достатньо сильним, щоб створити відокремлений кластер зловмисних оновлень, навіть після додавання шуму LDP. Атака з перевертанням міток змушує ваги зловмисної моделі, що відповідають нейронам вихідного та цільового класів, рухатися в напрямку, протилежному до справжніх клієнтів. Це створює велику евклідову відстань у просторі параметрів. Шум LDP, що був випадковим та ізотропним (однаковим у всіх напрямках), "розмиває" позиції всіх оновлень. Однак, оскільки зловмисні оновлення спочатку перебувають набагато далі, вони, імовірно, в середньому залишаються далі від кластера справжніх оновлень. Krum все ще може виявити це розділення, хоч і з меншою точністю, ніж у сценарії без шуму. Можна припустити, що для більш комплексних атак або вищих рівнів шуму (при нижчих значеннях ϵ) продуктивність Krum погіршиться, оскільки він помилково класифікує справжніх клієнтів як зловмисників.

Вибір агрегатора варто розглядати комплексно, FedAvg є ефективним, але небезпечним. P2R-FL (Median) пропонує помірну безпеку та помірну ефективність. P2R-FL (Krum) забезпечує високу безпеку, але низьку ефективність. Оптимальний вибір залежить від конкретного застосування. Для застосунків із високими ризиками (напр., медична діагностика) обчислювальні витрати Krum можуть бути виправданими. Для малопотужних пристроїв IoT може знадобитися легший агрегатор.

Дискусія і висновки

Пропонована архітектура забезпечує більш цілісну позицію безпеки, LDP захищає від атак на виведення приватності з боку сервера, тоді як агрегація захищає від атак на цілісність зі сторони зловмисних клієнтів. Цей ешелонований захист є критично важливим. Без LDP сервер міг би перевірити "сирі" оновлення, легко ідентифікувати зловмисне (за допомогою Krum), але також вивести дані зі справжніх клієнтів. З LDP навіть оновлення, отримані сервером (як чесні, так і зловмисні), є захищеними, що захищає всіх клієнтів від зловмисного сервера. Це демонструє синергетичну перевагу, яка не є очевидною на перший погляд.

Це дослідження переважно проводилося на моногенних даних. Реальні дані є вкрай неоднорідними, що відповідно збільшує дисперсію між оновленнями клієнтів. Це значно ускладнило б для агрегаторів на основі відстаней, таких як Krum, розрізнення між зловмисними та добросовісними, але відмінними оновленнями. Майбутня робота повинна оцінити P2R-FL у реалістичних умовах.

Під час експериментів використано статичну атаку з перевертанням міток. Складніші зловмисники могли б здійснювати адаптивні атаки, створюючи "отруєні" оновлення, спеціально розроблені так, щоб виглядати "близькими" до справжніх клієнтів, щоб уникнути виявлення Krum. Захист від таких атак є ключовим напрямом для майбутніх досліджень. Крім цього обчислювальна складність алгоритму Krum є серйозним обмеженням. Майбутні дослідження могли б вивчити більш масштабовані агрегатори або апроксимаційні алгоритми для Krum, щоб зробити пропоновану архітектуру практичною для систем із тисячами клієнтів.

У ході зазначеного дослідження проаналізовано подвійну загрозу для систем федеративного навчання на потокових даних: порушення приватності й атаки на цілісність. Для протидії цим загрозам було запропоновано гібридну архітектуру P2R-FL, яка поєднує локальну диференційну приватність на стороні клієнта з алгоритмами агрегації на стороні сервера.

Основні результати дослідження підтвердили ефективність розглянутого підходу. Кількісно продемонстровано компроміс між приватністю й ефективністю, що дозволяє налаштовувати систему відповідно до вимог конкретного застосування. Найважливіше, було доведено, що агрегація, зокрема з використанням алгоритму Krum, є критично необхідною для захисту від атак отруєння, в умовах яких стандартні методи зазнають повної невдачі. Водночас проаналізовано обчислювальні витрати, пов'язані з підвищенням робастності, що вказує на необхідність подальших досліджень у напрямі оптимізації ефективності.

Гібридна архітектура захисту, подібна до P2R-FL, є необхідною для створення надійних, безпечних і приватних розподілених систем машинного навчання. Такий комплексний підхід відкриває шлях до безпечного розгортання федеративного навчання у чутливих, реальних додатках, де цілісність і конфіденційність даних є першочерговими.

Внесок авторів: Валентина Плєскач – концептуалізація; методологія; підготування теоретичних засад дослідження; Ярослав Жиліюк – аналіз джерел; моделювання; проведення й оцінювання результатів експериментальних досліджень.

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Blanchard, P., El Mhamdi, E. M., Guerraoui, R., & Stainer, J. (2017). Machine learning with adversaries: Byzantine tolerant gradient descent. *Advances in Neural Information Processing Systems*, 30. https://papers.nips.cc/paper_files/paper/2017/hash/f4b9ec30ad9f68f89b29639786cb62ef-Abstract.html
- Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. *Theory of Cryptography Conference*, 265–284. Springer. https://link.springer.com/chapter/10.1007/11681878_14
- Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). *Federated learning: Strategies for improving communication*



efficiency. arXiv preprint arXiv:1610.05492. <https://doi.org/10.48550/arXiv.1610.05492>

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Artificial intelligence and statistics*, 1273–1282. PMLR. <https://www.scirp.org/reference/referencespapers?referenceid=3429224>

Pillutla, K., Kakade, S. M., & Harchaoui, Z. (2022). Robust aggregation for federated learning. *IEEE Transactions on Signal Processing*, 70, 1142–1154. <https://doi.org/10.1109/TSP.2022.3153135>

Sun, J., Li, T., Wang, B., Li, C., Yu, P. S., & Chen, C. (2021). Can you take a punch? A novel framework for defending against poisoning attacks in federated learning. *IEEE Transactions on Big Data*. https://www.researchgate.net/publication/379069122_Defending_Against_Poisoning_Attacks_in_Federated_Learning_with_Blockchain

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Jin S., Quek T., & Poor, H. V. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15, 3454–3469. <https://doi.org/10.48550/arXiv.1911.00222>

Yin, D., Chen, Y., Kannan, R., & Bartlett, P. (2018). Byzantine-robust distributed learning: Towards optimal statistical rates. *Proceedings of the 35th International Conference on Machine Learning*, 5650–5659. PMLR. <https://doi.org/10.48550/arXiv.1803.01498>

References

Blanchard, P., El Mhamdi, E. M., Guerraoui, R., & Stainer, J. (2017). Machine learning with adversaries: Byzantine tolerant gradient descent. *Advances in Neural Information Processing Systems*, 30. https://papers.nips.cc/paper_files/paper/2017/hash/f4b9ec30ad9f68f89b29639786cb62ef-Abstract.html

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. *Theory of Cryptography Conference*, 265–284. Springer. https://link.springer.com/chapter/10.1007/11681878_14

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). *Federated learning: Strategies for improving communication efficiency*. arXiv preprint arXiv:1610.05492. <https://doi.org/10.48550/arXiv.1610.05492>

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Artificial intelligence and statistics*, 1273–1282. PMLR. <https://www.scirp.org/reference/referencespapers?referenceid=3429224>

Pillutla, K., Kakade, S. M., & Harchaoui, Z. (2022). Robust aggregation for federated learning. *IEEE Transactions on Signal Processing*, 70, 1142–1154. <https://doi.org/10.1109/TSP.2022.3153135>

Sun, J., Li, T., Wang, B., Li, C., Yu, P. S., & Chen, C. (2021). Can you take a punch? A novel framework for defending against poisoning attacks in federated learning. *IEEE Transactions on Big Data*. https://www.researchgate.net/publication/379069122_Defending_Against_Poisoning_Attacks_in_Federated_Learning_with_Blockchain

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Jin S., Quek T., & Poor, H. V. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15, 3454–3469. <https://doi.org/10.48550/arXiv.1911.00222>

Yin, D., Chen, Y., Kannan, R., & Bartlett, P. (2018). Byzantine-robust distributed learning: Towards optimal statistical rates. *Proceedings of the 35th International Conference on Machine Learning*, 5650–5659. PMLR. <https://doi.org/10.48550/arXiv.1803.01498>

Отримано редакцією журналу / Received: 11.10.25

Прорецензовано / Revised: 15.11.25

Схвалено до друку / Accepted: 21.12.25

Valentyna PLESKACH, DSc (Econ.), PhD (Engin.), Prof.

ORCID ID: 0000-0003-0552-0972

e-mail: v_pleskach@ukr.net

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Yaroslav ZHYLIUK, PhD Student

ORCID ID: 0009-0008-9341-9164

e-mail: y.zhyliuk@gmail.com

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

PRIVACY IN STREAMING DATA: A HYBRID APPROACH WITH DIFFERENTIAL PRIVACY AND ROBUST FEDERATED LEARNING

Background. Federated learning (FL) has become a key paradigm for privacy-preserving machine learning on decentralized data in real-time distributed systems. However, it remains vulnerable to two distinct threats: attacks to infer private information from shared model updates and data poisoning attacks from malicious clients.

Methods. This paper presents a novel hybrid architecture designed to simultaneously address both threats in distributed streaming data processing environments, integrating client-side local differential privacy (LDP) to provide formal guarantees of privacy of model updates and server-side aggregation algorithms (coordinate median and Krum) to ensure robustness.

Results. An extensive empirical evaluation is performed using a simulated environment. The results quantify the trade-off between model utility and privacy, demonstrating a smooth decrease in accuracy as privacy (ϵ) increases. Furthermore, it is shown that the proposed architecture, especially with the Krum aggregator, maintains high model accuracy even in the presence of up to 30% malicious clients, a scenario in which standard federated averaging (FedAvg) completely fails.

Conclusions. The proposed architecture demonstrates that combining local privacy mechanisms with aggregation provides a comprehensive, layered security strategy for federated systems. This hybrid approach is critical for deploying secure and reliable real-time systems in real, adversarial environments.

Keywords: distributed systems, real-time, federated learning, differential privacy, streaming data, robust aggregation, Krum, poisoning attacks, cybersecurity.

Плескач Валентина є членом редакційної колегії цього видання, тому вона не брала участі в процесі рецензування та прийнятті рішення щодо публікації цієї статті.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

Pleskach Valentyna is a member of the Editorial Board of this publication; therefore, she did not participate in the peer review process or the decision-making regarding the publication of this article.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

УДК 004.056.5:004.7.056.53:004.89

DOI: <https://doi.org/10.17721/ISTS.2025.10.83-95>

Юрій КРАВЧЕНКО, д-р техн. наук, проф.

ORCID ID: 0000-0002-0281-4396

e-mail: kr34@ukr.net

Київський національний університет імені Тараса Шевченка, Київ, Україна

Олена ФІСУН, асп.

ORCID ID: 0009-0009-6640-3818

e-mail: lenafi@ukr.net

Київський національний університет імені Тараса Шевченка, Київ, Україна

МЕТОД ЗАБЕЗПЕЧЕННЯ ФУНКЦІОНАЛЬНОЇ СТІЙКОСТІ ПРОГРАМНО-КЕРОВАНИХ МЕРЕЖ В УМОВАХ ВПЛИВУ КІБЕРАТАК

Вступ. В умовах стрімкого розвитку мережевих технологій архітектура Software-Defined Networking (SDN) стала основою для побудови гнучких і керованих корпоративних мереж. Проте її централізована природа створює потенційні вразливості, що можуть бути використані зловмисниками для дестабілізації управління мережею. Забезпечення функціональної стійкості SDN в умовах кібератак різної інтенсивності є складною задачею, оскільки потребує динамічної адаптації параметрів управління для запобігання переходу системи в критичні стани.

Методи. Запропоновано метод підтримання функціональної стійкості SDN на основі поєднання стохастичного моделювання та рекурентних нейронних мереж типу LSTM (Long Short-Term Memory). Модель SDN формалізовано як систему з п'ятьма можливими станами та набором інтенсивностей переходів між ними. LSTM-модуль виконує функцію адаптивного регулятора, який у процесі роботи оцінює імовірності перебування системи в різних станах і коригує окремі інтенсивності переходів із метою підтримання цільового рівня функціональної стійкості. Для навчання LSTM застосовано поєднання імітаційного та підкріплювального підходів, що дає змогу враховувати як попередньо змодельовані сценарії атак, так і поточний стан мережі.

Результати. Розроблений метод продемонстрував ефективність у підтриманні стійкого функціонування SDN за рахунок своєчасного коригування параметрів системи у разі зміни інтенсивності загроз. Імітаційне моделювання для трьох сценаріїв (низький, середній і високий рівні загрози) показало, що без корекції параметрів стійкість зберігається лише за низького рівня атак, тоді як для середнього та високого рівнів у певні моменти часу спостерігаються критичні відхилення. Застосування LSTM-контролера дозволило вчасно виявляти ці моменти й адаптивно оновлювати інтенсивності переходів, відновлюючи стійкість мережі.

Висновки. Запропонований метод забезпечення функціональної стійкості SDN на основі LSTM дає змогу створити інтелектуальну систему адаптивного управління, здатну підтримувати стабільність мережі в умовах дестабілізуючих впливів кібератак. Результати моделювання підтверджують доцільність використання нейромережевих механізмів для прогнозування станів і корекції параметрів SDN у реальному часі. У перспективі планується розширити модель для багаторівневих SDN-архітектур і дослідити вплив просторової топології на ефективність адаптивного управління.

Ключові слова: програмно-керована мережа, функціональна стійкість, кібератака, стохастичне моделювання, рекурентна нейронна мережа.

Вступ

Сучасні програмно-керовані мережі (Software-Defined Networks, SDN) відіграють ключову роль у побудові гнучких, масштабованих і керованих інфраструктур передачі даних. Проте їхня централізована архітектура робить такі мережі вразливими до різноманітних кібератак, здатних призвести до деградації або повної втрати функціональної стійкості. Забезпечення безперервного функціонування SDN в умовах змінних загроз вимагає розроблення інтелектуальних методів управління, здатних адаптивно реагувати на аномалії та підтримувати стабільний стан системи (Кравченко, & Фісун, 2025).

Загальна постановка проблеми. Відокремлення площини управління від площини передачі даних у SDN забезпечує централізоване керування мережею, гнучке маршрутизаційне планування й ефективне використання ресурсів. Водночас така архітектура створює нові вектори кібератак на контролер і мережеві пристрої, що актуалізує проблему забезпечення функціональної стійкості SDN. Стохастичний характер динаміки мережі та часові залежності між атаками і реакцією системи ускладнюють підтримання стійкості, тому доцільним є використання оптимального управління інтенсивностями переходів між станами.

Отже, суть проблеми полягає у розробленні моделі та методу адаптивного управління станами SDN, що забезпечує її функціональну стійкість у динамічних умовах кібератак і мінімізує ризики деградації мережі.

Огляд літератури. У сфері забезпечення функціональної стійкості SDN під час кібератак проведено чимало досліджень, які розглядають загальні загрози, відмовостійкість і методи виявлення атак. Наприклад в роботі (Scott-Hayward, O'Callaghan, & Sezer, 2013) узагальнено безпекові аспекти площин SDN – контролера, комутаторів і протоколу OpenFlow – та класифіковано основні сценарії атак. Перевагою роботи є систематизація векторів загроз і контрзаходів, проте вона має описовий характер і не враховує кількісне моделювання функціональної стійкості чи адаптивне управління мережею.

В огляді (Silva et al., 2015) розглянуто відмовостійкість і резервування контролерів і мережевих елементів, що створює корисну основу для побудови архітектури SDN. Водночас дослідження зосереджується переважно на

інфраструктурних рішеннях і майже не охоплює динамічне моделювання стохастичних переходів між станами мережі під час цілеспрямованих атак.

Аналітичні підходи, зокрема й дослідження (Kotenko et al., 2023), поєднують марковські моделі атак, стохастичні мережі й елементи машинного навчання для оцінювання впливу атак на функціональність SDN. Перевагою роботи є формалізація взаємозв'язку між атаками та параметрами мережі, однак питання керування інтенсивностями переходів (λ) та розроблення адаптивних стратегій управління залишаються недостатньо опрацьованими.

Дослідження із застосуванням LSTM (Long Short-Term Memory) і гібридних глибоких мереж (Sahu, Choudhury, & Misra, 2022) демонструють високу ефективність у класифікації та виявленні атак у SDN, завдяки здатності відтворювати часові залежності. Водночас вони не розглядають питання керування інтенсивностями переходів між станами мережі – ключового для підтримання її стійкості. Методи Deep Reinforcement Learning (Sun et al., 2019) показують здатність агентів адаптивно керувати маршрутизацією чи QoS, проте орієнтовані вони переважно на продуктивність, а не на збереження працездатності під час атак.

Крім того, у роботі (Chen et al., 2021) використано глибокі нейронні мережі для виявлення DDoS-атак, але без подальшого управління після детекції. Аналогічно, у (Wang, Li, & Xu, 2022) розроблено RL-модель для оптимізації трафіка, проте без урахування стійкості до атак. Аналогічно, Liu, Zhao, & Yang (2023) запропонували CNN-attention архітектуру для точнішої детекції, однак без реалізації механізмів відновлення. Автори (Alharbi et al., 2023) запропонували блокчейн-підхід до безпеки SDN, який обмежений у застосуванні в реальному часі.

Інші сучасні роботи також підтверджують актуальність проблеми. Наприклад, Chhaytli, & Persikov (2021) розглядали безпечну маршрутизацію як засіб підвищення кіберстійкості. Автори Farooq, Riaz, & Alvi (2023) визначили ключові виклики безпеки SDN, а Hoang, Nguyen, & Pham (2024) показали ефективність LSTM у прогнозуванні мережевих маршрутів. Утім, усі ці підходи зосереджені на підвищенні QoS чи безпеки окремих компонентів, але не забезпечують адаптивного управління для підтримання функціональної стійкості SDN у динамічних умовах кібератак.

Отже, наведений огляд робіт показує, що попри значну кількість досліджень у сфері стійкості та безпеки SDN, більшість із них фокусуються або на виявленні атак, або на підвищенні продуктивності мережі. Недостатньо розробленими залишаються методи, які комплексно враховували б часову динаміку станів SDN, адаптивну зміну параметрів системи та критерії функціональної стійкості під час кібератак.

Метою нашої статті є забезпечення функціональної стійкості програмно-орієнтованих мереж (SDN) в умовах дії кібератак шляхом розроблення методу адаптивного керування параметрами мережі на основі оцінювання її поточного стану та рівня загрози.

Для досягнення поставленої мети визначено такі основні завдання дослідження:

1. Проаналізувати сучасні підходи до забезпечення стійкості та безпеки SDN у контексті кібератак різної інтенсивності.
2. Формалізувати модель функціонування SDN як стохастичної системи, здатної до переходу між різними станами під впливом зовнішніх загроз.
3. Розробити метод адаптивного управління, який забезпечує підтримання заданого рівня функціональної стійкості мережі.
4. Провести імітаційне моделювання процесу зміни ймовірностей перебування SDN у різних станах за різних рівнів загроз.

Методи

Обґрунтування підходу щодо забезпечення функціональної стійкості SDN. Концептуальна структура SDN включає прикладний рівень, API, рівень керування, протокол OpenFlow та рівень даних (рис. 1) (Karakus, & Durresi, 2017). Рівень даних, представлений програмними або апаратно-програмними комутаторами, забезпечує передачу трафіка L2/L3, виконуючи отримані від контролера правила через канал керування OpenFlow. Контролер є центральним елементом, що керує мережею, формує маршрути, збирає інформацію про стан пристроїв і надає інтерфейси API для взаємодії з мережевими застосунками.

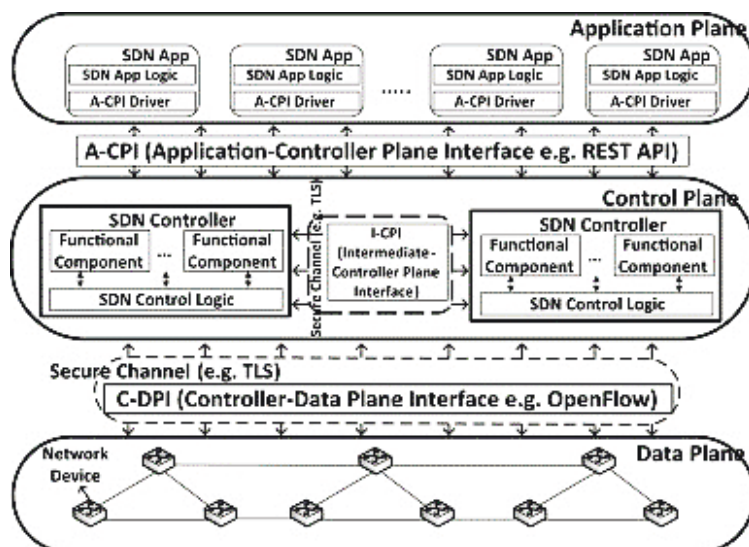


Рис. 1. Концептуальна структура SDN (Karakus, & Durresi, 2017)



Разом із перевагами, централізоване управління в SDN створює нові ризики. Основними вразливими елементами SDN є контролер, канал керування, користувацькі з'єднання та відкриті потоки комутаторів. Тому забезпечення функціональної стійкості SDN розглядають як здатність мережі підтримувати свої функції та продуктивність навіть під впливом кібератак.

Розглянемо модель функціональної стійкості SDN. Для оцінювання функціональної стійкості SDN необхідно визначити умови, за яких мережа втрачає здатність виконувати свої основні функції. Логічно її можна поділити на транспортну та локальну компоненти. Транспортну мережу вважають непрацездатною у випадку відмови або підміни контролера з метою перехоплення трафіка, виходу з ладу маршрутизаторів, що забезпечують транспортну складову, змін топології, які створюють "чорні діри" для трафіка, або порушення зв'язку між вузлами. З урахуванням цих факторів функціональну стійкість SDN із резервними комутаторами OpenFlow та контролером можна описати марковською моделлю з дискретними станами у неперервному часі, де час перебування в кожному стані підпорядковується експоненціальному закону розподілу (Kotenko et al., 2023). На рис. 2 наведено граф можливих станів системи та переходів між ними.

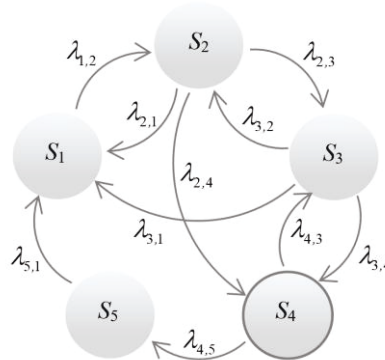


Рис. 2. Граф дискретних станів та умовних переходів в SDN

У загальному випадку для розподіленої корпоративної SDN під впливом кібератак можна виділити такі дискретні стани:

- S_1 – нормальна стійка робота мережі;
- S_2 – фаза розвідки, коли зломисник збирає інформацію про цільову інфраструктуру;
- S_3 – активні спроби реалізації атаки;
- S_4 – успішна компрометація з порушенням функціонування мережі (отримання зломисником доступу до контролера або інших критичних ресурсів);
- S_5 – відновлення після атаки (відновлення працездатності).

Ключові переходи між цими станами включають:

- λ_{12} – виникнення можливостей підключення зломисника (напр., через публічні мережі);
- λ_{21} – ухилення зломисника від атаки;
- λ_{23} – накопичення зломисником достатньої кількості розвідувальної інформації;
- λ_{24} – проведення непідготовленої атаки;
- λ_{31} – невдала атака з виявленням;
- λ_{32} – невдала атака без виявлення;
- λ_{34} – успішна компрометація;
- λ_{43} – запобігання доступу завдяки превентивним заходам;
- λ_{45} – виявлення аномалій у трафіку та поведінці пристроїв;
- λ_{51} – перезавантаження мережевих елементів із застосуванням оновлених параметрів.

Отже, початкові дані для задачі такі:

1. Граф агрегованих стійких станів SDN в умовах кібератаки $G = (S, \lambda)$ (рис. 2).

2. Набір станів SDN під час кібератаки:

$$S = (S_1, S_2, S_3, S_4, S_5). \quad (1)$$

3. Набір потоків подій, коли стани SDN змінюються в умовах кібератаки:

$$\Lambda = (\lambda_{12}, \lambda_{21}, \dots, \lambda_{ij}). \quad (2)$$

4. Вектор імовірностей початкових станів системи: $p_i(0) = (1, 0, 0, 0, 0)$.

5. Умова нормалізації:

$$\sum_{i=1}^5 p_i(t) = 1. \quad (3)$$

Інтенсивності потоку λ_{ij} є ключовим параметром, який визначає середню кількість подій, що виникають за одиницю часу. Числові значення інтенсивностей λ_{ij} можуть бути визначені на основі імітаційної моделі системи. Для аналізу динаміки переходів застосовують неперервний час $t \rightarrow \infty$, у межах якого розв'язується система лінійних диференціальних рівнянь із постійними коефіцієнтами, що відповідає опису однорідного марковського процесу.

Згідно з графом станів (рис. 2), складаємо систему диференціальних рівнянь Колмогорова з невідомими функціями $p_i(t)$ (Kravets et al., 2024):



$$\begin{aligned}
\frac{\partial p_1(t)}{\partial t} &= \lambda_{51}p_5(t) + \lambda_{31}p_3(t) - \lambda_{12}p_1(t), \\
\frac{\partial p_2(t)}{\partial t} &= \lambda_{12}p_1(t) + \lambda_{32}p_3(t) - \lambda_{23}p_2(t), \\
\frac{\partial p_3(t)}{\partial t} &= \lambda_{23}p_2(t) + \lambda_{43}p_4(t) - (\lambda_{31} + \lambda_{32} + \lambda_{34})p_3(t), \\
\frac{\partial p_4(t)}{\partial t} &= \lambda_{34}p_3(t) - (\lambda_{43} + \lambda_{45})p_4(t), \\
\frac{\partial p_5(t)}{\partial t} &= \lambda_{45}p_4(t) - \lambda_{51}p_5(t), \\
\sum_{i=1}^5 p_i(t) &= 1.
\end{aligned} \tag{4}$$

Кожне рівняння системи (4) відображає зміну ймовірності перебування SDN у певному стані залежно від можливих переходів між цим та іншими станами. На відміну від випадкових технічних збоїв, кібератаки мають цілеспрямований і багаторівневий характер, що може призвести до значних порушень у роботі мережі. Відтак забезпечення її функціональної стійкості передбачає не лише підвищення надійності компонентів, але й формування здатності системи своєчасно виявляти, локалізувати та нейтралізувати атаки за збереження працездатності.

Аналізуючи модель SDN, зображену на рис. 2, варто зазначити, що лише стан S_4 описує відмову системи внаслідок успішної атаки, тобто стан, якого необхідно уникнути. Тому показник функціональної стійкості доцільно визначати як сумарну ймовірність перебування системи в усіх інших станах $S_1 + S_2 + S_3 + S_5$, що характеризують її здатність зберігати працездатність за більшості типових умов.

Водночас перед оцінюванням цього показника необхідно визначити ризик кібератаки, що передбачає побудову її математичної моделі на основі вербального опису сценарію атаки.

Модель кібератаки на SDN базується на послідовності етапів, які відображають логіку дій зломисника – від збору інформації до приховування слідів. Такий підхід дозволяє описати кібератаку як послідовність станів і переходів між ними, що створює основу для математичного моделювання процесу (Kotenko et al., 2023).

Для кількісного опису цього процесу доцільно застосувати апарат *сигнальних графів*, який дає змогу подати кібератаку як сукупність взаємопов'язаних елементарних підпроцесів із певними ймовірностями та часовими характеристиками. Вузли графа відповідають станам системи, а ребра – можливим подіям чи переходам між ними. Використання правил топологічних перетворень і формули Мейсона дозволяє отримати узагальнену функцію передачі, що характеризує ймовірнісну динаміку розвитку атаки (Srikant, & Ying, 2014; Robert, 2013).

Розглянемо модель атаки "Підміна топології мережі". Щоб побудувати математичну модель кібератаки на зразок "Підміна топології мережі", подамо сценарій її реалізації у вигляді послідовності дій (табл. 1).

Таблиця 1

Сценарій кібератаки на зразок "Підміна топології мережі"

№	Опис етапу кібератаки	Позначення етапу	Час процесу	Функція розподілу часу
1	Перевірка доступності каналу зв'язку з атакованою мережею	$w(s)$	$\bar{t}_{test}$	$W(t)$
2	Обмін повідомленнями з контролером через протокол OpenFlow з одночасною підміною легітимного вузла шкідливим пристроєм	$m(s)$	$\bar{t}_{tac}$	$M(t)$
3	Передача контролеру зібраної мережевої статистики через протокол керування й аналіз отриманої відповіді	$l(s)$	$\bar{t}_{tr}$	$L(t)$
4	Зміна конфігурації мережі подачею фальшивих даних статистики	$z(s)$	$\bar{t}_{repl}$	$Z(t)$
5	Маніпулювання роботою мережі через введення контролера в оману фальшивими повідомленнями OpenFlow	$d(s)$	$\bar{t}_{re}$	$D(t)$

Наслідком атаки "підміна топології мережі" є проникнення зломисника під виглядом довіреного мережевого пристрою з використанням ідентифікаційних даних, здобутих у процесі розвідки. Опис цієї атаки за даними з табл. 1 наведено у вигляді сигнального графа (рис. 3) (Qin, Li, & Zhang, 2023).

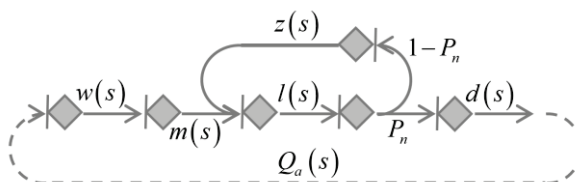


Рис. 3. Сигнальний граф для атаки "Підміна топології мережі"

Для обчислення еквівалентної функції використовують концепцію замкненого сигнального графа, що враховує наявність циклів першого та k -го порядків (Srikant, & Ying, 2014; Robert, 2013). Еквівалентна функція циклу k -го порядку

$$Q_k(k) = \prod_{i=1}^k Q_i(s). \tag{5}$$

де $Q_i(s)$ – еквівалентна функція i -го циклу першого порядку, що визначається як добуток еквівалентних функцій усіх гілок, які формують цей цикл.



Це також дає змогу застосувати топологічне рівняння Мейсона (Lu, Hsu, & Chang, 2020) для замкнених сигнальних графів із метою обчислення еквівалентної функції початкової мережевої моделі:

$$H = 1 + \sum_{k=1}^K (-1)^k Q_k(s) = 0, \quad (6)$$

де K – максимальний порядок циклів, що входять до сигнального графа.

Приєднання до сигнального графа допоміжної (фіктивної) гілки $Q_a(s) = \frac{1}{h(s)}$, де $h(s)$ позначає шукану еквівалентну функцію, дає можливість виділити та перерахувати всі цикли графа.

Цикли першого порядку

$$\frac{w(s)m(s)l(s)P_n d(s)}{h(s)}, (1 - P_n)z(s)l(s). \quad (7)$$

Оскільки в цьому сигнальному графі відсутні цикли другого та вищих порядків, рівняння (7) може бути спрощене та подане у такій формі:

$$1 - \frac{w(s)m(s)l(s)P_n d(s)}{h(s)} - (1 - P_n)z(s)l(s) = 0. \quad (8)$$

У цьому випадку еквівалентна функція набуває вигляду

$$h(s) = \frac{w(s)m(s)l(s)P_n d(s)}{1 - (1 - P_n)z(s)l(s)}. \quad (9)$$

Для визначення розрахованого відношення функції розподілу припустимо, що часові тривалості всіх елементарних процесів на ребрах графа є взаємно незалежними випадковими величинами з експоненціальним законом розподілу з параметрами

$$W(t) = 1 - e^{-wt}, M(t) = 1 - e^{-mt}, \\ L(t) = 1 - e^{-lt}, D(t) = 1 - e^{-dt}, Z(t) = 1 - e^{-zt}, \quad (10)$$

де $w = \frac{1}{\bar{t}_{test}}, m = \frac{1}{\bar{t}_{tac}}, l = \frac{1}{\bar{t}_{tr}}, d = \frac{1}{\bar{t}_{re}}, z = \frac{1}{\bar{t}_{repl}}$, а $\bar{t}_{test}, \bar{t}_{tac}, \bar{t}_{tr}, \bar{t}_{re}, \bar{t}_{repl}$ – середній час реалізації k -го етапу атаки.

Застосування перетворення Лапласа дає змогу отримати зображення функцій щільності розподілу часу виконання k -го етапу атаки. Таке зображення являє собою функцію у частотній області, отриману шляхом інтегрального перетворення Лапласа вихідної функції із часової області. Воно відображає поведінку процесу не у часі, а у спектральному вимірі, що дає змогу спростити аналіз складних систем (Amin, & Ahmad, 2022).

Зазначене перетворення визначається співвідношенням $F(s) = \int_0^\infty e^{-st} f(t) dt$, де $f(t)$ – функція в часовій області (залежність від часу t); $F(s)$ – її зображення у частотній області, отримане за допомогою перетворення Лапласа. Змінна s є комплексною і називається частотою Лапласа. Зазначений підхід дозволяє перетворювати диференціальні рівняння часової області на алгебраїчні рівняння у частотній області, що полегшує процес їхнього розв'язання. Наприклад, для експоненціальної функції $f(t) = e^{at}$, її зображення за перетворенням Лапласа матиме вигляд $F(s) = \frac{1}{s-a}$.

У нашому випадку

$$w(s) = \int_0^\infty e^{-st} d[W(t)] = \frac{w}{w+s}, \\ m(s) = \int_0^\infty e^{-st} d[M(t)] = \frac{m}{m+s}, \\ l(s) = \int_0^\infty e^{-st} d[L(t)] = \frac{l}{l+s}, \\ d(s) = \int_0^\infty e^{-st} d[D(t)] = \frac{d}{d+s}, \\ z(s) = \int_0^\infty e^{-st} d[Z(t)] = \frac{z}{z+s}. \quad (11)$$

Підставивши (10) у вираз (11), а отримані результати – у вираз (9), одержимо

$$h(s) = \frac{wmlP_n d(z+s)}{(w+s)(d+s)(m+s)[(l+s)(z+s) - (1 - P_n)zl]}. \quad (12)$$

Для спрощення подальших обчислень доцільно виконати заміну окремих коефіцієнтів при змінній s у знаменнику рівняння (12), що дозволить отримати зручнішу форму для аналітичного розв'язання:

$$A = d + l + m + w + z, \\ B = mw + mz + wz + d(l + m + w + z) + l(m + w + z - (1 - P_n)z), \\ C = mwz - lw(-1 + (1 - P_n)z) + lm(w + z - (1 - P_n)z) + d(wz + m(w + z) + l(m + l + z - (1 - P_n)z)), \\ D = -lmw(-1 + (1 - P_n)z) + (mwz - lw(-1 + (1 - P_n)z) + lm(w + z - (1 - P_n)z)), \\ E = dlmwz - dlm(1 - P_n)wz. \quad (13)$$

Щоб знайти оригінал еквівалентної функції (12), скористаємося дискретною формою розкладання Гевісайда (методом розкладання на прості дробі) (Malik, & Singh, 2023):

$$h(s) = \sum_{k=1}^n \frac{f(s_k)}{\varphi'(s_k)} \cdot \frac{1}{s-s_k} = \sum_{k=1}^5 \frac{wmlP_n d(z+s_k)}{5s_k^4 + 4As_k^3 + 3Bs_k^2 + 2Cs_k + D} \cdot \frac{1}{s-s_k}, \quad (14)$$

де $\varphi'(s_k)$ – похідна знаменника (12); s_k – розкладання полюсів (корені рівняння $s^5 + As^4 + Bs^3 + Cs^2 + Ds + E = 0$); w, m, l, d, z – визначають так само, як і в (10).

Тоді

$$h(t) = L^{-1}\{h(s)\} = \sum_{k=1}^5 \frac{wmlP_n d(z+s_k)}{5s_k^4 + 4As_k^3 + 3Bs_k^2 + 2Cs_k + D} \cdot e^{s_k t}. \quad (15)$$

Оскільки отриманий вираз описує функцію щільності ймовірностей, то відповідну інтегральну функцію розподілу ймовірностей визначають як її інтеграл від нуля до поточного моменту часу:

$$F(t) = \int_0^t h(t)dt = \sum_{k=1}^5 \frac{wmlP_n d(z+s_k)}{5s_k^4 + 4As_k^3 + 3Bs_k^2 + 2Cs_k + D} \cdot \frac{1-e^{s_k t}}{(-s_k)} \quad (16)$$

Середній час $\bar{t}_{attack}$, витрачений на реалізацію кібератаки, знаходять як

$$\bar{t}_{attack} = \int_0^\infty th(t)dt = \sum_{k=1}^5 \frac{wmlP_n d(z+s_k)}{5s_k^4 + 4As_k^3 + 3Bs_k^2 + 2Cs_k + D} \cdot \frac{1}{(-s_k)^2} \quad (17)$$

Отже, було визначено інтегральну функцію розподілу та обчислено середній час $\bar{t}_{attack}$ реалізації кібератаки. Аналогічні розрахунки можна виконати для кожного типу кібератаки, що дає змогу здійснити комплексне оцінювання їхнього впливу на стійкість SDN.

Приклад реалізації моделі кібератаки на SDN. Для ілюстрації підходу розглянемо атаку "Підміна топології мережі", використавши часові показники з табл. 1 як параметри елементарних підпроцесів: $\bar{t}_{test} = 4.3$ хв, $\bar{t}_{tac} = 2.5$ хв, $\bar{t}_{tr} = 5.0$ хв, $\bar{t}_{repl} = 5.5$ хв, $\bar{t}_{re} = 3.4$ хв. Імовірність успішного перехоплення управління в SDN змінюватимемо у межах $P_n = 0.1 \dots 1.0$.

На основі вихідних даних побудовано залежності $F(t)$ та T_a , (рис. 4–5), що дозволяють визначити: імовірність $P_n(t \leq T_a)$ реалізації кібератаки за час, не більший заданого T_a ; середній час реалізації атаки; а також часову межу, яка відповідає заданому рівню загрози реалізації кібератаки.

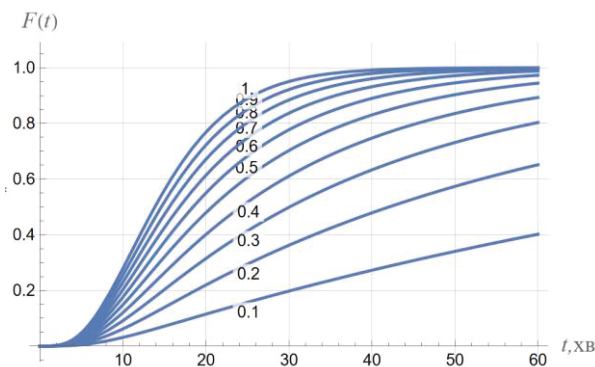


Рис. 4. Інтегральна функція розподілу ймовірності від часу реалізації кібератаки

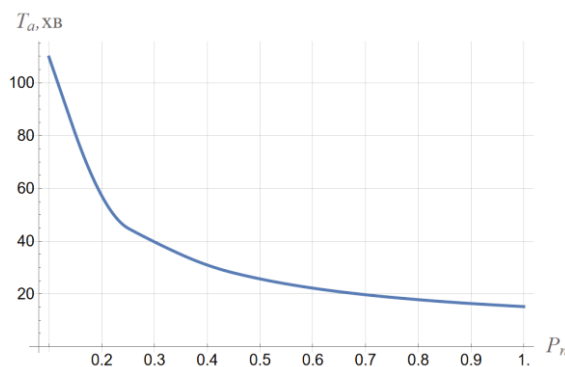


Рис. 5. Залежність середнього часу реалізації кібератаки від імовірності P_n

Аналіз побудованих залежностей (рис. 4–5) показує, що для реалізації кібератаки на зразок "Підміна топології мережі" за заданих часових характеристик етапів атаки необхідно близько 170 хв у разі ймовірності успішного перехоплення управління в SDN при $P_n = 0.1$, і приблизно 22 хв при $P_n = 0.9$. Отримані графічні залежності дають змогу оцінити, як зміна ймовірності успішного перехоплення управління в SDN у межах заданого часу впливає на функцію розподілу тривалості реалізації атаки. Спостерігається тенденція до зменшення середнього часу виконання атаки зі збільшенням значення P_n . Водночас із подальшим зростанням P_n ступінь його впливу на інтегральну функцію розподілу $F(t)$ поступово зменшується, і після досягнення певного порогу цей вплив стає незначним. Отже, отримано інтегральні функції розподілу та середній час $\bar{t}$ реалізації кібератаки, що може бути використано для аналогічних розрахунків щодо інших типів атак на SDN.

Метод оптимального управління для забезпечення функціональної стійкості SDN в умовах кібератак. Задачу забезпечення функціональної стійкості SDN в умовах деструктивного кібервпливу можна подати у вигляді задачі оптимального управління динамічною системою (4), у якій керуючі параметри $\lambda_{ij}(t)$ можуть мати безперервний або дискретний характер.

Формулювання задачі. Еволюція станів системи $p(t) = (p_1, \dots, p_5)$ описується системою звичайних диференціальних рівнянь $p'(t) = f(p(t), \lambda(t))$, що відповідає марковській моделі (4).

В межах поставленої мети – забезпечення функціональної стійкості SDN – завдання управління мережею полягає у максимізації інтегрального критерію ефективності

$$J[\lambda(\cdot)] = \int_0^T R(t)dt = \int_0^T (p_1(t) + p_2(t) + p_3(t) + p_5(t))dt \rightarrow \max_{\lambda(t)} \quad (18)$$



при обмеженнях на керування $\lambda_{ij}(t) \in [0, \lambda_{ij}^{max}]$, який відображає сумарну ймовірність перебування системи у працездатних станах протягом заданого інтервалу часу. Така постановка задачі дозволяє визначити оптимальну стратегію впливу на параметри системи з метою мінімізації наслідків кібератак і підтримання безперервності функціонування SDN.

Оскільки еволюція станів SDN характеризується часовою залежністю та наявністю "пам'яті" щодо попередніх подій – спроб кібератак і реакцій системи на них – для розв'язання задачі (18) та формування оптимального закону управління інтенсивностями переходів $\lambda_{ij}(t)$ найдоцільнішим є застосування рекурентних нейронних мереж (RNN). Цей клас моделей здатний ефективно відтворювати послідовні залежності, поєднуючи інформацію про поточний стан мережі з історією її функціонування, що дає змогу прогнозувати майбутні переходи між станами. Завдяки такій властивості RNN можуть адаптивно змінювати керуючі впливи $\lambda_{ij}(t)$ з метою максимізації інтегрального показника функціональної стійкості SDN у межах заданого часового інтервалу $[0, T]$. Це забезпечує підтримання безперервності та безпеки роботи мережі навіть за умов динамічних і непередбачуваних кібератак.

Параметризація закону управління SDN. Для практичної реалізації підходу потрібно формалізувати керування інтенсивностями переходів $\lambda_{ij}(t)$ у вигляді закону управління, який враховує поточний стан SDN та історію попередніх подій. Для цього доцільно застосувати рекурентні нейронні мережі типу LSTM, що дозволяють параметризувати закон управління як функцію від поточних спостережень і попередніх керувань.

На момент k (у час t_k) закон управління отримує вхідний вектор x_k , який зазвичай включає: поточні або апроксимовані ймовірності станів $p(t_k) = (p_1, \dots, p_5)$ (або інші метрики – ентропія потоків, лічильники аномалій тощо); поточний час t_k або нормалізований час $\frac{t_k}{T}$; попередні керування $\lambda_{ij}(t_{k-1})$; а також додаткові контекстні ознаки (стан контролера, завантаження CPU, індикатори зовнішніх загроз).

У такій концепції RNN (LSTM/GRU) потрібні тому, що вплив атак і захисних заходів має часову пам'ять: важливо врахувати темп наростання розвідки, повторні спроби чи застосовані контрзаходи. Рекурентна мережа зберігає прихований стан h_k , який акумулює цю історичну інформацію, і на його основі формує наступні значення інтенсивностей $\lambda_{ij}(t_{k+1})$.

Обґрунтування застосування мережі LSTM. LSTM – це спеціалізований тип рекурентних нейронних мереж (RNN), призначений для ефективного опрацювання послідовних даних із довготривалими залежностями. Архітектура LSTM уможливує подолання проблеми "згасання градієнта", характерної для класичних RNN. Завдяки своїй структурі LSTM забезпечує механізм вибіркового запам'ятовування або забування інформації, що дає змогу мережі адаптивно керувати часовим контекстом (Krichen, & Mihoub, 2025).

Послідовність кроків функціонування LSTM зображено на рис. 6.

Крок 1. Визначення інформації для видалення зі стану чарунки (рис. 6а). На цьому етапі сигмоїдальний шар забування (forget gate layer) визначає, які дані зі стану чарунки C_{t-1} слід зберегти, а які – відкинути. Він аналізує вхідний вектор x_t та попередній вихід h_{t-1} , після чого для кожного елемента стану формує значення від 0 до 1: 1 – означає "повністю зберегти інформацію"; 0 – "повністю видалити":

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (19)$$

Крок 2. Визначення нової інформації для збереження у стані чарунки (рис. 6б). На цьому етапі вхідний шар (input gate layer) із сигмоїдальною активацією визначає, які саме дані необхідно оновити. Після цього шар з активацією $\tanh$ формує новий вектор кандидатних значень $\tilde{C}_t$, які можна додати до стану чарунки для оновлення її змісту:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (20)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C). \quad (21)$$

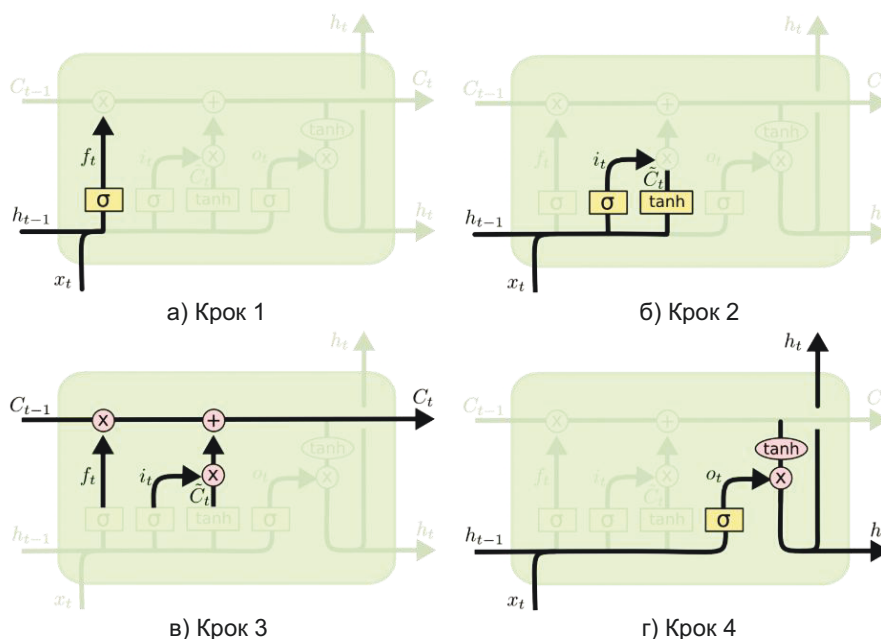


Рис. 6. Функціонування мережі LSTM (Olah, 2015)

Крок 3. Оновлення стану чарунки (рис. 6в). На цьому кроці попередній стан чарунки C_{t-1} перетворюється на новий C_t . Спочатку попередній стан множиться на f_t , що відповідає за "забуту" частину інформації. Потім до нього додається добуток Адамара $i_t \circ \tilde{C}_t$, який представляє нові кандидатні значення, зважені за ступенем їхньої важливості. У результаті формується оновлений стан чарунки C_t , що поєднує релевантну попередню інформацію з новими даними:

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t. \quad (22)$$

Крок 4. Формування вихідного сигналу (рис. 6г). На цьому етапі визначають, яку інформацію слід передати далі. Спочатку сигмоїдальний шар o_t вирішує, які частини оновленого стану чарунки C_t потраплять на вихід. Потім значення стану проходять через функцію активації $\tanh$, що обмежує їх у межах від -1 до 1 . Отриманий результат перемножується з вихідним сигналом сигмоїдального шару, завдяки чому на виході h_t залишається лише найбільш релевантна інформація для подальшого оброблення:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (23)$$

$$h_t = o_t \circ \tanh(C_t). \quad (24)$$

Застосування мереж LSTM для побудови методу оптимального управління функціональною стійкістю SDN в умовах кібератак має суттєві переваги завдяки їхній здатності моделювати часові залежності та запам'ятовувати попередні стани системи. LSTM забезпечують ефективне врахування історії динаміки атак і реакцій мережі, що дозволяє адаптивно коригувати інтенсивності переходів між станами SDN з урахуванням контексту попередніх подій.

Навчання LSTM. Для задачі забезпечення функціональної стійкості SDN в умовах кібератак найдоцільнішим є поєднання імітаційного навчання (offline learning) з подальшою онлайн-адаптацією. На першому етапі LSTM може бути попередньо натренована на симульованих сценаріях функціонування мережі, що охоплюють різні типи атак, збоїв і процесів відновлення. Такий підхід дозволяє безпечно сформувати базову політику управління інтенсивностями переходів без ризику впливу на реальну інфраструктуру. Імітаційне навчання дає змогу LSTM отримати початкові навички оптимального реагування на типові ситуації, зменшуючи час адаптації в реальних умовах.

Додатково, для підвищення ефективності процесу оптимізації може бути використаний елемент навчання з підкріпленням (RL) у контурі онлайн-адаптації. У цьому випадку LSTM отримує зворотний зв'язок у вигляді функції винагороди, що відображає рівень стійкості системи, і поступово вдосконалює свої дії. Така гібридна стратегія – "offline pretraining + online reinforcement fine-tuning" – забезпечує як безпечний старт, так і здатність до постійного самонавчання в умовах реальних кіберзагроз.

Схема управління для забезпечення функціональної стійкості SDN в умовах кібератак. Загальну схему управління SDN, побудовану на LSTM-мережі, зображено на рис. 7. Ця схема відображає роботу інтелектуального модуля управління функціональною стійкістю SDN, у якому LSTM-мережа виконує роль адаптивного прогностно-коригуючого елемента.

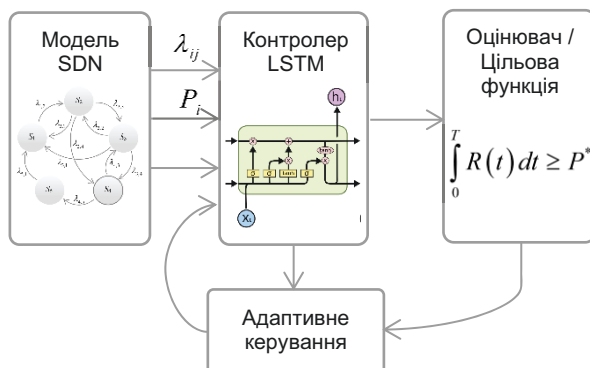


Рис. 7. Схема інтелектуалізованого управління SDN

У схемі на рис. 7 наведено:

1. Блок SDN-моделі (система станів $S_1, \dots, S_5$). Переходи між станами визначаються інтенсивностями λ_{ij} , окремі з яких ($\lambda_{12}, \lambda_{34}, \lambda_{45}, \lambda_{51}$) можуть змінюватися в часі.

2. LSTM-модуль. На вхід LSTM-модуля надходять часові ряди, що характеризують поточний стан SDN: поточні ймовірності станів ($P_1, \dots, P_5$); інтенсивності переходів λ_{ij} ; показники навантаження, затримок, аномалій трафіка; індикатори безпеки (кількість атак, типи подій тощо). LSTM, завдяки своїй пам'яті (cell state), враховує історію цих параметрів, що дозволяє відстежувати тенденції змін системи.

3. Робота LSTM-модуля управління. На основі попередніх і поточних станів LSTM-модуль прогнозує майбутню динаміку і формує керувальні дії – оновлені значення деяких інтенсивностей переходів: ($\lambda_{12}, \lambda_{34}, \lambda_{45}, \lambda_{51}$). У такий спосіб LSTM виконує оптимізацію параметрів системи, спрямовану на збереження функціональної стійкості.

4. Цільова функція (18) – це інтегральний показник стійкості, який відображає ймовірність того, що система перебуває у працездатному стані. LSTM коригує інтенсивності переходів таким способом, щоб підтримувати цей показник на рівні, що не нижчий від заданого порогу P^* .

5. Зворотний зв'язок. Отримані нові значення λ_{ij} подаються назад у модель SDN, яка еволюціонує в часі відповідно до скоригованих параметрів. Результат цього впливу знову надходить на вхід LSTM, утворюючи замкнений контур оптимального управління.



Отже, схема реалізує адаптивне управління на основі прогнозу, де LSTM безперервно навчається на поточних і минулих даних, забезпечуючи максимізацію показника функціональної стійкості SDN в умовах змінних кібератак і навантажень.

Результати

Для демонстрації запропонованого підходу розглянемо три сценарії, що відповідають низькому, середньому та високому рівням загрози кібератак (табл. 2), для яких задано відповідні значення інтенсивностей переходів λ_{ij} . Як видно з табл. 2, сценарії відрізняються співвідношеннями інтенсивностей переходів між станами. Зокрема, за низького рівня загрози спостерігається більша інтенсивність переходів до функціонально стійких станів (S_1, S_2, S_3, S_5), ніж до деградованого стану S_4 . За високого рівня загрози, навпаки, переважають переходи до стану S_4 , що відображає зростання ризику втрати стійкості. Середній рівень загрози характеризується більш збалансованими значеннями інтенсивностей переходів між усіма станами.

Таблиця 2

Початкові значення інтенсивностей переходів між станами SDN для різних сценаріїв реалізації загроз

Позначення	Рівень загрози		
	низький	середній	високий
λ_{12}	0.2	0.3	0.5
λ_{21}	0.5	0.3	0.05
λ_{23}	0.2	0.3	0.4
λ_{24}	0.05	0.3	0.15
λ_{31}	0.5	0.3	0.05
λ_{32}	0.4	0.3	0.02
λ_{34}	0.05	0.3	0.4
λ_{43}	0.5	0.3	0.05
λ_{45}	0.5	0.3	0.2
λ_{51}	0.3	0.3	0.3

Запропонований метод забезпечення функціональної стійкості SDN на основі застосування рекурентних нейронних мереж LSTM реалізується у чотири етапи:

- 1) формування початкових даних;
 - 2) отримання на основі початкових даних залежностей ймовірностей $P_1(t), \dots, P_5(t)$ перебування системи в станах $S_1, \dots, S_5$;
 - 3) запуск LSTM-модуля, прогнозування зміни $P_1(t), \dots, P_5(t)$ і визначення моментів часу, коли $P_{\Phi C}(t) = P_{\Phi C}^*$. Корекція заданих параметрів λ_{ij} шляхом визначення нових оптимальних значень λ_{ij} за деяким інтегральним максимумом $P_1(t) + P_2(t) + P_3(t) + P_5(t)$;
 - 4) оцінювання впливу скоригованих початкових даних на функціональну стійкість системи, яка досліджується.
- У результаті моделювання з використанням параметрів, наведених у табл. 2, отримано відповідні графічні залежності, подані на рис. 8–10.

З аналізу наведених графіків (рис. 8–10) видно, що за заданих інтенсивностей переходів властивість функціональної стійкості SDN, яка визначається ймовірністю перебування системи у працездатних станах $P_{\Phi C}$, зберігається протягом усього періоду лише за умов низького рівня загрози кібератаки. Для середнього та високого рівнів загроз ця властивість має обмежений у часі характер – близько 28 хв і 2 хв відповідно.

Як показує приклад на рис. 8, у сценарії низького рівня загрози система залишається стабільною і не потребує втручання. У випадку середнього рівня загрози (рис. 9) приблизно на 28-й хвилині спостерігається втрата стійкості $P_{\Phi C}(t) \leq P_{\Phi C}^*$, що вимагає коригування параметрів λ_{ij} . Після внесення змін до інтенсивностей переходів отримано нові значення $P_{\Phi C}(t)$, на основі яких побудовано оновлені залежності ймовірностей станів (рис. 11–12). Як показано на рис. 11, після регулювання у цьому сценарії спостерігається ще один момент часу – приблизно на 56-й хвилині, коли система знову переходить у критичний режим роботи. Це свідчить про необхідність повторного коригування параметрів для відновлення функціональної стійкості SDN і запобігання подальшому зниженню її працездатності.

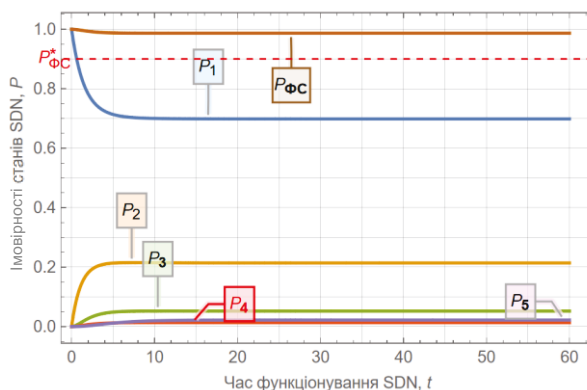


Рис. 8. Зміна ймовірностей станів SDN для низького рівня загрози атаки

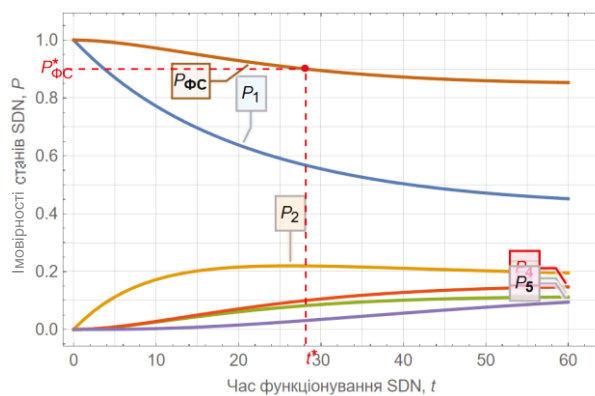


Рис. 9. Зміна ймовірностей станів SDN для середнього рівня загрози атаки

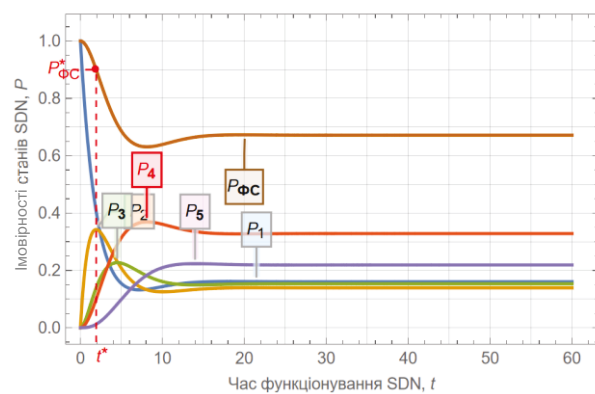


Рис. 10. Зміна ймовірностей станів SDN для високого рівня загрози атаки

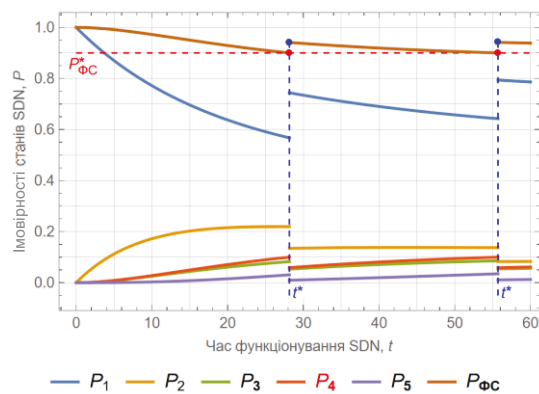


Рис. 11. Зміна ймовірностей станів SDN при середньому рівні загрози атаки з корекцією λ_{ij}

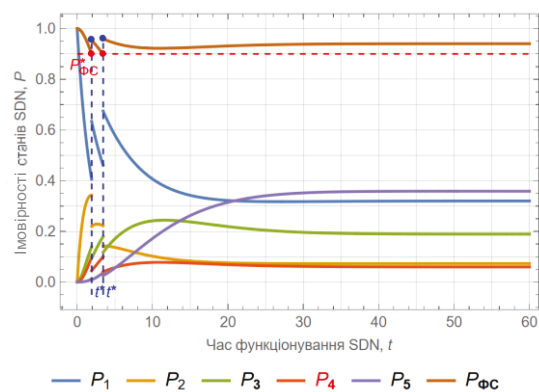


Рис. 12. Зміна ймовірностей станів SDN при високому рівні загрози атаки з корекцією λ_{ij}



У випадку сценарію з високим рівнем загрози (рис. 12) критичні стани системи спостерігаються вже на 2-й та 4-й хвилинах моделювання. Це вимагає оперативнішого втручання та корекції параметрів λ_{ij} у зазначені моменти часу, що дає змогу підтримати функціональну стійкість SDN навіть за умов інтенсивного зовнішнього впливу.

Таким чином, величина λ_{ij} зі стаціонарної стає змінною в часі $\lambda_{ij}(t)$ і підлягає корекції шляхом пошуку оптимальних значень за критерієм (18) на часовому інтервалі від t^* до t_{max} . Варто зазначити, що в наведених сценаріях (рис. 11–12) оптимізація здійснювалася не для всіх λ_{ij} , а лише для тих, що мають найбільший вплив на функціональну стійкість SDN. Зокрема: λ_{12} – ініціація розвідки та контроль підключень (вхід / доступ); λ_{34} – імовірність успішної атаки ("прорив"); λ_{45} – виявлення аномалій (перехід до відновлення); λ_{51} – швидкість відновлення системи.

Параметр λ_{12} впливає на частоту "початку проблем" – визначає частоту виникнення інцидентів: його зменшення знижує кількість атак і підвищує стабільність системи на початкових етапах роботи, що узгоджується із заходами контролю доступу та фільтрації трафіка (NAC). Параметр λ_{34} характеризує ефективність атаки, його зменшення (напр., через ізоляцію або обмеження привілеїв) безпосередньо підвищує показник стійкості $P_{FC}(t)$. Параметр λ_{45} відповідає за швидкість виявлення аномалій, чим він більший, тим швидше система реагує на загрозу. Нарешті, λ_{51} визначає темп відновлення після інциденту, його оптимізація суттєво впливає на інтегральний показник стійкості $P_{FC}(t)$.

Отже, запропонований підхід дає змогу реалізувати метод визначення рівня функціональної стійкості та побудову оптимального управління SDN в умовах кібератак. Отримані результати підтверджують ефективність моделі для оцінювання та підтримання стійкості мережі до типових кіберзагроз.

Дискусія і висновки

1. В результаті аналізу сучасних підходів до забезпечення стійкості програмно-керованих мереж (SDN) встановлено, що більшість існуючих рішень орієнтовано на виявлення атак і відновлення працездатності, проте вони не враховують часову динаміку функціональної стійкості системи. Це зумовлює необхідність розроблення методу, здатного не лише реагувати на інциденти, а й прогнозувати їхній вплив на стабільність SDN у режимі реального часу.

2. Формалізація моделі функціонування SDN як стохастичної системи, яка описується системою диференціальних рівнянь Колмогорова, дозволяє відобразити процеси переходів між станами мережі під впливом зовнішніх кібератак. Такий підхід забезпечує можливість кількісного оцінювання ймовірностей перебування SDN у функціонально стійких і нестійких станах залежно від інтенсивності атак.

3. Розроблений метод адаптивного управління, який ґрунтується на використанні рекурентної нейронної мережі LSTM для прогнозування динаміки станів SDN і коригування параметрів інтенсивностей переходів між станами системи з метою підтримання інтегрального показника функціональної стійкості, забезпечує своєчасне втручання в роботу мережі для запобігання переходу у критичний стан.

4. Результати імітаційного моделювання свідчать, що розроблена модель адекватно відображає поведінку SDN у різних умовах загроз. Для низького рівня кібератак мережа зберігає функціональну стійкість протягом усього періоду спостереження, тоді як для середнього і високого рівнів відбуваються короточасні втрати стійкості, які можуть бути компенсовані за рахунок своєчасного коригування параметрів моделі.

Отже, результати дослідження підтверджують ефективність розробленого методу забезпечення функціональної стійкості SDN в умовах кібератак і створюють основу для побудови інтелектуальних систем прогнозно-адаптивного управління мережею. У подальших дослідженнях планується розширити запропонований метод шляхом урахування просторової топології SDN і взаємодії між її підмережами, а також інтегрувати механізми глибинного навчання для підвищення точності прогнозування динаміки станів мережі. Перспективним напрямом є також розроблення адаптивного контролера, який у режимі реального часу здійснюватиме оптимізацію параметрів стійкості з урахуванням типу й інтенсивності кібератак.

Внесок авторів: Юрій Кравченко – концептуалізація, методологія, аналіз джерел, написання (перегляд і редагування); Олена Фісун – збір даних та їхня валідація, формальний аналіз, комп'ютерне моделювання, перевірка результатів, написання (оригінальна чернетка).

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Кравченко, Ю. В., & Фісун, О. С. (2025). Аналіз існуючих методів забезпечення функціональної стійкості складних систем для удосконалення комп'ютерних мереж. *Зв'язок*, 3, 29–35. <https://doi.org/10.31673/2412-9070.2025.025385>
- Alharbi, A., Alsubhi, K., Alghamdi, A., & Alotaibi, F. (2023). Blockchain-based decentralized security framework for software-defined networks. *IEEE Access*, 11, 36540–36555. <https://doi.org/10.1109/ACCESS.2023.3258792>
- Amin, M. B. M., & Ahmad, S. S. (2022). Laplace Transform for Solving System of Integro-Fractional Differential Equations of Volterra Type with Variable Coefficients and Multi-Time Delay. *Symmetry*, 14(5), 984. <https://doi.org/10.3390/sym14050984>
- Chen, Y., Zhang, L., Wang, Y., & Xu, J. (2021). Deep neural network-based anomaly detection in software-defined networking. *Computer Networks*, 195, 108180. <https://doi.org/10.1016/j.comnet.2021.108180>
- Chhaytli, Y., & Persikov, A. (2021). Cyber-resilient routing mechanisms for SDN. *Journal of Network and Computer Applications*, 187, 103097. <https://doi.org/10.1016/j.jnca.2021.103097>
- Farooq, M., Riaz, Q., & Alvi, S. (2023). Security challenges and solutions in software-defined networking: A systematic review. *Journal of Information Security and Applications*, 74, 103453. <https://doi.org/10.1016/j.jisa.2023.103453>
- Hoang, T., Nguyen, H., & Pham, Q. (2024). LSTM-based predictive routing for adaptive resource allocation in SDN environments. *Future Generation Computer Systems*, 154, 328–340. <https://doi.org/10.1016/j.future.2023.10.026>
- Karakus, M., & Durrezi, A. (2017). Quality of Service (QoS) in Software Defined Networking (SDN): A survey. *Journal of Network and Computer Applications*, 80, 200–218. <https://doi.org/10.1016/j.jnca.2016.12.019>
- Kotenko, I., Saenko, I., Privalov, A., & Laut, O. (2023). Ensuring SDN Resilience under the Influence of Cyber Attacks: Combining Methods of Topological Transformation of Stochastic Networks, Markov Processes, and Neural Networks. *Big Data Cogn. Comput.*, 7, 66. <https://doi.org/10.3390/bdcc7020066>
- Kravets, V., Kapitsa, M., Domanskyi, I., Kravets, V., Hryshechikina, T., & Zakurday, S. (2024). Devising an analytical method for solving the eighth-order Kolmogorov equations for an asymmetric Markov chain. *Eastern-European Journal of Enterprise Technologies*, 5(4(131)), 33–41. <https://doi.org/10.15587/1729-4061.2024.312971>
- Krichen, M., & Mihoub, A. (2025). Long Short-Term Memory networks: A comprehensive survey. *AI*, 6(9), 215. <https://doi.org/10.3390/ai6090215>



- Liu, X., Zhao, P., & Yang, K. (2023). Hybrid deep learning model for anomaly detection in SDN using CNN and attention mechanisms. *IEEE Transactions on Network and Service Management*, 20(2), 2331–2343. <https://doi.org/10.1109/TNSM.2023.3234456>
- Lu, W., Hsu, C., & Chang, Y. (2020). A matrix approach for analyzing signal flow graph. *Information*, 11(12), 562. <https://doi.org/10.3390/info11120562>
- Malik, A., & Singh, B. (2023). Application of Heaviside expansion method for solving fractional order differential equations in control systems. *Fractional Calculus and Applied Analysis*, 26(2), 541–562. <https://doi.org/10.1515/fca-2023-0034>
- Olah, Ch. (2015). Understanding LSTM Networks. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Qin, Y., Li, H., & Zhang, D. (2023). Fault diagnosability evaluation method based on multi-signal flow graph and similarity measure. *Systems Engineering and Electronics*, 45(1), 302–312. <https://doi.org/10.12305/j.issn.1001-506X.2023.01.35>
- Robert, P. (2013). *Stochastic Networks and Queues*. Springer Science & Business Media.
- Sahu, R., Choudhury, S., & Misra, S. (2022). LSTM-based anomaly detection in software-defined networking environments. *IEEE Transactions on Network and Service Management*, 19(3), 2471–2484. <https://doi.org/10.1109/TNSM.2022.3158729>
- Scott-Hayward, S., O'Callaghan, G., & Sezer, S. (2013). SDN Security: A Survey. In *2013 IEEE SDN for Future Networks and Services (SDN4FNS)* (pp. 1–7). <https://doi.org/10.1109/SDN4FNS.2013.6702553>
- Silva, A., Smith, P., & Mauthe, A., & Schaeffer-Filho, A. (2015). Resilience support in software-defined networking: A survey. *Computer Networks*, 92, 189–207. <https://doi.org/10.1016/j.comnet.2015.09.012>
- Srikant, R., & Ying, L. (2014). *Communication Networks: An Optimization, Control, and Stochastic Networks Perspective*. Cambridge University Press.
- Sun, Y., Peng, M., Zhou, Y., Huang, Y., & Mao, S. (2019). Application of deep reinforcement learning in mobile networks: A survey. *IEEE Communications Surveys & Tutorials*, 21(4), 3133–3174. <https://doi.org/10.1109/COMST.2019.2924243>
- Wang, J., Li, Z., & Xu, T. (2022). Reinforcement learning-based adaptive traffic control in SDN. *Computer Communications*, 188, 125–135. <https://doi.org/10.1016/j.comcom.2022.02.012>

References

- Alharbi, A., Alsubhi, K., Alghamdi, A., & Alotaibi, F. (2023). Blockchain-based decentralized security framework for software-defined networks. *IEEE Access*, 11, 36540–36555. <https://doi.org/10.1109/ACCESS.2023.3258792>
- Amin, M. B. M., & Ahmad, S. S. (2022). Laplace Transform for Solving System of Integro-Fractional Differential Equations of Volterra Type with Variable Coefficients and Multi-Time Delay. *Symmetry*, 14(5), 984. <https://doi.org/10.3390/sym14050984>
- Chen, Y., Zhang, L., Wang, Y., & Xu, J. (2021). Deep neural network-based anomaly detection in software-defined networking. *Computer Networks*, 195, 108180. <https://doi.org/10.1016/j.comnet.2021.108180>
- Chhaytli, Y., & Persikov, A. (2021). Cyber-resilient routing mechanisms for SDN. *Journal of Network and Computer Applications*, 187, 103097. <https://doi.org/10.1016/j.jnca.2021.103097>
- Farooq, M., Riaz, Q., & Alvi, S. (2023). Security challenges and solutions in software-defined networking: A systematic review. *Journal of Information Security and Applications*, 74, 103453. <https://doi.org/10.1016/j.jisa.2023.103453>
- Hoang, T., Nguyen, H., & Pham, Q. (2024). LSTM-based predictive routing for adaptive resource allocation in SDN environments. *Future Generation Computer Systems*, 154, 328–340. <https://doi.org/10.1016/j.future.2023.10.026>
- Karakus, M., & Durrresi, A. (2017). Quality of Service (QoS) in Software Defined Networking (SDN): A survey. *Journal of Network and Computer Applications*, 80, 2017, 200–218. <https://doi.org/10.1016/j.jnca.2016.12.019>
- Kotenko, I., Saenko, I., Privalov, A., & Laut, O. (2023). Ensuring SDN Resilience under the Influence of Cyber Attacks: Combining Methods of Topological Transformation of Stochastic Networks, Markov Processes, and Neural Networks. *Big Data Cogn. Comput.*, 7, 66. <https://doi.org/10.3390/bdcc7020066>
- Kravchenko, Yu. V., & Fisun, O. S. (2025). Analysis of existing methods for ensuring functional resilience of complex systems for improving computer networks. *Communication*, 3, 29–35 [in Ukrainian]. <https://doi.org/10.31673/2412-9070.2025.025385>
- Kravets, V., Kapitsa, M., Domanskyi, I., Kravets, V., Hryshechukina, T., & Zakurday, S. (2024). Devising an analytical method for solving the eighth-order Kolmogorov equations for an asymmetric Markov chain. *Eastern-European Journal of Enterprise Technologies*, 5(4(131)), 33–41. <https://doi.org/10.15587/1729-4061.2024.312971>
- Krichen, M., & Mihoub, A. (2025). Long Short-Term Memory networks: A comprehensive survey. *AI*, 6(9), 215. <https://doi.org/10.3390/ai6090215>
- Liu, X., Zhao, P., & Yang, K. (2023). Hybrid deep learning model for anomaly detection in SDN using CNN and attention mechanisms. *IEEE Transactions on Network and Service Management*, 20(2), 2331–2343. <https://doi.org/10.1109/TNSM.2023.3234456>
- Lu, W., Hsu, C., & Chang, Y. (2020). A matrix approach for analyzing signal flow graph. *Information*, 11(12), 562. <https://doi.org/10.3390/info11120562>
- Malik, A., & Singh, B. (2023). Application of Heaviside expansion method for solving fractional order differential equations in control systems. *Fractional Calculus and Applied Analysis*, 26(2), 541–562. <https://doi.org/10.1515/fca-2023-0034>
- Olah, Ch. (2015). Understanding LSTM Networks. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Qin, Y., Li, H., & Zhang, D. (2023). Fault diagnosability evaluation method based on multi-signal flow graph and similarity measure. *Systems Engineering and Electronics*, 45(1), 302–312. <https://doi.org/10.12305/j.issn.1001-506X.2023.01.35>
- Robert, P. (2013). *Stochastic Networks and Queues*. Springer Science & Business Media.
- Sahu, R., Choudhury, S., & Misra, S. (2022). LSTM-based anomaly detection in software-defined networking environments. *IEEE Transactions on Network and Service Management*, 19(3), 2471–2484. <https://doi.org/10.1109/TNSM.2022.3158729>
- Scott-Hayward, S., O'Callaghan, G., & Sezer, S. (2013). SDN Security: A Survey. In *2013 IEEE SDN for Future Networks and Services (SDN4FNS)* (pp. 1–7). <https://doi.org/10.1109/SDN4FNS.2013.6702553>
- Silva, A., Smith, P., & Mauthe, A., & Schaeffer-Filho, A. (2015). Resilience support in software-defined networking: A survey. *Computer Networks*, 92, 189–207. <https://doi.org/10.1016/j.comnet.2015.09.012>
- Srikant, R., & Ying, L. (2014). *Communication Networks: An Optimization, Control, and Stochastic Networks Perspective*. Cambridge University Press.
- Sun, Y., Peng, M., Zhou, Y., Huang, Y., & Mao, S. (2019). Application of deep reinforcement learning in mobile networks: A survey. *IEEE Communications Surveys & Tutorials*, 21(4), 3133–3174. <https://doi.org/10.1109/COMST.2019.2924243>
- Wang, J., Li, Z., & Xu, T. (2022). Reinforcement learning-based adaptive traffic control in SDN. *Computer Communications*, 188, 125–135. <https://doi.org/10.1016/j.comcom.2022.02.012>

Отримано редакцією журналу / Received: 05.11.25

Прорецензовано / Revised: 10.11.25

Схвалено до друку / Accepted: 19.12.25

Yurii KRAVCHENKO, DSc (Engin.), Prof.
ORCID ID: 0000-0002-0281-4396
e-mail: kr34@ukr.net
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Olena FISUN, PhD Student
ORCID ID: 0009-0009-6640-3818
e-mail: lenafi@ukr.net
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

METHOD FOR ENSURING FUNCTIONAL RESILIENCE OF SOFTWARE-DEFINED NETWORKS UNDER CYBERATTACK INFLUENCE

B a c k g r o u n d . In the context of the rapid development of network technologies, the Software-Defined Networking (SDN) architecture has become the basis for building flexible and manageable corporate networks. However, its centralized nature



creates potential vulnerabilities that can be used by attackers to destabilize network management. Ensuring functional stability of SDN under cyberattacks of varying intensity is a complex task, as it requires dynamic adaptation of control parameters to prevent the system from entering critical states.

M e t h o d s . *The paper proposes a method for maintaining functional stability of SDN based on a combination of stochastic modeling and recurrent neural networks of the LSTM (Long Short-Term Memory) type. The SDN model is formalized as a system with five possible states and a set of transition intensities between them. The LSTM module performs the function of an adaptive controller, which in the process of operation estimates the probabilities of the system being in different states and adjusts individual transition intensities in order to maintain the target level of functional stability. A combination of simulation and reinforcement approaches was used to train the LSTM, which allows taking into account both previously simulated attack scenarios and the current state of the network.*

R e s u l t s . *The developed method demonstrated its effectiveness in maintaining stable SDN operation by timely adjusting system parameters when the intensity of threats changes. Simulation modeling for three scenarios (low, medium, and high threat levels) showed that without parameter correction, stability is maintained only at a low level of attacks, while for medium and high levels, critical deviations are observed at certain points in time. The use of the LSTM controller made it possible to detect these moments in time and adaptively update the transition intensities, restoring network stability.*

C o n c l u s i o n s . *The proposed method for ensuring the functional stability of SDN based on LSTM allows creating an intelligent adaptive control system capable of maintaining network stability under the destabilizing effects of cyberattacks. The modeling results confirm the feasibility of using neural network mechanisms for predicting states and correcting SDN parameters in real time. In the future, it is planned to expand the model for multi-level SDN architectures and investigate the influence of spatial topology on the effectiveness of adaptive control.*

K e y w o r d s : *Software-Defined Network, functional stability, cyberattack, stochastic modeling, Recurrent Neural Network.*

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.



Володимир ДРУЖИНІН¹, д-р техн. наук, проф.
ORCID ID: 0009-0009-5049-0099
e-mail: volodymir.druzhynin@knu.ua

Євгеній БАТРАК², канд. техн. наук, доц.
ORCID ID: 0000-0003-3027-4507
e-mail: batrak.yevhenii@lil.kpi.ua

Наталія ЦЬОПА², канд. техн. наук, доц.
ORCID ID: 0000-0002-3215-0711
e-mail: tsopa.natalia@lil.kpi.ua

Ольга НІКОЛЕНКО², магістр
ORCID ID: 0009-0003-2304-7821
e-mail: nikolenko.olha@lil.kpi.ua

¹Київський національний університет імені Тараса Шевченка, Київ, Україна

²Національний технічний університет України

"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

МУЛЬТИДЖЕРЕЛЬНИЙ АЛГОРИТМІЧНИЙ АНАЛІЗ ЧИТАЦЬКИХ І КОМЕРЦІЙНИХ ТРЕНДІВ КНИЖКОВОГО РИНКУ

Вступ. Книжковий ринок зазнає швидких трансформацій під впливом цифрових платформ, а вчасне виявлення літературних трендів стає критично важливим для видавництва, що працюють із перекладною літературою. Мета дослідження – створення й обґрунтування інтегрованої моделі мультиджерельного аналізу ключових читацьких і комерційних показників за даними Goodreads, Amazon і The New York Times Bestsellers для забезпечення виявлення закономірностей формування літературних трендів та уточнення методологічних підходів до прогнозування їхньої динаміки.

Методи. Проаналізовано 280 книжок із Goodreads Choice Awards 2024, 225 видань зі списку Amazon Best Books 2024 і 949 книжок рейтингу NYT Bestsellers. Дані збирали Python-скриптами (Selenium, BeautifulSoup), очищали та нормалізували (pandas, numpy). Використано статистичний аналіз, PCA і кластеризацію K-Means для порівняння читацьких, комерційних і медійних показників.

Результати. Аналіз трьох джерел виявив виразні закономірності та спільні тенденції. На рівні читацьких уподобань (Goodreads) найширшу аудиторію зберігають жанри Romance і Fiction, тоді як History та Nonfiction стабільно демонструють високі рейтинги. Окремо виділяються "приховані перлини" – жанри World War II, Mental Health і Health, які поєднують високу оцінку з невеликим охопленням і відкривають можливості для цільового просування.

У комерційному сегменті (Amazon) переважають паперові формати Hardcover і Paperback, а Kindle Edition залишається найдоступнішим за ціною. Преміальні ніші, як-от Epic Fantasy та регіональна художня проза, утримують найвищі середні ціни, що підтверджує готовність читачів інвестувати у винятковий контент. Кореляційний аналіз засвідчив відсутність прямого зв'язку між вищою ціною і кращим рейтингом, проте показав помірну залежність між кількістю доступних форматів і загальною вартістю.

Медійна видимість (NYT Bestsellers) свідчить, що тривалий інтерес публіки забезпечують дитячі серії та нехудожня література: окремі видання залишаються у списку понад 300 тижнів. Сезонність дебютів має два піки – у першому та третьому кварталах, що підвищує шанси книжок на довготривалу присутність у рейтингах.

У підсумку лише вісім видань одночасно досягли високих показників за читацькими, ринковими та медійними критеріями, демонструючи збалансований успіх і найбільший потенціал для перекладу та міжнародного просування.

Висновки. Інтегрований аналіз трьох джерел показав, що комерційний успіх книжки не визначається одним фактором: потрібне поєднання стійкої читацької уваги, збалансованої цінової політики та здатності довго залишатися у публічному полі. Таке комплексне бачення дає змогу видавцям планувати портфель не лише за жанровими пріоритетами, а й за довгостроковою економічною доцільністю. Важливо враховувати й часові коливання попиту, адже ринкові цикли впливають на ефективність маркетингових і перекладацьких кампаній. Зіставлення читацьких, комерційних і медійних критеріїв створює основу для точнішого прогнозування та мінімізує ризик одноразових успіхів. У підсумку отримані дані формують методологічну базу, яка допомагає приймати стратегічні рішення щодо перекладу, інвестицій і розвитку міжнародних проєктів.

Ключові слова: інтегрований аналіз даних, кроссплатформне порівняння, відкриті дані, алгоритмічний аналіз, кластеризація, моделювання читацької поведінки.

Вступ

Книжковий ринок є важливим елементом культурного й економічного життя суспільства. З розвитком цифрових технологій змінюються формати і способи поширення літератури, однак інтерес до книг залишається стабільним, трансформуючись відповідно до сучасних умов. Для видавництва, що орієнтуються на якісну перекладну літературу, важливо вчасно розпізнавати актуальні літературні тренди, які забезпечують не лише комерційний успіх, але й тривалу читацьку підтримку. Важливим інструментом у цьому процесі є аналіз відкритих даних, що дає змогу системно оцінювати читацькі вподобання, ринкові тенденції та медійну видимість книжок.

Метою дослідження є створення й обґрунтування інтегрованої моделі мультиджерельного аналізу читацьких, комерційних і медійних показників книжкового ринку, що забезпечує виявлення закономірностей формування літературних трендів та уточнення методологічних підходів до прогнозування їхньої динаміки.

Огляд літератури. Goodreads і Amazon давно вважають багатим джерелом аналітики. У роботі (Kousha, Thelwall, & Abdoli, 2017) показано, що кількість рецензій і оцінок на Goodreads корелює з подальшим впливом книги,



проте дослідження не враховувало жанрову неоднорідність аудиторії. Подальші дослідження (Maity, Panigrahi, & Mukherjee, 2018) свідчать, що з точністю 89 % можна спрогнозувати Amazon-бестселери на основі поведінкових ознак із Goodreads (додавання книг на полиці, участь в обговореннях), проте представлена модель не аналізувала зміст відгуків і не охоплювала медійну видимість. Також показано, що середній рейтинг і кількість рецензій є основними драйверами потрапляння в Amazon-бестселери в період з 2009 по 2019 р., але не враховувались ціннові та жанрові фактори. У 2020 р. проаналізовано сезонні "вікна" дебютів і тривалість перебування художніх творів у NYT у період з 1931 по 2020 р., однак дослідження стосувалося лише твердої палітурки й не враховує змін критеріїв формування рейтингу (напр., перехід до врахування електронних продажів і нових категорій) (Pruett, 2020). В роботі (Kousha, & Thelwall, 2015) показано слабкі кореляції ($r \approx 0.2-0.4$) між Amazon-відгуками й науковими цитуваннями академічних монографій і бестселерів, проте ця робота обмежилася тільки двома жанровими підмножинами.

Попередні роботи доводять цінність кожного джерела, але їх вузькі вибірки, різні методики нормалізації й агрегації не дають цілісної картини. Отже, огляд публікацій зазначеної тематики показав відсутність єдиного підходу, здатного одночасно врахувати жанрово-форматну різноманітність, узгодити шкали оцінок і перевірити чутливість вагових коефіцієнтів у різних контекстах. Саме ці прогалини об'рунтовують ідею мультиджерельного аналізу: інтеграції Goodreads Choice Awards 2024, Amazon Best Books 2024 і NYT Bestsellers 2024 в єдину модель зі спільними метриками. Такий підхід не лише дає змогу зіставити емоційну, ринкову й медійну складові популярності, а й підвищує надійність рекомендацій для видавничих стратегій, адже базується на збалансованих даних трьох незалежних платформ.

Методи

Дослідження інтегрує три незалежні джерела, що репрезентують читацький, комерційний і медійний виміри популярності книжок. Використано 280 позицій фіналістів Readers' Favorite Books зі списку Goodreads Choice Awards 2024, 2 122 записів про формати та ціни для 225 унікальних видань зі списку Amazon Best Books 2024, а також 949 книжок, що входили до щотижневих рейтингів The New York Times Bestsellers 2024 р.

Для роботи з даними створено набір програмних інструментів на основі мови Python, які охоплюють повний аналітичний цикл: від збору й очищення – до глибокого оброблення, кластеризації й побудови візуалізацій. Збір даних здійснювався Python-скриптами з використанням Selenium, BeautifulSoup і requests для охоплення динамічних і статичних сторінок, збереження "сирих" даних у форматі CSV. Етап підготовки даних включав очищення текстових полів, нормалізацію жанрів і форматів та усунення дублікатів за допомогою pandas, numpy і networkx. Візуалізації та розподіли будували у matplotlib і seaborn. Для зниження розмірності застосовували StandardScaler і PCA, для сегментації – KMeans. Оцінювання кореляцій проводили за допомогою scipy.stats.pearsonr. Кожну платформу аналізували окремо, а потім результати зіставляли для виявлення збігів і розбіжностей між читацькою, ринковою та медійною оцінкою, після чого була побудована уніфікована модель для рекомендацій видавничих рішень.

На рівні Goodreads для кожної книги k зібрано назву, автора, середній рейтинг $r_k \in [1.0; 5.0]$, кількість оцінок $c_k \in N_0$ та жанрові теги, перетворені на бінарний вектор $\vec{g}_k \in \{0,1\}^{20}$. Після очищення даних і нормалізації жанрів ця структура дозволяє розглядати книжку як точку в тривимірному просторі "якість – популярність – жанрова структура". Для жанру j обчислювали середній рейтинг і середнє охоплення за формулами відповідно

$$\bar{r}_j = \frac{1}{|K_j|} \sum_{k \in K_j} r_k \text{ та } \bar{c}_j = \frac{1}{|K_j|} \sum_{k \in K_j} c_k. \quad (1)$$

Далі кожному жанру присвоювали три ранги: за потужністю $rank_{|K|}(j)$, за середнім рейтингом $rank_r(j)$ і за середнім охопленням $rank_c(j)$. Сума рангів утворює індекс трендовості

$$T_j = rank_{|K|}(j) + rank_r(j) + rank_c(j). \quad (2)$$

Щоб виявити, які жанрові поєднання найчастіше трапляються в межах однієї книжки і створюють "вузли" читацької уваги, розраховано "взаємопроникнення" жанрів. Цей показник вимірює, скільки разів пара жанрів i та j одночасно присутня в описі тієї самої книжки

$$s_{ij} = \sum_{k=1}^K g_{ki} \cdot g_{kj}. \quad (3)$$

Для виявлення прихованої (латентної) структури застосовано зменшення розмірності PCA з подальшою кластеризацією KMeans із параметром $K = 4$, і з мінімізацією суми квадратів відстаней до центрів:

$$V = \sum_{i=1}^K \sum_{x \in c_i} \|x - \mu_i\|^2. \quad (4)$$

Стабільність жанру оцінювалась як частка "дуже високих" проти суми "дуже високих" і "дуже низьких" оцінок за формулою

$$\text{Стабільність}_j = \frac{|K_j^{(r>4.4)}|}{|K_j^{(r>4.4)}| + |K_j^{(r<2.5)}|}. \quad (5)$$

Для жанрів із невеликою кількістю книжок, менше ніж п'ятнадцять, введено два індекси. Нішевий індекс

$$N_j = \frac{\bar{r}_j}{5} + \frac{\bar{c}_j}{\max_j \bar{c}_j} \quad (6)$$

та індекс перспективності

$$P_j = \frac{\bar{r}_j}{5} + \left(1 - \frac{\bar{c}_j}{\max_j \bar{c}_j}\right), \quad (7)$$

який виділяє "приховані перлини". Зв'язок "якість проти охоплення" перевірено коефіцієнтом Пірсона

$$r = \frac{\text{Cov}(r_k, c_k)}{\sigma_{r_k} \cdot \sigma_{c_k}}. \quad (8)$$

Для Amazon кожний запис містив назву, автора, дату першої публікації, середній рейтинг r_k , формат f (Hardcover, Paperback, Kindle Edition, Audiobook тощо), ціну p , тип ціни (нова, вживана, електронна тощо), основну категорію та підкатегорію. Такий підхід дозволяє описати не обсяги продажів, а структуру ринкової пропозиції: фізичні носії – їхню логістику, цифрові – жанрові сегменти, ціна – платоспроможний попит, рейтинг – сприйняття якості. Середні ціни та стандартні відхилення за форматом і типом визначено як



$$\mu_p^f = \frac{1}{n_{fp}} \sum_{j=1}^{n_{fp}} p_{fp}^{(j)} \quad (9)$$

та

$$\sigma_p^f = \sqrt{\frac{1}{n_{fp}} \sum_{j=1}^{n_{fp}} (p_{fp}^{(j)} - \mu_p^f)^2}. \quad (10)$$

Також обчислено кореляції між середньою ціною μ_p^f , середнім рейтингом μ_p^k та кількістю форматів n_f^k за формулою (8). Для порівняння "ціни проти якості" застосовувалося індекс value-for-money

$$V_k = \frac{\mu_p^k}{\mu_p^f}. \quad (11)$$

У просторі ознак $(\mu_p^k, \mu_p^f, n_f^k)$ виконано PCA з подальшою кластеризацією KMeans ($K = 5$).

Для NYT Bestsellers опис кожної книги включав тривалість присутності у списку w_k , середню позицію p_k , квартал дебюту $q_k \in \{Q_1, Q_2, Q_3, Q_4\}$, автора a_k і широту авторського портфеля s_a , що дорівнює кількості категорій, у яких автор був представлений. Для підмножини long-lived L_j , тобто книг із перебуванням у рейтингу не менше десяти тижнів, на рівні категорій оцінено середню тривалість

$$\bar{\omega}_j = \frac{1}{|L_j|} \sum_{k \in L_j} \omega_k, \quad (12)$$

і середню позицію

$$\bar{p}_j = \frac{1}{|L_j|} \sum_{k \in L_j} p_k. \quad (13)$$

Приховані (латентні) структури на рівнях категорій, авторів і книг отримано через PCA і KMeans у векторах (ω_k, p_k, s_a) .

Для інтеграції трьох аспектів успіху сформовано вибірку зі 126 творів, для яких доступні якісні показники на всіх платформах. Кожну книгу описано семивимірним вектором нормалізованих ознак:

$$\vec{x}_k = (\tilde{r}_k^{GR}, \tilde{c}_k^{GR}, \tilde{r}_k^{AM}, \tilde{c}_k^{AM}, \tilde{p}_k^{AM}, \tilde{w}_k^{NYT}, p\tilde{o}s_k^{NYT}). \quad (14)$$

Нормалізація виконувалася min-max для приблизно рівномірних розподілів. Для експоненційно скошених змінних, наприклад, кількість оцінок і ціна, застосовувалося логарифмічне перетворення. Для метрик, де менше означає краще, таких як позиція NYT, використано інверсну нормалізацію. Ваги блоків для читачького, ринкового та медійного вимірів визначено пропорційно до дисперсій:

$$D_x = \frac{1}{|s|} \sum (x_k - \bar{x})^2 \quad (15)$$

та

$$\omega_i = \frac{D_{x_i}}{D_{GR} + D_{AM} + D_{NYT}} \text{ за умови } \sum_i \omega_i = 1. \quad (16)$$

Отримано $\omega_{GR} = 0.29$, $\omega_{AM} = 0.37$, $\omega_{NYT} = 0.34$. Підсумковий індекс видавничої доцільності обчислено як

$$Total_Score_k = 0.29 \cdot \tilde{r}_k^{GR} + 0.37 \cdot \tilde{r}_k^{AM} + 0.34 \cdot \tilde{w}_k^{NYT}. \quad (17)$$

Додатково враховано показник Source Rank, що дорівнює кількості платформ, на яких присутня книга. В усіх обчисленнях використовувалася десяткова крапка і безперервна нумерація формул.

Результати

Розподіли Goodreads виявили виразну асиметрію середніх рейтингів із піком у діапазоні від 3.8 до 4.2, причому значення $\tilde{r}_k$ змінювалися від 3.43 до 4.63. Кількість оцінок варіювала від 630 до 1.1 млн, більшість книжок мали менше ніж 30 000 відгуків, що засвідчує нерівномірну увагу аудиторії та обґрунтовує потребу комбінованих показників, а не поодиноких метрик (рис. 1 і 2).

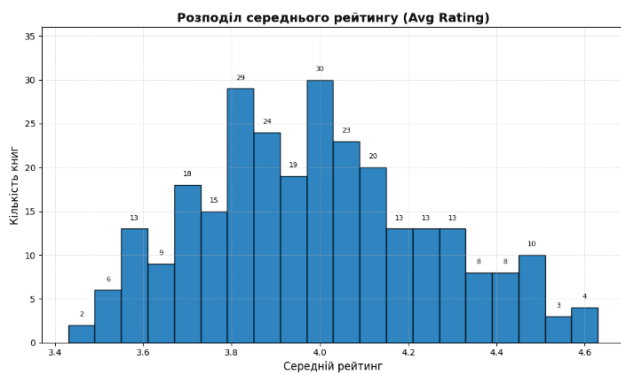


Рис. 1. Goodreads: розподіл середнього рейтингу

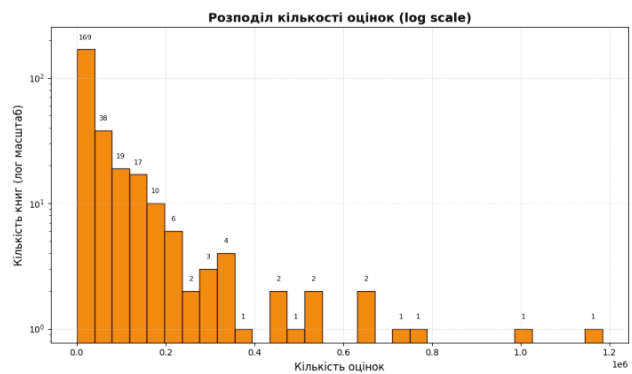


Рис. 2. Goodreads: розподіл кількості оцінок

Порівняння жанрів через індекс T_j вивело у лідери Book Club, Romantasy, Romance, Sports Romance та Audiobook, (рис. 3). Це означає, що поєднання масштабу вибірки, середніх оцінок і середнього охоплення формує стабільний попит саме в цих напрямках.

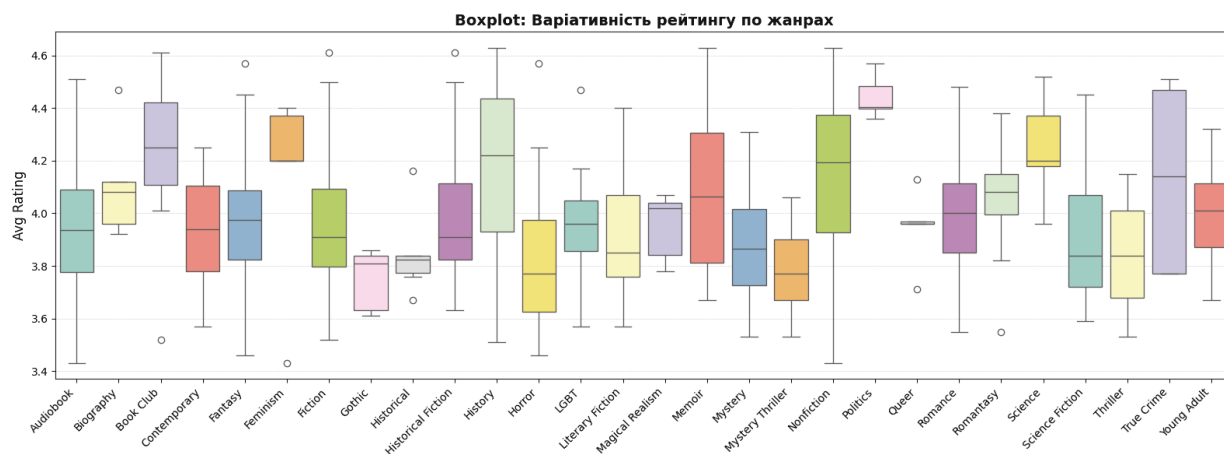


Рис. 3. Goodreads: варіативність рейтингу по жанрах

Матриця "взаємопроникнення" зафіксувала типові вузли читацьких уподобань, найчастіше разом траплялися пари Romance + Fantasy зі значенням 38, Mystery + Nonfiction зі значенням 36, Fiction + Historical Fiction зі значенням 32, а також Fiction + Fantasy зі значенням 20 (рис. 4). Такі комбінації відображають не лише тематичну близькість жанрів. Вони показують сталі очікування читача щодо набору рис твору. Наприклад, зв'язка Romance + Fantasy зазвичай означає магічний світ, емоційну лінію стосунків і високу ставку для героїв. Це корисно для планування міжжанрових лінійок і рекомендацій.

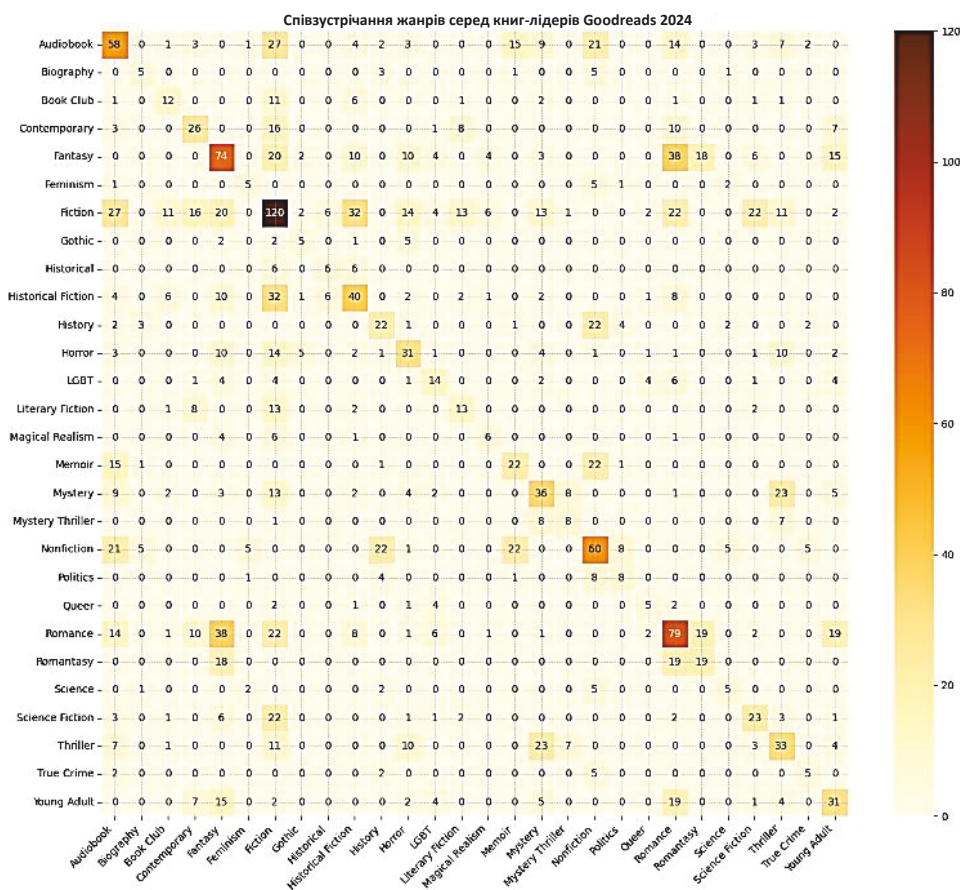


Рис. 4. Goodreads: теплова карта "співзустрічання"

Зменшення розмірності та кластеризація підтвердили сегментацію читацького поля на чотири групи. До групи складної прози увійшли History і Memoir із високими рейтингами та помірною популярністю. До емоційних жанрів належать Romance і Thriller із великими обсягами оцінок. Камерні напрями представлені Book Club і Small Town Romance з невеликою кількістю назв і стабільними показниками. Нішеві жанри, зокрема Paranormal і Poetry, мають невелику аудиторію та високі середні оцінки (рис. 5). Окрема кластеризація книг за векторами жанрової належності



дала п'ять архетипів читачьких уподобань, від преміальної прози до літнього бестселера, що дозволяє оперувати не лише жанрами, а стилістичними профілями.

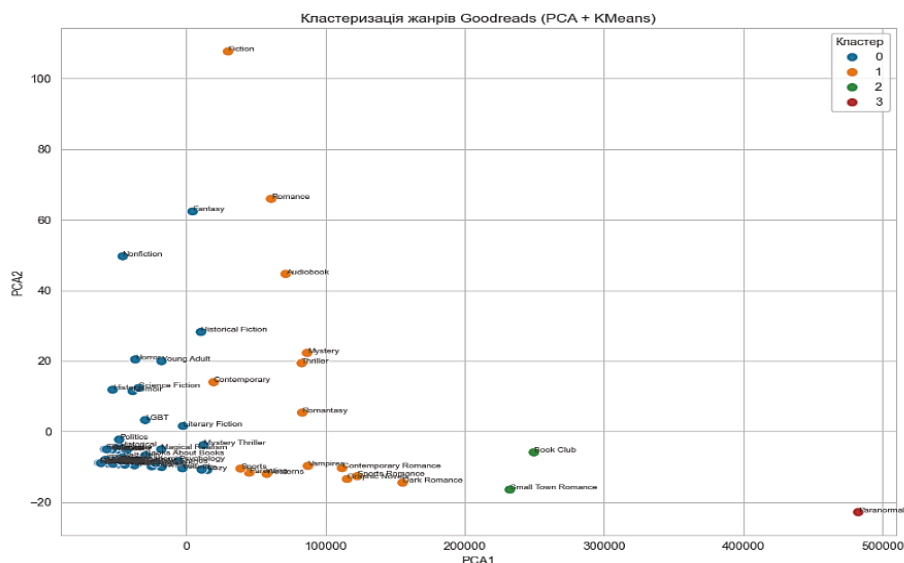


Рис. 5. Goodreads: кластеризація жанрів

Стабільність жанрів виявилася нерівномірною і це добре узгоджується з припущенням про різні механіки формування очікувань читачів у nonfictional і fictional напрямках. Найвищі значення спостерігалися у History зі значенням 0.64, у Nonfiction зі значенням 0.55. У цих сегментах автори спираються на верифіковані події й дати, архівні матеріали, статистику, наукові публікації та систему приміток і посилань. Фактчекінг і чіткі критерії точності роблять сприйняття більш передбачуваним, тому оцінки частіше концентруються у "зеленій зоні". Historical Fiction із 0.36 поводить як проміжна ланка між суворим Nonfiction і суто художньою прозою. Найнижчі значення були у Audiobook зі значенням 0.17 та у Fiction зі значенням 0.14 (рис. 6). Таке падіння очікуване з огляду на сильну різноманітність піджанрів і роль актора дубляжу у випадку аудіокниг, де інтерпретація тексту може як підняти, так і "згасити" сприйняття якості. Ризиковий профіль за нашою метрикою дає ще виразніший контраст. Mystery і Thriller мають показник 1.00. Це означає крайній розподіл з переважанням дуже високих або дуже низьких оцінок.

Для видавця це сигнал про "лотерейну" природу релізів у цих піджанрах. Невелика похибка у позиціюванні обкладинки, у темпі сюжету чи у фінальному твісті здатна різко зрушити сприйняття аудиторії в будь-який бік. Далі за рівнем ризику розміщуються Horror зі значенням 0.94, Romance зі значенням 0.89 і Fantasy зі значенням 0.89. Такий результат частково неочікуваний щодо Romance, який ми звикли вважати "масовим і безпечним", однак, зважаючи на "ефект завищених очікувань", абсолютно логічний і пояснюваний: коли читач іде на сильну емоційну обіцянку, розчарування карається різкіше, ніж у "спокійних" піджанрах. Це пояснює високий ризик за загальною позитивною тональністю ринку (рис. 7).

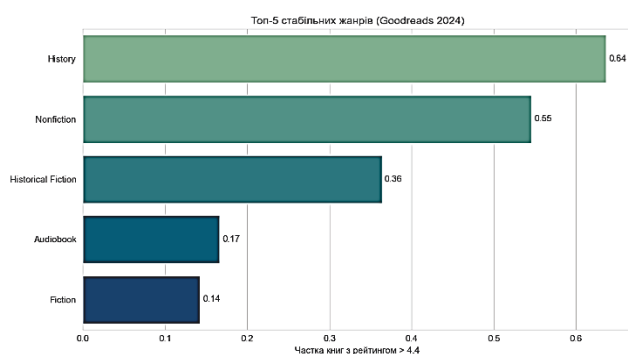


Рис. 6. Goodreads: п'ять лідерів стабільних жанрів

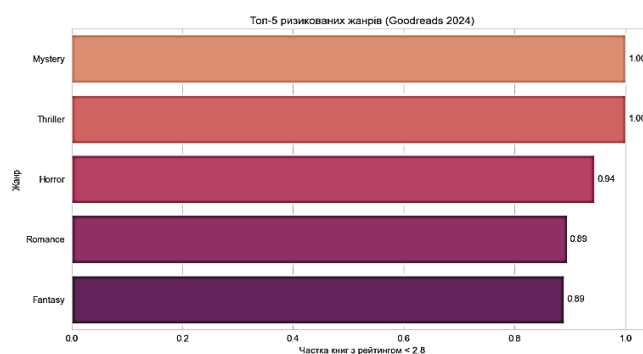


Рис. 7. Goodreads: п'ять лідерів ризикових жанрів

Для малих вибірок ми свідомо розділили на дві різні метрики нішевість і перспективність. Нішевий індекс N_j виділяє напрями з поєднанням добрих оцінок і відчутного, але не гігантського охоплення (жанри з менше ніж 15 книжками). Лідирують Paranormal зі значенням 1.8, Book Club зі значенням 1.4 і Small Town Romance зі значенням 1.3 (рис. 8). Це нішеві, але комерційно життєздатні сегменти з уже сформованим ядром аудиторії. Вони не масові, проте мають достатній попит, щоб підтримувати серійність і купівність, а за правильної стратегії можуть масштабуватися.

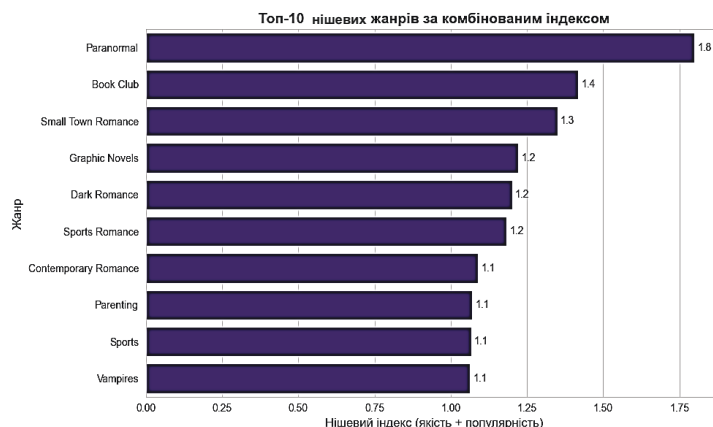


Рис. 8. Goodreads: десять лідерів нішевих жанрів за комбінованим індексом

Індекс перспективності P_j навмисно "підсвічує" те, що читачі вже високо оцінюють, але ще не встигли масово помітити. Тут у топі жанри World War II, Mental Health і Health. Вони всі вищі 4.0 за рейтингом і мають низьке охоплення – менше ніж 5 000 оцінок. Це класичні "приховані перлини", де додатковий маркетинговий імпульс і точне позиціонування можуть забезпечити тривале рівне споживання після стартового піку. Поділ є інструментальним для портфельних рішень. Нішеві, але життєздатні жанри слід масштабувати через серійність і розширення каналів. Жанри з високими оцінками та малим охопленням доцільно підтримувати фокусним промо і тестами позиціонування.

Окремо перевірено, чи можна спиратися на популярність як на проксі якості. Кореляція Пірсона, між середнім рейтингом і кількістю оцінок становить $r = 0.18$ при $p = 0.002$. Після логарифмування кількості оцінок тренд зберігається як слабкий додатний. Перевірка ранговою кореляцією також показує слабку залежність. Вилучення поодиноких точок з екстремально великим охопленням не змінює висновку про низьку силу зв'язку (рис. 9–10). Структура хмари точок підтверджує статистику. У діапазоні від 10^3 до 10^5 оцінок спостерігається найбільша концентрація творів із рейтингами приблизно 3.8–4.2. За дуже великих охоплень варіативність рейтингів зростає і більшість значень перебуває в інтервалі 4.0–4.3. Водночас присутні видання з рейтингами понад 4.4 у групах із невеликим і середнім охопленням. Є і протилежні випадки, коли твори з великою кількістю оцінок мають лише середні бали. Зв'язок нелінійний і гетероскедастичний. На лінійній шкалі видно вертикальну "стіну" спостережень для малих охоплень, а на лог-шкалі – слабкий нахил без вираженого зростання середніх значень. Окремі високовпливові точки на рівні близько мільйона оцінок не задають загального тренду. Межі шкали 1–5 створюють стелю близько 4.6, що додатково обмежує потенційний приріст рейтингу разом із ростом охоплення, Такий результат очікуваний, адже популярність іде хвилями і часто залежить від інформаційних приводів. На відміну від неї якість змінюється повільніше і може бути високою навіть у видань із невеликим охопленням. Тому вірусні сплески не замінюють системного розвитку лінійок і портфельного супроводу.

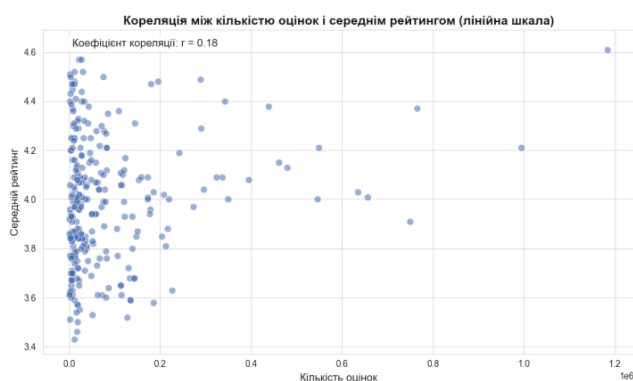


Рис. 9. Goodreads: кореляція між кількістю оцінок і середнім рейтингом (лінійна шкала)

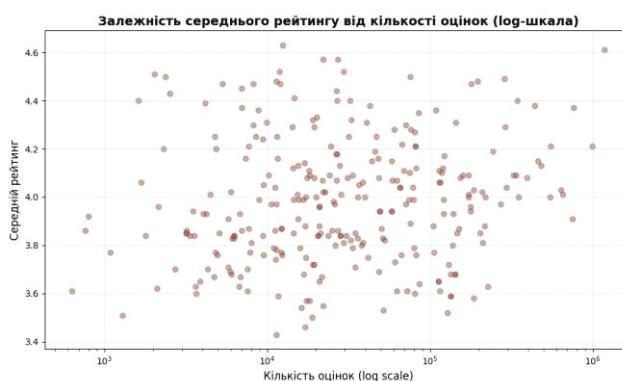


Рис. 10. Goodreads: залежність середнього рейтингу від кількості оцінок (логарифмічна шкала)

Комерційні дані Amazon підкреслили роль носіїв і цінових ніш. У вибірці домінують паперові формати із частками 36 % для Hardcover і 32 % для Paperback, тоді як Kindle має 10 %, а Audiobook 9 %. Це очікувано для підбірки "кращих книг року", де відбір зумовлює сильну присутність преміальних паперових релізів. Цифрові носії частіше трапляються у бізнесі та Nonfiction. Художня проза і біографії переважають на папері. Лише 8 % позицій виходять одразу у трьох і більше форматах. Найтипівша зв'язка – Hardcover разом із Paperback. Така структура підтверджує роль паперу як "якірного" каналу релізу і роль Kindle як каналу доступності. Узагальнені пропорції та перетини форматів згруповано на рис. 11–12.

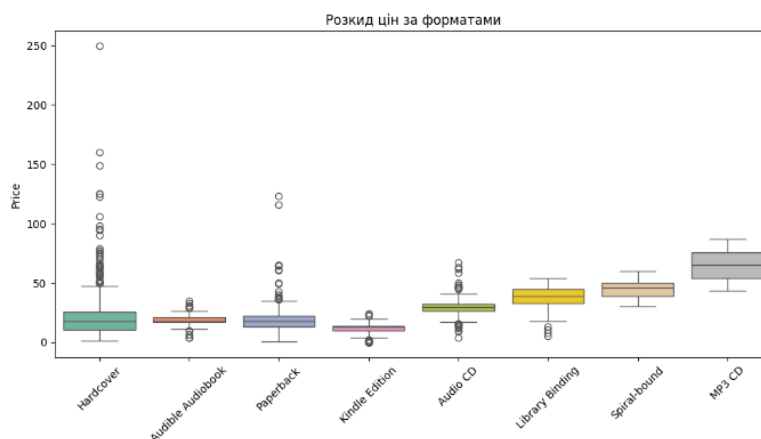


Рис. 11. Amazon: розкид цін за форматами

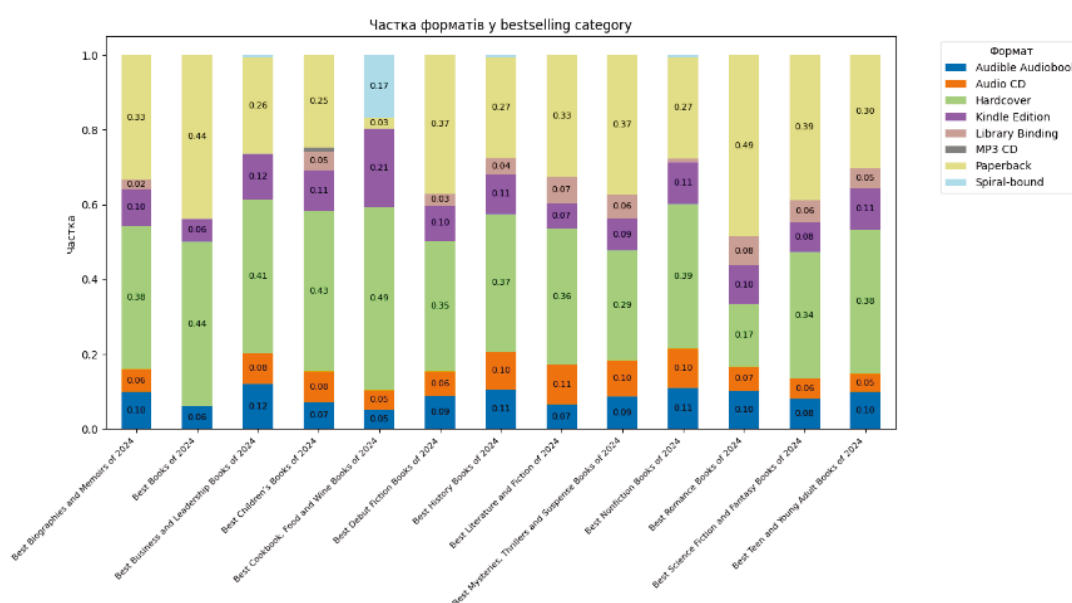


Рис. 12. Amazon: форматна насиченість книжкового каталогу

Розподіл цін "скошений праворуч" і має "довгий верхній хвіст". Медіана дорівнює 17.98 \$. Мінімум дорівнює 2.99 \$. Максимум дорівнює 249.52 \$. Середні значення формують "щаблі" між форматами: Kindle Edition ≈ 11.13 \$ < Paperback ≈ 18.01 \$ $\approx$ Audible Audiobook ≈ 18.34 \$ < Hardcover ≈ 21.49 \$ < Audio CD ≈ 29.91 \$ < Library Binding ≈ 37.19 \$ < Spiral-bound ≈ 44.52 \$ < MP3 CD ≈ 65.07 \$ (рис. 13–14). Асиметрія пояснюється тим, що у верхньому хвості зібрано колекційні та спеціальні видання, й окремі великі накладки з підвищеною собівартістю. Для Kindle діапазон цін вузький, що узгоджується з більш жорсткою ціновою межею електронних видань. Карту цін за форматами і типами продажу наведено у групі на рис. 13–14.

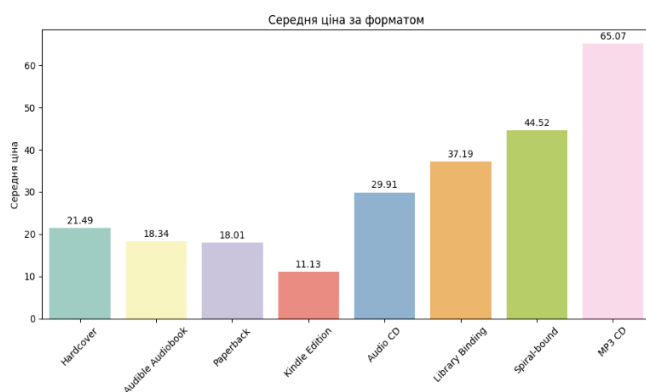


Рис. 13. Amazon: boxplot розкид цін за форматами

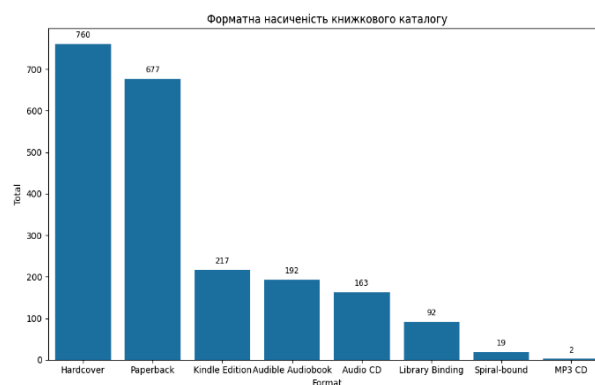


Рис. 14. Amazon: середня ціна за форматами



Профіль підкатегорій додатково розрізняє "дорогі" ніші та "бюджетне ядро". У межах bestselling-категорій найвищу середню ціну має Cookbook, Food & Wine Books – 26.07 \$. Далі йдуть блоки з вищим за середній рівнем: Nonfiction – 23.00 \$, Science Fiction and Fantasy – 22.59 \$, Literature and Fiction – 22.54 \$, History – 22.04 \$. Середній пояс формують Debut Fiction – 21.75 \$, Biographies and Memoirs – 20.95 \$, Business and Leadership – 20.73 \$, Mysteries, Thrillers and Suspense – 19.39 \$. Найнижчі значення спостерігаються у Romance – 18.20 \$ та Teen and YA – 18.43 \$, а Children's Books у цій вибірці опускається до 13.38 \$ (рис. 15). Отже, на рівні широких категорій дорожчими є напрями, близькі до Nonfiction і біографій, тоді як Romance, Teen / YA і частина дитячого сегменту тримають нижчі середні ціни. Це доповнює попередній зріз за підкатегоріями: там лідирують окремі преміальні ніші на кшталт Epic Fantasy та Literary Fiction LGBTQ+, а тут видно, що у грубішій агрегації найвищі середні утримують Biographies and Memoirs.

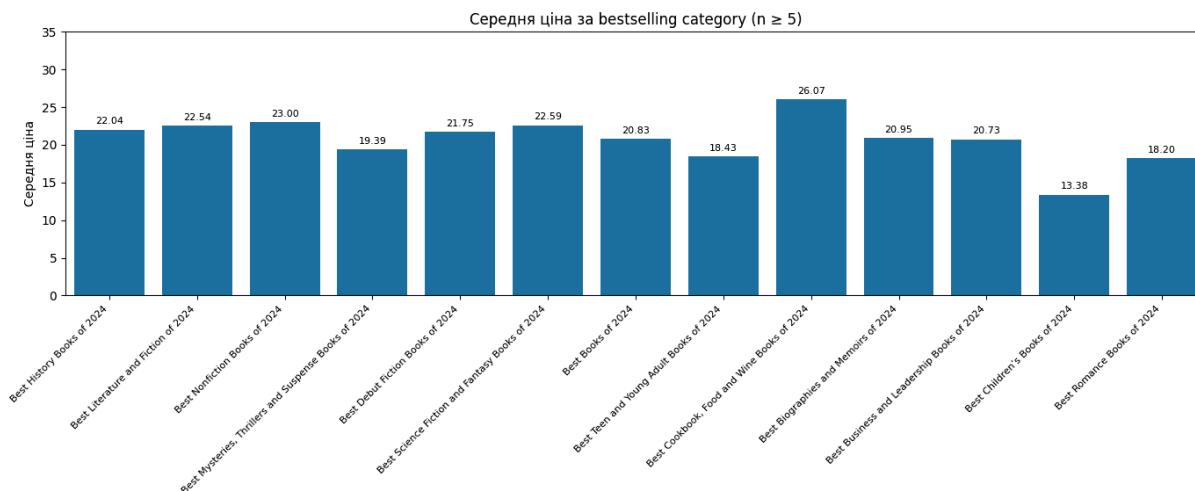


Рис. 15. Amazon: середня ціна за bestselling category

Перевірка зв'язків між основними показниками дає однозначну картину. Кореляція між ціною і рейтингом практично відсутня. Значення коефіцієнта дорівнює 0.06. Отже вищий прайс не супроводжується кращою середньою оцінкою. Кореляція між ціною і кількістю форматів слабка додатна. Значення коефіцієнта дорівнює 0.24. Це свідчить, що продукти із ширшою форматною лінійкою мають тенденцію до вищої середньої ціни. Кореляція між кількістю форматів і рейтингом слабка від'ємна. Значення коефіцієнта дорівнює мінус 0.14. Тобто розширення форматної палітри не пов'язане з підвищенням середнього рейтингу. Сумарні діаграми розкиду та коефіцієнти наведено на рис. 16–17.

Окремо оцінено співвідношення "ціна – оцінка" через індекс value for money. Для кожної книги обчислено відношення середнього рейтингу до середньої ціни. Найвище значення має Eyes That Kiss in the Corners у форматі Paperback. Значення дорівнює 9.06. Порівняння лідерів і "середняків" цього індексу показує, що окремі паперові позиції у популярних піджанрах формують найкращі відношення вартість проти оцінки.

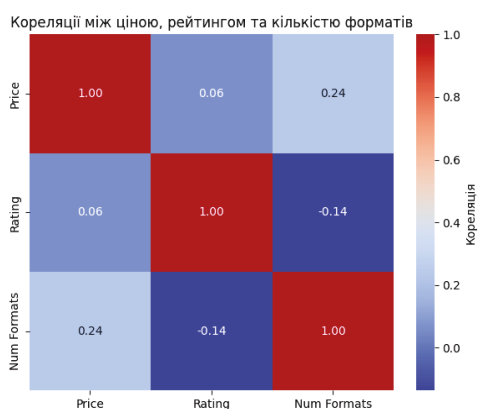


Рис. 16. Amazon: середня ціна за bestselling category

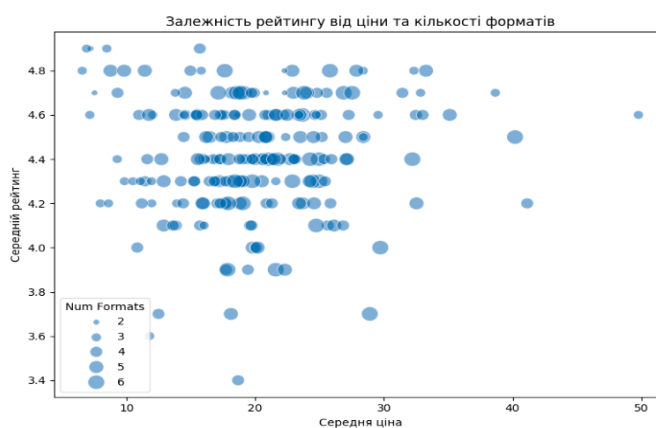


Рис. 17. Amazon: залежність рейтингу від ціни та кількості форматів

Кластеризація у просторі трьох ознак – середня ціна, середній рейтинг, кількість форматів – дає п'ять стійких груп (рис. 18). Перша група відповідає масовим бюджетним виданням, низькі ціни, один або два формати, рейтинги близько чотирьох. Друга група – масові з підвищеною ціною, два або три формати, рейтинги помірні. Третя група – збалансовані продукти середнього цінового рівня з трьома форматами. Четверта група – нішеві преміальні видання з високими цінами і невеликою кількістю форматів. П'ята група – "вітринні" позиції із широкою форматною палітрою



та цінами вище середніх, де рейтинг коливається в межах від 3.8 до 4.3. Таке розшарування узгоджується з очікуванням, що ціна і кількість форматів описують передусім маркетингову та виробничу стратегії.

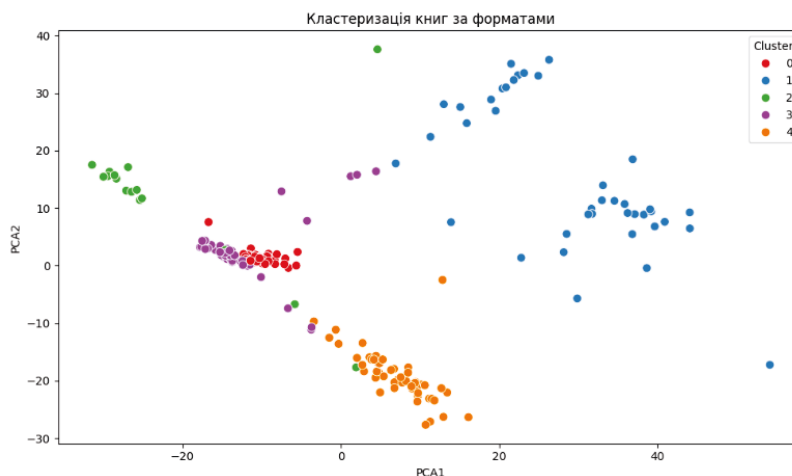


Рис. 18. Amazon: кластеризація за форматами

Медійна динаміка NYT віддзеркалила довготривалу видимість. Із 949 позицій 2024 р. у статусі long-lived залишилися 247 книжок. Найбільша частка припадає на Hardcover Nonfiction, Combined Print & E-Book Fiction та Hardcover Fiction (рис. 19). Це очікувано, бо саме ці три "опорні" таблиці NYT регулярно акумулюють найпотужніші релізи і мають стабільний попит у ритейлі та бібліотеках. Важливо, що різниця між ними невелика, отже тривала присутність у списку не є прерогативою лише одного сегменту і відображає загальний баланс між художньою і нехудожньою пропозиціями.

Порівняння середніх тривалостей і позицій для long-lived категорій демонструє різку асиметрію на користь серій для дітей і підлітків (рис. 20). Children's & Young Adult Series мають близько 307 тижнів середньої тривалості. Це узгоджується з логікою серійного споживання, коли кожна нова частина підтягує продажі попередніх, а також зі шкільними закупівлями і повторними перевиданнями. Такий контраст типовий для ринку, де "серійні" бренди забезпечують довге плато видимості, а одиночні романи частіше отримують хвилеподібну увагу. Для Hardcover Fiction і Combined Print & E-Book Fiction картина інша. Тут тривалість присутності коротша, а розкид позицій ширший. Більша частка релізів поводить як хвиля з яскравим стартом і швидким спадом. Hardcover Nonfiction тримається довше, ніж більшість художніх категорій, але коротше за дитячі серії. Це узгоджується з тим, що Nonfiction має триваліший життєвий цикл за рахунок тем із тривалою актуальністю. Перевидання у м'якій обкладинці та вихід оновлених видань продовжують обіг. У середніх позиціях видимі відтинки стабільності, які не характерні для більшості автономних романів. Розподіл усередині Children's Picture Books та Children's & Young Adult Series додає ще одну деталь. Дитячі книжки-картинки мають коротші хвилі окремих релізів, але часто повертаються до списку під святкові періоди і на початку навчального року. Серійні підліткові видання демонструють довші безперервні відрізки присутності. Це очікувано з огляду на відмінність каналів збуту та частоти перевидань. Додатково варто відмітити внесок бокс-сетів і омнібусів у дитячих і підліткових лінійках. Коли в серії виходить комплект або ювілейне перевидання, точка входу для нових читачів спрощується. Це підтримує довгі "поліці" присутності і підтягує середні позиції на кілька тижнів поспіль.

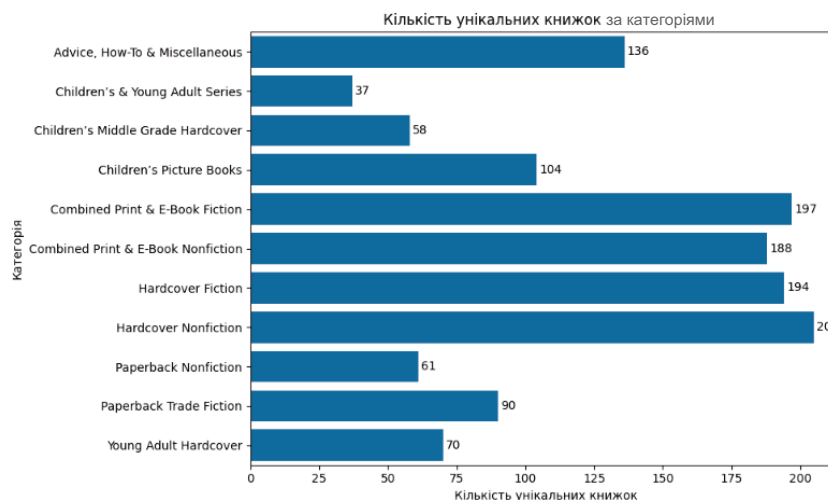


Рис. 19. NYT Bestseller: кількість унікальних книжок за категоріями

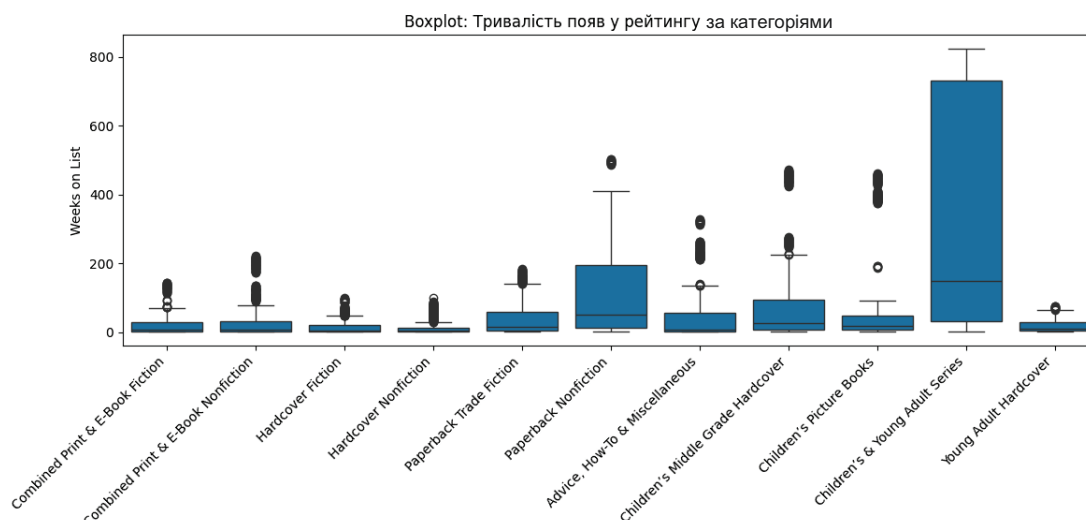


Рис. 20. NYT Bestseller: тривалість появ книжок у рейтингу за категоріями

Сезонність дебютів має два виразні піки. Найбільше успішних стартів у першому та у третьому кварталах. За тепловою картою найбільша кількість дебютів, що протрималися в рейтингах понад 20 тижнів, припала на перший квартал (рис. 21). У цей період різко зросли релізи в Children's Picture Books (16), Children's & Young Adult Series (14), Combined Print & E-Book Fiction (14) і Paperback Nonfiction (13). Другий підйом відбувся у третьому кварталі, де знову виділяється Children's Picture Books (9 дебютів) і помітно зросли показники Paperback Nonfiction (4) та кількох інших категорій. Четвертий квартал виявився спокійнішим, хоча поодинокі новинки з'являлися навіть у листопаді–грудні. Кумулятивні криві підтверджують таку динаміку (рис. 22). Children's Picture Books нарощували присутність найшвидше: від приблизно 10 позицій у січні до більш як 45 наприкінці року. Стійке, хоч і менш стрімке зростання демонстрували Paperback Nonfiction (приблизно із 17 до 28) та Combined Print & E-Book Nonfiction (з 13 до 20). Children's & Young Adult Series і Combined Print & E-Book Fiction досягли близько 20 "довгожителів", а їхні лінії на графіку помітно крутішають у серпні–вересні. Натомість Hardcover Fiction і Hardcover Nonfiction залишалися стабільно невеликими сегментами й завершили рік на рівні до 10 назв. Подібна структура випусків цілком очікувана: дитячі ілюстровані видання й серії традиційно виходять на початку року, коли плануються подарункові покупки, а комбіновані та підліткові формати активізуються ближче до осені, готуючись до святкового сезону. Nonfiction показав рівномірне нарощування впродовж усього року, що добре узгоджується з попередніми спостереженнями про стабільний попит на практичні та документальні книги.

Широта авторських портфелів різниться між блоками. Автори художньої та нехудожньої літератури частіше присутні у двох або трьох категоріях. Це очікувано для мультиформатних авторів, які виходять одночасно у твердій обкладинці, комбінованих (друк-цифра) списках і у форматах для різних аудиторій. Натомість у Children / Young Adult та "Other" домінує спеціалізація в одній категорії. Така спеціалізація природна для серій з усталеним позиціонуванням і цільовим віком.

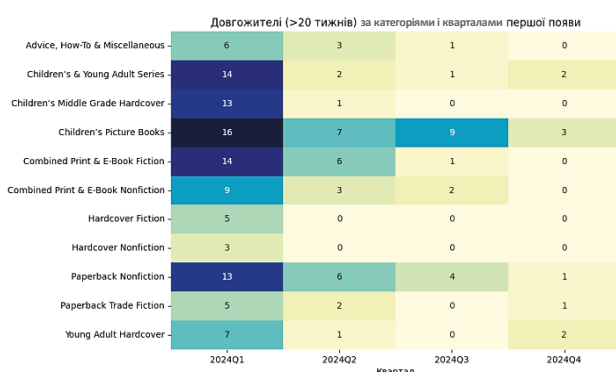


Рис. 21. NYT Bestseller: long-lived книжки за категоріями і кварталами першої появи в рейтингу

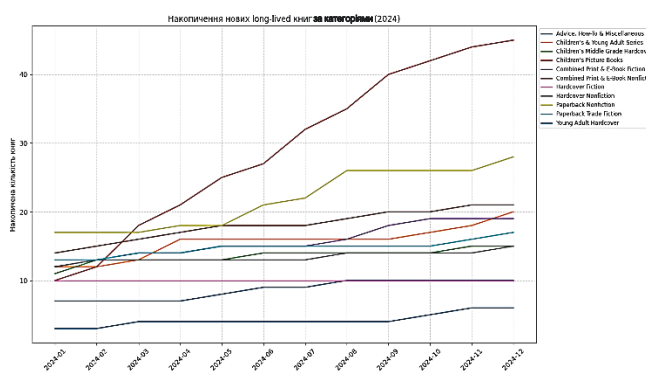


Рис. 22. NYT Bestseller: накопичення кількості long-lived книжок у ключових категоріях

Рейтинг long-lived авторів очолює Adam Wallace із шістьма релізами, далі Sarah J. Maas з п'ятьма і Colleen Hoover з чотирма (рис. 23). Це підтверджує, що стабільна видимість у NYT часто пов'язана або з активною серійністю, або з регулярними новими тайтлами в межах близького жанрового поля. Ідеться не про одиничний "постріл", а про повторюваний патерн присутності.

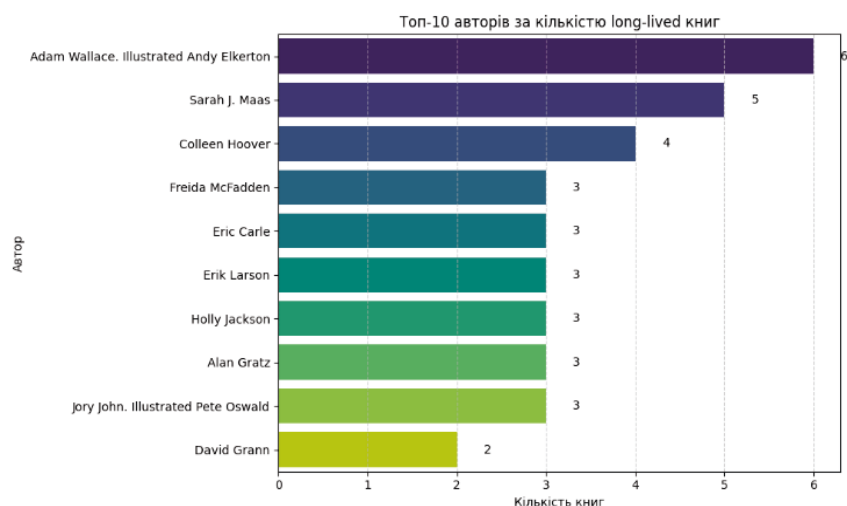


Рис. 23. NYT Bestseller: десять авторів – лідерів за кількістю long-lived книг 2024 року

Кластеризація за трійкою ознак виокремлює три типи категорій. Масові – з великою кількістю релізів і короткими періодами присутності. Збалансовані – з помірною кількістю позицій і стійкою середньою видимістю. Унікальні – з малою кількістю релізів, але з надзвичайно довгими періодами у списку. Остання група відповідає дитячим і Nonfiction -серіям, де працює ефект накопичення аудиторії. На рівні авторів виявлено вузькоспеціалізовані профілі, масових авторів із "короткоживучими" позиціями, збалансовані профілі та "лідерів довгої гри", у яких кожен реліз залишається надовго. На рівні окремих книжок виділяють чотири класи: новинки до 15 тижнів, стабільні Nonfiction -титли 15–50 тижнів, найстійкіші бестселери 100+ тижнів і сезонні явища з короткою видимістю після піку. Діаграми кластерів наведено на рис. 24–27.

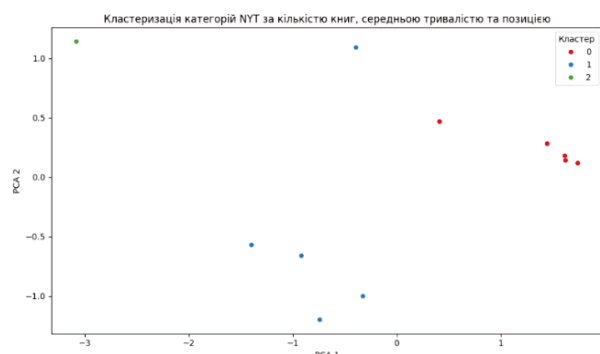


Рис. 24. NYT Bestseller: кластеризація категорій за кількістю книг, середньою тривалістю й позицією

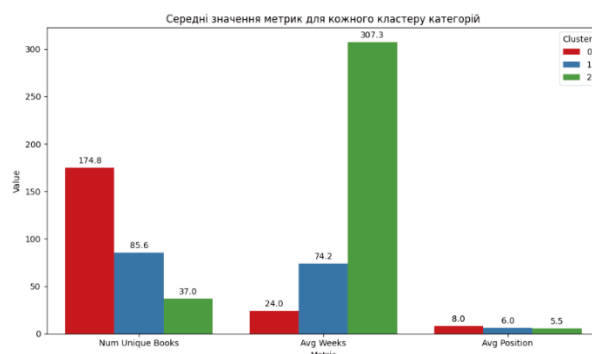


Рис. 25. NYT Bestseller: середні значення метрик для кожного кластеру категорій

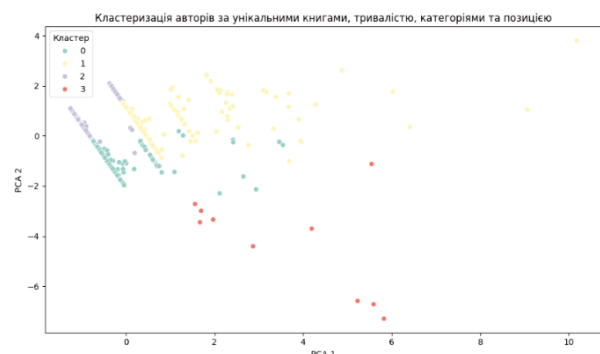


Рис. 26. NYT Bestseller: кластеризація авторів за унікальними книгами, тривалістю, категоріями й позицією



Рис. 27. NYT Bestseller: кластеризація книг за тривалістю, категоріями й позицією

Кількісний зріз long-lived дебютів показує виразну двогорбку сезонність: у першому та третьому кварталах зафіксовано по 142 релізи, тоді як другий квартал дав 118, а четвертий – 135 (рис. 28). Така різниця не випадкова, а лише підтвердила попередні отримані дані та добре вписується у типовий річний ритм книжкової індустрії, де пікові вікна



виходу припадають на початок року та період після літніх канікул, що узгоджується з відомими практиками планування великих прем'єр і маркетингових кампаній.

За сумарною тривалістю присутності лідирують довготривалі дитячі та підліткові серії, які формують "довгий хвіст" розподілу й забезпечують стаке оновлення попиту. Абсолютним рекордсменом є *Diary of a Wimpy Kid* із 822 тижнями у списку, далі *Harry Potter* із 821 тижнем і *Percy Jackson & The Olympians* із 755 тижнями (рис. 29). Такі приклади показують, що серійність і впізнаваний бренд здатні підтримувати інтерес читачів роками, створюючи ефект накопичення аудиторії й підтверджуючи стійкість франшиз як ключового драйвера довгострокової видимості в рейтингах NYT.

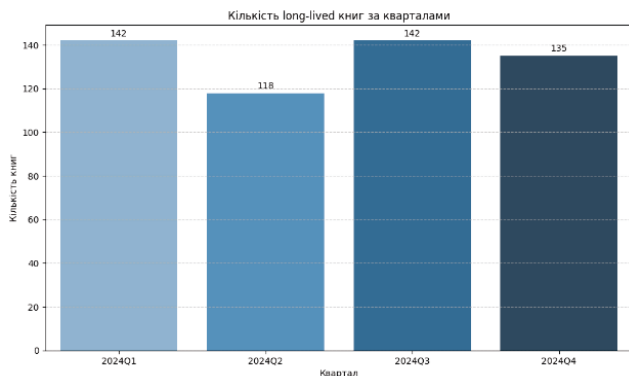


Рис. 28. NYT Bestseller: long-lived книжки за кварталами

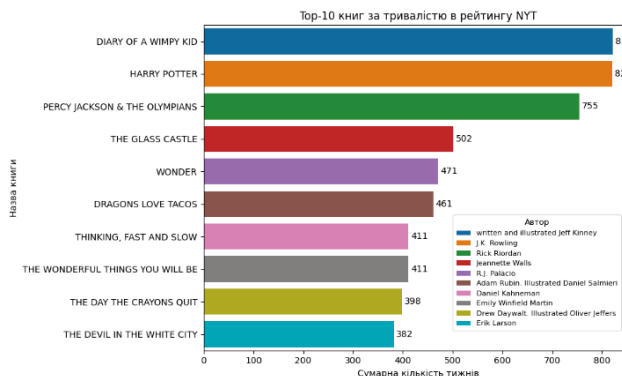


Рис. 29. NYT Bestseller: десять книг – лідерів за тривалістю в рейтингу

Інтегрований аналіз трьох джерел показав виразні відмінності у розподілах популярності та підтвердив очікувані ринкові закономірності. Для Goodreads основна маса нормалізованих значень сконцентрована у діапазоні 0.4–0.6, і лише 14 книжок перевищують 0.8 (рис. 30). Це добре узгоджується з попередніми спостереженнями щодо "довгого хвоста": навіть на платформі з активною читацькою аудиторією більшість видань утримуються на середньому рівні, а поодинокі релізи набирають виняткові оцінки. У каталозі Amazon максимум також припадає на діапазон 0.5–0.6, і лише вісім книжок мають значення вище 0.8 (рис. 31).

Така плоска форма розподілу збігається з дослідженнями великих комерційних маркетплейсів, де поєднується масова пропозиція та розпорошений попит. У рейтингах NYT концентрація ще вища: 64 книжки мають нормалізовану кількість тижнів, близьку до нуля, і лише шість перевищують 0.5 (рис. 32). Це очікувано й відображає жорсткий відбір і короткий життєвий цикл більшості бестселерів.

Дисперсійний аналіз для розрахунку вагових коефіцієнтів показав співвідношення 0.29 для Goodreads, 0.37 для Amazon і 0.34 для NYT. Такі пропорції логічні: комерційна компонента дещо переважає у підсумковій оцінці, що збігається з результатами досліджень про роль комерційних показників у загальному успіху видання. Підсумковий індекс Total Score має максимум близько 0.56 і виразну концентрацію у середньому сегменті (рис. 33). Це підтверджує, що досягнути одночасно високих значень за читацькими, ринковими та медійними метриками здатна лише невелика частина книжок, що цілком очікувано для змішаних ринкових даних і збігається з висновками міжнародних оглядів бестселерів.

Окремий аналіз показника Source Rank продемонстрував, що зі 126 книг, що представлені принаймні на двох платформах, – 78 представлених одразу на трьох, а решта – лише на двох. Для видань із Source Rank = 3 характерна цінова стабільність: міжквартильний розкид не перевищує 1.5, без екстремальних викидів (рис. 34). Це збігається з ринковими спостереженнями: багатоплатформні релізи зазвичай мають узгоджену політику ціноутворення. Частка таких релізів у кожному форматі становить близько 15–20 % (див. рис. 26), що підтверджує галузеві оцінки про обмежене коло "універсальних" бестселерів.



Рис. 30. Інтегрований аналіз: нормалізований рейтинг Goodreads

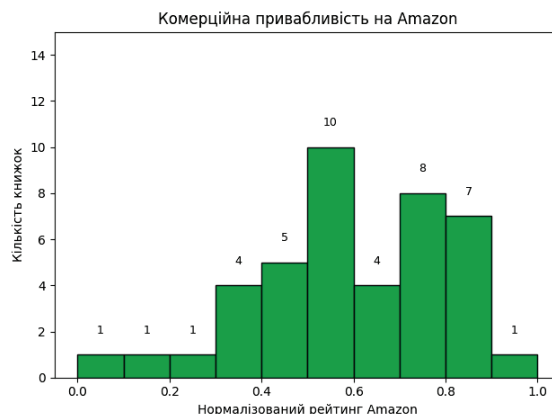


Рис. 31. Інтегрований аналіз: нормалізований рейтинг Amazon

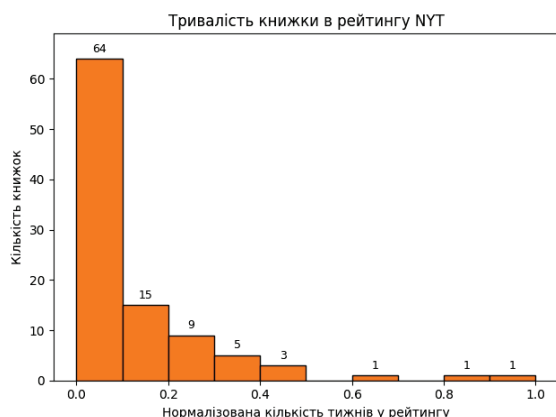


Рис. 32. Інтегрований аналіз: нормалізована кількість тижнів у рейтингу NYT Bestsellers

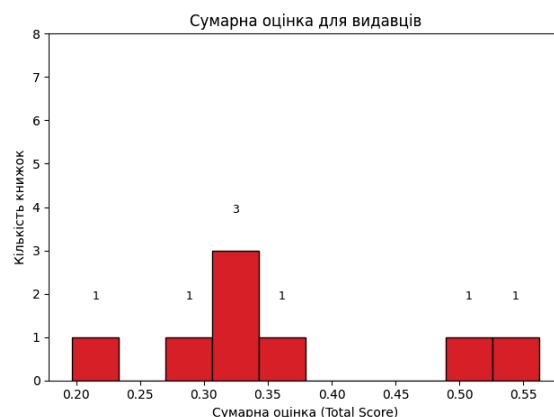


Рис. 33. Інтегрований аналіз: сумарна оцінка Total Score

Жанровий зріз виявив характерну структуру популярності. У групі з високим Source Rank найчастіше зустрічаються Romance і Fiction (по три позиції), а також Audiobook і Fantasy (по дві), що відображає стандартний набір для міжнародних бестселерів. Серед видань із середнім Source Rank домінують Romance і Fiction (по 38), далі йдуть Audiobook і Fantasy (по 28) та Nonfiction (27), що показано на рис. 34. Такі результати збігаються з іншими аналітичними роботами про стійкий глобальний попит на масові жанри. Профіль десяти книжок – лідерів (рис. 35), підтверджує цю картину: провідні позиції займають Fiction і Romance (по 7), далі Contemporary (5), а по 4 – Fantasy, Nonfiction і YA. Ця картина узгоджується з попередніми спостереженнями про гнучкість найуспішніших релізів, які поєднують масові жанри з елементами нішевих напрямів і завдяки цьому здатні довго утримувати увагу різних груп читачів.

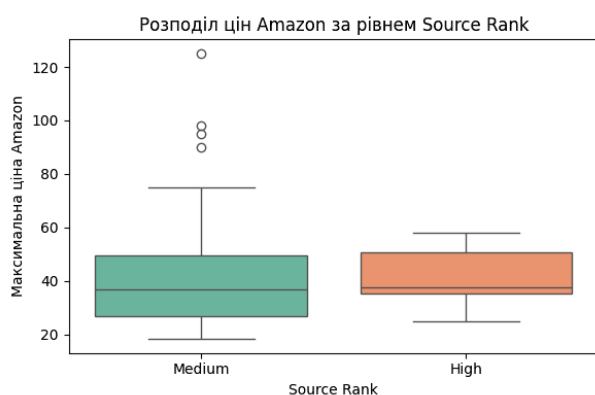


Рис. 34. Інтегрований аналіз: розподіл цін Amazon за рівнем Source Rank

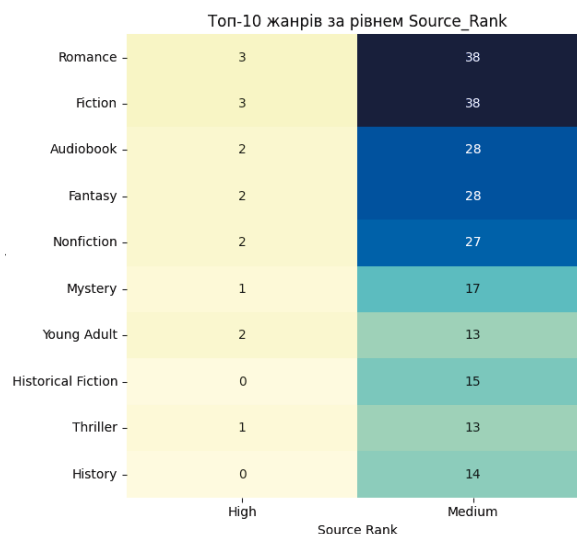


Рис. 35. Інтегрований аналіз: десяти книжок – лідерів жанрів за рівнем Source Rank

Фінальний список книжок із найвищим інтегральним потенціалом перекладу та просування охоплює *Fourth Wing*, *Yellowface*, *The Housemaid*, *Happy Place*, *Tomorrow*, and *Tomorrow*, *Remarkably Bright Creatures*, *The Wager* і *Demon Copperhead*. Усі вони демонструють переконливі результати на трьох платформах і поєднують сильні читачські оцінки, ринкову привабливість і медійну стійкість.

Дискусія і висновки

Отримані результати показують, що довготривала популярність книжки формується на перетині трьох взаємопов'язаних вимірів – читачького інтересу, ринкової привабливості та стабільної присутності в медіа. Поєднання даних Goodreads, Amazon і NYT Bestsellers доводить: успіх видання визначають не миттєві рейтингові сплески, а здатність утримувати увагу завдяки серійності, багатоформатній дистрибуції та продуманій ціновій стратегії. Ця динаміка добре корелює з моделями "довгого хвоста" та кумулятивного попиту, які пояснюють, як тривале споживання створює стійку додану вартість.

Порівняння з попередніми дослідженнями підтверджує і водночас розширює наявні уявлення. Висновки Kousha, Thelwall and Abdoli (2017) про зв'язок між кількістю рецензій і впливом книги доповнюються жанровою сегментацією та індексами нішевості й перспективності, показуючи, що високі рейтинги однаково можливі і в масових, і у вузьких сегментах. Дослідження Maity, Panigrahi and Mukherjee (2018) зосереджувалося на поведінкових сигналах із



Goodreads для прогнозу Amazon-бестселерів, тоді як запропонований підхід включає цінові та медійні параметри, доводячи, що мультиформатність і цінова стабільність є не менш впливовими факторами.

Спостереження Maloku and Maloku (2024) про значущість середнього рейтингу й кількості відгуків як ключових драйверів Amazon-бестселерів у представленій моделі отримали уточнення: фіксується відсутність кореляції між ціною та якістю, але водночас виявляється помірний позитивний зв'язок між кількістю форматів і комплексним індексом успіху. Аналіз сезонності (Pruett, 2020) стосувався лише друкованих видань, в той час як наші результати не лише відтворюють пікові періоди першого і третього кварталів, а й показують, як поява електронних та аудіоформатів змінила тривалість і характер присутності книжок у рейтингах після 2020 р. Нарешті, слабкі кореляції між відгуками Amazon і науковими цитуваннями, які виявили Kousha and Thelwall (2015), узгоджуються з визначенням у цій роботі показником слабкої залежності між кількістю оцінок і середнім рейтингом ($r \approx 0.18$), що ще раз підтверджує обмеженість популярності як показника якості.

Практичне значення представлених висновків виходить далеко за межі академічних дискусій. Видавці можуть використовувати інтегрований індекс і жанрові метрики для формування перекладацьких портфелів, вибору оптимального часу та формату релізів і визначення ніш, що здатні на довгу комерційну "життєздатність". Маркетингові команди та літературні агенти отримують інструменти для точнішого прогнозування попиту й сегментації аудиторій, а книжкові мережі та онлайн-платформи – аналітичну базу для керування асортиментом і персоналізації рекомендацій. Для дослідників культурних індустрій ці дані відкривають шлях до глибшого вивчення впливу серійності, мультиформатності й медійності на сталу популярність літератури.

Отже, мультиджерельний підхід не лише уточнює картину глобальних літературних трендів, а й надає практичну методологію для стратегічного планування – від добору авторів і жанрових напрямів до довгострокового управління видавничими інвестиціями та міжнародного просування.

Крім практичних висновків, дослідження відкриває і перспективи для подальших наукових пошуків. По-перше, це динамічний часовий вимір. Розширення вибірки на багаторічний період дасть змогу простежити еволюцію жанрових і цінових трендів у довгій перспективі й перевірити стійкість "довгого хвоста" в різні економічні та соціальні епохи, включно зі впливом глобальних криз чи різних технологічних зрушень.

По-друге, варто залучити текстову аналітику змісту. Інтеграція методів NLP для аналізу рецензій і самих текстів допоможе виявити стилістичні, тематичні або наративні фактори, що корелюють із довготривалою видимістю та високими рейтингами, розширюючи підхід від кількісних до якісних показників.

По-третє, перспективним є вивчення поведінкових сигналів читачів. Залучення даних із соціальних медіа, платформ на кшталт TikTok (BookTok), Wattpad чи Patreon дасть змогу дослідити, як фанатські спільноти та вірусні тренди впливають на комерційний успіх, серійність і життєвий цикл книжки.

Ще один напрям – моделювання цінових стратегій. Аналіз динамічного ціноутворення й акцій для друкованих, електронних та аудіоформатів дасть змогу точніше оцінити роль гнучкої цінової політики в підтримці стабільного попиту та довготривалої видимості.

Нарешті, перспективним кроком є розроблення прогнозних систем для видавців. Побудова моделей машинного навчання на основі інтегрованих метрик (жанр, формати, рецензії, медійність) може дати змогу прогнозувати ймовірність довготривалого потрапляння у списки бестселерів або визначати потенціал для міжнародного перекладу ще на етапі підготовки видання.

Зазначені напрями поглиблюють розуміння механізмів культурної та комерційної стійкості книжок і створюють підґрунтя для ще точніших інструментів стратегічного планування у видавничій сфері.

Внесок авторів: Володимир Дружинін – написання (перегляд і редагування); Євгеній Батрак – курування та валідація даних, написання (перегляд і редагування); Наталія Цьопа – концептуалізація, методологія, аналіз джерел, підготування огляду літератури; Ольга Ніколенко – формальний аналіз, збір емпіричних даних та їхня валідація, методологія, написання (оригінальна чернетка).

Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Kousha, K., & Thelwall, M. (2016). Can Amazon.com reviews help to assess the wider impacts of books? *Journal of the Association for Information Science and Technology*, 67(3), 566–581. <https://doi.org/10.1002/asi.23404>
- Kousha, K., Thelwall, M., & Abdoli, M. (2017). Goodreads reviews to assess the wider impacts of books. *Journal of the Association for Information Science and Technology*, 68(8), 2004–2016. <https://doi.org/10.1002/asi.23805>
- Maity, S. K., Panigrahi, A., & Mukherjee, A. (2018). *Analyzing Social Book Reading Behavior on Goodreads and how it predicts Amazon Best Sellers* (No. arXiv:1809.07354). arXiv. <https://doi.org/10.48550/arXiv.1809.07354>
- Pruett, J. (2022). *New York Times Hardcover Fiction Bestsellers (1931–2020)*. Post45 Data Collective. <https://doi.org/10.18737/CNJV1733p4520220211>

References

- Kousha, K., & Thelwall, M. (2016). Can Amazon.com reviews help to assess the wider impacts of books? *Journal of the Association for Information Science and Technology*, 67(3), 566–581. <https://doi.org/10.1002/asi.23404>
- Kousha, K., Thelwall, M., & Abdoli, M. (2017). Goodreads reviews to assess the wider impacts of books. *Journal of the Association for Information Science and Technology*, 68(8), 2004–2016. <https://doi.org/10.1002/asi.23805>
- Maity, S. K., Panigrahi, A., & Mukherjee, A. (2018). *Analyzing Social Book Reading Behavior on Goodreads and how it predicts Amazon Best Sellers* (No. arXiv:1809.07354). arXiv. <https://doi.org/10.48550/arXiv.1809.07354>
- Pruett, J. (2022). *New York Times Hardcover Fiction Bestsellers (1931–2020)*. Post45 Data Collective. <https://doi.org/10.18737/CNJV1733p4520220211>

Отримано редакцією журналу / Received: 17.10.25

Прорецензовано / Revised: 20.10.25

Схвалено до друку / Accepted: 19.12.25



Volodymyr DRUZHYNIN¹, DSc (Engin.), Prof.
ORCID ID: 0009-0009-5049-0099
e-mail: volodymyr.druzhynin@knu.ua

Yevhenii BATRAK², PhD (Engin.), Assoc. Prof.
ORCID ID: 0000-0003-3027-4507
e-mail: batrak.yevhenii@ill.kpi.ua

Nataliia TSOPA², PhD (Engin.), Assoc. Prof.
ORCID ID: 0000-0002-3215-0711
e-mail: e-mail: tsopa.nataliia@ill.kpi.ua

Olha NIKOLENKO², Master's Student
ORCID ID: 0009-0003-2304-7821
e-mail: nikolenko.olha@ill.kpi.ua

¹Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

²National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

MULTISOURCE ALGORITHMIC ANALYSIS OF READER AND COMMERCIAL TRENDS IN THE BOOK MARKET

Background. The book market is undergoing rapid transformation under the influence of digital platforms, and timely identification of literary trends has become critically important for publishers working with translated literature. The aim of this study is to identify key reader and commercial trends of 2024 based on data from Goodreads, Amazon, and The New York Times Best Sellers and to assess their potential for shaping effective publishing strategies.

Methods. The dataset included 280 books from the Goodreads Choice Awards 2024, 225 titles from the Amazon Best Books 2024 list, and 949 books from the NYT Best Sellers ranking. Data were collected using Python scripts (Selenium, BeautifulSoup), then cleaned and normalized with pandas and NumPy. Statistical analysis, Principal Component Analysis (PCA), and K-Means clustering were applied to compare reader, commercial, and media indicators.

Results. The integrated analysis of the three sources revealed distinct patterns and shared tendencies. On the reader-preference level (Goodreads), Romance and Fiction attract the broadest audience, while History and Nonfiction consistently achieve high ratings. "Hidden gems" such as the World War II, Mental Health, and Health categories combine strong ratings with limited reach, offering opportunities for targeted promotion. In the commercial segment (Amazon), printed formats – Hardcover and Paperback – dominate, whereas the Kindle Edition remains the most affordable. Premium niches such as Epic Fantasy and regional literary fiction maintain the highest average prices, reflecting readers' willingness to invest in exceptional content. Correlation analysis found no direct link between higher price and better rating but showed a moderate relationship between the number of available formats and overall price. Media visibility (NYT Best Sellers) demonstrates that sustained public interest is driven by children's series and nonfiction, with some titles remaining on the list for more than 300 weeks. Debut seasonality shows two peaks – during the first and third quarters – raising the likelihood of long-term chart presence. Overall, only eight titles achieved strong performance simultaneously across reader, market, and media criteria, demonstrating balanced success and the greatest potential for translation and international promotion.

Conclusions. The integrated analysis shows that a book's commercial success is not determined by a single factor but requires a combination of sustained reader engagement, balanced pricing strategy, and the ability to remain visible in the public sphere. This comprehensive perspective enables publishers to plan portfolios not only by genre priorities but also by long-term economic feasibility. It is equally important to account for temporal demand fluctuations, as market cycles influence the effectiveness of marketing and translation campaigns. Aligning reader, commercial, and media criteria provides a foundation for more accurate forecasting and reduces the risk of one-time successes. Ultimately, the findings create a methodological basis for strategic decisions on translation, investment, and international project development.

Keywords: integrated data analysis, cross-platform comparison, open data, algorithmic analysis, clustering, reader-behavior modeling.

Дружинін Володимир є членом редакційної колегії цього видання, тому він не брав участі в процесі рецензування та прийняття рішення щодо публікації цієї статті.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

Druzhynin Volodymyr is a member of the Editorial Board of this publication; therefore, they did not participate in the peer review process or the decision-making regarding the publication of this article.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

Наукове видання



БЕЗПЕКА ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ

№ 2(10)/2025

Редактор *Л. Магда*

Автори опублікованих матеріалів несуть повну відповідальність за підбір, точність наведених фактів, цитат, економіко-статистичних даних, власних імен та інших відомостей. Редколегія залишає за собою право скорочувати та редагувати подані матеріали.

Оригінал-макет виготовлено Видавничо-поліграфічним центром "Київський університет"
Виконавець *В. Філіпченко*



Формат 60x84^{1/8}. Обл.-вид. арк. 14.5. Ум. друк. арк. 13.0. Наклад 100. Зам. № 226-11674.
Гарнітура Arial. Папір офсетний. Друк офсетний. Вид. № Іт2.
Підписано до друку 29.12.25

Видавець і виготовлювач
ВПЦ "Київський університет",
б-р Тараса Шевченка, 14, м. Київ, 01601, Україна
☎ (38044) 239 32 22; (38044) 239 31 58; (38044) 239 31 28
e-mail: vpc@knu.ua
<http://vpc.knu.ua>

Свідоцтво суб'єкта видавничої справи ДК № 1103 від 31.10.02

