МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
«ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»

БЕЗПЕКА ІНФОРМАЦІЇ ТА ІНФРАСТРУКТУРИ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМ: МІЖДИСЦИПЛІНАРНИЙ ПІДХІД

Монографія

За редакцією
доктора економічних наук, професора Станіслава Горбаченка
та доктора технічних наук, професора Артема Соколова

1256 1233 1996 LIHA-PRES
Львів – Торунь
Liha-Pres
2025

УДК 004.056:004.7:005.8
Б40

За редакцією:

Станіслава Горбаченка, доктора економічних наук, професора, завідувача кафедри кібербезпеки Національного університету «Одеська юридична академія»;

Артема Соколова, доктора технічних наук, професора, професора кафедри кібербезпеки Національного університету «Одеська юридична академія»

Рецензенти:

Олена Нємкова – професор кафедри безпеки інформаційних технологій Національного університету «Львівська політехніка», доктор технічних наук, професор;

Сергій Євсеєв – завідувач кафедри кібербезпеки Національного технічного університету «Харківський політехнічний інститут», доктор технічних наук, професор;

Євген Васіліу – декан факультету інформаційних технологій та кібербезпеки Державного університету інтелектуальних технологій і зв'язку, доктор технічних наук, професор

Рекомендовано до друку рішенням Вченої ради
Національного університету «Одеська юридична академія»
(протокол № 2 від 24 вересня 2025 року)

Б40 **Безпека** інформації та інфраструктури інформаційно-комунікаційних систем: міждисциплінарний підхід : монографія / за редакцією доктора економічних наук, професора С. Горбаченка та доктора технічних наук, професора А. Соколова ; Національний університет «Одеська юридична академія». – Львів – Торунь : Liha-Pres, 2025. – 442 с.

ISBN 978-966-397-537-5
DOI https://doi.org/10.36059/978-966-397-537-5

Монографія присвячена всебічному дослідженню безпеки інформації та інфраструктури інформаційно-комунікаційних систем у контексті нових викликів цифрової епохи. Її концепція ґрунтується на міждисциплінарному підході, який поєднує методи управління, математичного моделювання, криптографії, стеганографії, штучного інтелекту й технічної інженерії. Монографія буде корисною для дослідників, викладачів, студентів, аспірантів і практиків, які формують майбутнє цифрової безпеки України.

УДК 004.056:004.7:005.8

© Авторський колектив, 2025
ISBN 978-966-397-537-5
© Національний університет «Одеська юридична академія», 2025

ПРОФЕСОРУ МИКОЛІ ДМИТРОВИЧУ ВАСИЛЕНКУ

Микола Дмитрович Василенко (1952–2024) – доктор фізико-математичних наук, доктор юридичних наук, професор, видатний вчений, педагог, засновник і перший завідувач кафедри кібербезпеки у Національному університеті «Одеська юридична академія», людина, що передбачила важливість міждисциплінарного підходу до захисту інформації в нову цифрову епоху. Його далекоглядність, відданість справі та щира підтримка колег і студентів залишили глибокий слід у серцях тих, хто мав честь з ним працювати.

Ця монографія створена у пам’ять про нього – як продовження справи, яку він започаткував.

ЗМІСТ

ВСТУПНЕ СЛОВО

Сучасний світ спирається на інформацію. Вона стала ресурсом і фундаментом функціонування економік, систем управління, соціальних інститутів, культури й навіть особистої ідентичності. Сьогодні інформаційно-комунікаційні системи є опорою глобального порядку, а забезпечення їх безпеки – справою, що виходить далеко за межі окремих технологій.

Кібербезпека сьогодні – це не лише справа технічних спеціалістів чи окрема галузь. Це складна і жива сфера, яка охоплює дуже багато – від математичних моделей до психології користувача, від правових норм до управлінських рішень. Щоб дійсно розуміти загрози і будувати ефективний захист, потрібно поєднувати знання з різних наук, мислити системно, бачити причинно-наслідкові зв'язки і вміти передбачати. Саме так, природно й закономірно, кібербезпека стає міждисциплінарною.

Ця монографія об'єднує провідних фахівців із найрізноманітніших сфер – від математики та криптографії до інтелектуальних систем, управління та менеджменту, стеганографії, інженерії та поведінкового аналізу. Вона народилася зі спільного усвідомлення: сучасна інформаційна безпека – це не тільки технологічні рішення, а складний симбіоз людського фактору, комунікації, контексту та соціальної довіри. Саме такий комплексний підхід дозволяє протистояти викликам цифрової епохи.

В межах факультету кібербезпеки та інформаційних технологій Національного університету «Одеська юридична академія» за особистої всебічної підтримки Президента університету, академіка Ківалова С.В. створено всі умови для проведення якісних досліджень та забезпечення іхньої междисциплінарності. Колектив кафедри кібербезпеки Національного університету «Одеська юридична академія» намагається відстежувати сучасні наукові тренди та інтегрувати їх як в навчальний процес, так і у власні наукові публікації.

Сучасний етап розвитку інформаційної безпеки характеризується посиленням складності та адаптивності кіберзагроз, які набувають властивостей поліморфізму, цілеспрямованості та довгострокового латентного впливу. Паралельно відбувається трансформація засобів

захисту, де традиційні реактивні механізми поступаються місцем системам, що ґрунтуються на предиктивній аналітиці, адаптивних криптографічних протоколах, методах маскування даних із застосуванням стеганографії, а також автономних механізмах прийняття рішень на основі штучного інтелекту.

У цих умовах дослідження інформаційної безпеки вимагає синтезу методологій із різних дисциплінарних полів: від теорії складних систем і квантових обчислень до когнітивної психології та соціальної інженерії. Необхідність такого підходу обумовлена тим, що ефективний захист повинен враховувати не лише технічні вразливості, а й антропогенні фактори, включаючи поведінкові патерни, механізми прийняття рішень в умовах невизначеності та динаміку довіри в цифрових середовищах.

Кібербезпека формує повноцінну міждисциплінарну систему, в якій поєднуються управлінські стратегії, математичне моделювання, криптографічно-стеганографічні механізми захисту та технікоприкладні рішення для критичних інфраструктур. Саме ці чотири компоненти визначають архітектоніку монографії, кожен із розділів якої фокусується на одному з ключових аспектів забезпечення безпеки в цифровому середовищі.

Перший блок – управлінський. У ньому розглядаються сучасні виклики, пов'язані з організацією кіберзахисту в умовах криз, динамічних загроз та складних інфраструктурних об'єктів. Акцент зроблено на необхідності координації дій, стратегічному плануванні, ризик-менеджменті та впровадженні інтелектуалізованих систем управління, здатних не лише реагувати на інциденти, а й запобігати їм.

Другий блок – математичний. Тут зібрані дослідження, що закладають теоретичний фундамент для створення алгоритмічних інструментів і прийняття рішень у кібербезпеці. Мова йде про стохастичне моделювання, машинне навчання, інформаційну геометрію та оптимізаційні підходи – ті інструменти, які дозволяють формалізувати ризики, аналізувати ймовірності атак і обирати оптимальні стратегії захисту в умовах невизначеності.

Третій блок – алгоритмічний. Увага приділена новим методам криптографічного і стеганографічного захисту, включно з еволюційним синтезом S-блоків, цифровими водяними знаками,

гібридними криптостего рішеннями й інструментами штучного інтелекту для аналізу прихованого контенту. Тут цифрова безпека постає як високоточна наука, де ефективність вимірюється стійкістю до атак, прихованістю і відтворюваністю сигналів.

Четвертий блок – прикладний. Останній розділ висвітлює технічні та організаційні аспекти захисту в галузях, що мають вирішальне значення для суспільства: енергетика, електромобільна інфраструктура, інформаційно-комунікаційні системи. Тут безпека розглядається як соціально відповідальний процес, що потребує не лише технологій, а й стратегічного мислення, далекоглядного управління та врахування етичного контексту.

Ми сподіваємося, що ця монографія стане корисною для дослідників, фахівців, викладачів, інженерів, управлінців, а також усіх тих, хто шукає відповідей на складні питання про безпеку інформації у світі, що змінюється. І водночас – ми віримо, що ця книга залишить у читача головне: розуміння того, що наука майбутнього, це наука заснована на міждисциплінарності.

Доктор економічних наук, професор Станіслав Горбаченко,
доктор технічних наук, професор Артем Соколов,
кафедра кібербезпеки,
Національний університет «Одеська юридична академія»

РОЗДІЛ 1. УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ В УМОВАХ ЗАГРОЗ ТА КРИЗ

Горбаченко Станіслав Анатолійович
д.е.н., професор, завідувач кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

Саврацький Олександр Олександрович
аспірант кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

УПРАВЛІНСЬКІ АСПЕКТИ ЗАБЕЗПЕЧЕННЯ КІБЕРЗАХИСТУ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

DOI https://doi.org/10.36059/978-966-397-537-5-1

В умовах сучасних геополітичних викликів жодна країна не в змозі забезпечити повний та абсолютний захист власної інфраструктури від впливу зовнішніх чинників. Відтак, єдиний альтернативний варіант полягає у зосередженні основних ресурсів на захисті критично важливих об'єктів. В цьому сенсі термін «критичність» вказує на вірогідність появи небезпеки чи певних скрутних обставин, які, в майбутньому можуть призвести до кризи або інших негативних наслідків.

На державному рівні саме критична інфраструктура є тим фундаментом, на якому тримається стабільність забезпечення основних послуг та функцій, необхідних для нормального функціонування суспільства. Отже кожна країна прагне сформувати власні унікальні стратегії, підходи та інструменти для забезпечення безпеки об'єктів критичної інфраструктури. При цьому орієнтація йде, насамперед, на ті специфічні загрози та вразливості, які генерує кіберсередовище. Особливо загрози кіберсередовища є актуальними для Україні, де з початком повномасштабної агресії спостерігається різке зростання і кількості цілеспрямованих кібератак на критичну інфраструктуру: енергетичну, транспортну, фінансову, урядову тощо.

В контексті посилення кіберзагроз забезпечення надійності, стійкості та безперебійності функціонування критичної інфраструктури є не лише результатом пошуку та використання відповідних технічних безпекових рішень. Одночасно із цим актуалізуються питання законодавчої підтримки, профільних наукових досліджень, впровадження управлінських інновацій, роботи з персоналом підприємств, що відносяться до критичної інфраструктури, організації міжнародної співпраці в сфері кібербезпеки тощо. З огляду на вищезазначене ефективне управління критичною інфраструктурою є неможливим без залучення та координації різних зацікавлених сторін, включаючи органи державної та місцевої влади, бізнес-структури та громадянське суспільство.

Однак на практиці існуючі управлінські підходи все ще недостатньо адаптовані до протидії кіберзагрозам. Зокрема, багато об'єктів критичної інфраструктури не мають ані регламентованих процедур, ані чіткої відповідальності за кіберзахист на рівні керівництва. Отже проблема збільшення ефективності кіберзахисту виходить за межі технічних змін і потребує, у тому числі, оновлення управлінських підходів.

Питання організації кіберзахисту критичної інфраструктури, в тих чи інших аспектах, виступали об'єктом досліджень значної кількості науковців, серед яких Д. Бірюков, І. Діордіца, О. Довгань, Д. Кобилкін, С. Кондратов, О. Суходоля, В. Ящук та інші. Однак, незважаючи на, значний науковий доробок управлінська проблема побудови ефективного кіберзахисту все ще не вирішена. Адже кіберзагрози постійно трансформуються та еволюціонують, стають все більш складними та витонченими. Як приклад можна навести використання штучного інтелекту при кібератаках, що, у свою чергу, вимагає системних та комплексних змін всієї управлінської ланки, особливо у сегменті критичної інфраструктури, який виступає головним об'єктом для кібератак. Тому дослідження управлінської складової кіберзахисту критичної інфраструктури України не втрачають власної актуальності.

Мета дослідження полягає у виявленні можливостей для збільшення ефективності захисту об'єктів критичної інфраструктури від кіберзагроз за рахунок організаційно-управлінських підходів.

Поставлена мета передбачає вирішення наступних завдань:

– проаналізувати сучасні кіберзагрози, що становлять небезпеку для об'єктів критичної інфраструктури;

– оцінити рівень захищеності об'єктів критичної інфраструктури на основі наявних нормативних, організаційних та технічних рішень;

– дослідити існуючі організаційно-управлінські підходи до забезпечення кібербезпеки критичної інфраструктури;

– визначити ключові проблеми в управлінні кіберзахистом на рівні об'єктів критичної інфраструктури;

– обґрунтувати доцільність впровадження менеджменту кібербезпеки в діяльність об'єктів критичної інфраструктури.

Об'єкт дослідження – управлінські процеси забезпечення кіберзахисту об'єктів критичної інфраструктури.

Предметом дослідження виступають управлінські підходи, методи та інструменти підвищення ефективності кіберзахисту об'єктів критичної інфраструктури України.

Критична інфраструктура може розглядатися як сукупність об'єктів, які є стратегічно важливими для економіки та національної безпеки, порушення функціонування яких може завдати шкоди життєво важливим національним інтересам [1]. Тобто мова йде про об'єкти, які є настільки важливими для функціонування держави, що їхнє знищення, критичне пошкодження або, навіть, тимчасове припинення функціонування можуть мати значний (а, в деяких випадках, й невідворотній) вплив на національну безпеку, економічний потенціал країни, конкурентоспроможність окремих галузей, якість життя населення тощо.

З функціонального погляду критична інфраструктура розглядається як сукупність об'єктів незалежно від форми власності, що реалізують функції, виробляють товари (послуги), які є життєво необхідними для людей і діяльності країни та порушення яких призведе до дестабілізації суспільних відносин [2].

З точки зору прив'язки до економічної діяльності критичну інфраструктуру також можна розглядати як сукупність об'єктів, систем, мереж, послуг, які є стратегічно важливими для економіки та безпеки країни, суспільства, населення і пошкодження, знищення або порушення діяльності яких може завдати шкоди життєво важливим інтересам України [3].

За ресурсним підходом критичною інфраструктурою України є системи та ресурси, фізичні чи віртуальні, що забезпечують функції та послуги, порушення яких призводять до найсерйозніших негативних наслідків для життя суспільства та соціально-економічного розвитку країн та національної безпеки [4].

Враховуючи сучасні тренди цифрової трансформації, що відбуваються у сучасному суспільстві в деяких визначеннях фігурує й інформаційна (нематеріальна) складова критичної інфраструктури. В цьому випадку, критичну інфраструктуру визначають як систему надзвичайно важливих матеріальних та нематеріальних об'єктів національної інфраструктури (а також їхню власність та результати діяльності), що забезпечують її стале функціонування, руйнація або пошкодження яких (наявними загрозами) може призвести до людських жертв і значних матеріальних збитків із найсерйознішими негативними наслідками для життєдіяльності суспільства, соціально-економічного розвитку країни та національної безпеки [5].

Виходячи з галузевого підходу критична інфраструктура – системний комплекс стратегічних матеріальних та інформаційних об'єктів виробничої, невиробничої, соціальної сфери, а також окремих складових частин цього комплексу, у тому числі ключових ресурсних, покликаних забезпечувати повноцінний життєвий цикл, безпеку, охорону здоров'я, добробут людини, сталий розвиток суспільства й економіки держави, підтримання її суверенітету з огляду на те, що зловмисне втручання у функціонування, а також пошкодження, руйнація або виведення з ладу системи або її елементів внаслідок диверсій, техногенних чи природних катастроф спроможне призвести до тяжких наслідків: жертв чи поневірянь, шкоди національним інтересам, знецінення надбань людини і держави [6].

Об'єкти критичної інфраструктури відзначаються значною різноманітністю за своїми характеристиками. По-перше, до критичної інфраструктури відносять як матеріальні (фізичні), так і нематеріальні (інформаційно-комунікаційні) компоненти. До фізичних об'єктів можна віднести електростанції, системи централізованого водопостачання, транспортні вузли, аеропорти, залізничні станції, мости та інші інженерні споруди, від яких залежить нормальне функціонування суспільства. Водночас інформаційно-комунікаційна інфраструктура

охоплює телекомунікаційні мережі, центри обробки даних, хмарні сервіси, системи зв'язку та управління, які забезпечують безперебійне передавання та збереження критично важливої інформації.

По-друге, зазначені об'єкти можуть належати до різних форм власності. Вони бувають державними (наприклад, об'єкти енергетичного сектору, морські та річкові порти, оборонні підприємства), муніципальними (заклади охорони здоров'я, комунальні підприємства, системи громадського транспорту), а також приватними (мобільні оператори, провайдери інтернету, логістичні структури, банківські установи). Вказана ситуація зумовлює потребу у координації між різними власниками та органами державної влади для забезпечення надійного функціонування інфраструктури в умовах криз чи надзвичайних ситуацій.

По-третє, діяльність таких об'єктів може здійснюватися на різних економічних засадах. Частина з них функціонує як неприбуткові установи (наприклад, системи очищення води, санітарно-епідеміологічні служби), інші мають змішану модель фінансування та умовно-прибутковий характер (наприклад, державні лікарні, освітні заклади). Нарешті, значна кількість об'єктів критичної інфраструктури працює як комерційні структури з чітко вираженою ринковою орієнтацією: банки, підприємства агропромислового комплексу, великі енергетичні підприємства тощо.

Тобто, критична інфраструктура охоплює за статусом широкий спектр об'єктів, що відрізняються за своєю природою, формою власності та економічною моделлю. З іншого боку можна виділити декілька ключових аспектів, що є спільними для переважної більшості об'єктів критичної інфраструктури (табл. 1).

Таким чином, незважаючи на внутрішню неоднорідність, основне призначення усіх об'єктів критичної інфраструктури полягає у створенні умов для реалізації базових потреб людини та найважливіших функцій держави. Відтак ієрархія об'єктів критичної інфраструктури ґрунтується, насамперед, на показнику критичності, який визначає ступінь (відносний рівень) важливості об'єкта і його вплив на суспільство. Згідно показника критичності до першої категорії відносяться об'єкти, що мають високу важливість для держави в цілому і здатні суттєво впливати на інші об'єкти критичної інфраструктури.

Таблиця 1 – Характерні риси об'єктів критичної інфраструктури [7]

№	Характеристика	Опис
1	Забезпечують надання надважливих послуг.	Надають послуги, від яких критично залежить не тільки економіка, а й усе суспільство: транспорт, енергетика, водопостачання, охорона здоров'я тощо.
2	Значна роль в національній безпеці.	Перебої у функціонуванні об'єктів критичної інфраструктури можуть мати серйозні наслідки для обороноздатності країни. Одночасно, є першими цілями для фізичних атак в умовах військових конфліктів.
3	Ризики кіберсередовища.	Є головною мішенню для кібератак, особливо в умовах військових конфліктів.
4	Активне державне втручання в діяльність.	Держава формує перелік об'єктів критичної інфраструктури, встановлює та, за необхідністю, змінює правила та норми для їхнього функціонування з метою підтримки надійності та безперебійності.
5	Залежність від міжнародної співпраці.	Забезпечення захисту об'єктів критичної інфраструктури потребує міжнародної координації та співпраці, особливо в ситуації, коли вплив критичної інфраструктура виходить за межі національних кордонів (транспорт, енергетика).

Якщо їхня робота буде порушена, виникне кризова ситуація державного значення. Друга категорія включає об'єкти, що є життєво важливими, а припинення чи порушення їх роботи спричинить кризову ситуацію регіонального значення. Третя категорія – важливі об'єкти, порушення чи припинення роботи яких спричинить кризову ситуацію місцевого значення. І, нарешті, четверта категорія складається з об'єктів, порушення чи припинення роботи яких спричинить кризову ситуацію локального значення [8].

В науковій літературі існує багато підходів до ідентифікації об'єктів критичної інфраструктури, зокрема: функціональний, системний, ризик-орієнтований, територіальний, економічний та нормативний, інтегрований підходи (рис. 1). Їх умовно можна поділити на дві основні групи – ті, що орієнтовані на характеристики об'єктів, та ті, що орієнтовані на аналіз ризиків і управління ними.

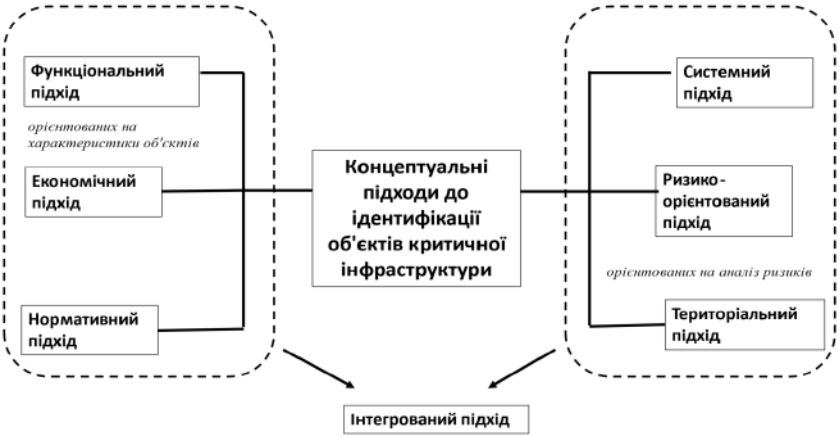


Рисунок 1 – Концептуальні підходи до ідентифікації об’єктів критичної інфраструктури

Обидві категорії підходів до управління критичною інфраструктурою – орієнтовані на характеристики об’єктів та орієнтовані на ризики – формують власні системи методів, інструментів та процедур, які застосовуються в залежності від мети аналізу та управлінських завдань. Підходи, що фокусуються на внутрішніх характеристиках об’єктів (структурі, функціональності, масштабах, забезпеченості ресурсами тощо), зазвичай використовуються при стратегічному плануванні, модернізації інфраструктури, розподілі фінансових та матеріальних ресурсів, а також для оцінки потужностей і можливостей реагування на підвищене навантаження.

Натомість підходи, що базуються на аналізі ризиків, є ключовими для забезпечення надійності та стійкості функціонування інфраструктурних об’єктів у надзвичайних умовах. Вони лежать в основі створення системи управління кризовими ситуаціями, розробки планів реагування на кіберінциденти, сценарного моделювання загроз, а також впровадження превентивних заходів для зниження уразливості інфраструктури.

Розмежування між цими підходами визначається передусім їхніми цільовими орієнтирами та застосовуваними методами аналізу. У кожному конкретному випадку вибір того чи іншого підходу – або

їх комбіноване застосування – залежить від характеру об'єкта, типу загроз, специфіки регіону, а також управлінських потреб і політик у сфері безпеки.

Основні критерії, за якими можна диференціювати вказані підходи, узагальнено в таблиці 2. Вони дозволяють структуровано оцінити, який підхід доцільніше застосовувати в тій чи іншій ситуації для досягнення максимальної ефективності в управлінні критичною інфраструктурою.

Отже, при організації системи управління критичною інфраструктурою необхідно усвідомлювати, що не існує єдиного універсального підходу, який би був ефективним у всіх умовах та для всіх типів об'єктів. Кожен елемент критичної інфраструктури є унікальним за своєю природою, має специфічні функціональні особливості, рівень взаємозалежності з іншими об'єктами, характерні загрози та вимоги до безпеки. Відповідно, підходи до забезпечення його захисту повинні адаптуватися до конкретного контексту.

Ефективне управління безпекою критичної інфраструктури потребує гнучкого поєднання різноманітних підходів – як тих, що базуються на структурно-функціональному аналізі об'єкта, так і тих, що орієнтовані на оцінку ризиків та реагування на загрози. Така комбінація дозволяє враховувати як технічні, так і організаційні аспекти захисту, створюючи умови для забезпечення високого рівня стійкості об'єктів до внутрішніх і зовнішніх викликів.

Під захистом критичної інфраструктури України слід розуміти комплекс заходів, реалізований у нормативно-правових, організаційних, технологічних інструментах, спрямованих на забезпечення безпеки та стійкості критичної інфраструктури [9]. В останні роки серед основних загроз безпеці та стійкості критичної інфраструктури вагоме місце займають кібератаки.

Одночасно відбувається й поступова зміна напряму загроз: з локального та галузевого на загальнонаціональний та наднаціональний [10]. А в умовах повномасштабної війни кібератаки, які здійснювалися на критичну інфраструктуру будь-якої іншої країни у світі, за регулярністю, масовістю, глибиною та витонченістю не йдуть у жодне порівняння з кібератаками, які відбуваються на об'єкти критичної інфраструктури України [11].

Таблиця 2 – Критерії розмежування підходів до ідентифікації об'єктів критичної інфраструктури

Критерії	**Підходи до ідентифікації об'єктів критичної інфраструктури**	
	1 група	**2 група**
Основна мета оцінки об'єктів критичної інфраструктури	Визначення якості та ефективності функціонування в контексті значення для економіки та суспільства.	Визначення потенційних загроз і розробка відповідного реагування.
Методи оцінки об'єктів інфраструктури	Допомагають дослідити основні аспекти функціонування інфраструктури для визначення впливу на соціально-економічну сферу та оточуюче середовище.	Дозволяють оцінити ризики, пов'язані з інфраструктурою, та прийняти необхідні заходи для забезпечення її стійкості та безпеки.
Орієнтація на вплив на економіку та суспільство	Економічний вплив і соціальний вплив є значущими для обох груп, оскільки вони визначають важливість об'єктів для економіки та суспільства.	
Регуляторні та нормативні вимоги	Відповідність нормативам та законодавча база є важливими для обох груп, оскільки вони впливають на визначення стандартів безпеки та відповідність законодавчим вимогам.	
Стратегії управління кризовими ситуаціями	Включають розробку планів надзвичайних ситуацій, постійний моніторинг стану інфраструктури, швидке реагування на кризові ситуації, відновлення функціонування після кризи і проведення аудиту для подальшого вдосконалення стратегій управління.	Включають заходи запобігання (профілактичні заходи), заходи мінімізації наслідків (плани евакуації, резервування даних), а також підготовку до реагування під час кризових ситуацій.

Вищевказані зміни здійснюють вплив і на управлінські дії в сфері кібербезпеки. Адже тривалий час ефективно знижувати рівень кіберзагроз, як ключових загроз безпеці, означало систематично знищувати бізнес-модель, за якою в умовах невисокого ризику нескладні для застосування інструменти використовуються для отримання значних прибутків [12]. У такій ситуації головною метою ефективного кіберзахисту було створення умов, за яких здійснення кібератаки ставало економічно невигідним для зловмисника, насамперед у фінансовому вимірі. Саме тому підприємства та інші суб'єкти господарювання зосереджувалися не лише на впровадженні технологічних рішень у сфері кібербезпеки, а й на системному аналізі власних цифрових активів з позиції їх уразливості та потенційних фінансових наслідків у разі компрометації. Така оцінка охоплювала як прямі, так і непрямі ризики: можливі репутаційні втрати, витік конфіденційної інформації, персональних даних працівників і клієнтів, злам фінансових систем, паралізацію роботи серверів або недоступність веб-ресурсів.

Однак упродовж останніх років кіберпростір став однією з ключових арен гібридної війни, що спричинило ескалацію кіберзагроз на якісно новий рівень. Кібератаки дедалі частіше набувають ознак цілеспрямованого впливу, ініційованого геополітичними гравцями або підтримуваного державними структурами. У таких випадках економічні міркування, як-от фінансова вигода чи рентабельність атаки, відходять на другий план або взагалі втрачають значення.

Метою подібних атак може бути дестабілізація функціонування критичної інфраструктури держави, порушення роботи урядових або військових систем, розповсюдження дезінформації, підрив довіри до державних інститутів чи порушення безпеки міжнародних партнерів. Відтак, рівень загроз значно зростає, і заходи з кіберзахисту мають адаптуватися до нових умов – передусім, шляхом посилення міжвідомчої координації, інтеграції національних і міжнародних стратегій безпеки.

Для формування пропозицій щодо покращення системи кіберзахисту критичної інфраструктури України, необхідно спочатку ідентифікувати увесь спектр потенційних кіберзагроз, а також їхні

джерела. Закон України «Про основні засади забезпечення кібербезпеки України» визначає кіберзагрози як наявні та потенційні явища і фактори, що становлять небезпеку життєво важливим національним інтересам України у кіберпросторі та негативно впливають на стан кібербезпеки держави [13].

В юридичній практиці під кіберзагрозою також розуміють протиправні, карані дії суб'єктів інформаційних правовідносин, які створюють небезпеку життєво важливим інтересам людини, суспільства та держави в цілому, реалізація яких залежить від належного функціонування інформаційних, телекомунікаційних та інформаційно-телекомунікаційних систем, а також відносинам щодо створення, збирання, одержання, зберігання, використання, поширення, охорони, захисту інформації [14].

З точки зору необхідності генерації відповідних управлінських рішень варто зрозуміти, що кіберзагрози виникають у результаті комплексу факторів, які створюють сприятливі умови для атак на інформаційні системи та інфраструктуру. В першу чергу мова йде про швидкий розвиток інформаційних технологій, що значно розширює цифровий простір та надає зловмисникам більше можливостей для впровадження шкідливих дій. Постійне зростання кількості підключених до мережі пристроїв (Інтернет речей) також створює нові вектори для атак, оскільки багато з них мають вразливості у системах безпеки [15; 16].

Другою умовою є людський фактор, який включає помилки користувачів, недостатній рівень кіберграмотності та невідповідне управління паролями. Важливу роль відіграє і соціальна інженерія, що спрямована на маніпулювання людьми з метою отримання доступу до конфіденційної інформації [17].

Третьою умовою є недостатній рівень захисту інформаційних систем. Вказана ситуація може бути викликана як застарілими технологіями, так і відсутністю сучасних засобів захисту, таких як міжмережеві екрани, системи виявлення та запобігання вторгненням, а також антивірусні програми. Відсутність регулярних оновлень програмного забезпечення також створює додаткові ризики [18–20].

Четвертою умовою є складність та взаємозалежність сучасних інформаційних систем, що ускладнює їхній захист і робить системи

більш вразливими до комплексних атак. Взаємозалежність означає, що уразливість в одній частині системи може призвести до компрометації інших частин [21].

П'ятою умовою є економічні та політичні мотиви зловмисників. Кіберзлочинці можуть бути мотивовані фінансовою вигодою, викраденням інтелектуальної власності або дестабілізацією роботи підприємств, що належать до критичної інфраструктури з політичних міркувань. Державні актори також можуть використовувати кібератаки для проведення шпигунських операцій або кібервійни [22–24].

Розгляд та систематизація існуючих кіберзагроз [25–29] в контексті впливу на діяльність об'єктів критичної інфраструктури надає можливість сформувати набір ознак для їхньої класифікації (табл. 3).

Таблиця 3 – Класифікація кіберзагроз критичній інфраструктурі

Класифікаційна ознака	**Види**
За типом загроз	Шкідливе програмне забезпечення. Мережеві атаки, такі як DoS, DDoS, та MitM. Фішинг та інші методи соціальної інженерії. Зловмисні дії інсайдерів.
За метою атак	Крадіжка даних. Знищення або пошкодження інфраструктури. Кібершпигунство з економічних чи політичних мотивів. Пряма фінансова мотивація.
За джерелом загроз	Зовнішні загрози. Внутрішні загрози.
За об'єктом впливу	Атаки, спрямовані на фізичні компоненти інфраструктури. Атаки на інформаційні системи, дані та мережі.

Зазначені класифікаційні ознаки дозволяють системно структурувати кіберзагрози не лише для глибшого розуміння їхньої сутності, але й для пошуку найбільш дієвих засобів та інструментів протидії. Створення ефективної системи кіберзахисту неможливе без усвідомлення реальних і потенційних загроз, постійного вдосконалення засобів оборони, а також інтеграції безпекових підходів у всі

операційні та управлінські процеси підприємств, що належать до критичної інфраструктури.

З огляду на вищезазначене побудова ефективної системи кіберзахисту критичної інфраструктури має ґрунтуватися на комплексному підході, який включає технічну, організаційну та правову аспекти складові, а його реалізація потребує відповідних управлінських рішень (рис. 2).

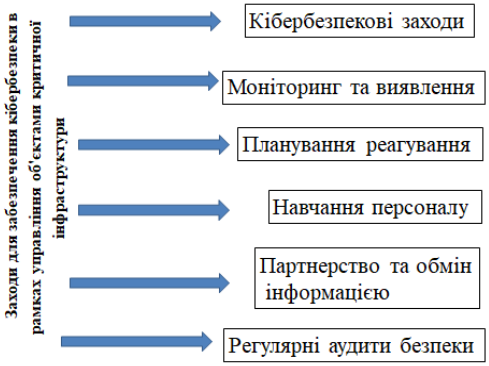


Рисунок 2 – Заходи щодо забезпечення кіберзахисту об'єктів критичної інфраструктури

Звісно, використання виключно управлінських інструментів не в змозі забезпечити захист критичної інфраструктури. Одночасно необхідно впроваджувати передові технології захисту, а також здійснювати регулярне оновлення та патчінг систем. Окремим напрямом можна виділити заходи мережевої безпеки, яка включає захист мережевого трафіку за допомогою інтрузивних систем виявлення та запобігання інтрузіям (IDS/IPS) та використання віртуальних приватних мереж (VPN) для захисту комунікацій. І, нарешті, захист даних, який включає шифрування чутливих даних під час зберігання та під час передачі, також забезпечення регулярних резервних копій та захист від втрати даних. Технічною складовою кіберзахисту об'єктів критичної інфраструктури опікується й держава, адже власник та/або керівник об'єкта критичної інфраструктури організовує невідкладне інформування урядової команди реагування на комп'ютерні

надзвичайні події України CERT-UA (у разі наявності – галузевої команди реагування на комп'ютерні надзвичайні події), а також функціонального підрозділу контррозвідувального захисту інтересів держави у сфері інформаційної безпеки Центрального управління СБУ (Ситуаційний центр забезпечення кібербезпеки СБУ) або відповідного підрозділу регіонального органу СБУ про кіберінциденти та кібератаки, які стосуються його об'єкта [30].

Серед організаційних заходів на підприємствах, що належать до критичної інфраструктури ключову роль відіграє робота з персоналом. Вона обов'язково має включати регулярне навчання співробітників, проведення тренувань і симуляцій кібератак (пентестів), а також розвиток корпоративної культури кібербезпеки. На рівні окремих працівників, особливо тих, хто має доступ до конфіденційної інформації, важливою є свідомість і готовність звітувати про потенційні кіберзагрози та кіберінциденти.

З огляду на вищевказане, в Україні трансформація управлінських систем та прийняття рішень у сфері захисту критичної інфраструктури від кіберзагроз має відбуватися в рамках наступних напрямків (рис. 3).

Рисунок 3 – Напрямки трансформації управлінських систем та прийняття рішень у сфері захисту критичної інфраструктури в контексті кіберзагроз

В управлінському контексті особливе значення має підготовка висококваліфікованих менеджерів, здатних не лише орієнтуватися в сучасних умовах цифрової загрози, а й ефективно застосовувати новітні інструменти управління для забезпечення безпеки критичної інфраструктури. З урахуванням зростаючої складності кіберзагроз, сучасний керівник повинен вміти приймати обґрунтовані рішення, аналізувати ризики, ініціювати впровадження технологічних рішень та координувати дії між різними суб'єктами безпеки. Для цього необхідним є запровадження спеціалізованих освітніх модулів з питань кібербезпеки у навчальні програми як технічного, так і управлінського спрямування. Такий міждисциплінарний підхід дозволить формувати керівників нового покоління, які володітимуть як глибокими фаховими знаннями, так і стратегічним баченням у сфері кіберзахисту.

Крім формальної освіти, важливим елементом підвищення управлінської спроможності є організація практико-орієнтованих заходів – тренінгів, семінарів, стратегічних сесій для керівного складу підприємств, що належать до критичної інфраструктури. Такі заходи мають бути спрямовані на опанування сучасних знань про типи кіберзагроз, способи їх виявлення, локалізації та нейтралізації, а також на розвиток навичок кризового управління в умовах кібератак. Додатковим ресурсом для підвищення наукової та практичної обґрунтованості рішень у сфері кібербезпеки має стати створення дослідницьких центрів, які б займалися розробкою інструментів управління, аналітикою ризиків і впровадженням інноваційних підходів у сфері цифрової безпеки. У контексті глобального характеру кіберзагроз важливо залучати до їхнього дослідження міжнародних експертів, які можуть передати передовий досвід, забезпечити міждержавну координацію та посилити інтеграцію в міжнародну систему кіберзахисту.

Паралельно з освітньо-науковими ініціативами, надзвичайно важливим є належне фінансування системи кібербезпеки, що в умовах обмежених ресурсів та постійного оновлення технологій вимагає чітко структурованого підходу на всіх рівнях управління. Інвестування у кібербезпеку варто розглядати не лише як витрати, а як стратегічну інвестицію у стійкість і життєздатність держави.

На рівні підприємств, що належать до критичної інфраструктури управлінські рішення мають бути спрямовані на впровадження новітніх технологій, оновлення програмного забезпечення, модернізацію систем захисту даних та мереж. Таке спрямування дозволяє значно знизити вразливість до зовнішніх атак і забезпечити безперервність функціонування в умовах кризових ситуацій.

На макрорівні, в межах ІТ-галузі, варто підтримувати дослідження та розробки, спрямовані на створення нових технологій виявлення, реагування та захисту від кіберзагроз. Розвиток науково-технічного потенціалу у цій сфері є запорукою підвищення ефективності захисних механізмів та конкурентоспроможності на міжнародному рівні. Водночас на державному рівні доцільним є створення спеціального фонду підтримки інновацій у сфері кібербезпеки. Такий фонд має стати інструментом стимулювання науково-дослідницької діяльності, комерціалізації технологій, а також підтримки стартапів, які розробляють рішення для захисту критичної інфраструктури. В залежності від специфіки кожного окремого проєкту кібербезпеки фінансування може надаватися на різних етапах – від формування ідеї та концептуального рішення до етапів тестування та впровадження технологій на практиці.

Серед фінансових інструментів, які можуть (хоча й в непрямій формі) сприяти зміцненню кіберзахисту об'єктів критичної інфраструктури, можна відзначити пільгове кредитування та надання грантової підтримки. Вказані інструменти дозволяють підприємствам здійснювати модернізацію наявної системи кіберзахисту без надмірного навантаження на бюджет, а також пришвидшують оновлення технічної бази. Крім того, ефективним підходом є використання моделі державно-приватного партнерства у сфері кібербезпеки. Об'єднання фінансових ресурсів держави та бізнес-структур не лише зменшує ризики для кожної зі сторін, а й створює сприятливе середовище для впровадження інновацій та масштабування рішень, особливо у разі реалізації складних та технологічно затратних проєктів.

Окремі підприємства, що відносяться до критичної інфраструктури можуть підвищити ефективність кіберзахисту на основі управлінської складової, а саме, за рахунок інтеграції менеджменту кібербезпеки. Під останнім слід розуміти процес планування, розробки, впровадження та

керування заходами, які спрямовані на захист комп'ютерних систем, мереж і даних від наявних та потенційних кіберзагроз.

Щоб довести доцільність застосування менеджменту кібербезпеки, насамперед, потрібно зупинитися на таких його категоріях, як результативність та ефективність. В загальному вигляді результативність менеджменту є мірою ступеня реалізації певної стратегії, а ефективність є оцінкою використання ресурсів підприємства у ході її реалізації [31]. Джерелом результативності менеджменту можуть виступати певний рівень спеціалізації менеджера, диференціація видів управлінської роботи, інформаційне забезпечення, автоматизація та алгоритмізація управлінських процесів, координація спільних і функціональних дій управлінської команди [32]. В контексті результативності менеджменту кібербезпеки, яка однаково залежить, як від управлінського фундаменту, так і від надбудови пов'язаної з небезпеками кіберсередовища, має бути визначена ступінь досягнення запланованих цілей у вигляді якісних та кількісних показників. У свою чергу, ефективність менеджменту, відображає якісну властивість результативності, а її розрахунок здійснюється відношенням результату до витрат на його отримання [33]. На прикладі менеджменту кібербезпеки це виявляється в економічній недоцільності здійснення кібервтручання внаслідок наявності належної системи кіберзахисту.

Застосування менеджменту кібербезпеки на підприємствах, що належать до критичної інфраструктури вимагає активного залучення їхнього керівництва. Насамперед, мова йде про прийняття низки управлінських рішень щодо придбання продуктів та послуг кібербезпеки, формування корпоративної культури безпеки та розробки комплексу заходів, що охоплюють організаційні та технічні аспекти кіберзахисту. Робота в цьому напрямі має здійснюватися одночасно на кількох рівнях – стратегічному, оперативному та тактичному – що дозволить забезпечити узгодженість дій, гнучкість реагування та стійкість до нових форм кіберзагроз.

По-перше, оцінка ризиків, яка дозволяє визначити, які активи критичної інфраструктури піддаються кібератакам, і яким ризикам вони піддаються. По-друге, розробка політик та процедур кібербезпеки. Вони мають визначити, як саме потрібно захищати від

кібератак кожен об'єкт критичної інфраструктури. По-третє, для реалізації політики та процедур кібербезпеки необхідно встановлення брандмауерів, антивірусного програмного забезпечення та систем виявлення вторгнень. По-четверте, персонал підприємств, що належать до критичної інфраструктури має бути обізнаний про кібербезпеку та вміти запобігати кібератакам [34].

Впровадження менеджменту кібербезпеки в межах окремих підприємств, що належать до критичної інфраструктури можна розглянути як певну послідовність етапів (табл. 4).

Таблиця 4 – Етапи впровадження менеджменту кібербезпеки

Назва етапу	Характеристика етапу
Кількісна та якісна оцінка ризиків	Оцінка ризиків за допомогою методології управління ризиками в галузі кібербезпеки OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation), міжнародного стандарту з управління ризиками інформаційної безпеки ISO/IEC 27005, методики кількісної оцінки ризиків FAIR (Factor Analysis of Information Risk) тощо
Формування внутрішньої політики кібербезпеки	Розробка нормативної бази з основними вимогами до кібербезпеки на підприємстві критичної інфраструктури. Це можуть бути документи із загальними правилами користування інформаційними ресурсами, визначення ролей, посад та відповідальних за кібербезпеку, правила доступу зберігання, обробки та передачі інформації тощо.
Вибір конкретних заходів забезпечення кіберзахисту	Визначення оптимальних заходів технічного чи організаційного характеру, формування бюджетів під вказані заходи.
Розробка процедур кібербезпеки	Визначаються та прописуються процедури з алгоритмом дій на випадок кіберінцидентів.
Впровадження та підтримка системи менеджменту кібербезпеки	Реалізація розроблених заходів: проведення регулярного моніторингу, оновлення програмних продуктів та безпекових політик, тренування персоналу (наприклад, за допомогою пентенстінгу), розробка навчальних програм тощо.

До початку впровадження повноцінної системи кіберзахисту на об'єктах критичної інфраструктури необхідним кроком є проведення попередньої категоризації. Категоризація передбачає не лише формальне визначення критичності, яка, як правило, закріплюється на державному рівні, але й детальніше оцінювання додаткових параметрів, що вказують на важливість конкретного об'єкта в контексті забезпечення кібербезпеки. Серед таких параметрів можна відзначити рівень залежності населення від функціонування об'єкта, його економічний вплив на регіон чи країну, роль в інформаційному обміні, стратегічне значення для суміжних галузей та інші релевантні характеристики.

Особливу увагу в процесі категоризації потрібно приділяти аналізу взаємозв'язків між об'єктами. Деякі з них можуть не виглядати критичними ізольовано, однак мати ключове значення через опосередкований вплив на функціонування інших систем або елементів інфраструктури. Таким чином, категоризація має носити не лише формальний, а й системний характер. Важливо, щоб вона залишалася динамічною, адже перелік критичних об'єктів потребує періодичної ревізії й оновлення залежно від змін у технологічному середовищі, соціально-економічній ситуації та загрозах у кіберпросторі.

Після завершення етапу категоризації можна переходити до ідентифікації та оцінки потенційних кіберзагроз, які можуть вплинути на ці об'єкти. Як уже зазначалося раніше, класифікація кіберзагроз включає низку ознак, що дозволяють глибше розуміти джерела, методи, умови та можливі наслідки атак. На цьому етапі важливо не лише провести одноразову оцінку загроз, а й запровадити механізми постійного аналізу тенденцій у сфері кібербезпеки. Вказані механізми передбачають застосування моніторингу змін у характері атак, виявлення нових інструментів кіберзлочинців, адаптацію до змін у законодавстві та міжнародних стандартах.

У практичному вимірі таке спостереження реалізується через аналіз звітів про кіберінциденти, участь у профільних конференціях та форумах, співпрацю з галузевими експертними спільнотами та кіберрозвідувальними організаціями. Водночас має здійснюватися регулярний аудит систем кіберзахисту, який включає оцінювання ефективності наявних політик, процедур, технічних засобів захисту та їх відповідності чинним стандартам.

Результати оцінки кіберзагроз та аудиту кіберзахисту слугують підгрунтям для розробки або адаптації стратегій управління ризиками. Зазначений процес охоплює планування оновлень програмного забезпечення, модернізацію технічних рішень, розробку нових внутрішніх регламентів, а також регулярне навчання персоналу, що дозволяє підвищити готовність до реагування на кіберінциденти.

Управління ризиками критичної інфраструктури має бути безперервним, з орієнтацією на мінімізацію потенційних втрат та забезпечення сталості функціонування ключових об'єктів в умовах постійно змінюваного кіберсередовища. Комплексний підхід до категоризації, ідентифікації загроз, аудиту безпеки та управління ризиками формує основу надійної системи кіберзахисту, здатної ефективно відповідати на сучасні виклики.

Вибір продуктів/послуг, заходів та інструментів для забезпечення кібербезпеки залежить від кожної конкретної ситуації і знаходиться більше в компетенції технічних фахівців, але, в будь-якому випадку, кінцеве управлінське рішення та подальше виділення ресурсів – це все ж прерогатива менеджменту (табл. 5).

Однак навіть при забезпеченні надійного кіберзахисту в технологічному аспекті, не можна забувати й про вплив людського чиннику. Часто працівники підприємств, що належать до критичної інфраструктури мають низький рівень обізнаності щодо кібербезпеки, що підвищує ризик успішних атак. На практиці можна спостерігати недостатнє розуміння основних принципів кібербезпеки, незнання актуальних загроз та методів захисту, працівники просто можуть не знати, як правильно реагувати на потенційні кіберінциденти, що може призвести до запізнілих дій і посилення наслідків атаки.

Крім того, в штатному розкладі підприємств, які належать до критичної інфраструктури нечасто можна спостерігати окрему керівну посаду відповідального за управління кіберзахистом. Зазвичай, такі функції виконує хтось з технічних фахівців – системний адміністратор, консультант з IT-безпеки, дизайнер систем мережевої безпеки, фахівець з безпеки тощо.

Хоча, з іншого боку, згідно «Переліку базових вимог із забезпечення кіберзахисту об'єктів критичної інфраструктури» кожен такий об'єкт повинен мати у своєму складі підрозділ або посадову особу з інформаційної безпеки, що відповідають за політику інформаційної безпеки,

прийняту на об'єкті критичної інфраструктури, та контроль за її дотриманням. Під час визначення відповідальних за інформаційну безпеку перевага має надаватися особам, які мають фахову освіту та досвід роботи у сфері технічного захисту інформації або інформаційної безпеки [35].

Таблиця 5 – Умови прийняття управлінських рішень щодо придбання продуктів та послуг кібербезпеки

Характеристика	Опис
Персоніфікація запитів.	Замовник бажає отримати індивідуальне рішення, наприклад, з урахуванням рівня критичності об'єкту, особливостей сфери діяльності, бізнес-процесів.
Варіативність рішень.	На ринку пропонується значна кількість універсальних та галузевих безпекових рішень на різний бюджет, а також послуги у сфері кіберзахисту пропонують окремі розробники та інтегратори.
Технічне погодження.	Хоча презентація безпекових рішень, зазвичай, проводиться перед керівництвом, підсумкове рішення щодо придбання обов'язково погоджується із технічними фахівцями.
Значні бюджети замовлень	Нормальне функціонування більшості об'єктів критичної інфраструктури напряму залежить від рівня кіберзахисту, відповідно і ціна продуктів та послуг кібербезпеки може бути значною.
Запит щодо комплексного захисту.	Причини виникнення та джерела кіберзагроз є різними, але замовник зазвичай вимагає універсальне рішення, яке гарантує максимально ефективний захист на будь-який випадок.
Бажання швидкого виконання	Керівництво здебільшого осмислює потребу у кібербезпеці лише після інциденту і підрахування втрат від нього. І вже тільки потім витрачає гроші на продукти/послуги кібербезпеки, щоб унеможливити або мінімізувати втрати у майбутньому.

Серед конкретних інструментів, що демонструють високу ефективність у практичному навчанні та підготовці персоналу до реагування на кіберінциденти, особливо виділяються симуляції. Вказані тренінги дозволяють моделювати різноманітні сценарії кібератак – від базових форм фішингу до складних ситуацій із витоком даних чи

зловмисним проникненням у мережі. Завдяки використанню спеціалізованих програмних засобів, симуляції створюють максимально реалістичні умови, які дають змогу не лише перевірити готовність персоналу, але й оцінити дієвість внутрішніх процедур безпеки. Після завершення імітації атаки проводиться детальний аналіз дій учасників та їх відповідності встановленим протоколам, що дозволяє виявити слабкі місця та вносити, за необхідністю корективи у навчальні програми і політики безпеки.

У контексті ефективного управління кіберзахистом критичної інфраструктури важливе значення має створення Центрів моніторингу безпеки – Security Operations Center (SOC). SOC забезпечує цілодобовий контроль за подіями у мережі, аналіз потенційних загроз та оперативне реагування на кіберінциденти. Такий центр функціонує як центральний вузол збору та обробки інформації про безпекові події, що дозволяє швидко локалізувати атаки і мінімізувати їхні наслідки для роботи критичних систем. Завдяки координації дій фахівців різного профілю SOC сприяє підвищенню загального рівня кіберстійкості інфраструктури.

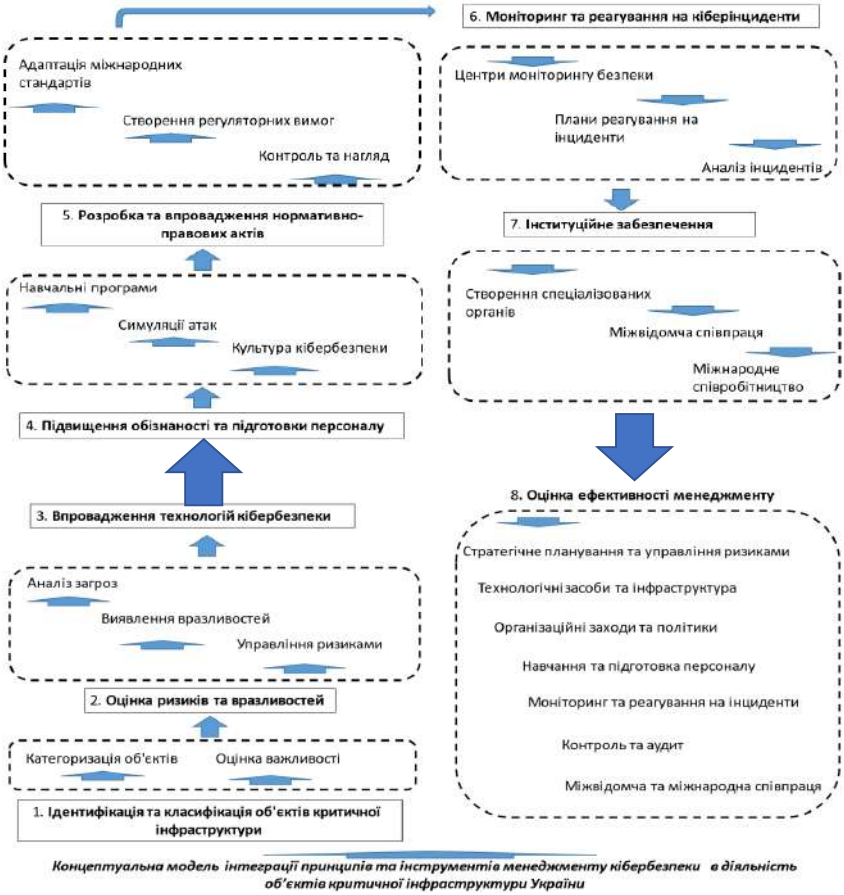
Не менш важливим елементом системи кіберзахисту є розробка та впровадження детальних тактичних планів реагування на кіберінциденти. Вказані документи регламентують послідовність дій у разі виявлення загрози чи атаки, включаючи процедури оповіщення відповідальних осіб, ізоляції уражених систем та заходи щодо їх відновлення. Планування таких заходів дозволяє зменшити час реагування, забезпечити безперервність роботи інформаційних ресурсів та мінімізувати збитки для організації.

З огляду на особливості функціонування об’єктів критичної інфраструктури, не менш важливою складовою їхнього кіберзахисту є наявність надійного нормативно-правового та інституційного фундаменту. Важливою складовою цього процесу є адаптація міжнародних стандартів кібербезпеки, таких як ISO/IEC 27001, до національних умов і законодавства за допомогою визначення конкретних стандартів та вимог щодо захисту інформаційних систем, що мають стратегічне значення для безпеки держави та економіки.

Розробка та впровадження відповідних нормативних актів забезпечує єдність підходів до захисту інформаційних активів на національному рівні, сприяє гармонізації українських стандартів із міжнародними та підвищує

загальний рівень кібербезпеки. В результаті створюється система, здатна ефективно протидіяти сучасним викликам у кіберпросторі та гарантувати стабільність функціонування критично важливих об'єктів.

На основі сформульованих в дослідженні припущень можна сформувати концептуальну модель інтеграції принципів та інструментів менеджменту кібербезпеки в діяльність об'єктів критичної інфраструктури України (рис. 4).



Концептуальна модель інтеграції принципів та інструментів менеджменту кібербезпеки в діяльність об'єктів критичної інфраструктури України

Рисунок 4 – Концептуальна модель інтеграції менеджменту кібербезпеки в діяльність об'єктів критичної інфраструктури України

Розробка та впровадження запропонованої концептуальної моделі забезпечить підвищення рівня захисту критичної інфраструктури України, зменшення ризиків кіберзагроз та в цілому сприятиме стійкості та безпеці суспільства в умовах сучасних викликів.

Висновки. Об'єкти критичної інфраструктури мають надзвичайне значення для національної безпеки, оскільки забезпечують безперебійну роботу ключових систем – енергопостачання, транспорту, зв'язку, водопостачання та водовідведення. Водночас, вказані об'єкти постійно стикаються з різноманітними ризиками, серед яких особливе місце займають загрози, пов'язані з кіберпростором. Сучасні виклики змушують розглядати питання кіберзахисту не лише з технічної точки зору, а й у контексті управлінських рішень, що включають формування адекватних бюджетів, створення ефективних організаційних структур, належну підготовку кадрів, а також впровадження механізмів планування, контролю і координації дій.

Для організації надійного кіберзахисту критичної інфраструктури необхідний комплексний підхід, який об'єднує зусилля на всіх рівнях – від урядових інституцій до окремих підприємств. Такий підхід дозволяє своєчасно запобігати кіберзагрозам і ефективно реагувати на кіберінциденти, що можуть суттєво вплинути на функціонування життєво важливих систем. Водночас комплексність означає не просто технічні заходи, а й продуману організаційну та управлінську стратегію.

У майбутньому зростання потреб у продуктах і послугах кібербезпеки буде стимулювати тіснішу співпрацю між державними органами, місцевою владою, бізнес-структурами, науковими установами, освітніми закладами, стартапами та окремими розробниками. Така взаємодія стане ключовою для обміну знаннями, впровадження передових технологій і кращих практик, що, в свою чергу, сприятиме загальному підвищенню рівня кіберзахисту критичної інфраструктури та забезпеченню стабільності функціонування важливих для держави систем у цифровому середовищі.

Список використаних джерел

1. Газдайка-Васильшин І. Б. Основні терміни проєкту Закону України «Про критичну інфраструктуру та її захист». *Право і суспільство*. 2019. Вип. 9. С. 15–20.

2. Франчук В. І., Пригунов П. Я., Мельник С. І. Безпека об'єктів критичної інфраструктури в Україні: організаційно-нормативні проблеми та підходи. *Соціально-правові студії.* 2021. Вип. 3. С. 142–148.

3. Лойко В. В., Храпкіна В. В., Маляр С. А., Руденко М. В. Економіко-правові засади забезпечення захисту критичної інфраструктури. *Фінансово-кредитна діяльність: проблеми теорії і практики.* 2020. № 4. С. 426–438.

4. Бірюков Д. С., Кондратов С. І., Насвіт О. І., Суходоля О. М. Зелена книга з питань захисту критичної інфраструктури в Україні. Київ : Національний інститут стратегічних досліджень, 2015. 35 с.

5. Єрменчук О. П. Поняття критична інфраструктура. *Інформаційна безпека людини, суспільства, держави.* 2018. № 1. С. 20–28.

6. Теленик С. С. Державна система захисту критичної інфраструктури України: концептуальні засади адміністративно-правового регулювання. Херсон : Видавничий дім «Гельветика», 2020. 602 с.

7. Ящук В. І. Принципи проєктування автоматизованих інформаційних систем управління об'єктами критичної інфраструктури. URL: http://surl.li/stkfad

8. Деякі питання об'єктів критичної інфраструктури: Постанова Кабінету міністрів України від 09.10.2020 р. № 1109. *Верховна рада.* URL: https://zakon.rada.gov.ua/laws/show/1109-2020-%D0%BF#Text

9. Зелена книга з питань захисту критичної інфраструктури в Україні : зб. матеріалів міжнар. експерт. нарад / упоряд.: Д. С. Бірюков, С. І. Кондратов ; за заг. ред. О. М. Суходолі. Київ : НІСД, 2015. 176 с.

10. Горбаченко С. А. Кібербезпека як складова економічної безпеки України. *Галицький економічний вісник*. 2020. Том 66. № 5. С. 180–186.

11. Пядишев В. Г. Кібербезпека критичних інфраструктур: закордонний досвід та українські реалії. *Теоретичні та практичні аспекти забезпечення національної безпеки*. 2022. Частина 2. С. 229–234.

12. Полторак А. С., Сухорукова А. Л., Бурковська А. І. Кібербезпека в системі трансформації управління бізнес-організацією. *Трансформація менеджменту бізнес-організацій: сучасні тренди та виклики* : колективна монографія / за заг. ред. Сагайдака М. П., Соболєвої Т. О., 2021. С. 158–176.

13. Про основні засади забезпечення кібербезпеки України: Закон України від 28.06.2024 р. № 2163-VIII. *Верховна рада*. URL: https://zakon.rada.gov.ua/laws/show/2163-19#Text.

14. Діордіца І. В. Поняття та зміст кіберзагроз на сучасному етапі. Goal, 2017. URL: http://goal-int.org/ponyattya-ta-zmist-kiberzagroz-na-suchasnomu-etapi/

15. Мінін Д. С. Підходи до визначення поняття «кібербезпека» URL: https://salo.li/87Ac1A4

16. Кіберзагрози і кібербезпека: чи здатні фахівці протистояти хакерам? 2016 URL: http://ukr.obozrevatel.com/news/34918-kiberzagrozi-i-kiberbezpeka-chi-zdatnifahivtsi-protistoyati-hakeram.htm

17. Лісовська Ю. Кібербезпека: ризики та заходи : навч. посіб. Київ : Кондор, 2019. 272 с.

18. Свердлик З. Кібербезпека та кіберзахист: питання порядку денного в українському суспільстві. *Український журнал з бібліотекознавства та інформаційних наук.* 2022. № 10. С.175–188.

19. Білявська Ю., Шестак Я. Кібербезпека та кібергігієна: нова ера цифрових технологій. *Товари і ринки*. 2022. № 3. С. 47–59.

20. Вишнівський В. В., Пампуха А. І. Кібербезпека в Україні. *Цифрова трансформація кібербезпеки*: науково-практична інтернет-конференція, 20 квітня 2022, Державний університет телекомунікацій Навчальнонаукового інституту захисту інформації. Київ, 2022. С. 31–33.

21. Кузьменко О., Маклюк О., Чернишова О. Кібербезпека бізнесу під час війни. *Економіка та суспільство*. 2022. № 44. С. 15–19.

22. Кібербезпека в інформаційному суспільстві: Інформаційно-аналітичний дайджест / відп. ред. О. Довгань ; упоряд. О. Довгань, Л. Литвинова, С. Дорогих ; Державна наукова установа «Інститут інформації, безпеки і права НАПрН України» ; Національна бібліотека України ім. В. І. Вернадського. Київ, 2023. № 7 (липень). 270 с.

23. Кіберризики: як розуміти та управляти. *10 Guards*. URL: https://10guards.com/ua/articles/cyber-risks/

24. Веселова Л. Ю. Адміністративно-правові основи кібербезпеки в умовах гібридної війни : дис. … доктора юридичних наук : 12.00.07 адміністративне право і процес; фінансове право; інформаційне право. Одеський державний університет внутрішніх справ. Одеса, 2021. С. 18.

25. Діордіца І. В. Класифікація кіберзагроз та їх легітимація у нормативно-правових актах України. *Підприємництво, господарство і право*, 2017. № 10. С. 206–211.

26. Діордіца І. В. Поняття і зміст кіберзагроз на сучасному етапі. *Підприємництво, господарство і право*, 2017. № 4. С. 99–107.

27. Нова стратегія кібербезпеки: як Україна захищатиметься в кіберпросторі? *Аналітичний центр ADASTRA*. URL: https://adastra.org.ua/blog/nova-strategiya-kiberbezpeki-yak-ukrayina-zahishatimetsya-v-kiberprostori.

28. Про затвердження Положення про організаційно-технічну модель кіберзахисту : Постанова Кабінету Міністрів України від 29.12.2021 № 1426. *Верховна рада України.* URL: https://zakon.rada.gov.ua/laws/show/1426-2021-%D0%BF#Text

29. Даник Ю. Г., Дупелич С. О. Стратегічні аспекти боротьби з робототехнічними комплексами. *Сучасні інформаційні технології у сфері безпеки та оборони.* 2017. № 2 (29). С. 16–25.

30. Про затвердження Загальних вимог до кіберзахисту об'єктів критичної інфраструктури : Постанова Кабінету міністрів України від 19.06.2019 р. № 518. *Верховна рада України*. URL: https://zakon.rada.gov.ua/laws/show/518-2019-%D0%BF#Text

31. Міська Г. Сутність поняття результативність та ефективність в менеджменті. *Науковий вісник Одеського національного економічного університет* : збірник наукових праць, 2020. № 3–4. С. 97–102. DOI: 10.32680/2409-9260-2020-3-4-276-277-97-102

32. Кузнєцов Е. А. Управлінський капітал: майстер-клас: матер. магіст. семінару. Одеса : Фенікс, 2020. 118 с.

33. Сахацький М. П., Казанджі А. В. Теоретико-методичні засади оцінки результативності управління виробничо-господарською діяльністю підприємства. *Фінансово-кредитна діяльність: проблеми теорії та практики.* 2017. № 1. С. 135–141.

34. Павук І. В., Кобилкін Д. С. Особливості формування концепції управління проєктними ризиками на об'єктах критичної інфраструктури. *Інновінг сучасних трендів в менеджменті безпеки*: матеріали всеукраїнської науково-практичної конференції. м. Львів. 26 травня 2023 року. Львів, 2023. С. 59–60.

35. Про затвердження Загальних вимог до кіберзахисту об'єктів критичної інфраструктури : Постанова Кабінету міністрів України від 19.06.2019 р. № 518. *Верховна рада України*. URL: https://zakon.rada.gov.ua/laws/show/518-2019-%D0%BF#Text

Бойко Віктор Дмитрович
к.т.н., доцент, доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ВИКЛИКИ УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ МЕРЕЖЕВИХ СИСТЕМ ТА ІДЕНТИФІКАЦІЇ ВТОРГНЕНЬ ПІД ЧАС КРИЗОВИХ ПОДІЙ

DOI https://doi.org/10.36059/978-966-397-537-5-2

1.1. Вступ

Починаючи з ARPANET, системи інформаційно-комунікаційних систем (information communication systems – ICS) спочатку проєктувалися, як стійкі, гнучкі системи націлені на максимальне збереження функціональності в умовах несприятливих впливів і факторів різного роду.

Будь-яка сучасна інформаційно-комунікаційна мережа (information communication network – ICN) з доступом до глобального інфопростору практично безперервно піддається атакам різного роду і характеру [1], та несприятливим впливам і вражаючим факторам. Сучасні ICN апріорі є вразливими – що підтверджується довгим списком прецедентів, описаних у [2]. Практика захисту та відновлення функціонування ICN показує, що загальна тенденція до ускладнення, централізації, ієрархізації та недостатньо продуманих рішень при плануванні архітектури глобальних ICN призвели до того, що сучасні ICN втратили істотну частину своєї гнучкості та здатності протистояти зовнішнім та внутрішнім впливам.

Ускладнення також призвело до того, що на сучасному етапі однією зі складних проблем захисту ICN є не максимізація стійкості ICN до несприятливих впливів і вражаючих факторів, а мінімізація часу відновлення функціональності та працездатності ICN після інциденту. Найбільш показовим є випадок «Київстар», який зміг відновити повну функціональність своєї мережі лише через 7 днів (з 12 до 20 грудня 2023 року [3; 4]).

У архітектурі безпеки ICN веб-сервери відіграють ключову, детермінуючу роль, оскільки вони є невід'ємними компонентами, що забезпечують безперебійне функціонування критично важливих

інформаційно-комунікаційних послуг. Вони є вузлами обробки запитів, сховищами даних та платформами для доступу до широкого спектру додатків та сервісів, від фінансових транзакцій до управління муніципальною інфраструктурою. Вихід з ладу цих критичних елементів призводить до каскадних наслідків, що поширюються по всій системі, мультиплікуючи збитки та провокуючи вторинні порушення функціональності, що ускладнює загальне відновлення.

Приклад цього явища ілюструється впливом збоїв мобільної мережі на банківський сектор, де значна частина банкоматів, POS-терміналів та систем двофакторної автентифікації залежать від функціонуючих веб-серверів операторів зв'язку. Втрата одного ключового елемента інфраструктури може призвести навіть до паралічу фінансових послуг та порушення роботи критичних комунальних систем, таких як вуличне освітлення.

Таким чином, підвищення живучості веб-серверів є критично важливим для забезпечення стійкості всієї мережевої інфраструктури. Це потребує інтегрованого підходу, що включає технічні рішення, організаційні заходи та постійне вдосконалення методологій, спрямованих на витримування зростаючих кіберзагроз та забезпечення безперебійного функціонування в кризових умовах.

З огляду на експоненційне зростання кількості та складності кібератак, включаючи такі загрози як руткіти, програми-вимагачі та цілеспрямовані атаки на промислові системи, захист веб-серверів є невід'ємною передумовою забезпечення живучості мережі. Це вимагає не лише впровадження передових технологій виявлення вторгнень та забезпечення цілісності даних, але й глибокого аналізу взаємозв'язків між мережевими компонентами та прогнозування потенційних каскадних відмов.

1.2. Загальні виклики стабільності мережі: погляд згори

1.2.1. Вразливість ICN у кризових умовах

Останнім часом почастішали збої глобальних інформаційних мереж. На думку авторів, це є наслідком кількох причин.

По-перше, спостерігається вибухове зростання обсягу та складності ICN. Forbes наводить наступні дані щодо мобільних операторів

в Україні: 19 мільйонів абонентів у Vodafone [5], 8,9 мільйонів у Lifecell [6], 24 мільйони у Київстар [7]. За цими ж даними Lifecell має близько 9000 базових станцій.

По-третє, досвід безперебійної роботи мобільних та інтернет-мереж призвів до того, що багато різних служб так чи інакше використовували їх як частину своєї інфраструктури. Насамперед це стосується банкінгу – банкомати, термінали поповнення, POS-термінали здебільшого використовують мобільний зв'язок для комунікації з банком і при виході з ладу мобільної мережі перестають функціонувати. Сюди ж відноситься широко розповсюджена та просувана як «передова технологія безпеки» двофакторна автентифікація, при якій підтвердження особи користувача відбувається за допомогою SMS-повідомлень. Згідно з джерелами Forbes, у ПриватБанку під час відключення «Київстар» не працювало до третини POS-терміналів та близько 5 % банкоматів. Це найбільш поширені випадки, однак, глобальний збій зв'язку показав, що існують й інші вразливі системи. Зокрема, ЛКП «Львівсвітло» було змушене здійснювати відключення ліній вуличного освітлення в ручному режимі [8].

У роботі [9], що вийшла ще до початку повномасштабного вторгнення, робився наступний висновок:

«В умовах впливу НВ важливим стає не тільки управління ліквідацією наслідків безпосередніх загроз, а й боротьба за функціональну живучість з метою не допустити розвалу системи. В результаті НВ системи можуть безпосередньо виходити з ладу не тільки системи, а й породжуватися»вторинні ефекти», впливаючи на суміжні з ними системи. Для загального розвалу інформаційної системи в результаті НВ часто не обов'язково, щоб було виведено 100 % її підсистем, а досить просто створити зниження працездатності одного або декількох ключових компонентів системи».

Подальші події (блекаути, втрата функціональності систем через бойові дії, нещодавнє відключення мобільного зв'язку «Київстар») підтвердили актуальність викладених у статті положень.

Все перелічене підкреслює актуальність проблеми забезпечення живучості сучасних ICN та гостро ставить питання про надійність і живучість інфраструктурних систем.

1.2.2. ICN та ICS: еволюція загроз та приклади вторгнень

Як зазначалося вище, більшість існуючих інформаційно-комунікаційних систем (ICN) після введення в експлуатацію піддаються атакам різного роду та характеру [1] і з великою ймовірністю можуть бути зламані або виведені з ладу. У випадку, якщо зламу або іншим несприятливим впливам піддаються глобальні інформаційні системи, або інформаційні системи, що обслуговують індустріальні та промислові комплекси (industrial control system – ICS), ризики та масштаби збитків багаторазово збільшуються.

Донедавна основним фокусом небезпеки були розробники промислових та інфраструктурних інформаційних систем, які, на відміну від розробників ПЗ «загального призначення», запізнюються з реакцією на нові загрози. Наприклад, при аналізі безпеки ICS багато фахівців схильні покладатися на так званий "air gap" – «повітряний зазор», що ізолює ICS від глобальних інформаційно-комунікаційних систем. Простіше кажучи – якщо керуюча система не підключена до Інтернету – звідти не може здійснюватися атака. При цьому один із найперших прикладів атаки – вірус Stuxnet, якраз був побудований на подоланні «повітряного зазору» [10].

При аналізі інцидентів мережевої інфраструктури найчастіше посилаються на аналіз розповсюдження та шкоди, завданої Stuxnet [10] – мережевого хробака, націленого на руйнування центрифуг фірми Siemens, задіяних у ядерній програмі Ірану.

Цей хробак розповсюджувався мережею, використовуючи зараження флеш-накопичувачів та чотири вразливості категорії **zero-day**. Наприклад, зараження машини відбувалося не за допомогою традиційного запуску autorun.inf, яке відоме всім і досить легко відстежується та блокується засобами операційної системи, а за допомогою .LNK-вразливості, яку складно виявити та ліквідувати.

Розповсюдження хробака відбувалося доти, доки він не опинявся на машині, яку ідентифікував як потрібну йому (тобто SCADA-систему Siemens, що має певну конфігурацію). Після цього відбувалося перехоплення управління та виконання руйнівних дій. Цікаво, що таким чином (приховано заражаючи флешки) хробак зміг дістатися до комп'ютерів ядерної програми, які не мали виходу в Інтернет та вважалися захищеними від зовнішніх атак.

Ще одним відомим прикладом цілеспрямованої шкідливої програми, націленої на ураження об'єктів промислової інфраструктури, є вірус **Triton** [11; 12], який був націлений на атаку «останнього рубежу оборони» – виведення з ладу приладових систем безпеки (SIS) Schneider Triconex.

Найбільш відомими прикладами атак в Україні можуть слугувати епідемії **WannaCry**, **Petya**, а також збої української енергосистеми у 2015 році, описані в [13] та в [14].

Про готовність комп'ютерної інфраструктури свідчить динаміка зараження під час епідемій програм-вимагачів Petya та WannaCry [15], коли більше третини комп'ютерних мереж (включаючи банківські та державні) виявилися виведеними з ладу. Загальний стан справ з еволюцією атак на об'єкти інфраструктури докладно розглянуто в [16] та [17].

Також одним з різновидів програм-вимагачів є шифрувальник **Snake**, виявлений компанією FireEye (див. [18]), націлений на вибіркову атаку промислових об'єктів, які використовують обладнання General Electric. Наприклад, програма-вимагач атакувала сервер GE Digital Proficy Process [19].

При цьому експерти FireEye стверджують, що якщо до цього моменту шифрувальники були націлені «в простір» – тобто заражали будь-який доступний користувацький комп'ютер, то у випадку Snake можна говорити про цілеспрямовану, таргетовану загрозу конкретним об'єктам інфраструктури.

1.2.3. Аналіз відновлення працездатності ICN: вторинні ефекти

Процес відновлення працездатності ICN являє собою поєднання різних факторів, що часто слабо піддаються оцінці. При відновленні працездатності обслуговуючий персонал може стикатися з проблемами різної природи та генези – недоліком запасних частин, дефіцитом дублюючих потужностей тощо, тощо. Можуть виникати погано передбачувані вторинні ефекти, які мультиплікують загальні збитки та викликають вторинні за своєю природою збої в роботі, що ще більше ускладнює відновлення системи. Це особливо характерно для так званих мереж масового обслуговування, окремим випадком яких є ICN.

Приклад такого вторинного ефекту виникає в таких мережах при виході з ладу деякої кількості вузлів мережі. Оскільки мережа є гнучкою за своєю природою, то вихід одного або декількох вузлів не призводить до загального виходу з ладу мережі безпосередньо, однак, призводить до того, що користувачі, які втратили точки обслуговування, перенаправляють свої запити на інші вузли мережі. Загальне навантаження на ці вузли зростає, що насправді є «мирним» варіантом атаки на відмову (DDoS атаки). Якщо мережа спроєктована без урахування цього ефекту, збільшення навантаження може призвести до відмови нових вузлів, що вціліли при первинному впливі. Вторинні відмови, у свою чергу, призводять до того, що зростає навантаження на вузли мережі, що залишилися – що веде до нової серії відмов. І так далі – процес може розвиватися аж до повної відмови мережі. У разі, якщо несприятливим впливам піддаються інформаційні системи, що обслуговують індустріальні та промислові комплекси, ризики та масштаби збитків багаторазово збільшуються [9].

У статті [2] було описано модель живучості мережі, яка ґрунтується на представленні мережі у вигляді сукупності віртуального або реального обладнання, які в рамках завдання розглядаються як граф вузлів, пов'язаних з'єднаннями різного характеру. Ці зв'язки між елементами можуть розглядатися як зв'язок по ресурсу, що відповідає категоріям живучості системи («енергія» – «інформація» – «речовина»), а сама модель має структурно/функціональний акцент («мережевий» за термінологією [1]). Також у моделі передбачено використання спеціальних, «віртуальних» вузлів, які моделюють причинно-наслідкові залежності та відносини між компонентами [20]). Це дозволяє доповнити модель елементами багатофакторного аналізу. Кожен із вузлів системи може бути частиною ICN, її елементом, або окремим робочим компонентом робочої підсистеми. При цьому цей вузол в рамках моделі має характеристики, що включають час і характер відновлення елемента. У статті пропонувалась методика оцінки часу відновлення використовує безрозмірні позамасштабні оцінки «важливості», «уразливості» елемента з точки зору загальної задачі мінімізації часу відновлення системи в цілому. Така оцінка традиційно спирається на критерій Бірнбаума [21] та ідеї, викладені в [22].

У роботі [22] розглянуто модель відновлення функціональності системи, в якій функція відновлення працездатності описана за допомогою диференціальних рівнянь. Рішення таких рівнянь для інтервалу часу, меншого, ніж чверть періоду рішення (0; π/2), в залежності від отриманих коефіцієнтів можна розділити на три основні категорії, кожна з яких відповідатиме режиму відновлення добре підготовленої, середньопідготовленої та поганопідготовленої системи. При цьому добре підготовлена система описується експоненційним законом, середньопідготовлена система – лінійним, погано підготовлена – тригонометричним – не дуже добре підготовлена система. При цьому під тригонометричним законом мається на увазі залежність, що відповідає першій чверті повного періоду синусоїдальної функції – на ділянці (0; π/2). Істотним моментом використання такої моделі є можливість як оцінки часу відновлення, і розгляд різних сценаріїв розвитку подій – зокрема каскадного виходу мережі з ладу, описаного вище.

Ми пропонуємо доповнити модель, щоб розширити її можливості щодо оцінки вторинних сценаріїв та наслідків кіберзагроз, додавши фактор контрольованого ризику виходу з ладу окремих компонентів системи. Для цього характеристики вузлів когнітивно-імітаційної моделі доповнюються спеціальною характеристикою, яка моделює ризик виходу з ладу компонента на кожному кроці моделювання процесу усунення несправностей та відновлення працездатності системи. Ця величина в залежності від цілей і завдань моделювання може характеризуватись різними за характеристиками розподілу (Нормальне, Вейбулла різних порядків), а в спрощеній моделі може бути представлена математичним очікуванням.

1.2.4. Оптимізація відновлення працездатності ICN: принципи ефективного резервного копіювання веб-серверів

Як зазначалось вище, на сучасному етапі, основною проблемою часто є не максимізація стійкості ICS, а мінімізація часу відновлення функціональності та працездатності ICS після інциденту [23, 1]. Тому на перший план виходить раціональна організація систем резервного копіювання та відновлення ICS (бекапів – backups). Одним із важливих елементів такого резервування є резервування інформації веб-серверів, причому у широкому розумінні цього

терміна – веб-сервер, база даних, операційна система, програмне забезпечення тощо. Нижче наведено рекомендації в організації системи ефективних бекапів, які розроблені в основному для веб-проєктів і систем, які надають інфраструктуру для цих проєктів, проте, ці рекомендації є порівняно універсальними і можуть бути використані для організації резервування більшості проєктів, так чи інакше пов'язаних з ICS.

Розглянуті системи можуть сильно відрізнятися за масштабом та задіяними ресурсами – залежно від проєктів. Різні проєкти вимагають різних стратегій відновлення, різної частоти збереження резервних копій тощо. Тому на попередньому етапі потрібен збір інформації та всебічне дослідження проєкту. Які саме частини проєкту є важливими та незамінними? Автор стикався із ситуацією, коли персонал резервував операційну систему цілком, не розуміючи, що саме треба зберігати на диск. Наскільки важливою є безперебійна робота проєкту? Який проміжок часу є на відновлення у найгіршому випадку?

Існує кілька основних принципів, що склалися в процесі експлуатації ICS та пройшли перевірку практикою:

1. Робоча система та система, що зберігає резервні копії, повинні бути розділені фізично.
2. Резервні копії проєкту мають шифруватися.
3. Після розгортання системи резервного копіювання має бути проведене хоча б одне тестове відновлення проєкту із резервної копії.
4. Запуск та всі проміжні етапи резервування повинні здійснюватися автоматично за розкладом. Процес резервування має документуватись у системних журналах.
5. Залежно від проєкту має зберігатися декілька копій. Повинен бути організований процес їхньої ротації.

Розглянемо ці принципи докладніше.

1. Робоча система та система, що зберігає резервні копії, повинні бути розділені фізично.

Резервні копії не повинні зберігатися на робочій станції проєкту. Це має бути окрема станція, дуже бажано фізично розділена з робочою. Різні приміщення, в ідеалі – різні будівлі. Ця вимога обумовлена здоровим глуздом. Якщо резервна копія зберігається на

робочій станції, то при виході станції з ладу доступ до бекапу буде втрачено. Фізичний поділ працює на додаткову безпеку бекапу. Слід пам'ятати, що крім кібераттак, робоча станція може піддаватися суто фізичним впливам: пожежа, затоплення приміщення, пошкодження внаслідок бойових дій. У цьому випадку бекап збережений на віддаленій дистанції має набагато більше шансів зберегтися, ніж якщо він буде в одному приміщенні з робочою станцією.

У межах невеликих та середніх проєктів реалізація цього принципу, всупереч першому враженню, необов'язково потребує ресурсів та витрат праці. Залежно від вимог до проєкту, в ролі сервера бекапу може виступати окремий комп'ютер з розгорнутим на ньому ssh-сервером – персональний комп'ютер, одноплатний комп'ютер. Автори зустрічали випадки організації систем бекапу на базі таких зовні «несерйозних» станцій, як списані ноутбуки та роутери з кастомною прошивкою. Повторимося – все залежить від вимог проєкту.

2. Резервні копії проєкту мають шифруватися.

Ця умова випливає із першого принципу. Бекап – це інформація про проєкт. Часто із конфіденційними даними. Якщо бекап зберігається окремо від робочого проєкту – у зловмисників з'являється потенційна можливість доступу до інформації в обхід систем захисту основного проєкту. Тому бекап, що зберігається, повинен бути захищений від доступу третіх осіб, як мінімум засобами сильної криптографії – тобто зашифрований.

3. Після розгортання системи резервного копіювання має бути проведене хоча б одне тестове відновлення проєкту із резервної копії.

Резервування часто виконується для досить складних структурою і великих за обсягом даних проєктів. У цьому випадку перевірити можливе відновлення зі збереженого бекапу чи ні, можна тільки проробивши цю процедуру на практиці. Інакше виникає ситуація, що отримала у фахівців назву «бекапу Шредінгера» – система резервування начебто працює, бекапи зберігаються, але чи все буде відновлено після збою, ніхто не знає.

Ще один неочевидний момент виникає тоді, коли проєкт змінює робочу конфігурацію, а бекапи зберігаються виходячи зі старої. Як було сказано вище, щоб така ситуація не виникала – потрібно щонайменше одноразове перевірочне відновлення з резервної копії. Бажано,

щоб воно проводилося регулярно і точно, воно вимагає повторення при кожній зміні конфігурації проєкту.

4. Запуск та всі проміжні етапи резервування повинні здійснюватися автоматично за розкладом. Процес резервування має документуватись у системних журналах.

Якщо процес резервування зав'язаний на людину – виникає людський фактор. Люди схильні плутати необхідні дії, забувати, хворіти тощо. Навіть якщо людині для збереження бекапу потрібно натиснути одну кнопку – велика ймовірність, що вона її забуде натиснути саме в день, коли знадобиться бекап.

Тому слід виключити людський чинник із процесу збереження бекапів. Навіть якщо не розглядати спеціалізоване стороннє програмне забезпечення, будь-яка сучасна операційна система має вбудовані можливості, які дозволяють планувати та здійснювати запуск програм на регулярній основі. Для ОС сімейства Windows – це task sheduler («Планувальник завдань»), для POSIX-сумісних систем (Linux, BSD) – це cron, або таймери служби systemd у більш сучасному варіанті.

Автоматизація процесу дозволяє отримати ще одну перевагу – процес резервування може запускатися у позаробочий час (наприклад о 3 або 4 годині ночі), не заважати робочим процесам та отримуючи максимальний доступ до ресурсів самостійно.

Однак, як у будь-якій іншій системі, автоматизація не завжди спрацьовує гладко спочатку, тому вона повинна обов'язково супроводжуватися звітами про минулі успішні бекапи, щоб це можна було проконтролювати.

5. Залежно від проєкту має зберігатися декілька копій. Повинен бути організований процес їхньої ротації.

Бекап рідко буває поодиноким. Залежно від проєкту, бекап може бути досить об'ємним. Це вимагає місця для зберігання та ефективного поводження з резервними копіями, що накопичуються. Оптимально організовувати стек та ротацію бекапів, а також зберігати бекапи за системою «рік – місяць – день тижня». Це дозволить з одного боку ефективно використовувати доступні ресурси зберігання даних, а з іншого – не втратити інформацію при збоях в роботі основної станції.

Дотримання перерахованих вище принципів дозволить організувати стійку і надійну систему резервування проєкту та уникнути втрат при можливих збоях, чим суттєво підвищить стійкість до несприятливих впливів та вражаючих факторів різного роду.

1.3. Питання безпеки веб-серверів як складової загальної мережевої стійкості

1.3.1. Веб-сервери, як основна мішень атак ботнетів

Веб-сервери є основною мішенню для атак, що здійснюються за допомогою ботнетів. Ця тенденція обумовлена їхньою критичною роллю у функціонуванні мережевої інфраструктури та широким розповсюдженням.

У широкому розумінні, веб-сервер охоплює не лише безпосередньо програмне забезпечення (наприклад, стеки LAMP/LAPP), але й усю пов'язану інфраструктуру. Він є основою та ключовим компонентом сучасної всесвітньої мережі.

Окрім очевидних великомасштабних рішень, значного поширення набули також різноманітні реалізації серверів для домашніх мереж, а також сервери, що забезпечують веб-інтерфейс для різних застосунків. Зокрема, у системах Linux веб-сервер використовується для конфігурації системи друку Common UNIX Printing System (CUPS). Широке впровадження цих рішень сприяло повсюдному поширенню веб-серверів, водночас перетворивши їх на найбільш очевидну мішень для кібератак.

Можна виділити наступні причини, з яких веб-сервери є бажаними та пріоритетними цілями для кібератак:

– Вебсервер за своєю природою є відкритою системою, яка постійно підключена до зовнішніх інформаційно-комунікаційних мереж. Відповідно, за невеликим винятком, для атаки на нього не потрібен фізичний доступ.

– З широким розповсюдженням веб-інтерфейсів веб-сервери набувають все більшого поширення та можуть інтегруватися у найнесподіваніші місця, наприклад, у системи конфігурації принтерів та іншого офісного обладнання.

– Веб-сервери все частіше виконують роль інтегруючого рішення для організації різноманітних систем колективної взаємодії

на рівні як зовнішніх зв'язків, так і роботи в інтранеті (внутрішніх локальних мережах організацій), а також функціонування промислових систем, які мають бути відокремлені від зовнішнього світу, проте на практиці ця умова рідко дотримується.

У разі успішної атаки шкідливе програмне забезпечення має високу ймовірність залишатися невиявленим протягом тривалого часу.

– Шкідливе програмне забезпечення, що впровадилося на сервер, потенційно здатне інфікувати велику кількість комп'ютерів, що взаємодіють з ним, за короткий термін.

– Шкідливе програмне забезпечення, що впровадилося на сервер, потенційно отримує доступ як до конфіденційної інформації на самому сервері, так і за його межами.

– Персонал, відповідальний за встановлення та налаштування веб-систем, часто є некомпетентним та некваліфікованим.

– Попередній факт ускладнюється тим, що існує безліч різних реалізацій програмного забезпечення веб-серверів, конфігурування та забезпечення безпеки яких можуть суттєво відрізнятися, що ускладнює навчання та перехід від однієї конфігурації до іншої.

– Функціонування веб-сервера забезпечується набором (стеком) взаємодіючих програм: власне веб-сервер, база даних, мова програмування та пов'язані з нею фреймворки, системи конфігурації та взаємодії, операційна система та протоколи доступу до неї. Кожен із цих компонентів, за умови неграмотного конфігурування та експлуатації, може становити потенційну або реальну вразливість.

– Зі складністю та різноманітністю систем роботи веб-сервера пов'язаний великий обсяг «сирої» інформації, в якій легко загубитися, якщо не володіти технічними засобами аналізу та обробки. Текстові логи відвідувань сервера за кілька днів цілком можуть досягати розміру порядку гігабайтів, що робить безглуздим їх «неозброєний перегляд».

Таким чином, забезпечення безпеки веб-серверів на всіх рівнях їх функціонування є і буде актуальним та важливим завданням, враховуючи ключову роль, яку веб-сервери відіграють у структурі сучасного інформаційного простору, а також все зростаючу роль інформаційних комунікацій (у тому числі через епідемію коронавірусу та пов'язане з нею підвищення інтересу до дистанційних форм роботи).

1.3.2. Руткіти як головний інструмент розповсюдження ботнетів

Основним джерелом мережевих атак в даний час є ботнети – мережі заражених malware комп'ютерів, що мають доступ до інтернету. Ботнети є серйозною проблемою – досить сказати, що більшість комп'ютерних атак здійснюється за їх допомогою [24; 25].

Одним з найбільш небезпечних різновидів шкідливого програмного забезпечення є руткіти (rootkit). Так називають ПЗ, націлене на приховане впровадження в систему (в даному випадку – веб-сервер) та отримання в ній прав адміністратора (ескалація привілеїв). Руткіт прагне залишатися прихованим, але при цьому йому, як правило, потрібен зв'язок із зовнішнім світом [26].

Ранні дослідники класифікували руткіти та аналогічне за призначенням ПЗ за класами прихованості [27]:

– **Програмне забезпечення, що демаскує себе – Рівень (–1)**: зараження тягне за собою порушення функціональності зараженого об'єкта, що робить виявлення зараження неминучим.

– **Відсутність прихованості – Рівень 0**: жодних спеціальних заходів для приховування присутності інфекції не вживається.

– **Елементарна прихованість – Рівень 1**: виконуються основні дії щодо запобігання виявленню, такі як збереження у змінюваних файлах метаданих дати та часу, щоб приховати факт їхньої зміни.

– **Середній ступінь маскування – Рівень 2**: частина образу (image) зараженого об'єкта в його початковому чистому стані до зараження зберігається для «показу» системі, щоб допомогти приховати «слід» вірусу.

– **Просунуте маскування – Рівень 3**: додаток використовує методи маскування, спеціально призначені для приховування зараження від спеціальних додатків, що забезпечують безпеку.

На ранній стадії основним об'єктом атак були фінансові та платіжні системи: підробка реквізитів банківських кредитних карток, злом кодувань супутникових телевізійних каналів тощо.

У міру збільшення рівнів прихованості впроваджуваний шкідливий код все більше адаптується не з метою «традиційного» вандалізму (видалення, псування файлів тощо), скільки для здійснення прихованої діяльності.

Такою діяльністю може бути:

– приховане зараження комп'ютерів, що контактують із сервером;
– збір конфіденційної інформації;
– організація та управління атаками.

На наступному етапі зломи стали інтелектуальнішими – зокрема, злам мережі банкоматів з метою отримання грошей та програм-вимагачів (ransomware) – шкідливих додатків, що шифрують диск та пропонують перевести гроші для дешифрування.

Наприклад, відомий мережевий хробак Stuxnet приховано розповсюджувався по системах, приховано заражаючи комп'ютери та уникаючи виявлення антивірусними системами доти, доки конфігурація зараженого комп'ютера не збіглася з потрібною. Після цього хробак спрацював та вивів з ладу центрифуги фірми Siemens, керуючі комп'ютери яких навіть не були підключені до інтернету (хробак потрапив туди із зараженого флеш-носія) [10].

Справжньою проблемою є атаки, що здійснюються для отримання приватної інформації, яка не підлягає розголошенню (наприклад, після атаки на Korea Hydro & Nuclear Power Co., Ltd. (KHNP) зловмисники отримали доступ до креслень та інструкцій для атомних реакторів [28]).

Все вищеперелічене визначає труднощі у виявленні не тільки діяльності руткіта, але часто і самого факту його вторгнення. Що опосередковано підтверджується наступними цифрами.

Згідно з Mandiant Security Effectiveness Report 2020 [19], 53% кібератак закінчуються невиявленим впровадженням в організаційні структури, і 91% всіх інцидентів не викликають спрацьовування систем безпеки.

У 2015 році середній час виявлення факту успішної кібератаки становив від 50 до 70 днів [29].

Згідно зі звітами IBM [30], тільки на ідентифікацію витоків йде до 280 днів, оцінки інших експертів варіюються в межах від сотень днів до року.

1.3.3. Аудит веб-серверів та виявлення фактів вторгнення у систему

Комп'ютер, підключений до всесвітньої павутини, безперервно піддається атакам та «випробуванню на міцність» його систем

захисту. При цьому кількість ботнетів, заражених комп'ютерів та атак зростає з кожним роком. Крім того, програмне забезпечення, техніки, тактики та стратегії сторони, що атакує, постійно вдосконалюються.

Визначальним фактором успішного функціонування сучасних ботнетів є часовий інтервал між вторгненням та зараженням комп'ютера та виявленням того, що це сталося. Цей інтервал визначає ступінь шкідливості malware – чим він довший, тим більше атак встигне здійснити заражена машина, до того, як це виявиться і будуть застосовані заходи у вигляді налаштування фаєрволу, антивірусного сканування тощо. аж до переустановки операційної системи. Однак, виявлення вторгнень у систему в сучасних умовах все ще недостатньо швидке. У 2015 році середній час виявлення факту успішної кібератаки складав від 50 до 70 днів [29]. Згідно з звітами IBM за 2020 рік [30], на ідентифікацію витоків витрачалося в середньому 280 днів, Madiant Security Effectiveness Report 2020 [19] вказує, що 53 % кібератак закінчується працювання систем безпеки.

Проблему намагаються вирішити системним підходом – використанням технології SIEM (Security information and event management), яка є синтезом існуючих раніше систем SEM (Security Event Management) і SIM (Security information management) [31; 32]. Однак, на практиці впровадження складних, витратних за ресурсами та потребують кваліфікації керуючого персоналу систем захисту, часто виявляється утрудненим для розробників та адміністраторів ICN середньої та нижчої ланки.

При цьому робота з ключовою для виявлення та управління кібербезпекою інформацією донедавна будувалася виходячи з парадигми того, що вся інформація в ICN буде доступна в будь-який момент часу, що функціонування ICN мінімально схильне до збоїв і без урахування зростання складності ICN. Ці питання докладно розглядалися у роботі [9] що вийшла ще до початку повномасштабного вторгнення, блекаутів та кількох глобальних відключень важливих структурних складових ICN («Київстар», facebook, microsoft тощо). Наступні події (блекаути, втрата функціональності систем через бойові дії, нещодавнє відключення мобільного зв'язку «Київстар») підтвердили актуальність викладених у статті положень та були додатково розглянуті у [2]. Сучасна ICN існує в умовах

все зростаючої щільності несприятливих впливів і факторів, що вражають [1]. При цьому структура самої ICN забезпечує досить великий ступінь автономності, що з одного боку дозволяє швидше відновлювати систему після збоїв або вимушеного простою (наприклад, в результаті блекауту), а з іншого ускладнює управління інформацією та ключовими подіями для кібербезпеки ICN. Наприклад, якщо блекаут викликав втрату зв'язності окремих частин ICN, це може призвести до збоїв у роботі SIEM, яка не розрахована на несинхронну роботу з розподіленою системою.

1.3.4. Засоби виявлення та пошуку руткітів

Класичний спосіб боротьби та виявлення шкідливого ПЗ полягав у побайтовому аналізі бінарних файлів з метою виявлення порушень цілісності файлу, невірних значень його метаданих. Пізніше до цього арсеналу додалися методи виявлення відомих версій шкідливого ПЗ за їхніми характерними сигнатурами (фрагментами коду), ще пізніше – евристичний аналіз файлів на пошук невідомих шкідливих додатків за характерними ознаками поведінки.

Серед сучасних засобів виявлення та пошуку руткітів в рамках локальної системи можна виділити три основні категорії:

– аналіз бінарних файлів;
– аналіз поведінки системи та користувача;
– аналіз системних журналів.

Традиційний класичний аналіз бінарних файлів системи, який, своєю чергою, включає сканування з метою виявлення сигнатур відомих шкідливих програм та евристичний аналіз. Як пошук сигнатур, так і евристичний аналіз, за всієї своєї практичної цінності, мають наступні недоліки.

Аналіз сигнатур, з одного боку, простий, інтуїтивно зрозумілий, не вимагає великого обсягу обчислювальних ресурсів і дозволяє визначати конкретну атаку (та використане для неї шкідливе ПЗ) з високою точністю та відносно низьким рівнем хибних спрацьовувань (false positives). З іншого боку, «на довгій дистанції» збір сигнатур та визначення вирішальних правил виявляється гранично трудомістким завданням – сучасні бази даних, наприклад, у базі даних антивірусу Comodo міститься понад 86 359 675 сигнатур – майже сто мільйонів.

Пошук та виявлення, підтримання актуальності та оновлення таких баз є трудомістким завданням. Сканування системи також має тенденцію затягуватися – як через зростання баз даних із сигнатурами, так і через зростання обсягів сканованого програмного забезпечення. Ця проблема може частково зніматися зростанням швидкодії системи та роботою з вибіркового сканування найбільш вразливих місць системи, однак швидкість роботи антивірусу стає все повільнішою. Крім того, пошук за сигнатурами не спрацьовує на поліморфних та самомодифікованих вірусах, на вірусах, що шифрують своє тіло тощо. Він не виявляє нових екземплярів шкідливого ПЗ і не здатний запобігти атаці з використанням щойно написаного програмного забезпечення.

Евристичний аналіз є в деякому сенсі протилежною аналізу сигнатур процедурою пошуку. Евристики для пошуку шкідливого ПЗ можуть виявляти нові та щойно написані екземпляри, база даних із правилами менш громіздка, ніж база із сигнатурами, відповідно, перевірки можуть теоретично відбуватися швидше та з меншими витратами ресурсів. Однак евристичні аналізатори зазвичай відрізняються високим рівнем хибних спрацьовувань. Евристичні правила досить легко обходяться, якщо вони відомі розробнику шкідливого ПЗ заздалегідь. Поширена практика, коли такі розробники заздалегідь закладають у ПЗ атаки спеціальні прийоми та системи, які дозволяють обходити евристичний аналіз – наприклад, фрагменти коду, що ламають або виводять з ладу відладчик при спробі трасування результатів виконання програми.

Як правило, існуюче антивірусне ПЗ використовує комбінований підхід, що поєднує пошук за сигнатурами та евристичний аналіз файлів. Однак ефективність такого ПЗ далека від заявленої.

Другий тип пошуку шкідливого ПЗ орієнтований на непрямий аналіз того, що відбувається в системі. Такий аналіз може включати аналіз мережевої активності (зміст та характер вихідного та вхідного трафіку), аналіз активності додатків, аналіз активності користувача.

Наприклад, найпростішим видом такого аналізу може бути аналіз стану портів системи, оскільки частина шкідливого ПЗ може використовувати прослуховування вільних портів для комунікації із зовнішнім світом.

Аналіз такого типу може бути досить ефективним, однак широке розповсюдження піринговых та криптографічних технологій комунікацій все більше ускладнює його роботу, оскільки все складніше стає аналізувати трафік, який не просто шифрується, а набуває такої властивості, як однорідність (наприклад, у разі використання технологій Tor та I2P – складно відокремити в потоці трафіку, що проходить через систему, нативний трафік самої системи від потоку транзитних даних, які проходять крізь систему, оскільки вона віддає частину своїх ресурсів загальній мережі).

З урахуванням усіх перелічених переваг та недоліків проблема захисту системи від вторгнень продовжує залишатися актуальною. При такому підході більша частина безпеки сервера визначається профілактичними заходами (hardening) та грамотною конфігурацією системи – встановленням брандмауера, використанням стійких паролів та системи відкритих ключів, правильним розподілом ролей та прав системних користувачів, зміною налаштувань за замовчуванням тощо.

Якщо розглядати існуючі технології виявлення вторгнень у сучасних інформаційно-комунікаційних мережах на більш високому рівні, у рамках технології SIEM для локальної мережі на рівні офісу, підприємства або технологічної мережі, то також можна виділити кілька основних категорій, які використовують різні методи виявлення. У найзагальнішому вигляді до них можна віднести:

– аналіз мережевого трафіку;
– аналіз вмісту пам'яті та жорстких накопичувачів;
– аналіз системних журналів.

Зазначимо, що системи виявлення вторгнень найчастіше будуються за гібридною моделлю, використовуючи кілька різних технологій.

Якщо розглядати ці схеми з точки зору виявлення джерела malware, то перша з цих технологій потребує серйозної підготовки та доступу до всіх ланок ієрархії ICN, через які проходить атака. При цьому аналіз мережного трафіку серйозно утруднений через фрагментацію ICN, яка у свою чергу пояснюється відставанням у впровадженні технологій IPv6, через що сучасний інфопростір є ієрархією підмереж з «сірими» IP-адресами, в якій існує багато рівнів

вкладеності і часто – підключення, що динамічно змінюються (найочевидніший приклад – смартфон, що перемикається між WiFi-мережами). Все це сильно ускладнює детектування та виявлення джерела атаки шляхом аналізу мережевого трафіку. Додатковими факторами, що погіршують, є все зростаючий обсяг переданих даних, низька грамотність як технічного персоналу, так і кінцевих користувачів, затримки з оновленням ОС і так далі.

Друга категорія технологій ґрунтується на аналізі вмісту пам'яті та жорстких накопичувачів. Найчастіше цю роль виконує антивірусний software різної спрямованості та методології (пошук по сигнатурах, пошук по евристикам). Цей метод також є досить повільним, особливо з урахуванням зростання обсягів інформації, що зберігається і передається, і обсягів самого software.

Одним з основних засобів як первинної, так і вторинної діагностики вторгнення є журнали подій (логи) програмного забезпечення та системні журнали (логи операційної системи) – набори діагностичних повідомлень, які генеруються штатно або нештатно функціонуючим програмним забезпеченням і можуть бути ефективно використані для виявлення факту вторгнення – з прийняттям подальших дій.

Логи часто є єдиним способом для роботи з розкриття інцидентів, їх аналіз може проводитися різними способами – як ручним, так і автоматизованим – шляхом підбору сигнатур та евристичних правил, однак застосовуваних не до аналізу бінарних файлів, а до аналізу повідомлень у логах.

Логи є цінним матеріалом для аналізу характеру атак, поведінки сторони, що атакує, та інформування у разі прориву захисту. Робота з логами та системними журналами є важливим компонентом безпеки системи, тому, якщо розгортається інтегрована система безпеки, до неї в обов'язковому порядку включається SIEM – Security Information and Event Management, який раніше являв собою поєднання SIM (Security Information Management) – системи управління інформацією про безпеку, та SEM (Security Event Management) – системи управління подіями безпеки.

SIEM забезпечує аналіз у реальному часі подій безпеки (тригерів), що надходять з різних пристроїв та додатків, і дозволяє завчасно реагувати на можливі загрози. Сучасні SIEM-системи можуть мати

досить складні та розгалужені архітектури та включати безліч різних джерел інформації. Однак, якщо серед великих корпорацій використання SIEM широко поширене, то для веб-серверів середнього рівня характерна, по-перше, відсутність будь-якої інтегрованої системи безпеки, по-друге, нехтування збором, захистом та обробкою логів.

Таким чином, аналіз системних журналів (логів) є одним з найбільш оперативних і доступних методів виявлення вторгнення malware. Логи є цінним матеріалом для аналізу характеру атак, поведінки атакуючої сторони та інформування у разі порушення захисту [25].

1.3.5. Журнали подій та виявлення вторгнень

Деякий час захист системних журналів не був пріоритетом при розробці систем безпеки – особливо серед системних адміністраторів середньої та нижчої ланки ICN, які часто не володіють ресурсами та знаннями для організації розвиненої інфраструктури безпеки, а іноді й поєднують роботу із забезпечення кібербезпеки з будь-якими іншими службовими обов'язками. Додатковою перешкодою була відсутність єдиних стандартів, тактики та політики захисту системних журналів. Впровадження комплексної інтегрованої системи, як правило, представляє складність і часто не є пріоритетом при розробці та експлуатації ICN. Відповідно, найчастіше експлуатація ICN поза великими корпораціями зводиться до набору hardening практик і «корисних порад», які далеко не завжди виконуються системними адміністраторами. Це вимагає організації інфраструктури, яка з одного боку забезпечувала б збирання, аналіз та зберігання системних журналів, з іншого боку дозволяла запобігти фальсифікації показань системних журналів і – з третьої сторони не була б ресурсозатратною як у розгортанні, так і в експлуатації. Останнє важливо, оскільки на будь-яких системах рівня нижче корпоративного, складність та витратність в експлуатації часто призводить до того, що рішення не запроваджується взагалі.

Адміністратори рідко беруть на себе працю організувати резервне копіювання, обробку та аналіз логів. При цьому резервна копія, навіть якщо вона реалізована, рідко перевіряється на цілісність та консистентність – чим цілком можна скористатися для маскування зламу системи шляхом підміни або спотворення логів, які можуть його виявити.

Причин такого нехтування декілька:

– не всі виділяють логи та системні журнали як значущий компонент безпеки системи (часто обмежуються встановленням брандмауера та антивірусу);

– логи займають великий обсяг (наприклад лог не найбільш навантаженого сервера за 8 днів становить понад 400 Мб);

– робота з логами вимагає витрат на організацію зовнішньої інфраструктури, при цьому, залежно від обсягу логів, така інфраструктура може виявитися досить дорогою та недостатньо ефективною;

– не склалося єдиного стандарту або практики роботи з логами;

– стандарт на повідомлення логів (Sigma – Generic Signature Format for SIEM Systems) ще не отримав загального визнання.

При цьому формат логів та їхня обробка часто залишається на розсуд організатора сервера – і йому не приділяється належної уваги.

Максимум захисту, який може здійснюватися в такому випадку, це:

– резервне копіювання логів;

– аналіз логів автоматичним аналізатором;

– аналіз логів за власними правилами.

Цього явно недостатньо – особливо з урахуванням зростання числа вторгнень та статистики щодо їхнього виявлення – навіть у великих корпораціях, що використовують SIEM. Показовою є поява нових типів шкідливого ПЗ, яке вже націлене на роботу з логами – хоча і досить примітивним чином.

На тлі всього вищезазначеного, логів як засіб виявлення malware поступово перетворюється на мету для проєктувальників malware, які чудово обізнані про ці технології і намагаються їх обійти.

У 2013 році корпорація MITRE анонсувала базу знань ATT&CK (Adversarial Tactics, Techniques & Common Knowledge – Тактики, техніки та загальновідомі знання про зловмисників) як спосіб опису та категоризації поведінки зловмисників, заснований на аналізі реальних атак. MITRE ATT&CK є структурованим списком відомих поведінок зловмисників, розділених на тактики та методи, і виражених у вигляді таблиць (матриць). Матриці для різних ситуацій та типів зловмисників публікуються на сайті MITRE. Також класифікація доступна в машиночитаних форматах STIX / TAXII. Оскільки

цей список дає комплексне уявлення про поведінку зловмисників під час зламу мереж, він вкрай корисний для різних захисних заходів, моніторингу, навчання та інших застосувань.

Відповідно до класифікації MITRE [33], атаки, пов'язані з модифікацією логів (Indicator Removal on Host: Clear Linux or Mac System Logs), належать до метакласу з ID T1070, який включає наступні техніки:

- T1070.001 Clear Windows Event Logs;
- T1070.002 Clear Linux or Mac System Logs;
- T1070.003 Clear Command History;
- T1070.004 File Deletion;
- T1070.005 Network Share Connection Removal;
- T1070.006 Timestomp.

Зокрема, туди входять техніки T1070.001 (Очищення журналів подій Windows) і техніка T1070.002 (Очищення системних журналів Linux або Mac). Ці дві техніки стосуються розглядуваного нами виду атак на систему повідомлень.

У 2020 році, в доповіді [34] ми припустили появу malware, яка маскуватиме факт вторгнення в систему шляхом заміни записів у системних журналах і запропонували технологію захисту, засновану на зв'язаних списках.

На той момент у базі даних MITRE було зареєстровано 2 різновиди malware, що використовують техніку T1070.001 (Очищення журналів подій Windows) та 1 різновид, що використовують техніку T1070.002 (Очищення системних журналів Linux або Mac).

У 2022-му році ця цифра склала 17 типів для T1070.001 і 3 типу для T1070.002.

На даний момент зареєстровано 28 типів, що використовують техніку T1070.001 та 4 різновиди, що використовують техніку T1070.002. Це тільки початок – зараз основною технологією маскування malware є просте видалення записів із системного журналу. Проте, з розвитком систем виявлення очікується появи malware, яке навчиться видаляти чи підробляти записи у системному журналі вибірково те щоб не видавати факт вторгнення відсутністю записів взагалі.

Розглянемо докладніше двох представників шкідливого ПЗ, які використовують техніку T1070.002.

1.3.6. Приклади шкідливого програмного забезпечення, орієнтованого на маскування дій та маніпуляції з логами

Перший з прикладів – це Proton. Він з'явився у 2017 році та використовував новий на той момент вектор зараження – коли зловмисник отримував доступ до реального сайту виробника програмного забезпечення та «троянізував» його додатки. Таким чином, зараження відбувалося під час встановлення цілком законного програмного забезпечення – при цьому не спрацьовував захисний периметр системи.

Показово, що поширення через злам сайту продемонструвало високий KPI, і програма на самому початку свого існування продавалася на чорному ринку за високими цінами: 100 BTC за необмежену кількість машин для зараження, потім 40 BTC та 2 BTC за одиничний комп'ютер, що навіть за мірками 2017 року становило високі суми порядку десятків та сотень тисяч доларів – і дає уявлення про попит на подібне ПЗ.

Троян пропонував широкий спектр дій на комп'ютері користувача – збір конфіденційної інформації, скріншоти робочого столу тощо.

Стирання логів для маскування дій троянця здійснювалося за допомогою команди, що видаляла системні логи і таким чином приховувала сліди дій на комп'ютері користувача:

```
sudo -k; echo '%@' | sudo -S rm -rf /var/log/* /Library/Logs/* && echo success;
```

Другий представник – криптомайнер Rocke – є характерним зразком нового покоління шкідливого програмного забезпечення.

Про групу зловмисників Rocke вперше стало відомо у 2018 році, коли в «медову пастку» (honeypot) потрапили перші зразки криптомайнера. Перші зразки були написані на Python, потім мову змінили на Go.

Для первинного впровадження Rocke використовує вразливості у веб-фреймворках Apache Struts 2, Oracle WebLogic та Adobe ColdFusion. Наприклад, у випадку Oracle використовується вразливість CVE-2017-10271, яка дозволяє отримати контроль над сервером та завантажити туди шкідливе програмне забезпечення.

Після впровадження запускається складний процес захоплення системи, що використовує мережеві сервіси paste.bin

(для завантаження додаткових шкідливих скриптів) та ідентифікатор ident.me для організації атак та розповсюдження на інші суміжні із захопленою системи.

Шкідлива програма виконує цілий набір дій:

- прописує себе в `crontab` так, щоб перезапускатися при кожному рестарті системи;
- ховає себе у списку процесів Linux за допомогою інструменту `libprocessehider`;
- завантажує запакований UPX майнер зі стороннього сайту. При цьому заголовок UPX спеціально модифікований, щоб зламати розпакувальник UPX. Замість "UPX!" рядок був замінений на "LSD!". Для розпакування зразків за допомогою розпакувальника, наданого командою UPX, необхідно вручну змінювати заголовки;
- зупиняє та деінсталює захисне програмне забезпечення Alibaba Cloud та Tencent Cloud – Alibaba Threat Detection Service та Tencent Cloud Host Security.

Цікава особливість шкідливої програми полягає в тому, що вона прагне позбутися інших криптомайнерів, оскільки за системні ресурси йде жорстка конкуренція. Для цього вона:

- модифікує дату та час шкідливих файлів так, щоб уникнути виявлення під час пошуку за метаданими;
- знаходить та знищує інші криптомайнери;
- змінює таблиці маршрутизації так, щоб заблокувати роботу інших криптомайнерів.

Нарешті, криптомайнер видаляє логи, надсилаючи нульове значення в наступні файли:

```
echo 0>/var/spool/mail/root
echo 0>/var/log/wtmp
echo 0>/var/log/secure
echo 0>/var/log/cron
```

Перші версії криптомайнера були написані мовою Python, згодом команда перейшла на використання дропера, написаного на Go. Системи виявлення на 2019 рік практично не помічали цей дропер.

Низький рівень виявлення шкідливого ПЗ у поєднанні з методами, що запобігають появі шкідливих процесів Rocke у запущених

процесах на машинах-жертвах, дозволив шкідливій програмі успішно розповсюджуватися протягом кількох тижнів.

Для перелічених вище випадків характерні наступні спільні моменти:

– зловмисники вже почали практикувати включення до шкідливого ПЗ інструментів маскування присутності шкідливих програм у системі, зокрема шляхом модифікації логів, які добре доповнюють інші інструменти маскування (timestomping, маскування процесів тощо);

– наявність навіть найпримітивнішої атаки – простого видалення логів – сильно знижує помітність наявності шкідливого ПЗ у системі, ускладнює виявлення та діагностику шкідливого ПЗ та дозволяє руткітам тривалий час функціонувати без виявлення – що, своєю чергою, збільшує масштаби зараження.

При цьому робота з логами та журналами системних подій перебуває на початковій стадії – як було зазначено вище, атака шляхом видалення логів спрацьовує лише за відсутності належної інфраструктури роботи з логами та журналами подій та низької грамотності системного адміністратора. За наявності елементарних засобів захисту, перелічених вище, віддалені логи будуть відновлені з резервної копії, а сам факт відсутності записів у системному журналі вказуватиме на наявність зовнішнього вторгнення та послужить приводом для вжиття всіх необхідних заходів щодо захисту системи.

Як аналіз логів з метою визначення вторгнення в систему, так і протидія йому все ще перебувають на початковому етапі розвитку. Нам видається, що прогрес у засобах виявлення потягне за собою появу шкідливого програмного забезпечення, яке матиме розвинений інструментарій роботи з логами та від простого знищення перейде до підробки логів з метою маскування своєї присутності в системі.

1.3.7. Забезпечення цілісності логів: від зв'язкових Списків до Verkle Trees

У роботах [35] було запропоновано таке рішення, що ґрунтувалося на технології зв'язкових списків. У цьому рішенні додатковий контроль за цілісністю системних журналів зводився до того, що в системний

журнал через деякі проміжки часу додавали спеціальне повідомлення, що представляє собою хеш-суму з повідомлень, що потрапили в системний журнал, включаючи передостаннє контрольне повідомлення. Таким чином, журнал з одного боку верифікувався наявністю ланцюжка контрольних сум, з іншого така схема не вимагала великих обчислювальних витрат, інфраструктури та гнучко налаштовувалась, оскільки інтервал можна було задавати як тимчасовим, так і за кількістю повідомлень у системному журналі. Така система добре показала себе на практиці на серверних системах середнього та низького рівня, проте, у зв'язку з широким поширенням хмарних систем, все гострішою стає необхідність в аналогічних заходах захисту, але вже дозволяють контролювати розподілені системи – де реалізація простої схеми зв'язкових списків не є можливою технічно.

У роботі [36] було запропоновано розвиток системи шляхом організації перехресних перевірок цілісності на базі технології дерев хешів, найвідомішою реалізацією яких є Merkle Tree. Використання Merkle tree виглядає наступним чином – на основі хешів вихідних записів вибудовується дерево хешів, в якому хеш верифікуються попарно, утворюючи собою ієрархію – дерево хеш, яке зі зростанням бази даних росте пропорційно `O(log N)`. Використання даної схеми роботи дозволить з одного боку зберегти всі переваги системи, заснованої на хеш-сумах (невимогливість до ресурсів, відсутність необхідності у побудові та обслуговуванні складної інфраструктури), з іншого, шляхом незначного збільшення надмірності дозволило отримати такі переваги, як можливість роботи з розподіленими ресурсами, що покращить загальну безпеку системи.

Подальшим розвитком ідеї стало використання Verkle Trees [37], в якому хеш-дерево скорочено до кількох рівнів, що серйозно впливає на швидкість та простоту логів інфраструктури перевірки. Крім інших переваг, Verkle Tree дозволяє гнучко використовувати алгоритми різної ефективності та роботи – від звичайного хешування до схеми поліноміального зобов'язання KZG [38]. В [36] використовувалася криптографічна хеш-функція SHA-256, щоб спростити реалізацію Verkle Tree у різних контекстах використання завдяки наявності інструментальної бази у вигляді бібліотек, що добре зарекомендували себе.

1.4. Адаптація систем безпеки до несприятливих умов функціонування ICN та ICS

Як було сказано вище, централізація збору системних журналів в даний час додатково утруднюється через все збоїв у мережі (блекаути, кібератаки, бойові дії і т. д.), що частішають, при яких мережа може частково або повністю втрачати зв'язність і розпадатися на незалежні сегменти з подальшим відновленням. Існуючі системи SIEM передбачають можливість практично миттєвого доступу до ключової інформації обсягом всього підконтрольного сегменту ICN. Короткочасна або тривала втрата взаємодії з підсистемами ICN може призвести до розладу та плутанини у існуючих системах контролю цілісності балок.

Отже, існує необхідність в інфраструктурі збору та верифікації системних журналів у рамках розподіленої ICN, яка не вимагала б складної інфраструктури та витрат на впровадження та дозволяла б функціонувати ICN в умовах часткової доступності її окремих «острівів» – напівавтономна або цілком автономна на деякі періоди часу. Існуючі рішення зазвичай зводяться до резервування та дублікації в тому чи іншому вигляді системних журналів по всій системі централізованим або централізованим чином [39; 40], що вимагає великих вкладень (наприклад, фактично подвоюється або потроюється необхідна ємність сховищ інформації) не забезпечує оперативного виявлення порушень цілісності системних журналів, оскільки у межах такої системи її можна виявити лише шляхом послідовної звірки записів оригіналу та копії. При цьому слід враховувати, що кібератаці або несприятливому впливу може бути схильний до будь-якого з рівнів системи.

У роботі пропонується схема асинхронної взаємодії систем верифікації логів. В рамках цієї системи низові ланки ICN працюють згідно зі схемою, наведеною в [35], де кожен журнал верифікується шляхом технології зв'язкових списків. Система розширюється за рахунок додавання до кожної контрольної суми метаданих: дати та внутрішнього часу обчислення дайджесту, кількості завірених повідомлень, а також ідентифікатора вузла ICN. Це є природним розширенням системи ведення логів і часто не вимагає перегляду схеми,

оскільки такі метадані додаються до запису автоматично самим програмним забезпеченням для ведення логів. Однак, замість регулярного збору інформації по системі, шляхом запитів від вищих вузлів ієрархії ICN, після кожного обчислення контрольної суми, низовий елемент виконує в термінології запропонованої системи дію PUSH, самостійно відправляючи запит «квитанцію» на рівень вище. Це може бути реалізовано як з підтвердженням отримання квитанції вищим вузлом, так і без неї. У свою чергу, кожен вищий вузол збирає отримані «квитанції». Після досягнення певної кількості квитанцій, він аналогічним чином обчислює свій дайджест, додає до нього метадані та надсилає отриману квитанцію ще на один рівень вище. Таким чином вибудовується ієрархічна система, подібна до Merkle Tree, яка дозволяє працювати асинхронно і без безперервного доступу до всіх мережевих сегментів ICN. У разі збою зв'язку повідомлення з конкретного вузла будуть просто надіслані пізніше.

При необхідності аудиту системи, вищестоящий вузол відправляється запит PULL, який нижчестоящі вузли дублюють ієрархії вниз. Після чого вузли, до яких прийшов PULL-запит знову виконують процедуру PUSH за всіма запитаними даними, а вузол, що надіслав запит PULL звіряє отримані повторно дайджести з централізовано. Така схема дозволяє з одного боку працювати в умовах тимчасових втрат зв'язку та зв'язності, з іншого боку за рахунок асинхронності дозволяє рівномірно розподіляти навантаження в ICN не перевантажуючи вузли-збирачі інформації великим обсягом обчислень.

1.5. Висновки

Аналіз поданих матеріалів дає змогу сформулювати низку ключових висновків щодо стану, вразливості, живучості та безпеки сучасних інформаційно-комунікаційних систем (ICN), зокрема веб-серверів.

Сучасні ICN, попри їхню складність та масштаб, апріорі вразливі до численних кібератак та несприятливих впливів. Тенденції до централізації та ієрархізації, часто без належного архітектурного планування, призвели до втрати гнучкості та здатності ефективно протистояти загрозам. У поточних умовах основним завданням

кібербезпеки ICN стає не стільки запобігання атакам (що є майже неможливим), скільки мінімізація часу відновлення функціональності після інциденту. Це підтверджується досвідом великомасштабних збоїв, таких як інцидент з «Київстар» у грудні 2023 року, коли повне відновлення зайняло близько семи днів. Веб-сервери є ключовими, детермінуючими компонентами ICN, що забезпечують безперебійне функціонування життєво важливих сервісів (від фінансових до комунальних). Їхній вихід з ладу призводить до каскадних відмов та мультиплікації збитків по всій мережі, як це було продемонстровано під час збоїв мобільних мереж, що вплинули на банківський сектор та комунальні служби. Події останніх років, включаючи блекаути та атаки на критичну інфраструктуру, підтвердили критичну важливість концепції живучості мережі. Це вимагає комплексного прогнозування та сценарного моделювання можливих збоїв та каскадного виходу з ладу систем, а також розробки моделей відновлення, що враховують вторинні ефекти та ризики.

Кількість, складність та інтелектуальність кібератак постійно зростають, зловмисники вдосконалюють свої техніки, тактики та стратегії. Руткіти та програми-вимагачі (наприклад, Stuxnet, Triton, WannaCry, Petya, Snake) еволюціонують від «традиційного» вандалізму до прихованої діяльності, спрямованої на збір конфіденційної інформації та організацію управління атаками. Традиційні захисні механізми, такі як «повітряний зазор», виявляються недостатніми для протидії сучасним загрозам, що підтверджено реальними інцидентами. Попри зусилля, середній час виявлення успішної кібератаки залишається надзвичайно великим (від 50–70 днів у 2015 році до 280 днів для ідентифікації витоків у 2020 році). Це дозволяє шкідливому програмному забезпеченню функціонувати непоміченим протягом тривалого часу, збільшуючи масштаби ураження та збитків. Класичні методи виявлення, такі як сигнатурний та евристичний аналіз бінарних файлів, хоч і цінні, мають значні недоліки: трудомісткість оновлення баз даних, вразливість до поліморфного та самомодифікованого ПЗ, а також високий рівень хибних спрацьовувань для евристик. Аналіз мережевої та системної активності також ускладнюється через використання шифрування та фрагментації мереж.

Завдяки своїй відкритості, постійному підключенню до зовнішніх мереж, повсюдному поширенню та ролі як інтегруючого елемента, веб-сервери стали основною мішенню для атак ботнетів та інших шкідливих програм. Їхня складна, багатокомпонентна архітектура та часто низька кваліфікація персоналу створюють численні вразливості. Журнали подій (логи) є основним засобом діагностики вторгнень та розкриття інцидентів. Проте, їхній величезний обсяг, відсутність стандартизації та недостатня увага до їхнього збору, захисту та обробки (особливо на середніх веб-серверах) є серйозною проблемою. Існуючі системи SIEM, попри свою ефективність, часто є надмірно складними та дорогими для широкого впровадження. З огляду на неминучість інцидентів, раціональна організація систем резервного копіювання (бекапів) є найважливішим елементом мінімізації часу відновлення працездатності ICN. Це стосується не лише даних, а й усієї інфраструктури веб-сервера (ОС, БД, ПЗ тощо). Для забезпечення стійкості та надійності систем резервного копіювання необхідно дотримуватися перевірених принципів: фізичне розділення робочої системи та сховища бекапів, шифрування резервних копій, обов'язкове тестове відновлення, автоматизація процесів резервування з документуванням у журналах, а також організація ротації кількох копій для ефективного використання ресурсів та збереження інформації.

Таким чином, для забезпечення живучості та стійкості сучасних ICN необхідний інтегрований, багатофакторний підхід. Він має включати не лише вдосконалення методів виявлення та протидії кіберзагрозам, а й розробку надійних стратегій відновлення, що базуються на ефективному резервному копіюванні, врахуванні вторинних ефектів збоїв та постійному аналізі еволюції загроз. Запропонована система контролю та управління розподіленими системами виявлення вторгнень з використанням логів у несприятливих умовах використовує асинхронний режим роботи та дозволяє з одного боку зберігати та підтримувати цілісність логів в умовах несприятливих впливів та вражаючих факторів, з іншого боку за рахунок асинхронності дозволяє рівномірно розподіляти навантаження в ICN, не перевантажуючи вузли-збирачі інформації великим обсягом обчислень. Впровадження та розгортання системи на запропонованих

принципах вимагає менше ресурсів у порівнянні з існуючими SIEM-системами і може бути здійснене як самостійно, так і з інтеграцією у вже розгорнуту SIEM, якщо вона є. Запропоновані заходи дозволять значно збільшити безпеку, оперативність реагування на інциденти та можливість відновлення існуючих ICN.

Список використаних джерел

1. Bi W., MacAskill K., Schooling J. Old wine in new bottles? Understanding infrastructure resilience: Foundations, assessment, and limitations. *Transportation Research Part D: Transport and Environment*. Elsevier BV, 2023. Vol. 120. P. 103–793. URL: http://dx.doi.org/10.1016/j.trd.2023.103793
2. Бойко В., Василенко М., Слатвінська В. Моделювання живучості та відновлення інформаційно-комунікаційних мереж в умовах дії кібер-загроз. *Інформаційні технології та суспільство*. 2024. № 1 (12). P. 13–19. URL: https://journals.maup.com.ua/index.php/it/article/view/3143
3. Олена К. Збій у роботі мережі Київстар – оновлення та рекомендації для абонентів. URL: https://biz.nv.ua/ukr/tech/zbiy-u-robo ti-merezhi-kijivstar-onovlennya-ta-rekomendaciji-dlya-abonentiv-50375659html (дата звернення: 2024–11–06).
4. Олена К. У Київстарі розповіли про наслідки кібератаки на 3,6 млрд грн. URL: https://biz.nv.ua/ukr/markets/u-kijivstari-rozpovili-pro-naslidki-kiberataki-na-3-6-mlrd-grn-50419968.html (дата звернення: 2024–11–06).
5. Інформація про Vodafone – Forbes Ukraine. URL: https://forbes.ua/profile/vodafone-251 (дата звернення: 2023–12–24).
6. Інформація про Lifecell – Forbes Ukraine. URL: https://forbes.ua/profile/lifecell-574 (дата звернення: 2023–12–24).
7. Інформація про Kyivstar – Forbes Ukraine. URL: https://forbes.ua/profile/kiivstar-244 (дата звернення: 2023–12–24).
8. Через несправність мобільного оператора Київстар відключення вуличного освітлення відбуваються в ручному режимі – Львівська міська рада. URL: https://city-adm.lviv.ua/news/city/housing-and-utilities/299531-cherez-nespravnistiu-mobilnoho-operatora-kyivstar-vid kliuchennia-vulychnoho-osvitlennia-vidbuvaiutsia-v-ruchnomu-rezhymi (дата звернення: 2023–12–24).

9. Boyko V., Vasilenko M., Slatvinska V. Survivability and sustainability of smart city information system components. *Municipal economy of cities*. O. M. Beketov National University of Urban Economy in Kharkiv, 2021. Vol. 6 (166). P. 20–27. URL: https://khg.kname.edu.ua/index.php/khg/article/view/5861

10. Barzashka I. Are Cyber-Weapons Effective? Assessing Stuxnet's Impact on the Iranian Enrichment Programme. *The RUSI Journal.* Taylor & Francis, 2013. Vol. 158(2). P. 48–56. URL: https://www.tandfonline.com/doi/abs/10.1080/03071847.2013.787735

11. Di Pinto A., Dragoni Y., Carcano A. TRITON: The first ICS cyber attack on safety instrument systems. *Proc. Black Hat USA*. 2018. P. 1–26.

12. Lee R. TRISIS Malware: Analysis of Safety System Targeted Malware. Dragos Inc. 2017. URL: https://www.dragos.com/wp-content/uploads/TRISIS-01.pdf

13. Case D. U. Analysis of the cyber attack on the Ukrainian power grid. *Electricity Information Sharing and Analysis Center (E-ISAC).* 2016. Vol. 388. URL: https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf

14. Slowik J. CRASHOVERRIDE: Reassessing the 2016 Ukraine electric power event as a protection-focused attack. Dragos, Washington, DC, USA, Tech. Rep., Aug. 2019. URL: https://www.dragos.com/wp-content/uploads/CRASHOVERRIDE.pdf

15. Fayi S. Y. A. What Petya / Not Petya ransomware is and what its remidiations are. Information Technology – New Generations. Springer, 2018. P. 93–100.

16. Slowik J. Evolution of ICS Attacks and the Prospects for Future Disruptive Events. Accessed: Feb, 2020. URL:https://www.dragos.com/wp-content/uploads/Evolution-of-ICS-Attacks-and-the-Prospects-for-Future-Disruptive-Events-Joseph-Slowik-1.pdf

17. Branquinho M. A. Ransomware in industrial control systems. what comes after wannacry and petya global attacks? *WIT Transactions on The Built Environment*. WIT Press, 2018. Vol. 174. P. 329–334.

18. Ransomware Against the Machine: How Adversaries are Learning to Disrupt. FireEye. 2020. URL: https://www.fireeye.com/blog/threat-research/2020/02/ransomware-against-machine-learning-to-disrupt-industrial-production.html

19. Mandiant. Mandiant Security Effectiveness Report. *FireEye.* 2020. P. 1–22. URL: https://www.fireeye.com/current-threats/annual-threat-report/security-effectiveness-report.html

20. *Бойко В. Д.* Моделювання кібернетичної складової живучості складних технічних систем. *Інформаційні управляючі системи та технології* : матеріали Міжнародної науково-практичної конференції (ІУСТ-Одеса-2014), 23–25 вересня 2014 року. Одеса, 2014. С. 239–241.

21. *Rausand M., Høyland A.* System Reliability Theory: Models, Statistical Methods, and Applications, Second Edition / 2nd ed. Wiley-Interscience, 2003. P. 636. URL: http://gen.lib.rus.ec/book/index.php?md5=5F7366820A70CEB471BD9BC7A29D4411

22. Cimellaro G. P., Reinhorn A. M., Bruneau M. Framework for analytical quantification of disaster resilience. Engineering Structures. Elsevier BV, 2010. Vol. 32 (11). P. 3639–3649. URL: http://dx.doi.org/10.1016/j.engstruct.2010.08.008

23. Sobhaninia S., Buckman S. T. Revisiting and adapting the Kates-Pijawka disaster recovery model: A reconfigured emphasis on anticipation, equity, and resilience. *International Journal of Disaster Risk Reduction. Elsevier BV.* 2022. Vol. 69. P. 102–738. URL: http://dx.doi.org/10.1016/j.ijdrr.2021.102738

24. Silva S. S. C., Silva R. M. P., Pinto R. C. G., Salles R. M. Botnets: A survey. Computer Networks. *Elsevier BV.* 2013. Vol. 57 (2). P. 378–403. URL: http://dx.doi.org/10.1016/j.comnet.2012.07.021

25. Barford P., Yegneswaran V. An Inside Look at Botnets. Advances in Information Security. *Springer US.* 2007. P. 171–191. URL: https://doi.org/10.1007%2F978-0-387-44599-1_8

26. *Blunden B.* The Rootkit arsenal: Escape and evasion in the dark corners of the system. Jones & Bartlett Publishers, 2012. P. 783.

27. Harley D., Lee A. The root of all evil? – Rootkits revealed. Online, 2007. P. 1–17.

28. Lee K., Lim J. The Reality and Response of Cyber Threats to Critical Infrastructure: A Case Study of the Cyber-terror Attack on the Korea Hydro & Nuclear Power Co., Ltd. KSII Transactions on Internet & Information Systems. 2016. Vol. 10 (2).

29. Johnson J. Average number of days to resolve a cyber attack on companies in the United States as of August 2015, by attack type.

URL: https://www.statista.com/statistics/193463/average-days-to-resolve-a-cyber-attack-in-us-companies-by-attack/ (дата звернення: 2022–06–10).

30. IBM. Cost of a Data Breach Report 2020. URL: https://www.ibm.com/security/digital-assets/cost-data-breach-report/#/ru (дата звернення: 2022–06–10).

31. Cinque M., Cotroneo D., Pecchia A. Challenges and Directions in Security Information and Event Management (SIEM) / 2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW). IEEE, 2018. P. 95–99. URL: https://doi.org/10.1109%2Fissrew.2018.00-24

32. González-Granadillo G., González-Zarzosa S., Diaz R. Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. Sensors. MDPI AG, 2021. Vol. 21 (14). P. 1–28. URL: https://doi.org/10.3390%2Fs21144759

33. Indicator Removal on Host: Clear Linux or Mac System Logs, Sub-technique T1070.002 – Enterprise MITRE ATT&CK. URL: https://attack.mitre.org/techniques/T1070/002 (дата звернення: 2022–06–10).

34. Бойко В., Василенко М. Система виявлення вторгненнь з використанням технології зв'язаних списків *Контроль і управління в складних системах (КУСС-2020)* : матеріали XV міжнародної конференції, м. Вінниця, 8–10 жовтня 2020 року, ВНТУ, 2020. С. 265–266. URL: http://ir.lib.vntu.edu.ua//handle/123456789/30577

35. Boyko V., Vasilenko M., Slatvinska V. Linked list systems for system logs protection from cyberattacks. *Information Technologies in Education, Science and Technology (ITEST-2022).* June 23–25, 2022 Cherkasy. 2022. P. 81–82. URL: https://er.chdtu.edu.ua/bitstream/ChSTU/4121/1/Збірник_тез_ІТОНТ-2022_макет_23_06.pdf#page=81

36. Boyko V., Vasilenko M., Slatvinska V. Linked List Systems for System Logs Protection from Cyberattacks. Information Technology for Education, Science, and Technics. ed. by Faure E., Danchenko O., Bondarenko M., Tryus Y., Bazilo C., Zaspa G. Springer Nature Switzerland, 2023. P. 224–234. URL: https://doi.org/10.1007%2F978-3-031-35467-0_15

37. Kuszmaul J. Verkle Trees. 2019. P. 11. URL: https://api.semanticscholar.org/CorpusID:218475793

38. Tas E. N., Boneh D. Vector Commitments with Efficient Updates. arXiv, 2023. P. 46. URL: https://arxiv.org/abs/2307.04085

39. Aßmuth A., Duncan R., Liebl S., Söllner M. A secure and privacy-friendly logging scheme. Cloud Computing 2021 / під ред. Bob Duncan, Yong Woo Lee, Manuela Popescu. International Academy, Research; Industry Association (IARIA), 2021. P. 8–12. URL: https://www.iaria.org/conferences2021/CLOUDCOMPUTING21.html.

40. Hangxia Z., Peng Z., Yong Y. Web log system of automatic backup and remote analysis. *International Conference on Computer Application and System Modeling (ICCASM 2010).* IEEE, 2010. P. 469–472. URL: https://doi.org/10.1109%2Ficcasm.2010.5620567.

Куклінова Тетяна Вікторівна
к.е.н., доцент, доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ІНТЕЛЕКТУАЛІЗАЦІЯ КІБЕРЗАХИСТУ ЕНЕРГЕТИЧНИХ ПІДПРИЄМСТВ: УПРАВЛІНСЬКИЙ ПІДХІД

DOI https://doi.org/10.36059/978-966-397-537-5-3

Сучасний етап цифрової трансформації суспільства супроводжується стрімким зростанням кіберзагроз, що особливо актуально для об'єктів критичної інфраструктури, зокрема енергетичного сектора. Енергетичні системи стають дедалі складнішими й водночас більш вразливими через активне впровадження цифрових технологій, систем автоматизації та дистанційного управління. Зловмисники дедалі частіше обирають енергетичні об'єкти як мішень для кібернападів, що може мати катастрофічні наслідки для економіки, національної безпеки та добробуту громадян. У зв'язку з цим зростає потреба у впровадженні інноваційних підходів до управління кіберзахистом енергетичних систем. Одним із найбільш перспективних рішень є використання технологій штучного інтелекту (ШІ), здатних забезпечити адаптивний, гнучкий та ефективний захист від динамічних кіберзагроз. Так, ШІ дає змогу в режимі реального часу аналізувати великі обсяги даних, виявляти аномалії, прогнозувати атаки та оперативно реагувати на кіберінциденти тощо.

У попередніх дослідженнях нами досліджувалися процеси системного впровадження ШІ в енергетичних компаніях, було надано визначення дефініції «системи ШІ», [1]. Також ми розглядали ключові переваги та інструменти ШІ в кібербезпеці енергогалузі, [2]. Незважаючи на зростаючу кількість досліджень у сфері кіберменеджменту таких вчених як Бойка В., Горбаченко С., Денисенко Г., Колодінського І., Федик В., Фроловой Я. та інших, існує потреба в системному науковому підході до інтеграціїі управління кібербезпекою в енергетиці з урахуванням можливостей ШІ.

Надійне енергопостачання є критично важливим для забезпечення національної безпеки. Вразливість енергетичної

інфраструктури призвела до серйозних соціально-економічних наслідків для України. Руйнування енергетичних об'єктів та фінансова нестабільність енергетичних компаній значно ускладнили ситуацію в енергетичному секторі України. Проблеми з енергетичними ресурсами спричиняють перебої в енергопостачанні, що, у свою чергу, негативно впливає на інші сектори економіки та населення. Розумні мережі дозволяють ефективно керувати енергетичними ресурсами, зменшують вплив на навколишнє середовище та підвищують енергетичну безпеку. Подальший розвиток у цій галузі вимагає синергії між урядом, бізнесом та академічними колами. Успішна цифровізація енергетичного сектору сприятиме створенню більш сталої, екологічно чистої та ефективної енергетичної системи у повоєнний період. Об'єктом дослідження є процеси управління кібербезпекою в енергетичному секторі.

Предметом дослідження виступають методи та моделі використання технологій ШІ для підвищення ефективності управління кіберзахистом енергетичних систем.

Метою дослідження є обґрунтування та розробка науково-методичних засад управління кіберзахистом енергетичного сектору з використанням технологій ШІ для підвищення ефективності виявлення, реагування та протидії сучасним кіберзагрозам.

Завдання дослідження:

1. Проаналізувати сучасний стан енергетичного сектора України.
2. Дослідити стан впровадження ШІ в енергомережі України.
3. Визначити основні вимоги до ефективної системи управління кібербезпекою в енергетиці.
4. Розробити шляхи управління кіберзахистом із застосуванням ШІ.

Енергетичні підприємства відіграють ключову роль в економічному розвитку будь-якої країни. Вони забезпечують безперебійне постачання електроенергії, тепла та палива, що є основою функціонування всіх галузей економіки. Без ефективної енергетичної системи неможливе стабільне виробництво, функціонування транспорт, зв'язку, сфери охорони здоров'я та інші сфери. Саме енергетика формує базу для індустріального, аграрного та цифрового секторів. Рівень розвитку енергетичних підприємств прямо впливає на конкурентоспроможність країни на світовому ринку.

Енергетичні підприємства належать до критичної інфраструктури держави. Їх стабільна робота є необхідною умовою життєдіяльності населення та функціонування державних інституцій. Порушення в роботі електростанцій, підстанцій чи газових магістралей можуть призвести до масштабних соціальних і економічних наслідків. Саме тому захист енергетичної інфраструктури є пріоритетом національної політики. Кібербезпека енергетичних підприємств – це не лише технологічне, а й стратегічне завдання. Використання цифрових технологій у енергетиці підвищує ефективність, але одночасно створює нові ризики. Тому енергетичний сектор потребує надійної системи кіберзахисту. Інвестиції в безпеку енергетичних підприємств – це інвестиції в стабільність України.

Крім того, енергетика забезпечує енергетичну незалежність та зміцнює національну безпеку. На тлі війни та гібридних загроз енергетичні об'єкти стали однією з головних мішеней для атак.

Таблиця 1 – Найпотужніші кібератаки на енергооб'єкти за період 2011–2024 р. [2]

Рік	Назва атаки	Ціль	Мета атаки
2011–2014	Dragonfly / Energetic Bear	Енергетичні компанії по всьому світу	Серія кібератак, зосереджених на тактиках, мотивах і наслідках для інтелектуальних мереж.
2017	NotPetya Ransomware	Енергетичні компанії по всьому світу	Порушення роботи систем енергоменеджменту, зв'язку та білінгу.
2016–2018	Industroyer / CrashOverride	Israel Electric Corporation	Серія атак на інфраструктуру електромереж, що спричинила масштабні збої.
2022–2024	Sandworm / Industroyer2	ДТЕК, енергетичні системи України	Багатоепізодні кібератаки з використанням нових методів впливу на системи керування й ОТ.

Кібератаки часто здійснюються злочинними угрупованнями, хактивістами або акторами, пов'язаними з державами, які мають конкретні цілі, такі як викрадення інтелектуальної власності, порушення роботи критичної інфраструктури або ведення шпигунської діяльності. Ці атаки можуть бути особливо руйнівними, оскільки вони спеціально адаптовані до конкретної цілі та часто залишаються непоміченими протягом тривалого часу. Починаючи з 2011 року, 84 % жертв, які постраждали від атак Dragonfly, належали до енергетичного та електроенергетичного сектору. Це свідчить про те, що більшість зусиль цієї групи були спрямовані на компрометацію критично важливої енергетичної інфраструктури, створюючи серйозні загрози для безпеки та стабільності енергосистем [3].

Так, Dragonfly (Energetic Bear) є однією з наймасштабніших і найтриваліших кібершпигунських операцій, спрямованих проти енергетичного сектору. Вперше активність цієї хакерської групи була зафіксована у 2011 році, з піковою інтенсивністю у період 2011–2014 рр. Основною метою атак були енергетичні компанії у США, Канаді та країнах Європи, з особливим акцентом на постачальників енергетичних технологій та операторів критичної інфраструктури. Так, група Dragonfly використовувала широкий спектр кіберзасобів, зокрема фішингові листи, шкідливі вебсайти, компрометацію оновлень програмного забезпечення та бекдори, вбудовані в легітимні програми. Однією з ключових особливостей було використання легальних інструментів адміністрування, що дозволяло хакерам залишатися непоміченими протягом тривалого часу.

Промисловість та ЖКГ споживають близько 80 % всіх енергоресурсів в країні і мають одні з найнижчих показників енергоефективності, а відповідно і найвищий потенціал її підвищення. Що також має ряд природоохоронних (екологічних) та соціальних ефектів. До екологічних можна віднести: зменшення спожитої енергії; зменшення викиду шкідливих речовин в навколишнє середовище; покращення загальної екологічної ситуації в регіоні. Серед соціальних впливів: можливість збільшення заробітних плат працівникам; збереження старих і створення нових робочих місць; покращення здоров'я та безпеки працівників; покращення корпоративної культури й зміцнення корпоративних цінностей.

Забезпечення стійкості енергетичної системи є необхідною умовою для швидкого повоєнного відновлення. Саме через енергетичну інфраструктуру реалізуються реформи зеленої трансформації та переходу до відновлюваних джерел енергії. Енергетичні підприємства – це також значне джерело бюджетних надходжень та зайнятості. Збереження та модернізація енергетичного комплексу – це фундамент для довгострокового економічного зростання.

Зазначимо, що виробниками електроенергії в Україні є 4 атомні електростанції, 15 теплових електростанцій (2 залишилися на неконтрольованій території), 43 теплові електростанції (10 розташовані на неконтрольованій території). Основою гідроенергетики України є каскад із 6 великих гідроелектростанцій на Дніпрі, а також Ташлицька гідроакумулююча електростанція на річці Південний Буг. Загалом працює 8 гідроелектростанцій та 3 гідроакумулюючі електростанції. Найбільшими приватними енергетичними компаніями в Україні є:

– ДТЕК (80 % теплової генерації країни) та інфраструктура збуту енергії та енергосервісна компанія.

– VS Energy – друга за величиною приватна енергетична компанія в Україні, яка контролює близько 1/3 ринку передачі енергії через локальні мережі в країні. Вона контролює «Київобленерго», «Рівнеобленерго», «Херсонобленерго», «Кіровоградобленерго», «Житомирobленерго», «Одесаобленерго», «Чернівціобленерго», «Севастопольенерго», група також має частку в «Ніколаєвобленерго» та «Хмельницькобленерго».

– «Укренергоконсалтинг» – холдингова компанія, яка контролює енергопостачальні компанії «Львівобленерго» та «Прикарпаттяобленерго».

Внаслідок російського вторгнення енергетичний сектор зазнав прямих збитків та непрямих фінансових збитків у розмірі 56,5 млрд доларів. Потреби у відновленні, що включають повну реконструкцію зруйнованих об’єктів за принципом «Відбудувати краще, ніж було», становлять 50,5 млрд доларів. Прямі збитки енергетичного сектору України станом на травень 2024 року становлять понад 16,1 млрд доларів. Найбільші збитки завдано руйнуванням об’єктів виробництва електроенергії (8,5 млрд доларів),

магістральних ліній електропередачі (2,1 млрд доларів), а також нафтогазової інфраструктури (3,3 млрд доларів) [4]. Мохор В. підкреслює, що з початком активних бойових дій енергетичний сектор України став об'єктом безперервних ракетних та артилерійських атак з боку російського агресора та станом на 30 квітня 2023 р. в секторі електроенергетики України загальна генеруюча потужність електростанцій зменшилася з 36 до 18,3 ГВт (на 51 %) [5, с. 75]. Під час нещодавніх обстрілів Україна втратила приблизно 25 % від загального споживання електроенергії. Основна маневрена генерація, включаючи гідро- та теплоелектростанції, втратила близько 80 % своєї потужності, тоді як атомна енергетика не в змозі оперативно задовольнити збільшений попит на електроенергію [6].

Загальновідомо, що процес перетворення енергії на електроенергію неминуче супроводжується викидами CO^2, особливо коли використовуються традиційні викопні джерела. Надмірна концентрація вуглекислого газу в атмосфері порушує природний баланс і сприяє глобальним кліматичним змінам. У той час як попит на електроенергію у світі постійно зростає, зростає і потреба в екологічно чистих, сталих джерелах її виробництва. Саме тому розвиток «зеленої» енергетики – на основі сонця, вітру, води та біомаси – стає критично важливим. Впровадження відновлюваних джерел енергії дозволяє не лише зменшити викиди CO^2, а й забезпечити довгострокову енергетичну безпеку, економічну ефективність і збереження довкілля для майбутніх поколінь.

За останні роки кількість промислових сонячних та вітрових електростанцій значно зросла. Більшість із них розташовані в південних регіонах України. Активний розвиток відновлюваної енергетики в Україні розпочався після прийняття «зеленого тарифу» у 2008 році, згідно з яким електроенергія, вироблена з енергії сонця, вітру тощо, закуповується в системі за ціною, що значно вища за ринкову. Чинна програма стимулювання розвитку відновлюваної енергетики розрахована до 2030 року. Водночас передбачається поступове зниження вартості «зеленої» електроенергії. Виробництво електроенергії з альтернативних джерел має зрости з 10 у 2023 році до 25–30 % до 2030 року [7].

В енергетичному секторі існують бар'єри, які уповільнюють впровадження змін. Один із головних – консервативне ставлення до нових

технологій, що часто зумовлене побоюваннями щодо їхньої ефективності та впливу на чинні бізнес-процеси. Багато енергокомпаній усе ще не завершили цифрову трансформацію, що потребує додаткових фінансових ресурсів та оновлення управлінських структур. Також значною проблемою залишається нестача кваліфікованих кадрів і залежність від застарілого обладнання, що ускладнює інтеграцію інновацій.

Кібербезпека і цифрова сталість енергогалузі взаємопов'язані, адже захист цифрових систем є основою їхньої стійкості. Проте на шляху впровадження технологій виникають труднощі, зокрема висока вартість, потреба у навчанні персоналу, стандартизації даних і забезпеченні конфіденційності.

Подолати цифрову інертність можна шляхом посилення внутрішньої комунікації, інвестування в підвищення кваліфікації працівників та активної підтримки з боку управлінців. Формування цифрової стійкості є необхідною умовою для модернізації енергетики та досягнення цілей сталого розвитку. Технології Smart Grid відкривають нові можливості для ефективного управління енергоресурсами, зменшення шкідливого впливу на довкілля та підвищення надійності енергосистем. Успішна цифровізація енергетичної галузі можлива лише за умов тісної взаємодії між державою, бізнесом та науковою спільнотою, що дозволить створити інноваційну, стійку та безпечну енергосистему. Атаки на Smart Grid можуть мати катастрофічні наслідки – від локальних збоїв до масштабних блекаутів.

Значна частина енергетичної інфраструктури України була збудована ще в радянський період і сьогодні перебуває у критичному стані, вимагаючи термінової модернізації. Застаріле обладнання функціонує з низькою ефективністю, спричиняючи суттєві втрати енергії та знижуючи надійність постачання. Фокусування на розвитку енергоефективного потенціалу сприятиме раціоналізації використання енергетичних ресурсів в Україні, зміцненню енергетичної безпеки, динамізації економічного зростання та формуванню передумов для сталого екологічного розвитку держави. В умовах постійного зростання вартості енергоносіїв і нестабільності їх постачання, раціональне та ефективне використання енергії стає визначальним чинником забезпечення довгострокової економічної

стабільності, стійкості та динамічного розвитку бізнес-структур. Низька продуктивність виробництва часто обумовлена неефективністю виробничих процесів, що значною мірою зумовлено нераціональним використанням енергетичних ресурсів. Оптимізація енергоспоживання, зокрема через впровадження сучасних технологічних рішень, зокрема ШІ, сприяє зростанню продуктивності праці та відповідному збільшенню доходів підприємства. Скорочення енергоспоживання позитивно впливає не лише на економічні показники, а й на екологічний стан довкілля, водночас покращуючи репутаційний капітал підприємства. У зв'язку з цим, впровадження енергоефективних технологій із використанням ШІ на підприємствах, в установах та організаціях є актуальним і стратегічно важливим кроком для сталого розвитку.

В умовах сучасних викликів головним завданням для України є зміцнення енергетичної безпеки та підвищення стійкості енергосистеми. Оскільки Україна інтегрована в загальноєвропейську енергетичну мережу, модернізація її енергетичного сектору має здійснюватися з урахуванням європейських стандартів і вимог.

Один із ключових напрямів трансформації української – цифровізація енергетичних мереж на базі децентралізації. Вона відкриває нові можливості для підвищення ефективності управління системами, зменшення втрат, швидкого реагування на аварійні ситуації та інтеграції відновлюваних джерел енергії. Цифрова модернізація є основою для створення гнучкої, розумної та екологічно орієнтованої енергетичної інфраструктури майбутнього.

На відміну від застарілих систем, Smart Grid інтегрує інформаційно-комунікаційні технології, а також рішення для збору й аналізу даних про виробництво, передачу та споживання електроенергії. Це дозволяє ефективно контролювати та управляти всіма рівнями енергетичної системи в режимі реального часу. Сучасне розумне місто неможливе без Smart Grid. Ця інфраструктура не лише підвищує енергоефективність та надійність постачання, а й створює умови для інноваційного розвитку, зменшення шкідливих викидів та підвищення безпеки. Її ключовим компонентом є інтелектуальні лічильники, які передають точні дані про споживання електроенергії майже в реальному часі. Завдяки цьому споживачі можуть:

– планувати використання енергії в залежності від тарифів (наприклад, уночі, коли електроенергія дешевша);

– оптимізувати витрати, використовуючи великі прилади – як-от пральні або посудомийні машини – у позапікові години;

– керувати зарядкою електромобіля, щоб зменшити навантаження на мережу та зекономити.

Для енергетичних компаній, операторів розподільчих мереж (колишніх обленерго) та на рівні держави впровадження Smart Grid відкриває нові можливості: зокрема, підвищується надійність енергопостачання завдяки своєчасному виявленню або попередженню аварій, покращується балансування навантажень, що дозволяє уникати відключень та втрат, забезпечується цифрове управління мережею зі зниженими експлуатаційними витратами, а також зростає загальна енергоефективність та екологічність енергосистеми.

В Україні робота з впровадження нових технологій в національну енергосистему розпочалася відносно недавно. Так, з 2014 року бельгійська компанія Tractebel розробляє та впроваджує низку пілотних технологій та проєктів Smart Grid на рівні системного оператора – НЕК «Укренерго». Оператори систем розподілу також поступово намагаються впроваджувати елементи розумних електричних мереж.

Товариство з впровадження технологій найбільш активно впроваджує ТОВ «ДТЕК Енерго». Компанія активно встановлює розумні лічильники та автоматизує систему електропостачання за допомогою сучасного програмного забезпечення. Це допоможе їй швидше «бачити» аварії та, як наслідок, набагато швидше їх усувати.

У 2019 році ТОВ «ДТЕК Енерго» почало цифровізацію своїх електростанцій, збільшуючи впровадження ШІ в управління тепловими електростанціями. SCADA-системи (Supervisory Control and Data Acquisition – диспетчерське управління та збір даних) використовуються ДТЕК для моніторингу та управління складними інфраструктурними або промисловими об'єктами в режимі реального часу. Вона допомагає диспетчерам бачити всю енергетичну систему області в режимі онлайн, контролювати режим її роботи та краще оцінювати критичну або аварійну ситуації. В разі аварії саме SCADA автоматично надсилатиме у чат-бот та на сайт повідомлення про відключення та повернення світла, то ж ви дізнаватиметесь про це за лічені

секунди. Ця система є частиною розвитку «розумних» електромереж, які мають забезпечити безперебійне електропостачання.

Але однією з головних вразливостей Smart Grid є SCADA-системи, які забезпечують управління та моніторинг об'єктів енергетичної інфраструктури. Хакери можуть отримати доступ до цих систем через несанкціонований віддалений доступ, фішинг або шкідливе програмне забезпечення.

MODUS X – це дочірня IT-сервісна компанія, виділена з IT-департаменту групи ДТЕК у 2022 році, для масштабного впровадження цифрових та AI-рішень как усередині холдингу, так і на ринку загалом. Вона має понад 500 інженерів, архітекторів, фахівців з кібербезпеки та дата-сайєнтистів, працює за принципом data-driven управління.

MODUS X реалізувала широкий спектр рішень:

– цифровий двійник шахт і енергоблоків;
– автоматизація сервісів (документообіг, HR, мобільні додатки);
– оптимізація логістики й прогнозування технічного стану з AI-алгоритмами;
– системи цифрової безпеки для енергетичної інфраструктури.

Також використовуючи моделі машинного навчання, MODUS дозволяє:

– виявляти складні та приховані атаки на IT- та OT-інфраструктуру;
– аналізувати поведінкові патерни користувачів і пристроїв;
– попереджати інциденти за допомогою предиктивної аналітики.

Так, у межах цифрової трансформації енергетичного сектору компанія MODUS X спільно з групою ДТЕК реалізує комплексні рішення на основі ШІ, машинного навчання та інструментів Microsoft Copilot Studio. Основна мета співпраці – забезпечення сталого розвитку та підвищення ефективності бізнес-процесів завдяки глибокій аналітиці, автоматизації та створенню інноваційного середовища. Microsoft Copilot – інструмент, що допомагає автоматизувати типові дії

Центральним елементом трансформації є перехід до data-driven моделі управління, де дані стають основою прийняття рішень. Такий підхід дозволяє не лише підвищити оперативність і точність

управлінських рішень, але й забезпечити гнучкість та адаптивність енергетичної компанії до викликів ринку.

Серед стратегічних цілей проєкту:

– Оптимізація процесів: використання алгоритмів машинного навчання для прогнозування, підтримки прийняття рішень, автоматизації рутинних задач і зниження операційних витрат.

– Масштабування експертизи: створення умов, за яких як технічні, так і бізнес-підрозділи можуть ефективно використовувати ШІ-інструменти у своїй діяльності.

– Формування культури data-driven: впровадження практик, коли аналітика та цифрові інструменти є обов'язковим елементом кожного етапу життєвого циклу бізнесу.

– Інституціалізація інноваційної спільноти: створення внутрішніх екосистем для обміну знаннями, менторства та підвищення цифрової зрілості персоналу.

Реалізація таких ініціатив потребує системного підходу: інтеграції гнучких цифрових технологій, розбудови внутрішньої експертизи та наявності стратегічного бачення розвитку. Таким чином ШІ розглядається не лише як технологічний інструмент, а як стратегічний ресурс, що здатен підвищити конкурентоспроможність, стійкість і інноваційність компанії [8].

Компанія «Українські розподільні мережі» здійснює корпоративне управління державними частками у семи обласних енергетичних компаніях, зокрема таких як Хмельницькобленерго, Миколаївобленерго, Запоріжжяобленерго та Тернопільобленерго. Ця компанія працює над 12 пілотними проєктами з іноземними інвестиціями для розвитку концепції Smart Grid серед компаній групи. Група телемеханізувала 65 % усіх високовольтних підстанцій та встановила 12 % розумних медичних пристроїв.

У жовтні 2022 року Кабінет Міністрів України затвердив Концепцію впровадження Smart Grid до 2035 року та ухвалив відповідний план заходів для її реалізації.

Ключова мета цього документа – зменшити втрати електроенергії в національних електромережах з 11,6 % до 7,5 %, що відповідає економії близько 6 млрд кВт·год електроенергії на рік. Такий обсяг еквівалентний спаленню приблизно 3 мільйонів тонн вугілля

на теплових електростанціях, що матиме значний екологічний та економічний ефект [9]. Впровадження Smart Grid розглядається як невід'ємна частина процесу відновлення та модернізації енергетичної інфраструктури України, яка зазнала масштабних руйнувань унаслідок російської військової агресії.

Повільне впровадження «розумних мереж» в Україні пов'язане, перш за все, з недостатнім фінансуванням. Регіональні компанії, що відповідають за експлуатацію, обслуговування та розвиток розподільчих мереж, мають обмежені бюджети. В середньому їхні інвестиційні програми становлять лише 300 мільйонів гривень. На державному та експертному рівнях обговорюються різні варіанти інвестування. Це включає міжнародне донорське фінансування конкретних проєктів, кредити європейських та світових банків, а також системи тарифів на основі стимулювання [10].

На нашу думку, повоєнна Україна повинна модернізувати свою енергетичну інфраструктуру на основі передових технологій, включаючи системи розумних мереж (Smart Grid), щоб досягти енергетичної незалежності та інтегрувати відновлювані джерела енергії. Інтеграція технологій ШІ, Інтернету речей та великих даних покращує прогнозування попиту, підвищує ефективність використання відновлюваної енергії та зменшує втрати в енергомережах. Крім того, цифрова стійкість енергетичного сектору є основоположною для зміцнення енергетичної безпеки та сприяння розвитку розумних міст.

Сьогодні основним джерелом інвестицій для будівництва Smart Grid є інвестиційні програми компаній, чого недостатньо для побудови Smart Grid. Процес впровадження ШІ в енергетичних компаніях часто супроводжується проблемою вибору розробника та інтегратора цифрових рішень і продуктів, а іноді, навпаки, ускладнюється відсутністю необхідних пропозицій на ринку.

На нашу думку, впровадження Smart Grid в Україні зіткнеться з перешкодами:

1. Будівництво та модернізація інфраструктури вимагають значних інвестицій.

2. Поєднання нових технологій із застарілими енергетичними системами може бути складним завданням.

3. Сонячна та вітрова енергетика залежать від погодних умов, що вимагає рішень для балансування навантаження та накопичення енергії.

Розвиток технології Smart Grid дозволить оптимізувати виробництво, передачу та використання електроенергії, що стимулюється для зменшення споживання енергії. Оскільки впровадження енергоефективності зменшить втрати енергії, необхідно також стимулювати інших виробників електроенергії, включаючи іноземних, до підтримки монополій. Зокрема, впровадження технологій «мікромереж» зможе підвищити стійкість української системи до військових та стихійних лих.

Україні потрібні заходи щодо стимулювання створення малих енергетичних компаній з метою побудови цивілізованого європейського ринку електроенергії з привабливою нормою прибутку для компаній, забезпечення доступу до дешевих довгострокових банківських кредитів, збалансування та оптимізації структури вітчизняної генерації в напрямку відновлюваної електроенергії. Проблеми декарбонізації пов'язані з передачею електроенергії та необхідністю створення нових ліній, технології дозволяють розширити потужність існуючих мереж для сонячної та вітрової енергії.

Децентралізація енергосистеми є ключовим напрямом трансформації енергетичного сектору України в умовах післявоєнного відновлення та стійкого розвитку. Вона передбачає перехід від централізованої моделі генерації до гнучкої мережі малих і середніх виробників, інтегрованих у єдину цифрову інфраструктуру. Такий підхід дозволяє зменшити залежність від великих централізованих об'єктів, які є вразливими до кібератак та військових дій, а також посилити енергетичну стійкість громад і регіонів.

Особливої актуальності набуває впровадження Smart Grid як інтелектуальної енергосистеми нового покоління, яка забезпечує двосторонній обмін енергією між споживачами та генераторами, враховує динамічні потреби ринку й дозволяє в режимі реального часу здійснювати балансування навантаження, моніторинг споживання та виявлення відхилень.

Децентралізовані системи на базі ВДЕ (сонячної, вітрової енергії, біомаси) у поєднанні з накопичувачами енергії та енергетичними

кооперативами дають змогу не лише забезпечити енергетичну автономію громад, а й активізувати локальний бізнес, створити нові робочі місця та підвищити інвестиційну привабливість регіонів.

Крім того, інтеграція ШІ у в управління децентралізованими мережами дозволяє автоматизувати процеси оптимізації виробництва та споживання, прогнозування аварій, і тим самим – мінімізувати технічні втрати та забезпечити надійність системи.

Таким чином, децентралізація енергосистеми України, з опорою на Smart Grid, цифрові технології та інноваційне управління, не лише підвищить енергетичну безпеку держави, а й сприятиме її інтеграції до європейського енергетичного простору на засадах сталості, відкритості та інновацій.

Технології ШІ в Smart Grid ефективні у використанні, але їх поширення створює серйозні виклики для енергетичної галузі. Системи ШІ сьогодні виступають не лише технологічним рушієм, а й стратегічним партнером у досягненні цілей сталого розвитку. Їхнє впровадження в різні сфери суспільного життя сприяє створенню більш ефективних і екологічно безпечних рішень. Водночас автоматизація, що базується на ШІ, здатна змінити структуру ринку праці. Зростання технологічного безробіття стає однією з ключових проблем, з якими стикається індустрія інформаційних технологій та цифрових послуг.

Попри побоювання щодо конкуренції між ШІ та людською працею, основне призначення таких систем – не замінити людину, а підсилити її можливості. ШІ ефективно виконує повторювані, аналітичні завдання, тоді як людина зосереджується на креативних, стратегічних і емоційно забарвлених аспектах роботи. Співпраця з інтелектуальними системами дозволяє фахівцям розвивати унікальні навички, підвищуючи продуктивність і задоволення від професійної діяльності. Крім того, розвиток ШІ відкриває нові можливості працевлаштування в галузях, пов'язаних з його розробкою, впровадженням і управлінням.

Також інтеграція ШІ в робочі процеси сприяє переосмисленню ролі працівника у цифровій економіці. Людина стає не просто виконавцем, а куратором інтелектуальних систем, аналітиком даних, стратегом інновацій. Це зумовлює потребу в нових компетенціях, таких як цифрова грамотність, навички роботи з великими даними, етичне управління ШІ та кросдисциплінарне мислення.

В енергетичному секторі, впровадження ШІ дозволяє зменшити ризики, пов'язані з людським фактором, підвищити безпеку об'єктів критичної інфраструктури, а також скоротити витрати за рахунок автоматизації складних технічних процесів. Проте це не усуває потреби в людському контролі – навпаки, зростає роль фахівців, здатних забезпечити прозорість, надійність і соціальну відповідальність у взаємодії «людина – машина».

Більше того, розвиток етичних стандартів і нормативної бази у сфері ШІ є критично важливим для запобігання дискримінації, помилок систем і зловживань. Саме людина повинна задавати цінності, пріоритети та межі використання ШІ, щоб гарантувати, що технології працюють на благо суспільства.

Таким чином, ШІ не витісняє людину, а виводить її роль на якісно новий рівень, де креативність, гнучкість, моральний вибір і стратегічне мислення залишаються незамінними. Це відкриває перед суспільством унікальні можливості для трансформації ринку праці, освіти та моделей управління, спрямованих на сталий і людяний розвиток.

ШІ вже відіграє важливу роль у трансформації економічних, соціальних та екологічних процесів, наближаючи сталий розвиток. У післявоєнної відбудови Україна має унікальний шанс модернізувати енергетичну інфраструктуру, орієнтуючись на інноваційні технології, зокрема ті, що базуються на ШІ. Це робить інтелектуальні системи незамінним інструментом для підприємств, які прагнуть залишатися гнучкими, інноваційними й конкурентоспроможними у динамічному глобальному середовищі.

Крім того, значну загрозу становлять внутрішні загрози – як ненавмисні, так і навмисні дії співробітників. Нерідко колишні працівники зберігають доступ до критичних систем після звільнення, що створює додаткові ризики. Низький рівень обізнаності щодо інформаційної безпеки серед співробітників, відсутність регулярних тренінгів та навчання посилюють проблему. Часто працівники не усвідомлюють важливості кібербезпеки або сприймають вимоги безпеки як зайву формальність. В умовах перевантаження, стресу або рутини знижується уважність, імовірність допущення помилок зростає.

В енергетичному секторі, де системи управління мають пряме відношення до об'єктів критичної інфраструктури, наслідки людських помилок можуть бути особливо серйозними. Через недбалість або неуважність персоналу зловмисники можуть отримати контроль над процесами, які впливають на енергозабезпечення регіонів. Часто керівництво підприємств недооцінює важливість інвестування в розвиток культури кібербезпеки. Замість впровадження ефективних програм навчання, відповідальність перекладається виключно на ІТ-відділ, хоча кібербезпека – це спільна відповідальність усього колективу.

Надзвичайно важливо створити на підприємстві середовище, де кожен працівник усвідомлює свою роль у забезпеченні безпеки. Впровадження автоматизованих систем контролю дій користувачів, регулярне оновлення паролів, обмеження доступу, системи виявлення аномалій – усе це може зменшити ризики, пов'язані з людським фактором. Проте навіть найсучасніше програмне забезпечення не буде ефективним без належної поведінки з боку персоналу. Навчання, симуляції атак, регулярні інструктажі, інформаційні кампанії – усе це має стати частиною щоденної роботи.

Розвиток водневої енергетики посідає провідне місце в глобальній стратегії переходу до відновлюваних джерел енергії та формування низьковуглецевої економіки. Інтеграція ШІ у цю сферу відкриває нові перспективи для оптимізації процесів виробництва, зберігання та споживання водню. В умовах зростаючих викликів, пов'язаних з енергетичною безпекою, необхідністю зменшення викидів парникових газів і пошуком стійких енергетичних рішень, роль інновацій стає визначальною.

Зелений водень, що виробляється з використанням відновлюваних джерел енергії, є екологічно чистим видом палива – при його спалюванні утворюється лише вода. Це робить його важливим інструментом у процесі декарбонізації, особливо в секторах, де електрифікація є складною або економічно невигідною, таких як важка промисловість, авіація чи морські перевезення. ШІ відіграє ключову роль в оптимізації процесів у водневій енергетиці, зокрема під час електролізу, знижуючи енергоспоживання та підвищуючи загальну ефективність виробництва. Системи на основі машинного навчання

дозволяють динамічно налаштовувати параметри роботи обладнання в реальному часі відповідно до змін зовнішніх умов, що підвищує гнучкість і стабільність процесів.

Крім того, ШІ використовується для аналізу ринкової ситуації та прогнозування попиту на водень, що дозволяє більш точно планувати обсяги виробництва, логістику та зберігання ресурсу. Завдяки інтелектуальній аналітиці можливе своєчасне виявлення технічних аномалій і запобігання поломкам, що знижує ризики аварій і продовжує термін служби обладнання.

Також алгоритми ШІ сприяють балансуванню енергетичних систем, забезпечуючи ефективну інтеграцію водневих технологій з відновлюваними джерелами енергії, такими як сонячна та вітрова. Це відкриває нові горизонти для розвитку сталої, гнучкої та безпечної енергетичної інфраструктури.

Для України розвиток водневої енергетики є надзвичайно актуальним і стратегічно важливим. Це нове джерело енергії може суттєво посилити енергетичну незалежність держави та сприяти декарбонізації економіки. Водночас для його впровадження необхідно вирішити низку інфраструктурних і фінансових питань. Передусім, потрібне забезпечення фінансування на модернізацію обладнання, будівництво електролізерів та розвиток супутньої інфраструктури.

Особливу увагу слід приділити адаптації української газотранспортної системи до транспортування водню. Оптимальним варіантом є переорієнтація наявної газової інфраструктури для створення повноцінного ланцюга – від виробництва до логістики та кінцевого споживання. Крім того, важливо розробити проєкти з переоснащення вугільних електростанцій для їх переходу на водневе паливо, а також ініціювати запуск муніципального транспорту, що працює на водневій енергії.

Інтеграція ШІ у водневу енергетику є не просто технологічною новацією, а критично важливим елементом стратегії переходу до низьковуглецевої економіки. ШІ забезпечує підвищення ефективності всіх етапів – від виробництва до транспортування та споживання водню, сприяючи оптимізації ресурсів, зниженню витрат та підвищенню надійності енергетичних систем. Такий підхід дозволить

Україні не лише модернізувати свою енергетичну інфраструктуру, а й зайняти гідне місце на глобальному енергетичному ринку майбутнього.

Але високий рівень цифровізації енергогалузі призводить до кіберзагроз, які спрямовані на безпеку електромереж і створюють серйозну проблему через численні атаки. Тому кібербезпека стає важливим елементом енергетичної безпеки. Повномасштабна збройна агресія Російської Федерації проти України спричинила безпрецедентну хвилю кібератак, скерованих передусім на енергетичну інфраструктуру та інші об'єкти критично важливого значення.

Так, до повномасштабного вторгнення Україну атакували окремі угруповання хакерів. Підкреслюється, що зараз хакерів координує один центр. Росія витрачає на кожну кібератаку мільйони доларів. У 2023 році було зафіксовано близько 55 кібератак на енергооб'єкти. Кібератаки комбінуються з ракетними ударами та атаками дронів по енергооб'єктах. З 24 лютого 2022 року кількість кіберинцидентів на держкомпанії «Укренерго» зросла втричі, порівняно з попереднім роком. До повномасштабної війни російських хакерів цікавила фінансова вигода: викуп за повернення систем у лад чи крадіжку даних із метою продажу. Зараз дестабілізація критичної інфраструктури, щоб припинити енергопостачання. Російські хакери використовують різні підходи: від сканування IT-периметра до DDoS-атак. Під час однієї з DDoS-атак на Укренерго кількість запитів могла перевищувати 5 млн за кілька годин [11].

Крім того, кібератаки можуть впливати не лише на енергетичну систему України, а й на енергосистеми сусідніх країн. Подібні кібератаки мають ризик повторення в інших державах. Це обумовлено тим, що структура українських мереж є типовою для багатьох країн. Типи обладнання, що використовуються в українській енергосистемі, також широко розповсюджені за її межами.

Враховуючи набутий практичний досвід, Україна здатна зробити суттєвий внесок у формування глобальної архітектури кібербезпеки, зокрема через обмін знаннями, розробку стандартів протидії гібридним загрозам та участь у міжнародних ініціативах з кіберзахисту критичної інфраструктури.

Українські енергетичні компанії впровадили багаторівневу систему кіберзахисту, яка охоплює постійний моніторинг, програму підготовки персоналу, міжнародну технічну підтримку та використання адаптивних технологій. Такий підхід забезпечив високий рівень стійкості до скоординованих гібридних атак з боку Росії. Продовження цих зусиль, розширення захисту на малих та середніх операторів, а також активніше впровадження ІТ-рішень на основі ШІ дозволить зміцнити кіберщит національної енергетичної інфраструктури.

Згідно з дослідженнями, ефективність російських атак залишається низькою саме завдяки високому рівню кіберстійкості України. Попри численні масштабні спроби втручання, жодна з атак за останні роки не призвела до порушення роботи автоматизованих систем управління підстанціями.

Українські енергокомпанії щороку витрачають понад 100 мільйонів гривень на ліцензійне програмне забезпечення для кіберзахисту, навіть за умов обмеженої ліквідності. Частина витрат компенсується завдяки донорській міжнародній підтримці. При цьому великі державні компанії – зокрема «Укренерго», «Енергоатом» та «Нафтогаз» – мають кращий рівень захищеності через доступ до значно більших фінансових ресурсів.

Поєднане використання кібератак, дронів та ракет стало визначальною ознакою тактики російських військ у зимовий період. Такий гібридний підхід спрямований на максимальне порушення роботи критичної інфраструктури, зокрема енергетичної системи, шляхом одночасного фізичного та цифрового впливу.

Система захисту енергетичної інфраструктури повинна охоплювати механізми виявлення, запобігання та реагування на дії зловмисників, з метою мінімізації впливу кібератак на мережу та пов'язані економічні наслідки. Вона також має забезпечувати автоматичне виявлення, усунення або зменшення наслідків технічних несправностей як на локальному, так і на загальносистемному рівні. У критичних ситуаціях система повинна мати змогу впроваджувати вимушені обмеження споживання електроенергії, а також стимулювати управління попитом, щоб збалансувати навантаження та зберегти стабільність енергопостачання.

Ця затяжна кібервійна стала справжнім випробуванням для національної системи кіберзахисту, водночас сформувавши унікальний масив практичного досвіду, що має стратегічну цінність у глобальному контексті. Посилення кіберзагроз унаслідок бойових дій суттєво актуалізувало необхідність впровадження інтелектуальних, адаптивних та автономних систем захисту, здатних забезпечувати стійку та безперебійну роботу критичної інфраструктури навіть в умовах кризових ситуацій. Україна, внаслідок прямого зіткнення з кібервійною, здобула безцінні знання щодо методів виявлення, нейтралізації та запобігання масштабним кіберінцидентам, зокрема у сферах енергетики, зв'язку та фінансів. Цей досвід набуває особливого значення у світлі зростаючої цифровізації економіки та суспільства, яка, у свою чергу, підвищує потребу в надійних засобах кіберзахисту як у державному, так і в приватному секторах. Системна загроза, що постала перед Україною, стимулювала не лише розвиток внутрішнього ринку послуг у сфері кібербезпеки, а й викликала зростаючий інтерес міжнародної спільноти до підтримки кіберстійкості країни шляхом надання технічної, експертної та матеріально-логістичної допомоги.

Найпопулярніші види кібератак: DDoS, Ransomware, Phishing 2020. Найбільша частка ринку припадає на США, Китай, Німеччину та Велику Британію. Країни, на які здійснюється найбільша кількість кібератак: США, Україна, Південна Корея, Китай. UAC-0010 (Gamaredon/ФСБ) залишається найбільш активним російським.

Кіберінциденти у Smart Grid можуть призводити до порушення балансу навантаження, пошкодження обладнання та зниження якості електропостачання. В умовах воєнного конфлікту кіберзагрози набувають гібридного характеру і можуть бути частиною стратегічних атак на державну інфраструктуру.

Аналіз наведених даних про кібератаки на різні сектори України у 2024 році показує загальне зростання кіберзагроз у другій половині року порівняно з першою половиною. Найбільш динамічне зростання спостерігається у військовому секторі (+82 %) та в структурах місцевої влади (+53 %). Водночас енергетичний сектор демонструє найменшу зміну – лише +2 %, що потребує окремого аналізу. Незначне зростання кількості атак на енергетичний сектор (+2 %) не означає зниження інтересу до нього з боку зловмисників.

Таблиця 2 – Динаміка кібератак на ключові сектори України у 2024 році

Сектор	І півріччя 2024 року	ІІ півріччя 2024 року	Темп приросту, %
Військовий сектор	276	502	+82 %
Державні організації	473	665	+41 %
Енергетичний сектор	124	127	+2 %
Телеком та ІТ-сектор	26 (19+7)	37 (33+4)	+42 %
Місцева влада	542	831	+53 %

Джерело: складено за даними [12]

Це, швидше за все, свідчить про вже високу базову активність та постійну присутність загроз. Енергетика залишається однією з головних цілей кібератак, особливо в умовах війни, оскільки атаки на енергосистему мають високий дестабілізуючий потенціал. Незначна динаміка може бути результатом покращення кіберзахисту в галузі після масових атак попередніх періодів. Водночас можлива і зміна тактики атак: менше шумних або масових атак, натомість точкові, приховані, націлені на порушення SCADA-систем або ланцюги постачання. Порівняно з державними органами (+41 %) чи телекомунікаційним та ІТ-сектором (+42 %), де спостерігається значне зростання атак, енергетика виглядає «стабільною». Водночас місцева влада (+53 %) та військові структури (+82 %) стали новими центрами загострення, що може відволікати ресурси уваги від енергетичної сфери. Незважаючи на невелике зростання, ризики залишаються високими, оскільки енергетика тісно інтегрована з іншими секторами – зокрема з місцевою владою, телекомом та держструктурами. Атаки на один із сегментів можуть мати ланцюговий ефект.

SWOT-аналіз показує, що попри наявність суттєвих слабких сторін та зовнішніх загроз, енергетичні підприємства України мають потужний потенціал для формування ефективної системи кіберзахисту. Реалізація можливостей, пов'язаних із впровадженням інтелектуальних технологій та підвищенням рівня цифрової грамотності персоналу, дозволить значно зміцнити кіберстійкість галузі. При цьому необхідно приділяти особливу увагу управлінським, освітнім та нормативно-правовим аспектам. Кіберстійкість енергосистеми має стати пріоритетом державної політики у сфері національної безпеки

Таблиця 3 – Свот-аналіз енергетичної галузі України

Сильні сторони	Слабкі сторони
• Високий рівень усвідомлення критичності кібербезпеки серед керівництва підприємств. • Початок впровадження систем моніторингу та реагування на інциденти. • Співпраця з міжнародними партнерами у сфері кіберзахисту. • Досвід реального протистояння кіберзагрозам в умовах війни.	• Застаріла ІТ-інфраструктура на багатьох енергетичних об'єктах. • Обмежене фінансування заходів із кібербезпеки. • Низький рівень кваліфікації персоналу у сфері ІТ-захисту. • Відсутність комплексної стратегії кіберменеджменту на рівні підприємств.
Можливості	**Загрози**
• Інтеграція ШІ, машинного навчання та big data для підвищення ефективності захисту. • Державні ініціативи щодо цифрової трансформації та захисту критичної інфраструктури. • Можливість отримання грантів та технічної підтримки від міжнародних донорів. • Стандартизація та сертифікація кібербезпеки.	• Зростання складності кібератак та використання кібервійни як інструменту гібридного впливу. • Зовнішні атаки на інфраструктуру з боку ворожих держав або організованих хакерських угруповань. • Поширення шкідливих програм, зокрема через уразливі IoT-пристрої. • Недостатній правовий захист у випадку кібератак на приватні енергетичні компанії.

Джерело: складено особисто автором

Сьогодні кібербезпека перестала бути виключно технічною категорією – вона трансформувалася в критичний елемент національної безпеки, стійкості бізнесу та довіри громадян до державних і корпоративних систем. Особливо гостро ця проблема постає для України, яка з початку повномасштабного вторгнення Росії стала мішенню для безпрецедентної хвилі кібератак, зокрема на критичну інфраструктуру та енергетичні підприємства. Кіберзахист повинен розглядатися не як реакція на інциденти, а як інтегрована складова стратегії стримування, аналогічна до традиційної оборонної політики.

Одним із нових викликів є децентралізація кіберзлочинності. Сучасні загрози стають дедалі більш автономними, мобільними та технологічно оснащеними, тоді як нормативно-правове регулювання, як правило, не встигає за темпами інновацій. Якщо країни не розроблятимуть гнучкі механізми кіберрегулювання, вони ризикують постійно відставати у протидії новим вразливостям.

Технології подвійного призначення, криптографічні рішення та ШІ одночасно слугують інструментами інновацій та джерелами загроз. Все залежить від того, хто контролює ці інструменти та в якій системі управління вони функціонують. ШІ може бути як щитом, так і зброєю – і саме ефективна модель кіберменеджменту визначає, якою стороною буде реалізований потенціал ШІ.

Цифрова трансформація енергетичної галузі виводить концепцію Smart Grid на передній план модернізаційних процесів. Інтелектуальні енергетичні мережі забезпечують оптимізацію споживання, підвищення надійності систем та інтеграцію відновлюваних джерел енергії. Однак, поряд з перевагами, зростає спектр кіберзагроз, які можуть мати катастрофічні економічні, технологічні та соціальні наслідки. Критична інфраструктура, до якої належать енергетичні мережі, вимагає пріоритетного захисту на всіх етапах життєвого циклу систем.

Одним із ключових принципів безпеки у Smart Grid має стати "Security by Design" – вбудована з початку розробки концепція безпеки, яка враховує моделювання загроз, визначення вимог до кожного компонента мережі, постійний аудит та дотримання стандартів безпеки впродовж усього життєвого циклу проєкту. Такий підхід дозволяє мінімізувати необхідність пізньої перебудови небезпечних елементів, яка може призводити до затримок у впровадженні інновацій.

Найбільш серйозні наслідки кібератак на Smart Grid включають: порушення енергопостачання, пошкодження обладнання, витік персональних даних користувачів, репутаційні втрати та вимушені надзвичайні інвестиції у енергетичні системи. Цільові атаки стають дедалі складнішими, зокрема на SCADA-системи, що керують об'єктами критичної інфраструктури.

Багаторівнева стратегія захисту Defense in Depth є фундаментом кібербезпеки Smart Grid. Вона передбачає логічну сегментацію мережі відповідно до критичності компонентів, створення

контрольованих точок доступу між сегментами, розгортання незалежних захисних механізмів на кожному рівні, а також регулярне тестування ізоляції сегментів для виявлення вразливостей. Важливим залишається баланс між безпекою і функціональністю: надмірна ізоляція може знижувати ефективність системи, а недостатній рівень сегментації створює ризики для масштабних кібератак.

Важливою складовою сучасної кібербезпеки є управління ризиками третіх сторін. Екосистема Smart Grid охоплює обладнання, програмне забезпечення та сервіси від численних постачальників. Аналітичні дані вказують, що понад 60 % кіберінцидентів у критичній інфраструктурі пов'язані з уразливостями сторонніх компонентів. Стратегія управління цими ризиками включає встановлення чітких вимог до постачальників, проведення незалежного аудиту, моніторинг вразливостей та наявність планів реагування на інциденти, пов'язані з ланцюгами постачання.

Таблиця 4 – Види ШІ для протидії кіберзагрозам

Категорія	**Застосування у сфері кібербезпеки енергетики**
1	2
Аналітичний ШІ	Надання експертних порад щодо захисту та виявлення аномалій у трафіку. Прогнозування атак та розробка превентивних стратегій.
Функціональний ШІ	Автоматичне виявлення та реагування на загрози через IoT-пристрої. Постійний моніторинг і захист інтернет-пристроїв.
Інтерактивний ШІ	Взаємодія з операторами в реальному часі під час інцидентів. Проведення симуляцій та тренінгів для персоналу.
Текстовий ШІ	Автоматичний аналіз логів і текстових даних для виявлення аномалій. Застосування NLP для аналізу зовнішніх повідомлень про загрози.
Візуальний ШІ	Біометричний контроль доступу до критичних зон.Аналіз відео з метою виявлення несанкціонованих дій.
Роботизована автоматизація процесів	Автоматичне оновлення ПЗ, патч-менеджмент, моніторинг мережі. Фізичний контроль інфраструктури за допомогою роботів.

Продовження таблиці 4

1	2
Розпізнавання мови та комп'ютерний зір	Ідентифікація голосових команд і контроль доступу. Аналіз відеопотоків для виявлення підозрілої активності.
Data Science	Обробка великих даних для аналізу трафіку та логів. Побудова моделей ML для прогнозування атак.
Reactive Machine та Limited Memory AI	Реакція в реальному часі на кіберзагрози. Виконання базових завдань моніторингу та виявлення аномалій.
Theory of Mind AI та Self-Awareness AI	Розуміння намірів кіберзлочинців для покращення захисту. Автономне прийняття рішень щодо кібербезпеки на основі самосвідомості.

Джерело: складено за даним [13]

Кібербезпека все частіше розглядається як інтегрована складова ESG-стратегії (екологічні, соціальні та управлінські фактори). Інвестиції в захист Smart Grid прямо корелюють із довгостроковою стійкістю компаній. Захист критичної інфраструктури сприяє уникненню екологічних інцидентів (Environmental), забезпечує стабільне енергопостачання для спільнот (Social) та демонструє відповідальне управління ризиками (Governance). Компанії з високим рівнем кіберзахисту мають вищу інвестиційну привабливість та стійкість у кризових ситуаціях.

Таким чином, безпека Smart Grid виходить за межі IT-дисципліни і постає як стратегічне завдання для енергетичних компаній, державних органів та глобальних партнерств. В умовах гібридної війни, технологічної еволюції та діджиталізації саме проактивний, архітектурно вбудований і багаторівневий підхід до кіберзахисту здатен забезпечити надійність, стійкість та довіру до енергетичної інфраструктури майбутнього.

Але ШІ підвищує економічну ефективність, зменшуючи ймовірність людської помилки – однієї з основних причин порушень кібербезпеки. Для забезпечення ефективного управління, моніторингу та захисту Smart Grid ми виокремлюємо наступні категорії ШІ.

Кіберменеджмент є ключовим елементом управління сучасною цифровою організацією. Він охоплює планування, координацію,

впровадження та контроль заходів з кібербезпеки. Ефективний кіберменеджмент забезпечує стійкість ІТ-інфраструктури до внутрішніх і зовнішніх загроз. Він передбачає формування політик інформаційної безпеки на рівні підприємства. До завдань кіберменеджера входить управління ризиками, інцидентами та реагуванням на атаки. Важливою складовою є постійний моніторинг систем і аналіз аномалій. Роль кіберменеджменту зростає в умовах гібридної війни та цифровізації. Він тісно пов'язаний з ІТ-менеджментом, кризовим управлінням та стратегічним плануванням [14]. Ефективні протоколи доступу та контроль ідентифікації є його невід'ємною частиною. Кіберменеджмент передбачає інтеграцію ШІ для виявлення та прогнозування загроз. Навчання персоналу є критично важливим для формування кіберстійкої культури. Кіберменеджмент має бути адаптивним до змін середовища загроз. До функцій також належить управління зовнішніми підрядниками та провайдерами. Важливе значення має аудит систем безпеки та проведення тестів на проникнення. Кіберменеджери повинні володіти навичками як у сфері технологій, так і управління. Кіберменеджмент охоплює захист конфіденційності, цілісності та доступності інформації. Він потребує належного фінансування та підтримки з боку топменеджменту. Ефективна комунікація під час кіберінцидентів – ще один критично важливий аспект. Успішний кіберменеджмент сприяє підвищенню довіри клієнтів і партнерів. З огляду на критичну важливість енергетичного сектору для національної безпеки, кіберменеджмент набуває стратегічного значення. Інтеграція кіберменеджменту з енергоменеджментом є критично важливою для формування енерго- та кіберстійкості. Успішна реалізація цієї інтеграції вимагає міжгалузевої співпраці, державної політики кіберзахисту критичної інфраструктури та інвестицій у новітні технології безпеки.

Людський фактор є одним із ключових елементів, що визначають рівень кібербезпеки енергетичних підприємств. Досвід показує, що більшість кібератак стаються не через технічні недоліки, а через помилки або необережні дії персоналу. Найчастіше зловмисники використовують методи соціальної інженерії, зокрема фішинг, коли працівник самостійно надає доступ до систем, вважаючи повідомлення або запит легітимним. Однією з головних проблем залишається

недотримання працівниками політик безпеки – використання слабких паролів, ігнорування багатофакторної автентифікації, відкритий доступ до конфіденційної інформації.

Критичний дефіцит кадрів, які здатні працювати на перетині цифрових та виробничих сфер, створює додаткові ризики. Однією з актуальних проблем у сфері кібербезпеки є гостра нестача кваліфікованих фахівців. Зі зростанням кількості та складності кіберзагроз зростає й потреба в професіоналах, здатних ефективно реагувати на атаки, забезпечувати безпеку інформаційних систем і впроваджувати сучасні технології захисту. Проте темпи підготовки спеціалістів не відповідають викликам часу: багато освітніх програм залишаються теоретичними, а молоді фахівці не завжди мають достатній практичний досвід для роботи в умовах реальних кіберінцидентів.

Таким чином, фахівці повинні мати можливість працювати з реальними кіберзагрозами в умовах, максимально наближених до реальності. Це суттєво зменшує ризики, пов'язані з навчанням, що базується виключно на теорії, та сприяє формуванню практичних навичок. Використання ШІ у симуляційних платформах дозволяє моделювати сценарії з реальними кіберінцидентами – від фішингових атак і DDoS-ударів до шкідливого програмного забезпечення. Завдяки цьому здобувачі освіти мають змогу взаємодіяти з цифровим середовищем, приймати управлінські рішення, випробовувати різні методи захисту та аналізувати їхню результативність. Такий підхід поєднує глибокі теоретичні знання з цінним практичним досвідом, необхідним для протидії кіберзагрозам у сфері енергетики. Тому особливої важливості набуває розвиток освітніх ініціатив, симуляційних платформ, дуальної освіти та тісної співпраці між навчальними закладами, бізнесом і державними структурами для оперативної підготовки нової генерації кіберфахівців.

Також віддалена робота, використання незахищених Wi-Fi-з'єднань, домашніх пристроїв у корпоративних середовищах створюють нові вектори для проникнення, які особливо активно експлуатуються фішинговими кампаніями. Ці атаки все частіше використовують елементи ШІ для створення персоналізованих повідомлень у реальному часі, що значно підвищує їхню ефективність і складність виявлення. Статистикою підтверджено, що 96 % кібератак і витоку

даних розпочинається з e mail. E-mail захищений так само, як поштова листівка, адже електронні листи «проходять» через вразливі та потенційно небезпечні поштові сервери. При цьому, SSL/TSL не гарантує безпеки. Сьогодні перехоплення даних та «взлом» можливі всього за $ 200, а мережеві імпланти для перехоплення трафіку коштують всього від $ 60. Не дивно, що браузери назвичайно вразливі до «взломів». E-mail інфраструктура не верифікує відправника. Засоби захисту периметру компанії не захищають e-mailпісля відправки листів [15].

Навчання персоналу та підвищення кіберграмотності є ключовими для зменшення ризиків людського чинника. Кібербезпека повинна бути інтегрована в проєктування смарт-мереж на всіх етапах життєвого циклу. Впровадження стандартів допомагає уніфікувати політики безпеки. Умови децентралізації енергогенерації, зокрема використання ВДЕ, створюють додаткові виклики для забезпечення захисту від кібератак. Співпраця між операторами енергосистем, IT-компаніями та урядом є необхідною для створення єдиного захисного середовища. Важливу роль відіграє створення систем раннього попередження та обміну інформацією про інциденти.

Втім, стрімкий розвиток ШІ породжує дотримання етичних норм під час їх використання. Також одним із ключових напрямів впровадження ШІ в Україні, на нашу думку, є аналіз мережевого трафіку та моніторинг поведінки енергетичних систем у режимі реального часу з використанням інтелектуальних алгоритмів. Такий підхід дозволяє не лише оперативно виявляти потенційні кіберзагрози, але й підвищувати загальну стійкість критичної енергетичної інфраструктури до зовнішніх впливів.

Захист Smart Grid вимагає комплексного підходу, що включає як технічні, так і організаційні заходи. Необхідно впроваджувати багаторівневу автентифікацію, шифрування даних та безпечні протоколи комунікації. Кібербезпека повинна охоплювати як рівень генерації та передачі енергії, так і кінцевих споживачів. Особливо важливо забезпечити ізоляцію критичних компонентів від загальнодоступних мереж. Постійний моніторинг аномалій у мережевому трафіку дозволяє виявляти потенційні загрози на ранніх стадіях. Інтеграція систем ШІ у кіберзахист Smart Grid забезпечує вищий рівень

адаптивності до нових типів атак. Використання машинного навчання сприяє покращенню виявлення аномальної поведінки пристроїв. Необхідно проводити регулярні тести на проникнення для виявлення вразливих місць. Законодавча база повинна регулювати вимоги до кіберзахисту у Smart Grid та передбачати відповідальність за нехтування безпекою.

Постійне оновлення програмного забезпечення та патч-менеджмент повинні бути частиною щоденної практики управління кіберризиками. Розвиток цифрових двійників (digital twins) у Smart Grid дає нові можливості для моделювання атак і тестування захисту. Кібератаки на лічильники та пристрої користувачів можуть використовуватися для підриву довіри до всієї системи. Впровадження кіберменеджменту на підприємствах енергетичного сектору має здійснюватися в рамках загальної стратегії цифрової трансформації. Майбутнє смарт-енергетики тісно пов'язане з розвитком ефективної, гнучкої та прогнозованої системи кібербезпеки. Саме поєднання технологічних інновацій із проактивною безпековою політикою дозволить досягти високого рівня захищеності критичної енергетичної інфраструктури.

Таким чином, в сучасних умовах глобальної цифровізації та зростання кіберзагроз питання забезпечення кібербезпеки енергетичних підприємств набуває критичного значення. Особливо це актуально для України, де енергетична інфраструктура регулярно стає об'єктом цілеспрямованих атак з боку державних та приватних акторів. У відповідь на зростання складності та частоти кібератак виникає потреба у новій парадигмі управління кібербезпекою, що базується на застосуванні інтелектуальних технологій. Інтелектуалізація кіберзахисту передбачає інтеграцію ШІ, машинного навчання, аналітики великих даних та автоматизованих систем ухвалення рішень у процеси управління кіберризиками. Такий підхід дозволяє не лише виявляти загрози в режимі реального часу, а й прогнозувати їх розвиток, оцінювати вразливості та формувати проактивні стратегії захисту. Управлінський аспект інтелектуалізації кіберзахисту охоплює стратегічне планування, розподіл ресурсів, навчання персоналу, впровадження стандартів та забезпечення безперервності бізнес-процесів в умовах цифрових загроз. Ключовими бар'єрами

є низька обізнаність управлінців у сфері цифрової безпеки, недостатнє фінансування інновацій та обмежений доступ до висококваліфікованих фахівців. Разом із тим, міжнародний досвід демонструє, що впровадження ШІ в систему кібербезпеки суттєво знижує рівень втрат від інцидентів та підвищує надійність критичних інфраструктур.

В умовах зростаючої нестабільності в Чорноморському регіоні та посилення гібридних загроз, необхідним кроком є розширення міжнародної співпраці, зокрема з ЄС і НАТО. Спільне реагування на інциденти, обмін аналітикою та уніфікація стандартів захисту критичної інфраструктури повинні стати пріоритетами кіберстратегії України.

Основними шляхами мають бути:

Інтелектуальний моніторинг – використання алгоритмів машинного навчання для аналізу мережевого трафіку та виявлення аномальної активності.

Проактивне управління ризиками – прогнозування можливих сценаріїв атак із урахуванням історичних даних та поведінкових патернів.

Автоматизоване реагування – впровадження систем, здатних автономно виявляти та ізолювати шкідливі впливи до моменту ескалації.

Інтеграція кіберменеджменту в стратегічне управління енергетичною галузю – включення питань кіберзахисту до стратегій розвитку, інвестиційних планів.

Освітній компонент – формування культури кібербезпеки серед персоналу, включаючи регулярні тренінги, симуляції атак та сертифікацію фахівців.

Список використаних джерел

1. Дегтярьова О. О., Куклінова Т. В., Куклінова С. І. Системи ШІ в оптимізації енергетичних процесів:інноваційні підходи до ефективного управління сталим розвитком підприємства. *Вісник Хмельницького національного університету.* 2024. № 1. С. 359–362.

2. Degtiareva O., Shyriaieva N. Kuklinova T. Artificial Intelligence Solutions for Cybersecurity in Energy Systems. 2024 IEEE International Workshop on Technologies for Defense and Security (*TechDefense*). Naples, Italy. 2024. P. 177–182. DOI: 10.1109/TechDefense63521.2024.10863745.

3. Khan F., Mohsin S., Durad H. Dragonfly Cyber Threats: A Case Study of Malware Attacks Targeting Power Grids. *Journal of Computing & Biomedical Informatics*. 2023. Vol. 04. Is. 02. P. 172–185.

4. Збитки та втрати енергетичного сектору України внаслідок повномасштабного вторгнення Росії перевищили $56 млрд. Оцінка KSE Institute станом на травень 2024 року. Kyiv School of Economics. 2024. URL: https://kse.ua/ua/about-the-school/news/zbitki-ta-vtrati-energetichnogo-sektoru-ukrayini-vnaslidok-povnomasshtabnogo-vtorgnennya-rosiyi-perevishhili-56-mlrd-otsinka-kse-institute-stanom-na-traven-2024-roku/

5. Мохор В. В. Цифрова трансформація енергетики як запорука забезпечення її стійкості : Стенограма доповіді на засіданні Президії НАН України 4 жовтня 2023 року. *Вісник Національної академії наук України.* 2023. № 12. С. 74–79.

6. Омельченко В. Чому насправді відключають світло. *Центр Разумкова.* 2024. URL: https://razumkov.org.ua/napriamky/videokomentari/chomu-naspravdi-vidkliuchaiut-svitlo-ekspert

7. Особливості вітчізняного виробництва електроенергії. *Ukrainian Energy Exchange*. 2024. URL: https://www.ueex.com.ua/rus/presscenter/news/osobennosti-otechestvennogo-proizvodstva-elektroenergii/

8. Як Modus X розробляє та впроваджує рішення на ML та Copilot у DTEK. *Speka*. 2024. URL: https://speka.media/yak-modus-x-rozroblyaje-ta-vprovadzuje-risennya-na-ml-ta-copilot-u-dtek-936wmx

9. Держава почала створювати розумні електричні мережі – уже є 12 пілотних проєктів. Liga.net. 2024. URL: https://biz.liga.net/ua/all/all/novosti/derzhava-pochala-stvoriuvaty-rozumni-elektrychni-merezhi-uzhe-ie-12-pilotnykh-proiektiv

10. Голіздра О. Розбудова Розумна мережа – шлях до підвищення стійкості енергосистеми України. Укрінформ. 2024. URL: https://www.ukrinform.ua/rubric-economy/3863598-rozbudova-smart-grid-slah-do-pidvisenna-stijkosti-energosistemi-ukraini.html.

11. Чайка О. Російські хакери координують дії з військовими та посилюють атаки напередодні зими. Як Україна протистоїть кібератакам на енергосистему. November 15. 2023. *Forbes Ukraine.* URL: https://forbes.ua/company/rosiyski-khakeri-koordinuyut-dii-z-

viyskovimi-ta-posilyuyut-ataki-naperedodni-zimi-yak-ukraina-protistoit-kiberatakam-na-energosistemu-08112023-17242

12. Russian cyber operations: analytics for the H2 2024. State Service of Special Communications and Information Protection of Ukraine. Kyiv: State Service of Special Communications and Information Protection of Ukraine, 2024. 22 p.

13. Дегтярьова О. О. Соціально-економічні перспективи застосування ШІ в бізнес-середовищі: переваги та ризики. *Вісник соціально-економічних досліджень* : зб. наук. праць. Одеса : Одеський національний економічний університет. № 1–2. 2023. С. 118–130.

14. Горбаченко С. Місце менеджменту кібербезпеки у сучасній управлінській науці та практиці. *Сталий розвиток економіки*. 2024. № 1 (48). С. 144–149. URL: https://doi.org/10.32782/2308-1988/2024-48-19.

15. Краус К. М., Краус Н. М., Штепа О. В. Цифрова трансформація кібербезпеки на мікрорівні в умовах воєнного стану. URL: https://elibrary.kubg.edu.ua/id/eprint/42325/1/Kraus_Tsyfrova_transformatsiia_kiberbezpeky_2022.pdf

РОЗДІЛ 2. МАТЕМАТИЧНІ ІНСТРУМЕНТИ В КІБЕРБЕЗПЕЦІ

Чепурна Олена Євгенівна
к.ф.-м.н., доц., доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ РИЗИКІВ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ: ЗАСТОСУВАННЯ СТОХАСТИЧНОГО АНАЛІЗУ ТА МАШИННОГО НАВЧАННЯ ДЛЯ ОПТИМІЗАЦІЇ СТРАТЕГІЙ

DOI https://doi.org/10.36059/978-966-397-537-5-4

1.1. Актуальність проблеми оцінки ризиків інформаційної безпеки у сучасних інформаційних системах

Критична важливість інформаційно-комунікаційних систем (ІКС) для функціонування державного управління, бізнес-структур, енергетичних мереж і соціальної інфраструктури зростає рік від року. Водночас стрімка цифровізація супроводжується зростанням складності загроз і ризиків інформаційної безпеки (ІБ), які мають стохастичну природу, різнорідну динаміку та високий ступінь невизначеності. В умовах гібридних і цільових кібератак класичні підходи до оцінювання ризиків, що базуються на статичних моделях або експертних таблицях, виявляються недостатньо гнучкими та масштабованими [1]. Згідно з даними звіту IBM X-Force Threat Intelligence Index 2024, середній збиток від інцидентів витоку даних у стратегічних секторах економіки перевищив 5 млн дол. США, а середній час виявлення загрози перевищив 200 днів [2]. За статистикою Cisco Annual Cybersecurity Report (2023), понад 50 % успішних атак відбуваються внаслідок складних багатоступеневих сценаріїв, що включають елементи соціальної інженерії, експлуатацію вразливостей нульового дня, а також атак на канали віддаленого доступу [3].

Такі умови вимагають застосування формалізованих підходів до оцінювання ризиків, здатних охопити як часові залежності, так

і ймовірнісну природу подій. Стохастичні моделі, зокрема, марковські процеси, пуассонівські потоки подій, методи Монте-Карло – дають змогу моделювати зміну станів системи, інтенсивність атак, а також вірогідність критичних інцидентів у часовому просторі [4]. Водночас методи машинного навчання (ML), що базуються на історичних даних і розподілах аномалій, дозволяють виявляти складні, латентні залежності у поведінці систем і порушників [5].

Об'єктом дослідження є процеси оцінювання та прогнозування ризиків інформаційної безпеки в інформаційно-комунікаційних системах із високим ступенем невизначеності.

Предметом дослідження – математичні моделі на основі стохастичного аналізу та методів машинного навчання, що використовуються для оцінювання, прогнозування та оптимізації стратегій інформаційної безпеки.

Мета дослідження – розробка теоретичних і прикладних підходів до моделювання ризиків ІБ шляхом інтеграції стохастичних процесів і машинного навчання для формування адаптивних стратегій управління ризиками.

Основні завдання дослідження:

1) проаналізувати актуальні методи стохастичного аналізу, застосовувані у сфері ІБ;

2) дослідити можливості алгоритмів ML для прогнозування та виявлення інцидентів ІБ;

3) побудувати комбіновані моделі для комплексної оцінки ризиків з урахуванням часової динаміки та поведінкових характеристик;

4) оцінити ефективність розроблених моделей на основі даних звітів провідних компаній (Cisco, AWS, Microsoft);

5) Сформулювати практичні рекомендації щодо впровадження моделей у різних класах організацій.

Методологічну основу дослідження становить міждисциплінарний підхід, що поєднує інструменти теорії ймовірностей, математичної статистики, теорії ризику, кібернетики та інтелектуального аналізу даних.

Методи дослідження включають: моделювання дискретних випадкових процесів (марковські ланцюги, пуассонівські потоки); чисельне моделювання з використанням методів Монте-Карло;

supervised та unsupervised ML-алгоритми (Random Forest, XGBoost, Isolation Forest); методи багатокритеріальної оптимізації; аналіз нормативних документів (ISO/IEC 27001, 27005) і порівняльну оцінку практичних кейсів.

Таким чином, дослідження базується на обґрунтованій необхідності створення гібридних моделей оцінювання ризиків, здатних адаптуватися до реальних загрозових сценаріїв та мінімізувати наслідки інцидентів шляхом прийняття оптимальних управлінських рішень у сфері інформаційної безпеки.

Управління ризиками інформаційної безпеки (ІБ) в умовах зростаючої складності загроз вимагає системного використання математичних методів, здатних формалізувати невизначеність, враховувати часові залежності й адаптуватися до змін середовища. Математичне моделювання дозволяє перейти від інтуїтивних або експертно-емпіричних оцінок до кількісного аналізу ризиків, зокрема, через побудову формалізованих моделей атак, обчислення ймовірностей настання інцидентів, оцінку очікуваних збитків та оптимізацію стратегій протидії.

Однією з ключових моделей є використання марковських процесів для представлення змін стану об'єкта ІБ у часі. Наприклад, система може переходити між станами «нормальний», «загроза виявлена», «атака успішна», «відновлення». Ймовірності переходу між цими станами дозволяють побудувати матрицю переходів та обчислити довгострокові характеристики безпеки [6]. Такий підхід застосовувався, зокрема, у дослідженнях моделей проникнення та вразливостей у хмарних середовищах [7].

Іншою поширеною моделлю є пуассонівський процес, який використовується для моделювання частоти настання інцидентів безпеки, зокрема DoS-атак або спроб несанкціонованого доступу. Якщо інциденти виникають випадково, але з відомою інтенсивністю λ, можна визначити ймовірність k інцидентів за фіксований час T, використовуючи розподіл Пуассона (1) [8]:

$$P(k;\lambda T) = \frac{(\lambda T)^k e^{-\lambda T}}{k!}. \quad (1)$$

Для систем зі складною топологією та взаємозв'язками між компонентами доцільно застосовувати методи Монте-Карло. Вони

дозволяють згенерувати велику кількість сценаріїв атак або відмов і визначити розподіл втрат або рівень ризику з урахуванням випадкових факторів [9].

Крім того, математичне моделювання слугує інструментом для перевірки ефективності контрзаходів. Наприклад, впровадження багатофакторної аутентифікації або засобів мережевої ізоляції може бути оцінено шляхом перерахунку ймовірностей успішного проходження зловмисника через модельовану систему за змінених умов [10].

Таким чином, математичне моделювання виконує функцію аналітичного ядра процесів управління ризиками ІБ: від попередньої оцінки до моніторингу й адаптації стратегії. Його поєднання з даними, отриманими внаслідок спостережень (логи, телеметрія, статистика атак), дозволяє здійснювати кількісну оцінку й приймати оптимальні рішення в умовах високої динаміки загроз.

1.2. Побудова інтегрованих моделей для аналізу та оптимізації ризиків: інтеграція стохастичних моделей і машинного навчання для оцінки ризиків ІБ

Розвиток інформаційно-комунікаційних систем супроводжується зростанням складності ландшафту кіберзагроз, що потребує використання інтегрованих моделей аналізу ризиків. Актуальність таких підходів обумовлена не лише динамікою появи нових векторів атак, а й необхідністю врахування множинних, взаємопов'язаних факторів, які впливають на ефективність захисту. Стохастичні процеси дозволяють відобразити часову структуру та невизначеність подій, тоді як алгоритми машинного навчання (ML) здатні автоматично виявляти приховані залежності та адаптуватися до нових шаблонів атак. Інтеграція цих підходів створює передумови для побудови адаптивних моделей інформаційної безпеки (ІБ), здатних до самонавчання та прогностичного аналізу [11].

Важливо підкреслити, що ефективність традиційних моделей, заснованих лише на статистичних оцінках, часто обмежується у випадках швидкої еволюції загроз: нові типи атак, способи обходу захисту чи вразливості у складних розподілених системах важко враховувати без залучення інструментів ML, які можуть навчатися на

потоках даних у режимі реального часу. Саме тому синергія стохастичних підходів та штучного інтелекту дозволяє не просто фіксувати факти минулих інцидентів, а й прогнозувати розвиток ситуацій, формувати сценарії реагування до їхнього виникнення. Структурно, такі моделі поєднують: стохастичні моделі – марковські ланцюги, пуассонівські процеси, моделювання Монте-Карло – для кількісного опису випадкових подій, визначення функцій розподілу та обчислення коефіцієнтів ризику для складних багаторівневих систем; алгоритми машинного навчання – ансамблеві дерева, глибокі нейромережі, методи кластеризації й підсилення – для виявлення аномалій, класифікації загроз, формування моделей поведінки та прогнозування ймовірних сценаріїв розвитку подій [12].

Методологічна інтеграція відбувається у двох напрямках: ML-алгоритми використовуються для оцінки параметрів стохастичних моделей. Наприклад, на основі історичних інцидентів кібератак формується матриця переходів марковського ланцюга, що дозволяє змоделювати ймовірності наступних кроків атак, враховуючи різні типи векторів та динаміку змін у поведінці зловмисників [13], стохастичні моделі коригують роботу ML – наприклад, шляхом формування обмежень на допустимі значення ознак, визначення ступеня ризику чи зважування ваги ознак відповідно до інтенсивності подій, змодельованих за пуассонівським розподілом, що підвищує достовірність автоматичних висновків [14].

Гібридні підходи вже знайшли своє практичне застосування у низці важливих напрямів: раннє виявлення APT-кампаній та цільових атак на критичні інфраструктури, коли класичні методи часто неефективні; моделювання ризиків для хмарних платформ і розподілених обчислювальних середовищ, де взаємозв'язки між компонентами надзвичайно складні та змінювані; оптимізація архітектури кіберзахисту, визначення пріоритетів для впровадження контрзаходів і розробка стратегій реагування, що враховують як модельовані сценарії, так і реальні дані з телеметрії чи журналів безпеки.

Такі моделі можуть адаптивно реагувати на зміну поведінки зловмисників, швидко оновлюючи свою структуру та параметри завдяки безперервному навчанню на потоках даних, що надходять із різних джерел – від сенсорів до систем відеонагляду чи логів

мережевої активності [15]. Їх реалізація можлива за допомогою сучасних бібліотек Python/Scikit-learn, TensorFlow, PyTorch, а також засобів обробки потокових даних (Kafka, Spark), що дозволяє досягати продуктивності в режимі наближеного реального часу навіть у масштабних корпоративних середовищах.

Додатково, інтегровані моделі на основі стохастичного аналізу й ML відкривають нові перспективи для впровадження проактивних стратегій захисту, коли система не лише ідентифікує вже здійснені атаки, а й оцінює ймовірність появи нових типів загроз, автоматично формуючи рекомендації для адміністраторів з урахуванням поточного стану мережі, доступних ресурсів і рівня критичності об'єктів. Таким чином, ці підходи формують ядро сучасних адаптивних стратегій ІБ – систем, здатних не лише виявляти порушення безпеки, а й прогнозувати критичні сценарії, а також пропонувати оптимальні дії в умовах високої динаміки та невизначеності цифрового середовища.

Таким чином, ці підходи формують ядро сучасних адаптивних стратегій інформаційної безпеки – систем, здатних не лише виявляти порушення безпеки, а й прогнозувати критичні сценарії розвитку подій у складних цифрових середовищах. Завдяки використанню поєднання стохастичного моделювання та алгоритмів машинного навчання, аналітичні інструменти отримують нову якість: вони можуть не просто реагувати на вже відомі інциденти, а й розпізнавати нові, раніше не зафіксовані загрози, виявляти аномалії у поведінці користувачів чи автоматизованих систем, а також оперативно перебудовувати свою структуру відповідно до динаміки атак.

Ключовою перевагою таких моделей стає їхня здатність адаптуватися – постійно оновлювати параметри на основі потоків актуальних даних, отриманих із різноманітних джерел: від сенсорних мереж, систем моніторингу мережевого трафіку до журналів подій у хмарних сервісах чи промислових системах. Це відкриває можливість не тільки для глибшого розуміння природи загроз та вразливостей, але й для проактивного формування сценаріїв реагування, коли система сама пропонує оптимальні шляхи нейтралізації ризиків ще до того, як вони матеріалізуються у реальні інциденти.

Більше того, інтеграція стохастичного аналізу зі штучним інтелектом дозволяє здійснювати багатофакторну оптимізацію

архітектури захисту: враховувати не лише технологічні параметри та поточну структуру мережі, а й економічні аспекти, людський фактор і специфіку галузі. Таким чином, сучасні математичні моделі стають не просто інструментом оцінки ризику, а повноцінною платформою для підтримки прийняття управлінських рішень в умовах невизначеності, забезпечуючи стійкість організацій до нових викликів цифрової реальності.

1.3. Нагальні методологічні проблеми у моделюванні ризиків інформаційної безпеки

Становлення ефективних підходів до оцінювання ризиків інформаційної безпеки (ІБ) відбувається в умовах глибоких трансформацій, зумовлених як еволюцією кіберзагроз, так і ускладненням цифрової інфраструктури. На перший план виходить проблема невизначеності поведінки як зловмисників, так і інформаційних систем, що унеможливлює використання статичних моделей для управління ризиками. Ризики ІБ є не лише мультифакторними, але й мають часову динаміку, де ймовірності порушень безпеки змінюються залежно від середовища, контексту та активності атакувальних об'єктів.

Однією з ключових аналітичних проблем є відсутність загальноприйнятої узгодженої моделі оцінювання ризику, що враховує складність взаємозв'язків між технічними, поведінковими й організаційними чинниками. Існуючі підходи або базуються на експертних оцінках, які є суб'єктивними, або застосовують евристичні правила, які не масштабуються в умовах змінного середовища. Постає необхідність в обґрунтуванні системи моделей, яка дозволяє гнучко адаптуватися до нових типів загроз і структур даних.

Проблемним є також інтеграція моделей різного рівня абстракції – від низькорівневого аналізу подій до стратегічного планування заходів захисту. Стохастичні моделі, зокрема марковські ланцюги й пуассонівські процеси, добре описують частотні характеристики загроз, проте не враховують змінну логіку поведінки зловмисників. З іншого боку, методи машинного навчання дозволяють виявляти складні шаблони у даних, але часто не мають формалізованої інтерпретації результатів у термінах ризику. Внаслідок цього виникає потреба в поєднанні

аналітичної строгості стохастичних моделей із прогностичною потужністю ML, що відкриває нову методологічну площину.

Ще однією критичною проблемою є нестача репрезентативних відкритих даних для моделювання ризиків у контексті реальних інфраструктур. Більшість досліджень базується або на симульованих вибірках, або на обмежених за обсягом інцидентах. Це ускладнює калібрування моделей, знижує достовірність прогнозів і обмежує застосування отриманих результатів. Потреба в інтеграції реальних даних зі звітів провідних компаній (Cisco, AWS, Fortinet) у формальні моделі є актуальним напрямом прикладної аналітики ІБ.

Окремим напрямом невирішених питань залишається проблема обґрунтування ефективності стратегій захисту на основі моделювання ризиків. Багато існуючих підходів не дозволяють визначити, які саме заходи дають найкращий ефект у конкретному контексті, особливо за обмежених ресурсів. Оптимізація стратегій ІБ потребує аналітичного апарату, що дозволяє здійснювати багатокритеріальну оцінку заходів захисту з урахуванням змінної поведінки середовища, вартості впровадження та ймовірності загроз.

Таким чином, сукупність методологічних викликів, пов'язаних із невизначеністю, багаторівневою структурою ризиків, дефіцитом відкритих даних і складністю обґрунтування захисних дій, формує контекст, у якому постає необхідність нових підходів до математичного моделювання. Пропонована у цьому дослідженні інтеграція стохастичного аналізу з методами машинного навчання покликана надати обґрунтовану основу для вирішення означених проблем у системах ІБ нового покоління.

2.1. Формалізація ризиків інформаційної безпеки як стохастичних процесів

Оцінювання ризиків інформаційної безпеки (ІБ) ґрунтується на розумінні їх як імовірнісних подій із потенційно значущими наслідками для цілісності, конфіденційності або доступності інформаційних ресурсів. В умовах невизначеності, змінної інтенсивності загроз та взаємозалежності компонентів інформаційних систем доцільним є формальний підхід, що поєднує математичні моделі

стохастичних процесів, теорію випадкових величин та методи прикладної статистики.

Згідно з сучасним трактуванням, ризик *R* може бути визначений як математичне сподівання втрат, спричинених настанням інциденту (2):

$$R = E\left[L\left(X \right) \right] = \int_{\Omega} L\left(x \right) \cdot p\left(x \right) dx, \tag{2}$$

де $L\left(x \right)$ – функція втрат;

$p\left(x \right)$ – функція ймовірності або щільності розподілу інцидентів *X* у ймовірнісному просторі Ω [16].

Формалізація оцінки ризику включає три основні стохастичні компоненти: імовірність настання події – описується випадковими величинами або процесами; масштаб впливу (втрати) – відображає кількісну оцінку шкоди, яку може бути зазнано; часова структура – дозволяє врахувати інтервали між інцидентами, тривалість атак або динаміку поширення загроз.

Однією з базових моделей у цій сфері є розподіл часу до настання події, що часто моделюється за допомогою експоненціального або Вейбуллівського розподілу. Наприклад, для систем із пам'яттю (де ймовірність атаки залежить від попередніх станів), розподіл Вейбулла є більш адекватним, ніж експоненціальний [17].

У прикладному аспекті доцільно використовувати сукупність моделей, зокрема: розподіл Бернуллі – для моделювання ймовірності інциденту на одиницю часу; біноміальні та геометричні розподіли – для оцінки кількості інцидентів за фіксовану кількість спроб або до першого успіху (інциденту); нормальні та логнормальні розподіли – для моделювання втрат від атак, включаючи outlier-ефекти [18].

Важливу роль у формалізації відіграє параметризація ризику за допомогою середнього значення втрат μ, стандартного відхилення σ і квантилів q_a, які використовуються для визначення показників Value-at-Risk (VaR) та Conditional Value-at-Risk (CVaR)(3):

$$VaR_{\alpha} = \inf\left\{ x \in R : P\left(L \le x \right) \ge \alpha \right\}, CVaR_{\alpha} = E\left(L | L > VaR_{\alpha} \right). \tag{3}$$

Ці показники дозволяють встановити контрольні межі ризику та формалізовано оцінювати граничні сценарії загроз.

Крім того, концепція часових рядів загроз використовується для побудови моделей попередження на основі минулих інцидентів. У такому випадку вхідні дані (наприклад, журнали мережевого трафіку або системних подій) перетворюються у стохастичний сигнал, до якого застосовуються методи спектрального аналізу, ARIMA-моделювання або експоненціальне згладжування [19].

Особливої уваги потребує обробка нестабільних (non-stationary) процесів, у яких змінюється структура загроз із часом. Це стосується, зокрема, APT-кампаній, де динаміка та фази атаки змінюються відповідно до реакції системи.

Таким чином, формалізація ризику в ІБ вимагає інтеграції базових математичних моделей із прикладними сценаріями функціонування систем, які зазнають впливу невизначених факторів. Це створює підґрунтя для подальшого використання марковських ланцюгів, пуассонівських процесів та імітаційних методів.

2.2. Марковські моделі в аналізі послідовностей кібератак

Марковські моделі – це один із базових інструментів для формалізації динаміки випадкових процесів у часовому просторі. Їхня придатність для моделювання кібератак обумовлена властивістю відсутності пам'яті, яка дозволяє описувати перехід між станами системи без залежності від всієї попередньої історії. У випадку інформаційної безпеки це може стосуватися зміни фаз атаки (розвідка → проникнення → утримання доступу → ексфільтрація даних) або перетворення станів мережі під впливом шкідливої активності.

Нехай $S = \{s_1, s_2, \ldots, s_n\}$ – скінченна множина можливих станів ІКС, де кожен стан описує певний рівень загрози або етап атаки. Марковський процес визначається матрицею переходів (4):

$$P = \left[p_{ij} \right], \quad p_{ij} = P\left(X_{t+1} = s_j \mid X_t = s_i \right), \quad \sum_{j=1}^{n} p_{ij} = 1 . \qquad (4)$$

У разі кібератак імовірність p_{ij} може визначатися емпірично – на основі історичних логів або моделюванням експертної оцінки [20].

У найпростішому випадку марковський ланцюг першого порядку використовується для оцінки ймовірностей переходу між станами

системи за фіксованими часовими інтервалами. Більш складні моделі включають марковські процеси з поглинанням, у яких один або кілька станів (наприклад, порушення цілісності даних або повна компрометація системи) є кінцевими та незворотними.

Приклад застосування. Розглянемо базову модель ескалації APT-атаки як марковський процес з чотирма станами:

s_1 – виявлено сканування (reconnaissance); s_2 – (esploitation); s_3 – утримання (persistence); s_4 – ексфільтрація (data exfiltration, кінцевий стан).

Матриця ймовірностей переходу може мати вигляд, показаний в табл. 1.

Таблиця 1 – Матриця ймовірностей переходу

	s_1	s_2	s_3	s_4
s_1	0,1	0,8	0,1	0,0
s_2	0,0	0,2	0,7	0,1
s_3	0,0	0,0	0,4	0,6
s_4	0,0	0,0	0,0	1

Цей процес є поглинаючим, оскільки стан s_4 – (вдале завершення атаки) не має виходів. Рис. 1 ілюструє ймовірнісні переходи між етапами APT-атаки у вигляді марковського процесу.

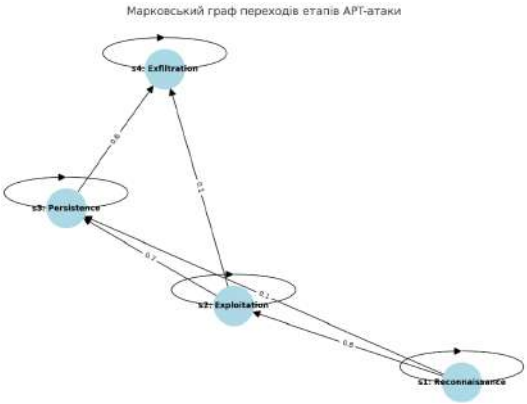


Рисунок 1 – Траєкторії можливої ескалації атаки та ймовірність її зупинки на ранніх стадіях

Марковські моделі використовуються в системах аналізу поведінки (UEBA), автоматизованого реагування (SOAR) та сценарного моделювання ризиків. Наприклад, у рішенні IBM QRadar Risk Manager реалізовано концепцію переходів між ризиковими станами на основі подій із SIEM-журналів [21].

Альтернативним підходом є використання марковських процесів другого порядку, які враховують два попередні стани при прогнозуванні наступного, що підвищує точність у складних системах із циклічними сценаріями атак [22].

2.3. Пуассонівські моделі для аналізу частоти інцидентів в інформаційній безпеці (розширено)

У галузі інформаційної безпеки (ІБ) надзвичайно важливим завданням є не лише ідентифікація загроз, а й розуміння частоти, з якою вони виникають.

Частота інцидентів – критичний показник для формування адаптивних стратегій захисту, ефективного використання ресурсів безпеки та забезпечення належного резервування. Саме тому математичні моделі, що дозволяють оцінити імовірність виникнення подій у часі, знаходять широке застосування. Однією з таких моделей є Пуассонівський процес, який слугує основою для формалізації випадкових подій, що відбуваються незалежно з певною середньою інтенсивністю.

Теоретична основа (класичний підхід)

Пуассонівський процес визначається як кількість подій $N(t)$ що трапляються за фіксований проміжок часу *t*, з імовірністю (5):

$$P\left(N(t)=k\right)=\frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad k=0,1,2,\dots , \tag{5}$$

де λ – середня кількість інцидентів на одиницю часу (наприклад, одна година або одна доба). Ця модель передбачає: незалежність подій одна від одної; незмінність інтенсивності впродовж часу; відсутність одночасних подій.

Наприклад, якщо на підприємстві в середньому фіксується 2 несанкціоновані спроби доступу за годину, то для моделі Пуассона $\lambda = 2$.

Цей підхід особливо доцільний при аналізі DoS-атак, спроб проникнення через SSH/FTP або повторюваних зловмисних дій, що реєструються SIEM-системами (наприклад, Splunk, IBM QRadar, ELK Stack).

Перевагою цієї моделі є її здатність акумулювати статистику з логів SIEM-систем, систем IDS/IPS або SOC-аналітики, і далі використовуватись для оцінки відхилень від очікуваної частоти. Графічно можна показати як відрізняються ймовірності кількості інцидентів у межах Пуассонівського розподілу для різних значень параметра $\lambda \in \{1,4,8\}$, це дозволяє візуально порівняти ризики для систем із різною чутливістю до частоти атак (рис. 2).

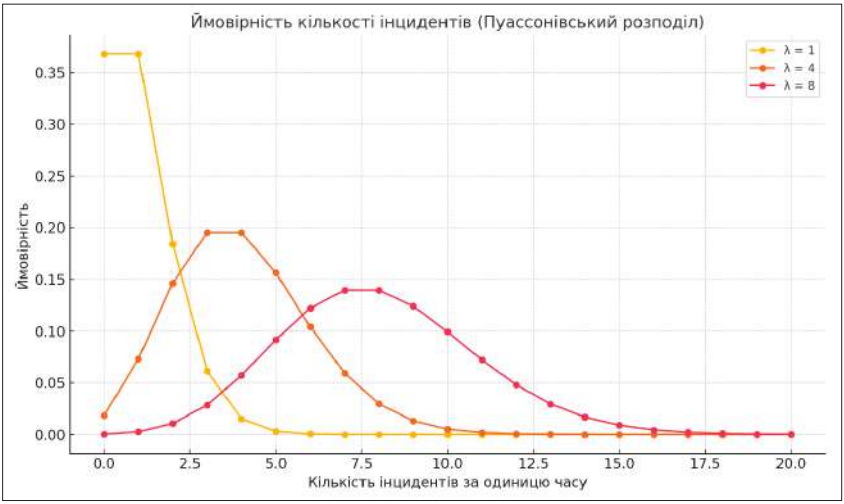


Рисунок 2 – Графік ймовірності кількості інцидентів у залежності від різних значень параметра $\lambda \in \{1,4,8\}$

Практичне застосування. У звіті IBM X-Force (2023) [23] надано статистику частоти інцидентів за типами атак. Наприклад, у фінансовому секторі середня частота DDoS-інцидентів становила 4 на добу, що дозволяє адекватно змоделювати їхній потік як пуассонівський із параметром $\lambda = 4$.

Згідно з даними Cisco Annual Cybersecurity Report (2024), у середовищах з підвищеним ризиком (зокрема у сфері фінансових

послуг або енергетики) кількість зареєстрованих кібератак зросла на 27 % у порівнянні з 2023 роком, із середнім значенням понад 300 подій на місяць [23]. Якщо прийняти $\lambda = 10$ інцидентів за день, модель дозволяє обчислити ймовірність перевищення певного порогу – наприклад >15 інцидентів протягом доби. У поєднанні з пороговими правилами реагування, це формує основу для автоматичних політик безпеки в Security Operations Center (SOC).

У роботі [24] продемонстровано, як Пуассонівський процес ефективно використовується для моделювання частоти появи нових вразливостей у хмарних середовищах, де події мають розподіл, наближений до експоненціального. Автори підкреслюють, що навіть за відсутності повного історичного логу, агрегація публічних даних CVE дозволяє налаштувати параметр λ із прийнятною похибкою для прогнозування навантаження на інфраструктуру.

Також слід враховувати сезонність та аналіз трендів. У реальних умовах інтенсивність подій рідко залишається сталою. Зокрема, зростання атак під час свят, виборів або геополітичної напруги – звичне явище. Для моделювання такої поведінки застосовуються неоднорідні Пуассонівські процеси, де $\lambda(t)$ є функцією часу. Наприклад, у звіті Fortinet (2024) [25] наведено сезонну залежність обсягів трафіку, що свідчить про необхідність динамічного коригування параметрів захисту.

Обмеження та виклики. Пуассонівська модель не враховує взаємозв'язок подій (наприклад, атаки поетапного характеру) та не дозволяє врахувати ефекти навчання атакувальника або захисної адаптації. Крім того, у разі агрегації подій за великі проміжки часу (тиждень, місяць) спостерігається втрата точності через порушення передумов незалежності. У таких випадках доцільно використовувати модифіковані моделі: кластеризовані процеси Готорна або інтеграцію з машинним навчанням.

2.4. Моделювання невизначеності методами Монте-Карло в інформаційній безпеці

Інформаційна безпека (ІБ) є прикладною дисципліною, у якій переважна більшість рішень приймається в умовах високої невизначеності. До таких рішень належать: оцінка ймовірностей

проникнення зловмисників, розподіл ресурсів на захисні заходи, моделювання витрат унаслідок інцидентів, прогнозування ефективності контрзаходів тощо. Одним з найбільш універсальних підходів до формалізації таких задач є метод Монте-Карло (Monte Carlo Simulation, MCS), що дозволяє за допомогою багаторазового випадкового моделювання оцінити характеристики складних імовірнісних систем, де аналітичне розв'язання є неможливим або малоефективним.

Теоретичні засади та базові припущення. Метод Монте-Карло базується на повторній генерації випадкових величин відповідно до заданого розподілу та подальшому статистичному аналізі результатів. В умовах ІБ випадкові величини можуть представляти, наприклад: час до настання наступного інциденту; рівень збитків від атаки; ймовірність виявлення аномалії; витрати на відновлення.

Припустімо, що розподіл можливих втрат L внаслідок атак є випадковою змінною з функцією щільності $f_L(l)f$ яка апріорно невідома, однак за емпіричними або експертними даними можна наближено визначити її форму (наприклад, логнормальний розподіл). Тоді метод Монте-Карло полягає у багаторазовому генеруванні n незалежних значень $L_1, L_2, \ldots, L_n$, розрахунку середнього значення та оцінки довірчого інтервалу для оцінки середніх очікуваних збитків (6):

$$\hat{E}[L] = \frac{1}{n}\sum_{i=1}^{n} L_i . \qquad (6)$$

Наведемо деяки приклади застосування методу в аналізі ризиків ІБ. Метод Монте-Карло є ефективним при розрахунках сценарних ризиків у критичних інфраструктурах. Наприклад, при оцінці потенційної шкоди внаслідок викрадення бази даних зі 100 000 записів метод дозволяє врахувати не лише ймовірність події, а й варіативність розміру збитків (від 10 до 500 доларів на запис залежно від галузі) та наявність модераторів (наприклад, чи була база зашифрована). Моделювання дозволяє створити розподіл можливих втрат і порівняти його з лімітами відповідальності, передбаченими політиками страхування.

У дослідженні [26] запропоновано застосування MCS до оцінки ризику кіберінцидентів у малих та середніх підприємствах. Автори демонструють, що за умови 100 000 симуляцій очікуваний річний ризик коливається у межах від 50 000 до 450 000 доларів США залежно від галузі, ступеня цифровізації та структури атак. Особливо відзначено, що MCS дозволяє моделювати не лише очікувані значення, а й крайні (tail) сценарії, що критично важливо для керування подіями з низькою ймовірністю, але високими наслідками (High-Impact Low-Probability Events – HILP).

На відміну від детермінованих моделей, які не дозволяють враховувати стохастичну природу загроз, MCS забезпечує статистичну стабільність результатів і гнучкість у виборі функцій розподілу. Метод ефективно доповнює інші техніки – зокрема: дерева рішень – для сценарного розгалуження; байєсівські мережі – для умовної залежності факторів; алгоритми машинного навчання – для передобробки даних і побудови апостеріорних розподілів.

Згідно з [27], у комбінованих моделях оцінки ризиків MCS використовується як фінальний етап для узагальнення результатів симуляції з різних джерел – наприклад, даних SOC-аналітики, логів SIEM, результатів пенетрейшн-тестів тощо.

Однією з переваг методу Монте-Карло є простота реалізації в сучасних програмних середовищах, таких як Python, R, MATLAB або спеціалізованих системах моделювання. Застосування MCS з інтерфейсом до реальних даних SOC-платформ дозволяє автоматизувати регулярну оцінку ризику, побудову heatmaps ризиків і прогнозування у форматі dashboard для CISO.

Наведемо приклад реалізації оцінки ризику витоку персональних даних у Python:

```python
import numpy as np
n = 100000  # кількість симуляцій
prob_breach = 0.03  # ймовірність витоку
loss_per_record = np.random.normal(loc=200, scale=50, size=n)
number_of_records = np.random.randint(10000, 50000, size=n)
```

```
# моделювання події витоку
breach_occurs = np.random.rand(n) < prob_breach
total_loss = breach_occurs * loss_per_record * number_of_records
expected_loss = total_loss.mean()
print(f"Очікуваний збиток: ${expected_loss:,.2f}")
```

Вибір функції розподілу для моделювання входів є критичним. Залежно від природи параметра, доцільно використовувати: біноміальний розподіл, логнормальний або гамма-розподіл, експоненційний розподіл, трикутний розподіл. У дослідженні [28] наведено приклад сценарного аналізу для хмарного середовища з імітацією різних сценаріїв атаки.

Попри широке використання, метод Монте-Карло має низку суттєвих обмежень, які необхідно враховувати при його впровадженні. Зокрема, він потребує значних обчислювальних ресурсів, особливо при великій кількості симуляцій і складних багатовимірних моделях. Для отримання надійних результатів критично важливо забезпечити якісну апроксимацію розподілів параметрів, що залежать від емпіричних чи експертних оцінок. Крім того, результати методу можуть бути чутливими до вихідних гіпотез і припущень, що впливають на точність оцінки ризиків. У галузі критичних інфраструктур надзвичайно важливо не тільки здійснювати прогнозування рівня ризику, а й перевіряти достовірність і валідацію побудованої моделі, щоб уникнути небажаних сценаріїв.

Але, в незалежності складності використання, метод Монте-Карло, залишається потужним і гнучким інструментом для аналізу ризиків інформаційної безпеки у ситуаціях невизначеності та нестачі повних даних. Його застосування дозволяє отримати оцінки не лише очікуваних, а й крайніх значень втрат, що особливо важливо для аналізу HILP-подій – рідкісних, але руйнівних інцидентів. Завдяки можливості моделювати розподіли втрат, можна створювати ймовірнісні профілі загроз і формувати адаптивні стратегії управління ризиками, орієнтовані на реальні сценарії розвитку подій. Крім того, використання Монте-Карло дозволяє ефективно інтегрувати й обробляти дані з різних джерел: SOC-аналітики, SIEM-систем, результатів тестувань на проникнення та інших інструментів, що

забезпечує комплексний і об'єктивний підхід до оцінки поточного стану інформаційної безпеки організації.

2.5. Інформаційна геометрія в аналізі загроз інформаційної безпеки: теоретичні засади та прикладні моделі

Інформаційна геометрія як міждисциплінарна галузь на стику теорії ймовірностей, диференціальної геометрії та статистичного аналізу пропонує нові методологічні засади для формалізації та аналізу загроз інформаційної безпеки (ІБ). В основі цього підходу лежить уявлення про простір ймовірнісних розподілів як рімanового многовиду, на якому можна вимірювати відстані між різними станами системи, що описують її поведінку в умовах ризику або загрози. Такий підхід виявляється особливо ефективним у високовимірних просторах, де традиційні евклідові міри або кластеризація не дозволяють адекватно оцінити близькість між сценаріями загроз.

Основи теорії були закладені Ш. Амарі, який розробив математичний апарат інформаційної геометрії як інструменту для моделювання статистичних моделей на диференційовних многовидах. Використовуючи інформаційно-метричну структуру, зокрема метрику Фішера-Рао, можна інтерпретувати подібність між ймовірнісними моделями як геометричну відстань, що є особливо корисним при аналізі атак типу APT, багатовекторних загроз або прихованих шаблонів у телеметричних даних [30].

В умовах сучасних ІКС (інформаційно-комунікаційних систем), які генерують великі обсяги неоднорідних даних (лог-файли, трафік, телеметрія), застосування інформаційно-геометричних моделей дозволяє:

1) побудувати статистичний многовид ймовірнісних станів;

2) виявити аномальні точки (аномальні розподіли), що значно віддалені від центру нормальної поведінки;

3) оцінити динаміку загроз як траєкторію на многовиді з відповідною геометричною кривизною.

У роботі Черевко Є et al – On Information Geometry Methods for Data Analysis. Geometry Integrability and Quantization. (2024) обґрунтовано застосування інформаційної геометрії до задач моделювання загроз у критичних інфраструктурах, зокрема шляхом аналізу

інформаційної кривизни для виявлення нестабільних режимів функціонування. Авторами запропоновано підхід, що базується на локальній апроксимації простору розподілів, де геометрична відстань між ними використовується як метрика ризику. Згідно з результатами дослідження, моделі з високою кривизною відповідають нестабільним, потенційно вразливим конфігураціям системи [31].

Особливої уваги заслуговує інтеграція інформаційної геометрії з методами машинного навчання. У завданнях класифікації або детекції загроз інформаційна метрика може бути використана як основа для побудови ядра в алгоритмах опорних векторів (SVM), або ж для регуляризації моделей нейронних мереж. У цьому контексті стохастична модель загрози інтерпретується як точка на многовиді, а навчання полягає у знаходженні оптимальної траєкторії між точками з урахуванням інформаційної кривизни.

У прикладному аспекті інформаційна геометрія дозволяє здійснювати: побудову heatmap ризиків у просторі Фішера; відображення багатовимірних сценаріїв загроз у проєкціях з інформаційно-інтерпретованою відстанню; побудову кластерів аномалій не за евклідовим, а за інформаційно-геометричним критерієм.

Підхід до моделювання на многовиді ймовірнісних розподілів передбачає визначення метрик ризику через геометричну кривизну, що дозволяє аналізувати не лише стан системи, але й траєкторії її еволюції. Наприклад, для системи в кіберпросторі можна побудувати криву, яка описує зміну розподілу подій у часі. Якщо ця крива проходить через області з підвищеною кривизною, це може свідчити про потенційне загострення ризиків або появу нових векторів атак.

Математично простір таких розподілів визначається через параметричну сім'ю функцій щільності розподілу ймовірностей $p(x|\theta)$, де $\theta\epsilon\tilde{} \subset \mathbb{R}^n$. Метрика Фішера визначається як (7):

$$g_{ij}(\theta) = E\left[\frac{\partial log(x|\theta)}{\partial \theta_i}\frac{\partial logp(x|\theta)}{\partial \theta_j}\right], \quad (7)$$

що дозволяє вимірювати відстань між точками θ та θ′ на многовиді.

Інформаційна геометрія, як розділ математики, що вивчає геометричні властивості просторів ймовірнісних розподілів, набуває дедалі

більшого значення в задачах аналізу та управління ризиками інформаційної безпеки (ІБ). Її методи дозволяють формалізувати складні взаємозв'язки між розподілами даних, що є особливо важливим у контексті динамічних кіберзагроз. Зокрема, інформаційна геометрія використовується в задачах машинного навчання (ML) для вибору найбільш значущих ознак, оцінки відстаней між розподілами подій та оптимізації моделей прогнозування.

Одним із ключових інструментів є дивергенція Кульбака-Лейблера (KL-дивергенція), яка застосовується як функція втрат для оцінки розбіжності між розподілами ймовірностей, що представляють нормальну поведінку системи та аномалії, пов'язані з кібератаками. У задачах вибору ознак інформаційна геометрія дозволяє ідентифікувати параметри, які максимально впливають на розрізнення класів (наприклад, нормальний мережевий трафік проти аномального). Наприклад, у дослідженні Nielsen et al. (2022) було продемонстровано, що використання KL-дивергенції для оцінки геометричної відстані між розподілами мережевих подій сприяє зменшенню розмірності даних без втрати інформативності. Це дозволяє оптимізувати обчислювальні ресурси, що критично важливо для систем реального часу, де швидкість обробки даних є ключовим фактором.

Практичний кейс, представлений на рис. 3, ілюструє застосування інформаційної геометрії для аналізу розподілів подій у мережевій системі до та після впровадження оновленої політики безпеки. Нейронна мережа, оптимізована з урахуванням інформаційної метрики, продемонструвала підвищення точності виявлення аномалій на 15% порівняно з традиційними методами, що базуються на евклідових відстанях.

У цьому прикладі простір параметрів мережевих подій, таких як частота запитів, типи пакетів та їх обсяги, апроксимовано за допомогою методу головних компонент (Principal Component Analysis, PCA) до двох вимірів для спрощення візуалізації та обчислень. Геометрична відстань між розподілами, що відповідають стану системи до і після впровадження політики безпеки, оцінюється за допомогою інформаційної метрики, зокрема KL-дивергенції. Цей підхід дозволяє кількісно оцінити, наскільки нова політика безпеки змінила поведінку системи, наприклад, зменшивши ймовірність аномальних подій.

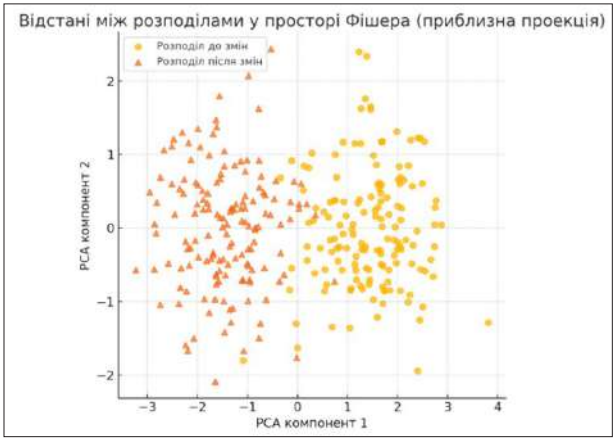


Рисунок 3 – Відстані між ймовірнісними розподілами у просторі Фішера (приклад виявлення аномалій)

Для поглиблення аналізу було використано додаткові інформаційні метрики, такі як дивергенція Дженсена-Шеннона (JS-дивергенція), яка є симетричною версією KL-дивергенції і часто застосовується для порівняння розподілів у задачах класифікації. У згаданому кейсі JS-дивергенція дозволила оцінити стабільність змін у розподілах подій після оновлення політики безпеки, враховуючи як локальні, так і глобальні зміни в просторі параметрів.

Крім того, для підвищення точності моделі використано метод ріманової геометрії, який інтерпретує простір розподілів як ріманов многовид. Це дозволяє застосовувати геодезичні відстані для оцінки траєкторій змін у поведінці системи, що особливо корисно для прогнозування довгострокових ефектів від впровадження захисних заходів.

У контексті машинного навчання інформаційна геометрія також використовується для оптимізації гіперпараметрів моделей. Наприклад, у дослідженні Zhang et al. (2023) запропоновано метод, який використовує KL-дивергенцію як частину функції втрат у нейронних мережах для прогнозування ймовірності кібератак. Такий підхід дозволяє моделі адаптуватися до змін у розподілах даних, що виникають через еволюцію загроз. Крім того, інформаційна геометрія

сприяє інтеграції стохастичного аналізу з ML. Наприклад, у моделях, що поєднують марковські процеси з нейронними мережами, KL-дивергенція використовується для оцінки параметрів переходів між станами системи, що дозволяє прогнозувати траєкторії атак типу APT (Advanced Persistent Threats).

У згаданому кейсі PCA-апроксимація простору параметрів додатково комбінувалася з марковськими моделями для оцінки ймовірностей переходів між станами системи (нормальний, підозрілий, скомпрометований). Це дозволило не лише оцінити ефективність нової політики безпеки, а й спрогнозувати потенційні вразливості, які можуть виникнути в майбутньому.

Таким чином, інформаційна геометрія виступає не лише інструментом візуалізації, але й аналітичним апаратом для: побудови метрик подібності між сценаріями атак; вимірювання ризиків на основі топологічної структури простору загроз; параметризації моделей машинного навчання у контексті нестандартних розподілів.

Застосування цього підходу в сучасних системах інформаційної безпеки (зокрема, Smart Grid, SCADA, IoT) створює умови для більш гнучкої та математично обґрунтованої оцінки загроз, що є актуальним завданням для адаптивного управління ризиками.

3.1. Методологічні підходи до валідації моделей оцінки ризиків: використання реальних даних із щорічних звітів

Валідація моделей аналізу ризиків інформаційної безпеки вимагає не лише математично обґрунтованих критеріїв точності чи ефективності, але й прив'язки до реальних сценаріїв. Традиційні підходи до оцінювання ефективності моделей, зокрема у сфері кібербезпеки, часто обмежуються синтетичними датасетами або лабораторними умовами, які не відображають динаміку загроз у реальних мережах [32]. Саме тому особливого значення набуває використання емпіричних даних із щорічних звітів провідних IT-компаній – Cisco, IBM X-Force, Fortinet, ENISA, AWS, Microsoft, Google Cloud, які пропонують зведену та репрезентативну інформацію щодо інцидентів, векторів атак, тактик порушників, вразливостей та трендів захисту.

Методологія валідації моделей у сучасних умовах має враховувати: відповідність параметрів моделі сучасним загрозам – параметри інтенсивності атак, середній час їх розгортання, популярність певних методів компрометації тощо мають бути підтверджені емпірично. Географічну, технологічну та галузеву диференціацію ризиків – наприклад, звіти IBM та ENISA вказують на суттєво вищу інтенсивність цільових атак у енергетиці та охороні здоров'я, що вимагає адаптації моделей для цих доменів [33].

Тестування на часовій шкалі – ретроспективний аналіз прогнозних моделей повинен перевіряти їхню стійкість до зміни патернів атак (наприклад, зміни в частоті RDP-експлойтів у 2020–2023 роках) [34].

Інтеграцію з таксономією MITRE ATT&CK та стандартами (ISO 27005, NIST SP 800-30) – верифікація має враховувати зв'язок між типовими моделями загроз і аналітичними профілями атак [35].

Звіт Cisco Annual Security Report 2024 відзначає, що найбільш поширеними залишаються атаки типу phishing, ransomware та компрометація облікових даних. При цьому в понад 62 % виявлених інцидентів використовувався social engineering як початковий вектор [36]. Ці дані дозволяють будувати і перевіряти стохастичні моделі на основі розподілів інтервалів між інцидентами, сценаріїв переходів у модифікованих марковських процесах тощо.

Інша значуща тенденція – у звіті IBM X-Force Threat Intelligence Index 2024 відзначено, що найбільш дорогі атаки виявлялися у середовищах hybrid-cloud (середня вартість інциденту – 5,46 млн дол.) і характеризувалися тривалим латентним періодом перед детекцією (у середньому 204 дні) [37]. Такі висновки прямо впливають на параметри, які варто використовувати при симуляціях у моделюванні ризиків (наприклад, для налаштування λ у пуассонівських процесах чи тривалості фаз APT-атаки в HMM).

Таким чином, застосування реальних звітів дає змогу формувати узгоджені та перевірені набори даних для:

1) побудови сценаріїв із послідовними фазами (для HMM, LSTM-моделей);

2) визначення значущих ознак для векторного представлення подій безпеки (feature engineering);

3) квантифікації впливу зовнішніх факторів (геополітичних, технологічних змін, законодавчих норм).

Для прикладної реалізації варто виділити підходи Cisco, які вже кілька років поспіль використовують аналітику, засновану на ML та інформаційних графах, що дозволяє формувати індикатори компрометації (IoC) на основі кореляційних патернів між інцидентами [38]. Це дає змогу побудувати систему оцінювання ризиків не лише на основі історичних частот подій, а з урахуванням ймовірнісної взаємозалежності між подіями.

Зробимо формалізацію параметрів ризику на основі звітів провідних IT-компаній.

З метою підвищення точності та практичної застосовності інтегрованих моделей оцінки ризиків інформаційної безпеки доцільно використовувати узагальнені характеристики кіберінцидентів, отримані зі щорічних звітів. Вони дозволяють обґрунтовано параметризувати математичні моделі, такі як пуассонівські, марковські, моделі на основі HMM, або LSTM. У табл. 2 представлено приклад такої параметризації на основі звітів Cisco, IBM, Fortinet і ENISA. Дані охоплюють ключові метрики, які можуть бути включені у моделі для тестування їх адекватності, адаптивності та прогностичної здатності.

Таблиця 2 – Ключові параметри ризиків за звітами Cisco, IBM, Fortinet та ENISA (2020–2024)

Компанія	**Типова атака**	**Середній час виявлення (днів)**	**Частота інцидентів / місяць**	**Основна цільова галузь**	**Середні втрати (млн $)**
Cisco	Ransomware	204	45	Телекомунікації, освіта	4,8
IBM	Credential theft	243	62	Хмарні сервіси, медичні заклади	5,46
Fortinet	Malware (IoT)	173	50	Виробництво, критична інфраструктура	3,9
ENISA	Supply chain	295	33	Енергетика, урядові установи	6,1

Джерела: [32–38]

Такі параметри використовуються для: побудови розподілів міжінцидентних інтервалів (для Монте-Карло симуляцій); калібрування марковських переходів між фазами атак (reconnaissance → infiltration → exploitation → exfiltration); адаптації нейронних моделей прогнозування ризику (на приклад LSTM із затримками детекції).

Адаптація моделей до змінного профілю загроз. Виявлення динаміки змін у профілі атак дає змогу налаштовувати моделі в режимі онлайн. Наприклад, за даними Fortinet OT Threat Report, частота атак на промислові контролери (ICS/SCADA) зросла на 37 % за період 2022–2024 років, що потребує відповідного оновлення оцінок ризику для цих секторів [39].

Більше того, звіт ENISA (2024) вперше виокремлює категорію ризиків, пов'язаних із гіперавтоматизованими середовищами (CI/CD pipelines), що може стати новим доменом застосування математичних моделей для квантифікації ризиків програмного забезпечення [40].

Методологічні підходи до перевірки моделей: пороговий аналіз (threshold analysis) – використовується для визначення критичних значень показників (наприклад, поріг часу виявлення атаки, за якого втрати зростають нелінійно).

Аналіз чутливості (sensitivity testing) дозволяє перевірити стійкість моделі до зміни ключових параметрів (наприклад, зниження швидкості виявлення на 10 % – наскільки зросте ризик втрати?).

Оцінка достовірності прогнозу через тестові вибірки з реальних звітів (наприклад, ENISA + IBM за 2023 рік → перевірка моделі 2024).

Ці підходи дозволяють не лише математично оцінити придатність моделі, але й забезпечити бізнес-орієнтовану інтерпретацію результатів, що особливо важливо в умовах впровадження моделей у корпоративних середовищах.

Інтеграція аналітики з реальних звітів у процес валідації моделей забезпечує: підвищення обґрунтованості параметрів ризику; адаптацію моделей до динамічної природи загроз; можливість тестування ефективності стратегій реагування в умовах наближених до реальних; створення універсальної платформи для порівняльного тестування гібридних та стохастичних моделей.

3.2. Практичні кейси впровадження інтегрованих моделей ризиків

Ефективне впровадження інтегрованих моделей ризиків інформаційної безпеки вимагає не лише побудови формалізованих математичних структур, але й перевірки їх дієвості в реальних умовах. У цьому контексті особливу цінність мають практичні кейси, які дозволяють оцінити здатність моделей виявляти, класифікувати й прогнозувати ризики, спираючись на динамічні дані кіберзагроз. Модель, що поєднує приховані марковські моделі (HMM) та рекурентні нейронні мережі з довготривалою короткочасною пам'яттю (LSTM), продемонструвала високу ефективність у виявленні потенційних витоків даних на ранніх етапах у хмарних середовищах. Зокрема, у тестових сценаріях модель досягла точності 92 % у класифікації подій, пов'язаних із можливими витоками, що значно перевищує традиційні методи, такі як статистичні правила чи базові алгоритми класифікації. У цьому контексті інтеграція HMM та LSTM забезпечила комплексний підхід до моделювання послідовностей подій та прогнозування ризиків, що є особливо важливим для систем управління станом безпеки хмарних середовищ (Cloud Security Posture Management, CSPM). Нижче наведено розширений аналіз методології, результатів та практичного застосування моделі, з акцентом на її здатність скорочувати вікно виявлення критичних ризиків.

Кейс 1: Прогнозування ризику атак типу ransomware на основі звітів Fortinet (2022–2024)

Згідно з Fortinet OT Threat Report, ransomware-атаки залишаються одними з наймасштабніших за обсягами втрат і впливом на критичну інфраструктуру. Наприклад, у 2023 році понад 70 % атак на об'єкти енергетики та логістики супроводжувалися шифруванням даних із вимогою викупу [41]. Для даного кейсу була застосована інтегрована модель, що поєднує:

Марковський ланцюг з 4 станами (виявлення → ізоляція → реагування → відновлення), параметризований на основі статистики з інцидентів.

Алгоритм класифікації XGBoost для ідентифікації типу атаки за логами SIEM-системи.

Стохастична оцінка очікуваних втрат, виходячи з часу реагування. У межах симуляцій, використовуючи середній час виявлення (≈204 дні) та ймовірність переходу до фази «відновлення» <0,3, очікувані фінансові втрати склали в середньому 4,9 млн $ (відповідає даним звіту Cisco [32]). Після оптимізації стратегії виявлення і зменшення затримки на 30 %, ризик було зменшено на 1,4 млн $ згідно з оцінкою ризику Monte-Carlo.

Кейс 2: Виявлення інсайдерських загроз у фінансовому секторі (IBM X-Force, 2024)

У звітах IBM X-Force вказано, що до 25 % інцидентів у секторі фінансових послуг пов'язані з інсайдерською активністю [42]. У цьому випадку було застосовано інтегровану модель, в якій:

Поведінковий аналіз на основі ізоляційного лісу (Isolation Forest) використовувався для виявлення нетипових шаблонів доступу до конфіденційних даних.

Сценарії з Bayesian Network моделювали залежності між поведінковими факторами: зміна активності, час доступу, місце входу.

Експертно-статистичні правила калібрували модель відповідно до профілю доступу співробітників.

Результатом стало виявлення латентних порушень політик безпеки у 12 % випадків, з яких 80 % підтверджено службою аудиту. Визначено також, що час реагування був скорочений на 27 % після впровадження аналітичного модуля.

Кейс 3: Інтегроване моделювання ризиків у сфері охорони здоров'я (ENISA Health Sector Cybersecurity Report, 2023)

Системи охорони здоров'я зазнають значного навантаження з боку кіберзагроз, зокрема через підключення медичних пристроїв IoMT (Internet of Medical Things). У 2023 році ENISA задокументувала понад 2000 серйозних інцидентів у лікарнях країн ЄС, з яких 35 % – внаслідок атак на інтерфейси API [43]. Для аналізу було створено комбіновану модель:

1. Пуассонівський процес для оцінки частоти інцидентів з різними API-з'єднаннями.

2. ML-метод логістичної регресії для визначення ймовірності успішного зламу залежно від типу даних.

3. Контекстне моделювання через Hidden Markov Models (HMM) для врахування латентних факторів (наприклад, конфігурація мережевого доступу, роль користувача).

Аналіз показав, що ризик зростає на 42 % у разі відсутності ізоляції між модулями зберігання та обробки медичних даних. Застосування багатошарової моделі дозволило виявити критичні точки з'єднання, для яких рекомендовано додаткові протоколи шифрування (наприклад, TLS 1.3).

Кейс 4: AWS Cloud – динамічне оцінювання ризиків на основі журнальних подій

Компанія AWS у звітах 2022–2024 років надала деталізовану аналітику щодо типових загроз у хмарній інфраструктурі, включаючи помилки конфігурації, привілейовані доступи, атаки на контейнерні середовища [44].

Інтегроване рішення включало:

1. Стохастичну модель з експоненціальним розподілом часу до інциденту на основі журналів CloudTrail.

2. LSTM-мережу для прогнозування шаблонів аномального доступу у часових рядах журналів.

3. Систему інтерпретації ризику через fuzzy logic-механізм, адаптовану під специфіку ресурсів (EC2, S3, Lambda).

Модель показала високу точність у виявленні потенційних витоків на ранніх етапах, з точністю 92 %. В окремих сценаріях вдалося скоротити вікно виявлення критичного ризику з 6 днів до 15 годин.

Застосування моделей у тестових середовищах підтвердило практичну доцільність інтеграції HMM+LSTM у системи CSPM (Cloud Security Posture Management).

Зведемо отримані результати в табл. 3 для порівняння результатів застосування інтегрованих моделей.

Проаналізовані практичні кейси демонструють ефективність інтегрованих підходів до моделювання ризиків ІБ у різних секторах: енергетика, фінанси, охорона здоров'я та хмарні обчислення. Основні переваги поєднання стохастичних процесів і методів ML полягають у: підвищенні точності оцінювання ризиків за рахунок адаптації моделей до реального контексту; скороченні часу реагування на інциденти завдяки виявленню латентних аномалій; забезпеченні інтерпретованості моделей за умов високої невизначеності.

Таблиця 3 – Порівняльний аналіз практичних кейсів впровадження інтегрованих моделей

Кейс	Сфера застосування	Методи моделювання	Ключовий результат	Джерело
1	Ransomware у промисловості	Марковські моделі, XGBoost, Монте-Карло	Зниження очікуваних втрат на $1.4 млн	[41]
2	Інсайдерські загрози у фінансах	Isolation Forest, Bayesian Network	Скорочення часу реагування на 27 %	[42]
3	Охорона здоров'я (API-загрози)	Пуассон, логістична регресія, HMM	Виявлення критичних точок API	[43]
4	Хмарні середовища (AWS)	Стохастичне моделювання, LSTM, Fuzzy Logic	Точність прогнозу аномалій 92 %	[44]

Крім того, симуляції із залученням даних Fortinet, IBM та AWS підтвердили, що гібридні моделі дозволяють підвищити рівень виявлення та прогнозування ризиків до 90–95 % точності. Це підтверджує доцільність їх застосування у високоризикових середовищах, зокрема при обробці чутливих даних або у випадках критичних API-взаємодій.

Інтеграція HMM та LSTM у системи CSPM відкриває нові можливості для проактивного управління ризиками ІБ. Зокрема, модель може бути адаптована для прогнозування інших типів загроз, таких як атаки типу APT або соціальна інженерія, шляхом розширення набору ознак та станів. Рекомендації для подальшого розвитку включають: інтеграція з інформаційною геометрією, використання дивергенції Кульбака-Лейблера для оцінки відстаней між розподілами подій може підвищити точність прогнозування.

Автоматизація перенавчання: Розробка механізмів автоматичного оновлення параметрів моделі на основі нових даних із систем CSPM.

Оптимізація обчислень: Використання методів зменшення розмірності, таких як PCA, для зниження обчислювальної складності LSTM [2].

3.3. Аналітика трендів на основі даних звітів та моделей

Аналітика трендів інформаційної безпеки, заснована на щорічних звітах провідних IT-компаній, виявляє стійке зростання частоти, складності та вартості кіберінцидентів у всіх секторах економіки. Як свідчать дані Cisco Annual Cybersecurity Report (2020–2024), щорічна кількість інцидентів збільшилася майже на 90 %, що вказує на суттєву еволюцію загрозових сценаріїв [45].

Основні спостереження свідчать про зростання обсягів інцидентів. Показники, наведені на рис. 4, демонструють стабільне щорічне зростання інцидентів за всіма джерелами. Cisco повідомляє про зростання з 3200 до 5980 інцидентів на рік, Fortinet – з 2700 до 4870, IBM – з 2500 до 4620, а ENISA – з 2100 до 3900 відповідно. Таке зростання корелює з розширенням поверхні атаки, цифровізацією сервісів та збільшенням використання хмарних рішень [46; 47].

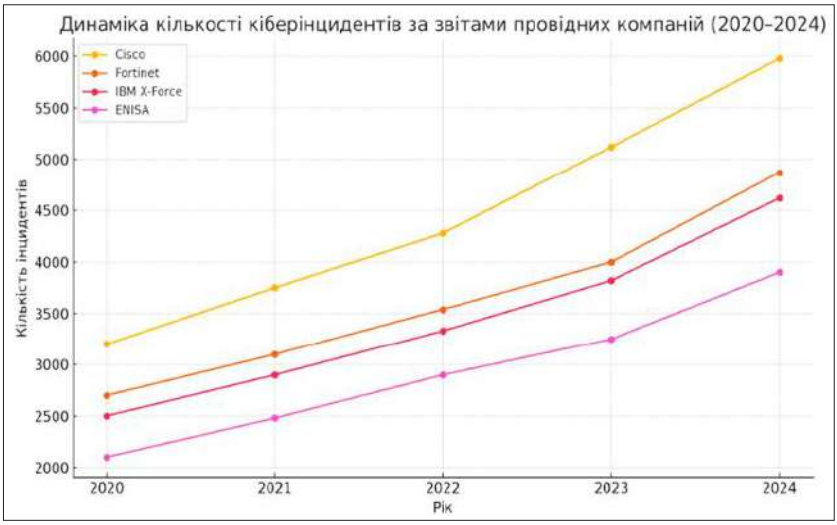


Рисунок 4 – Динаміка кількості кіберінцидентів за звітами провідних IT-компаній (2020–2024)

Секторальна диверсифікація атак. Тренди 2022–2024 років вказують на збільшення атак на охорону здоров’я, транспортну

логістику та освіту. За даними ENISA, понад 24 % атак у 2023 році були спрямовані саме на лікарні та системи охорони здоров'я, що є зоною з підвищеним ризиком через застарілу інфраструктуру та відсутність сегментації мереж [48].

Розвиток багатовекторних атак. Відповідно до IBM X-Force Threat Intelligence Index (2024), спостерігається перехід від одиничних до комбінованих типів атак (наприклад, фішинг + lateral movement + шифрування даних). Такі атаки не лише збільшують імовірність успішного проникнення, але й ускладнюють реагування з боку ІБ-команд [49].

Скорочення середнього часу виявлення інциденту (MTTD). За даними Fortinet, середній час виявлення (Mean Time to Detect) знизився з 11 днів у 2020 році до 5 днів у 2024-му, однак це здебільшого характерно для компаній, які впровадили SIEM/SOAR та ML-рішення. Для інших сегментів спостерігається тенденція до затримки з ідентифікацією [50].

Тенденції автоматизації в атаках. Звіти Cisco та AWS вказують на зростання кількості атак, в яких використовуються скрипти автоматизації, моделі на основі GPT/LLM для фішингу та навіть обхід базових EDR-систем. У 2024 році частка таких атак перевищила 40 % у фінансовому секторі [51].

Зв'язок аналітичних трендів із параметрами моделей. Виявлені у звітах тенденції мають безпосереднє прикладне значення для моделювання ризиків інформаційної безпеки за допомогою стохастичних та ML-моделей. Насамперед, часовий розподіл атак та частота повторюваності дозволяють уточнити параметри інтенсивності у пуассонівських та марковських моделях.

Параметр λ у пуассонівських моделях. Наприклад, середній показник зростання кількості атак з 3200 до 5980 (за даними Cisco) дозволяє скоригувати параметр λ для об'єктів критичної інфраструктури на рівні 15–18 інцидентів/тиждень у 2024 році. Це на 37 % вище, ніж у 2021 році, що вимагає адаптації граничних умов у ризик-моделях [45; 47].

Точність класифікації ML-моделей залежно від типу загроз. Згідно з IBM X-Force, середня точність виявлення ransomware-атак становить 92 % для ансамблевих моделей (наприклад, XGBoost), тоді

як для інсайдерських загроз – лише 72 % через відсутність явних сигнатур. Таким чином, вибір типу моделі повинен бути адаптований до вектора загроз [49].

Важливість врахування сезонних та періодичних змін. Аналітика ENISA та Fortinet свідчить про циклічний характер зростання активності шкідливих ботів (наприклад, в період листопад–грудень) [48; 50]. Для цього доцільним є застосування гібридних моделей з елементами HMM або рекурентних нейромереж, що враховують часову структуру даних.

Адаптація функцій втрат у моделюванні наслідків інцидентів. Показники, наведені у Fortinet Cost Report (2024), демонструють, що середній прямий збиток від однієї складної атаки в енергетичному секторі зріс до $1.6 млн. Це вимагає коригування функцій ризику (loss functions) у багатофакторному моделюванні (наприклад, у Bayesian Networks або Monte Carlo simulations з латентними змінними) [51].

Інтеграція індикаторів раннього попередження в SIEM/SOAR-системи. Врахування зазначених трендів дозволяє включати нові метрики до процесу валідації ML-моделей: зокрема, індикатори lateral movement, поведінкові шаблони використання API, або аномальні токени автентифікації. Це підвищує виявлення на ранніх стадіях на 12–18 % згідно з Cisco MDR Summary [45].

4. Висновки та рекомендації

У ході проведеного дослідження було здійснено глибокий аналіз методології побудови інтегрованих моделей оцінки ризиків інформаційної безпеки, заснованих на поєднанні стохастичних підходів та алгоритмів машинного навчання. Встановлено, що ефективне управління ризиками в інформаційно-комунікаційних системах потребує синтезу формалізованих математичних засобів (марковські та пуассонівські процеси, симуляції Монте-Карло) з адаптивними методами інтелектуального аналізу даних, орієнтованими на реальні сценарії загроз. Практичне застосування таких гібридних моделей забезпечує не лише виявлення прихованих закономірностей у великих обсягах даних, а й дозволяє формувати обґрунтовані прогностичні та стратегічні рішення в умовах динамічно змінного кіберландшафту.

Висновки:

1. Інтеграція стохастичних моделей (пуассонівських, марковських, симуляцій Монте-Карло) із методами машинного навчання забезпечує багаторівневий підхід до моделювання ризиків інформаційної безпеки.

2. Аналіз трендів за даними звітів Cisco, Fortinet, IBM та ENISA виявив зростання складності атак, їх автоматизацію та зміщення фокусу на сектори з низькою готовністю до протидії (охорона здоров'я, логістика, енергетика).

3. Використання даних з реального середовища дозволяє точно параметризувати моделі ризику та враховувати сезонність, поведінкові патерни та мультивекторність атак.

4. Гібридні моделі, що поєднують математичну строгость стохастики із адаптивністю ML, демонструють найкращу ефективність при виявленні аномалій і прогнозуванні наслідків інцидентів.

5. Незважаючи на позитивні результати, застосування таких моделей обмежується складністю впровадження, вимогами до якості даних та обчислювальними ресурсами.

Рекомендації.

1. Розробляти адаптивні моделі, які можуть оновлювати параметри в режимі реального часу на основі нових індикаторів загроз.

2. Застосовувати ансамблеві ML-методи (Random Forest, XGBoost) для класифікації загроз та раннього попередження у SIEM-системах.

3. Використовувати інформаційну геометрію як основу для побудови простору ознак для оцінки аномалій у поведінці користувачів та пристроїв.

4. Проводити регулярну ревізію параметрів стохастичних моделей з урахуванням щорічних звітів провідних IT-компаній.

5. Забезпечити інтеграцію моделей у захищені середовища з підтримкою інтерпретованості, що критично важливо для критичної інфраструктури.

Перспективи подальших досліджень у сфері інтегрованого математичного моделювання ризиків інформаційної безпеки пов'язані з необхідністю розширення методологічної бази на складні, багатокомпонентні середовища з гетерогенними мережами, зокрема на

архітектури IoT, SCADA та Smart Grid, які мають високий рівень розподіленості та специфічні протоколи взаємодії.

Особливий інтерес становить формування моделей, що поєднують індуктивне машинне навчання з апаратом інформаційної геометрії для виявлення структурних зрушень у поведінці об'єктів кібер-фізичних систем, що дозволяє виявляти глибинні закономірності та аномалії в просторах ознак з високою розмірністю.

Крім того, доцільним є поглиблене дослідження можливостей підкріплювального навчання (reinforcement learning) для формування адаптивних стратегій реагування в режимі реального часу, особливо в умовах невизначеності та обмеженості обчислювальних ресурсів. Одним із ключових векторів подальшої роботи є розробка інтегральних метрик ефективності гібридних моделей, які враховують середній час виявлення й реагування на інциденти, частоту хибно-позитивних спрацьовувань, стійкість до змін середовища. Важливим етапом валідизації запропонованих рішень є апробація моделей у рамках реальних кейсів пілотних проєктів, зокрема в контексті державно-приватного партнерства, що дозволяє врахувати специфіку галузевих ризиків і забезпечити практичну цінність досліджень.

Список використаних джерел

1. ENISA. Threat Landscape for Information and Communication Technologies : звіт, 2024. European Union Agency for Cybersecurity, 2024. 120 p. URL: https://www.enisa.europa.eu/publications/threat-landscape-2024 (дата звернення: 10.07.2025).

2. Palo Alto Networks. Unit 42 Cloud Threat Report : звіт, 2024. Palo Alto Networks, Inc., 2024. 85 p. URL: https://www.paloaltonetworks.com/unit42/cloud-threat-report (дата звернення: 20.06.2025).

3. Cisco. Annual Cybersecurity Report : звіт, 2023. San Jose : Cisco Systems, Inc., 2023. 90 с. URL: https://www.cisco.com/c/en/us/products/security/security-reports.html (дата звернення: 01.07.2025).

4. NIST. Guide to Industrial Control Systems Security (SP 800-82 Rev. 2). Gaithersburg : National Institute of Standards and Technology, 2015. 247 p. DOI: https://doi.org/10.6028/NIST.SP.800-82r2

5. Microsoft. Digital Defense Report : звіт, 2023. Redmond : *Microsoft Corporation.* 2023. 110 p. URL: https://www.microsoft.com/

en-us/security/business/security-intelligence-report (дата звернення: 20.07.2025).

6. Wang P., Lu W., Zhuang W. A Markov model for secure state transitions in cloud computing. *IEEE Access*. 2020. Vol. 8. P. 12567–12577. DOI: https://doi.org/10.1109/ACCESS.2020.2966264

7. Doynikova E., Yevsieieva O. Markov chain analysis of cloud attack vectors. *Eastern-European Journal of Enterprise Technologies.* 2022. Vol. 2. No. 9. P. 55–63. DOI: https://doi.org/10.15587/1729-4061.2022.255943

8. Ross, S. M. Introduction to Probability Models. 11th ed. London : Academic Press, 2014. 767 c.

9. Khand R., Zio E. A Monte Carlo simulation framework for cyber-physical risk analysis in power systems. *Reliability Engineering & System Safety.* 2021. Vol. 211. P. 107558. DOI: https://doi.org/10.1016/j.ress.2021.107558

10. NIST. Security and Privacy Controls for Information Systems and Organizations (SP 800-53 Rev. 5). *Gaithersburg : National Institute of Standards and Technology.* 2022. 492 p. DOI: https://doi.org/10.6028/NIST.SP.800-53r5

11. Wozniak M., Graña M., Corchado J. M. A survey of multiple classifier systems as hybrid models for big data classification. *Information Fusion.* 2020. Vol. 44. P. 1–15. DOI: https://doi.org/10.1016/j.inffus.2017.11.008

12. Rios R. A., Lopez J. Analysis and modeling of adaptive cyber risk with ML and probabilistic techniques. *Computers & Security.* 2021. Vol. 108. P. 102335. DOI: https://doi.org/10.1016/j.cose.2021.102335

13. Kotenko I., Chechulin A. Markov game model for cyber attack and defense simulation. *Future Generation Computer Systems.* 2018. Vol. 79. P. 1027–1039. DOI: https://doi.org/10.1016/j.future.2017.08.043

14. Sgandurra D., Muñoz-González L., Mohsen R., Lupu E. Automated dynamic analysis of ransomware. Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. New York : ACM, 2016. P. 377–388. DOI: https://doi.org/10.1145/2897845.2897886

15. Buczak A. L., Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE*

Communications Surveys & Tutorials. 2016. Vol. 18, No. 2. P. 1153–1176. DOI: https://doi.org/10.1109/COMST.2015.2494502

16. Alcaraz C., Zeadally S. Critical infrastructure protection: Requirements and challenges for the 21st century. *International Journal of Critical Infrastructure Protection.* 2015. Vol. 8. P. 53–66. DOI: https://doi.org/10.1016/j.ijcip.2014.12.002

17. Ross S. M. Introduction to Probability Models. 12th ed. London : Academic Press, 2019. 842 c. DOI: https://doi.org/10.1016/C2017-0-02578-1

18. Jang-Jaccard J., Nepal S. A survey of emerging threats in cybersecurity. *Journal of Computer and System Sciences.* 2014. Vol. 80, No. 5. P. 973–993. DOI: https://doi.org/10.1016/j.jcss.2014.02.005

19. Li X., Liao H., Zhuang J. Modeling cyber risks with time-varying threat data: A stochastic programming approach. *Computers & Security.* 2021. Vol. 105. P. 102251. DOI: https://doi.org/10.1016/j.cose.2021.102251

20. Wang L., Islam T., Long T., Singhal A. Using Markov models for security risk assessment of smart grid. *Electric Power Systems Research*. 2020. Vol. 181. P. 106–194. DOI: https://doi.org/10.1016/j.epsr.2019.106194

21. Chen P. P., Desmet L., Joosen W. Adaptive security monitoring using Markov models. Journal of Information Security and Applications. 2014. Vol. 19/ No. 3. P. 167–181. DOI: https://doi.org/10.1016/j.jisa.2014.02.001

22. Kotenko I., Chechulin, A. A cyber attack modeling and impact assessment framework based on attack graphs and Markov models. *Future Generation Computer Systems.* 2013. Vol. 29/ No. 8. P. 2024–2037. DOI: https://doi.org/10.1016/j.future.2013.01.010

23. Check Point Software Technologies. Cyber Security Report 2024 : звіт, 2024. Check Point Software Technologies Ltd., 2024. 95 с. URL: https://www.checkpoint.com/cyber-security-report/ (дата звернення: 05.07.2025).

24. Kotecha K., Shrivastava R., Tanwar S. Security risk analysis of cloud systems using Poisson modeling. *Journal of Information Security and Applications.* 2021. Vol. 61. P. 102894. https://doi.org/10.1016/j.jisa.2021.102894

25. Chen T. M., Abu-Nimeh, S. Lessons from Stuxnet. *Computer.* 2011. Vol. 44. No. 4. P. 91–93. DOI: https://doi.org/10.1109/MC.2011.115

26. Pavlidou M., Katsikas S. K. A Monte Carlo Simulation Framework for Cybersecurity Risk Assessment. *Information Systems Frontiers*. 2022. Vol. 24. P. 1019–1035. DOI: https://doi.org/10.1007/s10796-021-10152-y

27. Zhuang Z., Zhao J. Integrating Monte Carlo and Bayesian Networks in Cyber Risk Quantification. *Journal of Cybersecurity*. 2023. Vol. 9. No. 1. P. tyaa029. DOI: https://doi.org/10.1093/cybsec/tyaa029

28. Khajuria A., Reddy G. T. Simulation of Cyber Risk in Cloud Environments Using Monte Carlo. *International Journal of Cloud Computing.* 2022. Vol. 11. No. 2. P. 105–117. DOI: https://doi.org/10.1504/IJCC.2022.122034

29. Dutt V., Ahn Y. S. Towards Reliable Monte Carlo-Based Simulation in Cybersecurity Analytics. *Computers & Security*. 2023. Vol. 127. P. 102706. DOI: https://doi.org/10.1016/j.cose.2023.102706

30. Amari S. Information Geometry and Its Applications. Tokyo : Springer, 2016. 374 c. DOI: https://doi.org/10.1007/978-4-431-55978-8

31. Chepurna O., Cherevko Y., Kuleshova Y. On Information Geometry Methods for Data Analysis. *Geometry, Integrability and Quantization.* 2024. Vol. 29. P. 11–22. DOI: https://doi.org/10.7546/giq-29-2024-11-22

32. *Cisco*. Annual Cybersecurity Report : звіт, 2024. San Jose : Cisco Systems, Inc., 2024. 100 p. URL: https://www.cisco.com/c/en/us/products/security/security-reports.html (дата звернення: 21.06.2025).

33. ENISA. Threat Landscape Report for Industrial Control Systems : звіт, 2024. European Union Agency for Cybersecurity, 2024. 105 с. URL: https://www.enisa.europa.eu/publications/industrial-control-systems-2024 (дата звернення: 22.06.2025).

34. CrowdStrike. Global Threat Report : звіт, 2024. CrowdStrike Holdings, Inc., 2024. 80 с. URL: https://www.crowdstrike.com/global-threat-report/ (дата звернення: 10.07.2025).

35. NIST. Guide for Conducting Risk Assessments (SP 800-30 Rev. 1). Gaithersburg : National Institute of Standards and Technology, 2022. 96 p. DOI: https://doi.org/10.6028/NIST.SP.800-30r1

36. Fortinet. Global Cybersecurity Threat Landscape Report : звіт, 2024. Fortinet, Inc., 2024. 70 p. URL: https://www.fortinet.com/resources/threat-reports/global-cybersecurity-2024 (дата звернення: 10.07.2025).

37. IBM. Cost of a Data Breach Report : звіт, 2024. IBM Corporation, 2024. 85 p. URL: https://www.ibm.com/reports/data-breach (дата звернення: 20.07.2025).

38. Cisco Talos. Threat Intelligence : звіт, 2024. San Jose : Cisco Systems, Inc., 2024. 75 p. URL: https://blog.talosintelligence.com/ (дата звернення: 10.07.2025).

39. Sophos. Active Adversary Playbook : звіт, 2024. Sophos Group plc, 2024. 60 с. URL: https://www.sophos.com/en-us/content/active-adversary-playbook (дата звернення: 11.07.2025).

40. ENISA. Threat Landscape for Software Supply Chains : звіт, 2024. European Union Agency for Cybersecurity, 2024. 90 p. URL: https://www.enisa.europa.eu/publications/software-supply-chains-2024 (дата звернення: 11.07.2025).

41. Fortinet. OT and ICS Security Threat Landscape : звіт, 2024. Fortinet, Inc., 2024. 65 p. URL: https://www.fortinet.com/content/dam/fortinet/assets/threat-reports/ot-threat-report-2024.pdf (дата звернення: 20.07.2025).

42. Trend Micro. Annual Cybersecurity Report : звіт, 2024. Trend Micro Incorporated, 2024. 88 с. URL: https://www.trendmicro.com/en_us/research/24/annual-cybersecurity-report.html (дата звернення: 11.07.2025).

43. ENISA. Cybersecurity in the Healthcare Sector : звіт, 2023. European Union Agency for Cybersecurity, 2023. 95 p. URL: https://www.enisa.europa.eu/publications/health-sector-cybersecurity (дата звернення: 09.07.2025).

44. Amazon Web Services. AWS Security Best Practices : звіт, 2024. Amazon Web Services, Inc., 2024. 120 с. URL: https://docs.aws.amazon.com/security/ (дата звернення: 09.07.2025).

45. Cisco. Cybersecurity Threat Trends : звіт, 2024. San Jose : Cisco Systems, Inc., 2024. 92 p. URL: https://www.cisco.com/c/en/us/products/security/threat-trends.html (дата звернення: 20.07.2025).

46. AWS. AWS Security Trends Report : звіт, 2024. Amazon Web Services, Inc., 2024. 100 с. URL: https://aws.amazon.com/security/ (дата звернення: 09.07.2025).

47. Symantec. Internet Security Threat Report : звіт, 2024. Broadcom Inc., 2024. 80 p. URL: https://www.broadcom.com/products/cyber-security/symantec/threat-report (дата звернення: 09.07.2025).

48. ENISA. Threat Landscape for Healthcare : звіт, 2024. European Union Agency for Cybersecurity, 2024. 85 p. URL: https://www.enisa.europa.eu/publications/healthcare-threat-landscape-2024 (дата звернення: 02.07.2025).

49. IBM Security. X-Force Threat Intelligence Index : звіт, 2024. IBM Corporation, 2024. 90 p. URL: https://www.ibm.com/reports/threat-intelligence (дата звернення: 07.07.2025).

50. McAfee. Enterprise Threat Report : звіт, 2024. McAfee, LLC, 2024. 75 p. URL: https://www.mcafee.com/enterprise/threat-report (дата звернення: 11.07.2025).

51. Fortinet. The Cost of a Data Breach Report : звіт, 2024. Fortinet, Inc., 2024. 70 p. URL: https://www.fortinet.com/resources/reports/cost-of-data-breach-2024 (дата звернення: 11.07.2025).

Черевко Євген Володимирович
к.ф.-м.н.,доц., доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

Чепурна Олена Євгенівна
к.ф.-м.н., доц., доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

Слатвінська Валерія Миколаївна
PhD в галузі «Право», асистент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ІНФОРМАЦІЙНА ГЕОМЕТРІЯ: КІБЕРБЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ

DOI https://doi.org/10.36059/978-966-397-537-5-5

Актуальність теми. Із зростанням складності цифрових систем та широким використанням високовимірних моделей обробки інформації актуалізується проблема формалізованого підходу до аналізу безпеки інформаційних процесів. Більшість прикладних систем захисту базується на евристичних або сигнатурних методах, які демонструють обмежену ефективність за умов змінних розподілів даних, нових типів атак або нестабільного інформаційного середовища. У цьому контексті виникає необхідність пошуку альтернативних методів аналізу інформації таких, наприклад, як інформаційна геометрі, яка виступає в якості засобу побудови внутрішньо узгоджених моделей інформаційного простору.

Згідно з аналітичними звітами NIST, кібератаки, такі як розподілені відмови в обслуговуванні, цільові вторгнення та атаки на ланцюги постачання, становлять значну загрозу для інформаційних систем [1].

Інформаційна геометрія як міждисциплінарна галузь формалізує поняття статистичного многовиду, наділеного метричною структурою Фішера та узгодженою з нею афінною зв'язністю. Застосування таких структур дозволяє описувати розподіли ймовірностей у термінах диференціальної геометрії, аналізувати кривину параметричних

просторів, а також визначати напрямки локальних змін систем, індукованих зовнішнім впливом. У статистичних задачах, пов'язаних з виявленням аномалій, оцінкою ризиків, оптимізацією байєсівських процедур та моделюванням стійкості до атак, інформаційна геометрія надає інструменти для кількісного аналізу з урахуванням як локальних, так і глобальних властивостей розподілів.

Попри теоретичну опрацьованість тематики в класичних роботах Амарі, Нагаоки та Ватанабе, прикладне застосування інформаційної геометрії в галузі кібербезпеки залишається недостатньо розвиненим. Існують поодинокі дослідження, що демонструють ефективність використання геометричних метрик для виявлення змін у мережевому трафіку, оцінки відстаней між гіпотезами в задачах класифікації, а також формалізації криптографічних моделей у термінах рімановых многовидів. Однак системна інтеграція цих методів у задачі захисту інформації потребує розробки нових математичних конструкцій, адаптованих до структурного і стохастичного характеру кіберзагроз.

В Україні законодавство, зокрема Закон «Про основні засади забезпечення кібербезпеки України» (2017), визначає кібербезпеку як необхідну умову захисту національних інтересів у цифровому просторі [2]. У цьому контексті інформаційна геометрія, як математична дисципліна, що поєднує диференціальну геометрію, теорію інформації та статистичний аналіз, пропонує перспективні інструменти для моделювання інформаційних процесів, аналізу кіберзагроз і розробки стійких систем захисту.

Огляд сучасного стану досліджень. Теоретичні основи інформаційної геометрії сформовано в працях Амарі, де вперше систематизовано концепцію статистичного многовиду та інформаційних дивергенцій, таких як дивергенція Кулбека-Ляйблера [3]. У контексті кібербезпеки ці методи застосовуються для аналізу інформаційних потоків і виявлення аномалій. Наприклад, Картер і Стрейле досліджували використання метрик Фішера для ідентифікації відхилень у мережевому трафіку, що дозволяє виявляти кібератаки з високою точністю [4]. Черевко Є., Чепурна О. та Кулешова Є. розвинули ці ідеї, запропонувавши геометричні методи для аналізу складних даних [5].

У криптографії інформаційна геометрія знаходить застосування в аналізі квантових систем, як показано в роботі Нільсена та Чуанга,

де геометричні підходи використовуються для оцінки стійкості квантових протоколів [6]. Сучасні дослідження, такі як робота Лі та співавторів, демонструють потенціал інформаційної геометрії для розробки постквантових криптографічних алгоритмів [7].

Аналіз великих даних є ще одним важливим напрямом. Ватанабе запропонував геометричні методи для оптимізації байєсівських алгоритмів, які застосовуються в задачах кібербезпеки [8]. Чен та співавтори показали, як інформаційна геометрія сприяє виявленню прихованих закономірностей у великих масивах даних, що підвищує ефективність систем раннього попередження атак [9].

В Україні Коваленко та співавтори наголошують на необхідності математичних методів, зокрема інформаційної геометрії, для захисту критичної інфраструктури [10], тоді як Петренко досліджує геометричне моделювання кібератак у умовах невизначеності [11].

Інтеграція інформаційної геометрії з алгоритмами штучного інтелекту, як показано в роботі Сміта та співавторів, відкриває нові можливості для адаптивного захисту IoT-систем [12]. Проте залишаються виклики, пов'язані з обчислювальною складністю геометричних алгоритмів і відсутністю стандартизованих підходів до їх реалізації в системах кібербезпеки.

Обґрунтування вибору тематики. Інформаційна геометрія вибрана як основа дослідження через її здатність представляти інформаційні процеси у вигляді геометричних структур, що забезпечує точний аналіз складних даних і розробку адаптивних методів захисту. Традиційні підходи, такі як статистичний аналіз або сигнатурні методи, часто є недостатньо ефективними для протидії новим типам кіберзагроз, зокрема атакам нульового дня чи цілеспрямованим вторгненням. Геометричні методи дозволяють моделювати розподіли ймовірностей як точки на многовидах, вимірюючи інформаційні відстані, що є особливо цінним для аналізу великих даних, криптографії та оцінки ризиків. В Україні актуальність дослідження посилюється потребою в захисті критичної інфраструктури та протидії гібридним загрозам у цифровому просторі.

Постановка задачі та дослідницька проблема. Основна проблема полягає в обмеженій інтеграції методів інформаційної геометрії в системи кібербезпеки, що знижує їхню здатність протистояти

складним кіберзагрозам. Відсутність систематизованих підходів до використання геометричних методів для аналізу інформаційних потоків, моделювання ризиків і розробки криптографічних систем створює потребу в нових теоретичних і практичних розробках. Задача дослідження полягає в розробці математичних моделей і алгоритмів на основі інформаційної геометрії для підвищення ефективності захисту інформації, зокрема в задачах аналізу даних, виявлення аномалій і створення стійких криптографічних систем.

Об'єкт і предмет дослідження. Об'єктом дослідження є процеси забезпечення кібербезпеки в інформаційних системах, включаючи аналіз даних, виявлення кіберзагроз і захист інформації. Предметом дослідження є методи інформаційної геометрії, зокрема статистичні многовиди, дивергенції та метрики, застосовані до моделювання інформаційних процесів і розробки систем кібербезпеки.

Мета дослідження. Метою даного дослідження є формалізація підходів інформаційної геометрії у задачах моделювання та аналізу кібербезпеки.

Завдання дослідження.

1. Систематизувати математичні основи інформаційної геометрії, включаючи теорію статистичного многовиду та інформаційних дивергенцій, для їх застосування в кібербезпеці.
2. Розробити геометричні методи аналізу інформаційних потоків щоб в перспективі виявляти аномалій у мережевому трафіку.
3. Оцінити потенціал інтеграції інформаційної геометрії з алгоритмами штучного інтелекту.
4. Проаналізувати обчислювальні обмеження геометричних методів і розробити способи їх оптимізації.

1. Основи Рімановоої геометрії

1.1. Тензори

Нехай задано тривимірний векторний простір V, в якому ми оберемо ортонормований базис e_1, e_2, e_3. Розглянемо довільний вектор $\vec{a}$. Його можна розкласти за векторами базису:

$$\vec{a} = a^1 e_1 + a^2 e_2 + a^3 e_3 = \sum_{i=1}^{3} a^i e_i .$$

Координатні індекси для векторів, з причин, що пояснимо нижче – будемо писати згори справа від головного символу. Зокрема, це дозволяє застосовувати так зване *правило Ейнштейна*. Воно полягає в тому, що поява однакового індексу згори і внизу означає підсумовування за цим індексом в межах його змінення. Тобто розкладання вектору $\vec{a}$ за базисом можна записати так:

$$\vec{a} = a^i e_i, (i = 1,2,3).$$

Тепер розглянемо скалярний добуток двох векторів $\vec{a} = a^i e_i$ та $\vec{b} = b^i e_i$:

$$\vec{a} \cdot \vec{b} = a^1 b^1 + a^2 b^2 + a^3 b^3 = \sum_{i=1}^{3} a^i b^i .$$

Зауважимо зокрема, що в цьому випадку ми не можемо застосовувати правило Ейнштейна, оскільки обидва індекси є верхні. Втім, можна використати одиничну матрицю, так звану *дельту Кронекера*:

$$\delta_{ij} = \begin{cases} 1, якщо\, i = j \\ 0, якщо\, i \neq j. \end{cases}$$

Тоді скалярний добуток векторів $\vec{a} = a^i e_i$ та $\vec{b} = b^i e_i$ може бути записаний так:

$$\vec{a} \cdot \vec{b} = a^i \delta_{ij} b^j \, (i, j = 1,2,3).$$

Тепер виконаємо лінійне перетворення нашого базису, використовуючи невироджену матрицю $C = \left(C_i^j\right)$, $\det C \neq 0$:

$$e_1' = C_1^1 e_1 + C_1^2 e_2 + C_1^3 e_3;$$
$$e_2' = C_2^1 e_1 + C_2^2 e_2 + C_3^3 e_3;$$
$$e_1' = C_1^1 e_1 + C_1^2 e_2 + C_1^3 e_3.$$

Або коротше:

$$e_1' = C_i^j e_j \, (i, j = 1,2,3).$$

Надалі ми не будемо кожного разу виписувати діапазон зміни індексів, якщо останній буде очевидним з контексту. Також, зауважимо, що можна з міркувань зручності перейменовувати індекс за яким ведеться підсумовування, наприклад $\vec{a} = a^i e_i = a^k e_k$. Про такий індекс іноді кажуть що він «*німий*».

В новому базисі розкладання вектору $\vec{a}$ виглядатиме таким чином:

$$\vec{a} = a^{i'} e_i' .$$

Але вектор $\vec{a}$ (вектор $\vec{b}$ також) лишився тим самим, тобто

$$a^i e_i = a^{i'} e_i' = a^{i'} C_i^j e_j .$$

Це означає що відповідно мають змінитися і координати його розкладання.

Перейменовуючи німі індекси маємо:

$$a^i e_i = a^k e_k = a^{i'} C_i^j e_j = a^{i'} C_i^k e_k ,$$

або

$$a^k e_k = a^{i'} C_i^k e_k .$$

Враховуючи лінійну незалежність векторів базису, маємо:

$$a^k = a^{i'} C_i^k .$$

Фактично це є скороченим записом матричного добутку:

$$\begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} C_1^1 & C_2^1 & C_3^1 \\ C_1^2 & C_2^2 & C_3^2 \\ C_1^3 & C_2^3 & C_3^3 \end{pmatrix} \begin{pmatrix} a^{1'} \\ a^{2'} \\ a^{3'} \end{pmatrix} .$$

Зауважимо, що на відміну від матричної форми, у скороченій формі запису, так званій *тензорній* – ми можемо міняти порядок множників, не боячись втрати сенсу:

$$a^k = C_i^k a^{i'} .$$

З цього випливає, що

$$\begin{pmatrix} a^{1'} \\ a^{2'} \\ a^{3'} \end{pmatrix} = \begin{pmatrix} C_1^1 & C_2^1 & C_3^1 \\ C_1^2 & C_2^2 & C_3^2 \\ C_1^3 & C_2^3 & C_3^3 \end{pmatrix}^{-1} \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix},$$

де C^{-1} є оберненою матрицею до C.

$$a^{i'} = \left(C_k^i \right)^{-1} a^k .$$

Тобто ми бачимо, що координати векторів при зміні базису змінюються *контраваріантно* до зміни векторів базису. Такі координати ми будемо називати *контраваріантними, і* відповідні верхні індекси також називатиме *контраваріантними.* Нижні індекси мають назву *коваріантних*. Надалі індекси що відповідають координатам в новому базисі будемо позначати як штриховані. Матрицю заміни базису позначатимемо:

$$C = C^{i}_{k'}.$$

Зворотню до неї, відповідно:

$$C^{-1} = C^{i'}_{k}.$$

Зрозуміло, що

$$C^{i}_{k'} C^{k'}_{j} = \delta^{i}_{j}, C^{i'}_{k} C^{k}_{j} = \delta^{i'}_{j'}.$$

Тут δ^{i}_{j} – вже знайомий нам *символ Кронекера*:

$$\delta^{i}_{j} = \begin{cases} 1, \text{ якщо } i = j \\ 0, \text{ якщо } i \neq j. \end{cases}$$

Пояснимо, чому в попередньому випадку, коли ми записували через дельту скалярний добуток, то обидва індекси були представлені як коваріантні. Справа в тому, що при заміні базису, скалярний добуток двох векторів

$\vec{a} \cdot \vec{b} = a^{i} \delta_{ij} b^{j}$ має лишатися незмінним, навіть у випадку, коли новий базис не буде ортонормованим. Введемо позначення для нашого початкового базису:

$$g_{ij} = \delta_{ij},$$

тоді:

$$\vec{a} \cdot \vec{b} = a^{i} g_{ij} b^{j}.$$

В новому базисі скалярний добуток має виглядати аналогічним чином:

$$\vec{a} \cdot \vec{b} = a^{i'} g_{i'j'} b^{j'},$$

але матриця $g_{i\,j}$, взагалі кажучи, вже не буде одиничною. Продовжуємо:

$$\vec{a}\cdot\vec{b} = a^{i'}\acute{g}_{i'j'}b^{j'} = a^k C_k^{i'}\acute{g}_{i'j'}C_l^{j'}b^l = a^k g_{kl}b^l.$$

Отже ми бачимо, що

$$C_k^{i'}\acute{g}_{i'j'}C_l^{j'} = g_{kl},$$

або

$$\acute{g}_{i'j'} = C_{i'}^k g_{kl}C_{j'}^l.$$

Об'єкт позначається як g_{kl} називають двічі *коваріантним тензором*. Якщо такий об'єкт визначає скалярний добуток, то він має назву *метричного тензору*, або просто *метрики*. У випадку ортонормованого базису:

$$g_{ij} = \delta_{ij}.$$

Взагалі кажучи, враховуючи властивості скалярного добутку g_{ij} є невиродженою симетричною матрицею, що визначає додатньо означену квадратичну форму, тобто таку, що для будь-якого ненульового вектора $\vec{a}$:

$$a^i g_{ij} a^j > 0.$$

Узагальнюючи поняття вектору, метричного тензору для *n*-вимірного простору, можна ввести більш загальне поняття *p-контраваріантного* та *q-коваріантного тензора,* або *тензору типу* (p,q).

Означення. *Тензором типу* (p,q) *ми називаємо геометричний об'єкт, що визначається в кожному координатному базисі лінійного простору сукупністю чисел:*

$$T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$$

(кожен з індексів $i_1 i_2 \ldots i_p; j_1 j_2 \ldots j_q$ *приймає незалежно один від одного усі значення від* 1 *до n) і змінюється при переході від одного базису* $e_1, e_2, \ldots e_n$, *до іншого* $e_{1'}, e_{2'}, \ldots e_n$ $\left(e_i = C_i^{i} e_i\right)$ *відповідно до наступного закону:*

$$T'^{\,i'_1 i'_2 \ldots i'_p}_{j'_1 j'_2 \ldots j'_q} = T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} C_{i_1}^{i'_1} C_{i_2}^{i'_2} \ldots C_{i_p}^{i'_p} C_{j'_1}^{j_1} C_{j'_2}^{j_2} \ldots C_{j'_q}^{j_q}. \qquad (1.1)$$

Фактично, якщо, наприклад, g_{ij} є квадратною матрицею розмірності $(n \times n)$, то тензор $T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$ є $(p+q)$-мірною «гіперкубічною» матрицею розмірності $(\underbrace{n \times n \times \ldots \times n}_{p+q\ разів})$.

Варто додати, що ми можемо для матриці g_{ij}, оскільки вона не є виродженою, знайти обернену. Її позначають як g^{ij}. Очевидно, що

$$g_{ik}g^{kj} = \delta_i^j.$$

За допомогою цих двох матриць можна робити операції «опускання» та «піднімання» індексів. Наприклад, взявши контраваріантний вектор a^i, опустивши індекс ми отримуємо *коваріантний вектор*, або просто *ковектор*:

$$a^k g_{kj} = a_j.$$

Ковектор можна витлумачити як лінійну функцію на множині векторів. Наприклад, якщо нам заданий ковектор f_i, то це означає, що на множині векторів n – вимірного простору задано лінійну функцію, що побудовано наступним чином:

$$f(x) = (f, x) = f_i x^i = f_1 x^1 + f_2 x^2 + \ldots + f_n x^n.$$

Піднімання індексу робиться за допомогою матриці g^{ij}. Наприклад для ковектору f_i ми можемо побудувати відповідний вектор:

$$f_i g^{ij} = f^j.$$

Якщо ми називаємо об'єкт «вектором», то найчастіше маємо на увазі саме контраваріантний вектор. Зрозуміло, що у випадку ортонормованого базису, контраваріантні і коваріантні компоненти одного того ж самого об'єкту співпадатимуть.

Також, надалі нам може стати в пригоді операція так званої *згортки* тензору. Згортка перетворює типу (p,q) у тензор типу $(p-1, q-1)$. Наприклад,

$$S^{i_2 i_3 \ldots i_p}_{j_1 j_3 \ldots j_q} = T^{\alpha i_2 i_3 \ldots i_p}_{j_1 \alpha j_3 \ldots j_q} = \sum_{\alpha=1}^{n} T^{\alpha i_2 i_3 \ldots i_p}_{j_1 \alpha j_3 \ldots j_q}.$$

Як ми бачимо, новий тензор $S^{i_2 i_3 \ldots i_p}_{j_1 j_3 \ldots j_q}$ отримано шляхом згортки за першим контраваріантним та другим коваріантним індексами. Взагалі кажучи, операцію згортки можна здійснювати по будь-якій парі контраваріантного та і коваріантного індексів. Головне, щоб в парі індексів – один був коваріантним а другий контраваріантним. Якщо цю вимогу порушувати – то отриманий внаслідок згортки об'єкт вже не буде тензором, оскільки закон перетворення координат тензору (1.1) вже не буде виконуватись.

1.2. Гладкі многовиди

Перш за все зазначимо, що вважатимемо усі функції тут і надалі диференцьованими потрібну кількість разів, або навіть аналітичними, тобто такими, чиї степеневі ряди збігаються до саме цих функцій, принаймні локально.

Для подальшого засвоєння цілком достатньо розуміти, що об'єктом дослідження ріманової геометрії є n-вимірний простір, тобто множина, на якій положення кожної точки задається набором з n чисел. Цей параграф призначений лише для обґрунтування можливості введення системи координат на цьому просторі, тому при першому читанні його можна пропустити. Додамо, що в подальшому викладення терміни «простір» і «многовид» використовуватимуться як синоніми.

Як пише французький математик Елі Картан [16], дуже важко дати загальне означення поняття многовиду. Але воно є фактично центральним поняттям геометрії, тому зараз ми спробуємо надати деяке уявлення про нього, так щоб не обтяжувати викладення аксіоматикою. Як зауважують Амарі і Нагаока, многовид M – це множина на якій задано координатну систему. Елементами цієї множини є точки. Наявність координатної системи на многовиді M означає взаємно-однозначне відображення M (або його підмножини) на R^n, яке дозволяє кожній точці p співставити впорядкований набір n дійсних чисел – координат точки (рис. 1).

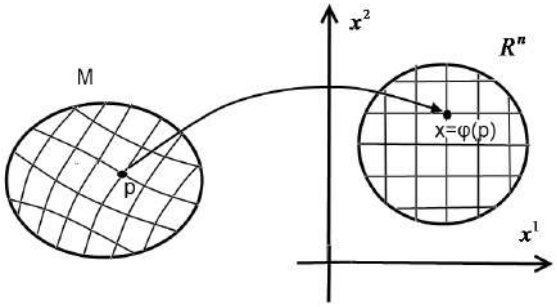


Рисунок 1 – Локальна координатна система на многовиді M

Нехай M є многовидом і $\varphi : M \to R^n$ є координатною системою на M. Тоді, φ відображує кожну точку $p \in R^n$ на n дійсних чисел [13]:

$$\varphi(p) = \left[x^1(p), \ldots, x^n(p) \right] = \left[x^1, \ldots, x^n \right].$$

Вони є координатами точки p. Кожне x^i може розглядатися як функція $p \to x^i(p)$, що відображує точку p на її i-ту координату. Ми будемо називати функції $x^i : M \to R$ *координатними функціями*. Відображення φ має назву *карти многовиду*. Але на жаль на завжди вдається обійтися однією картою, так щоб воно було взаємно однозначним. Наприклад, візьмемо систему географічних координат земної поверхні, а саме: θ – географічна широта, ψ – географічна довгота то раптом виявиться, що точка полюса, наприклад північного описується як координатами $\theta = \frac{\pi}{2}$; $\psi = 0$, так і наприклад, $\theta = \frac{\pi}{2}$; $\psi = \frac{\pi}{4}$. Тобто, точка полюса відображується у цілий відрізок площини R^2: $\theta = \frac{\pi}{2}$; $\psi \in [0; 2\pi]$. Тому множину M покривають відкритими множинами U_a, тобто

$$M = \bigcup_a U_a.$$

На кожній множині U_a ми задаємо функцію φ_a, так щоб відображення $\varphi_\alpha : U_\alpha \to R^n$ було взаємно однозначним. Разом вони утворюють карту $(U_\alpha, \varphi_\alpha)$. Множина усіх карт має назву *атласу* многовиду. Як ми бачимо, в загальному випадку ми маємо справу не з однією картою, а *множиною*, що складається з багатьох карт. Власне, тому M і має назву *многовиду*.

Також, атлас многовиду має відповідати ще одній вимозі, а саме: якщо $U_\alpha \cap U_\beta \neq \varnothing$, то існує гладке відображення:

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha \left(U_\alpha \cap U_\beta \right) \to \varphi_\beta \left(U_\alpha \cap U_\beta \right)$$

однієї області R^n на іншу. Фактично, остання формула описує перехід від однієї карти многовиду до іншої. Зауважимо, що надалі ми будемо працювати *локально,* тобто розглядати геометричні властивості на відкритій множині $U \subset M$, у межах, в яких

відображення $\varphi: M \to R^n$ на відкриту множину $\varphi(U)$ є неперервним. Зрозуміло, що такою множиною із обраною системою координат $(U_\alpha, \varphi_\alpha)$ може бути одна з карт атласа многовиду *M*.

Далі розглянемо функції на многовиді. Нехай $f: M \to R$ є функцією, що є визначеною на многовиді. Якщо ми вже обрали деяку координатну систему $\varphi = \left[x^1, \ldots, x^n\right]$ для *M*, то ця функція має бути представленою як функція координат, тобто якщо $\left[x^1, \ldots, x^n\right]$ є координатами точки *p*, то $f(p) = \overline{f}\left(x^1, \ldots, x^n\right)$, де $\overline{f} = f \circ \varphi^{-1}$. Таким чином, $\overline{f}$ є такою, що приймає дійсні значення, а областю визначення є якась відкрита підмножина R^n. Якщо $\overline{f}\left(x^1, \ldots, x^n\right)$ є диференційовною в кожній точці $\varphi(U)$, то частинна похідна $\frac{\partial}{\partial x^i}\overline{f}\left(x^1, \ldots, x^n\right)$ буде також функцією на $\varphi(U)$. Таким чином, ми можемо визначити частинні похідні для *f*:

$$\frac{\partial f}{\partial x^i} \overset{\text{def}}{=} \frac{\partial \overline{f}}{\partial x^i} \circ \varphi : M \to R.$$

Аналогічно можна визначити похідні вищих порядків для функцій заданих на многовиді *M*.

1.3. Дотичні вектори і дотичні простори

Нехай $\left[x^1, \ldots, x^n\right]$ – деяка система координат на многовиді *M*. Розглянемо *i*-ту координатну криву, що утворюється при фіксації усіх координат x^j крім x^i, тобто $i \neq j$, і проходить крізь точку *p* (рис. 2).

Позначимо дотичний вектор до координатної кривої x^j як e_i. Таким чином ми отримуємо дотичний простір T_p, породжений векторами $e_1, e_2, \ldots e_n$. І якщо точка p' є «дуже близькою» до *p*, $\left[x^i\right]$ і $\left[x^i + dx^i\right]$ є відповідно координатами *p* та p', то з'єднуючи ці точки отримуємо вектор $\overrightarrow{pp'} = dx^i e_i$. Він є елементом отриманого дотичного простору T_p.

Тепер розглянемо заміну системи координат на *M*:

$$x^i = x^i\left(x^{1'}, x^{2'} \ldots x^{n'}\right), (i = 1, 2, \ldots n).$$

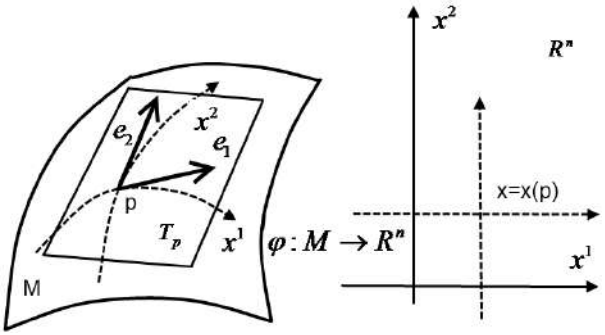


Рисунок 2 – Відображення дотичного до многовиду M простору на $\mathbb{R}^n$

Тоді лінійні перетворення векторів базису будуть здійснюватися за допомогою матриці Якобі:

$$\left(\frac{\partial x^i}{\partial x^{i'}}\right) = \begin{pmatrix} \frac{\partial x^1}{\partial x^{1'}} & \cdots & \frac{\partial x^1}{\partial x^{n'}} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial x^{1'}} & \cdots & \frac{\partial x^n}{\partial x^{n'}} \end{pmatrix}.$$

Тобто, ці вектори базису e_i будуть перетворюватися за формулою:

$$e_{i'} = \frac{\partial x^i}{\partial x^{i'}} e_i \,.$$

Очевидно, що матриця $\left(\frac{\partial x^i}{\partial x^{i'}}\right)$ є невиродженою, тобто $\det\left(\frac{\partial x^i}{\partial x^{i'}}\right) \neq 0$.

Обернена до неї матриця виглядатиме так:

$$\left(\frac{\partial x^{i'}}{\partial x^i}\right) = \begin{pmatrix} \frac{\partial x^{1'}}{\partial x^1} & \cdots & \frac{\partial x^{1'}}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^{n'}}{\partial x^1} & \cdots & \frac{\partial x^{n'}}{\partial x^n} \end{pmatrix}.$$

Легко бачити, що

$$\frac{\partial x^k}{\partial x^{i'}}\frac{\partial x^{i'}}{\partial x^j}=\delta^k_j,\ \ \frac{\partial x^{k'}}{\partial x^i}\frac{\partial x^i}{\partial x^{j'}}=\delta^{i'}_{j'}.$$

Якщо в дотичному лінійному просторі T_p заданий вектор $\vec{a}=a^i e_i$, то його координати змінюватимуться контраваріантно:

$$a^{i'}=\frac{\partial x^{i'}}{\partial x^i}a^i.$$

Врахуємо, що оскільки наш многовид M є гладким, то дотичний простір існує в кожній точці що лежить принаймні в межах деякої карти многовиду. Також, в цих межах може бути задане векторне поле $a^i(x)$, де $x=\left(x^1,x^2\ldots x^n\right)$. Тоді

$$a^{i'}(x')=\frac{\partial x^{i'}}{\partial x^i}a^i(x),$$

де $x'=\left(x^{1'},x^{2'}\ldots x^{n'}\right)$.

Також, у кожній точці для відповідного дотичного простору задамо метричний тензор:

$$g_{ij}(x)=e_i(x)\cdot e_j(x).$$

Його компоненти є скалярними добутками базисних векторів дотичного простору у точці x.

Для цього двічі коваріантного тензору маємо:

$$g'_{i'j'}(x')=\frac{\partial x^i}{\partial x^{i'}}g_{ij}(x)\frac{\partial x^j}{\partial x^{j'}}.$$

Узагальнюючи маємо таке означення.

Означення 2. *Тензорним полем (або просто тензором) типу* (p,q) *ми називаємо геометричний об'єкт, що визначається в кожному координатному базисі лінійного простору сукупністю чисел*

$$T^{i_1i_2\ldots i_p}_{j_1j_2\ldots j_q}(x)$$

(кожен з індексів $i_1i_2\ldots i_p;\ j_1j_2\ldots j_q$ *приймає незалежно один від одного усі значення від 1 до n) і змінюється при переході від однієї*

системи координат $x^1, x^2 \ldots x^n$ *до іншої* $x^{1'}, x^{2'} \ldots x^{n'}$ *(за формулами* $x^i = x^i\left(x^{1'}, x^{2'} \ldots x^{n'}\right)$ *відповідно до наступного закону:*

$$T'^{\,i'_1 i'_2 \ldots i'_p}_{j'_1 j'_2 \ldots j'_g}\left(x'\right) = T^{\,i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}\left(x\right)\frac{\partial x^{i'_1}}{\partial x^{i_1}}\frac{\partial x^{i'_2}}{\partial x^{i_2}}\ldots\frac{\partial x^{i'_p}}{\partial x^{i_p}}\frac{\partial x^{j_1}}{\partial x^{j'_1}}\frac{\partial x^{j_2}}{\partial x^{j'_2}}\ldots\frac{\partial x^{j_q}}{\partial x^{j'_g}}. \quad (1.2)$$

1.4. Коваріантне диференціювання та зв'язність

Нехай на многовиді *M* задано векторне поле $\vec{a}\left(x\right)$. Це означає, що в кожній точці воно може бути представленим як $\vec{a}\left(x\right) = a^i\left(x\right)e_i\left(x\right)$. Отже, як координати вектора так і вектори базису дотичного простору є функціями координат $x = \left(x^1, x^2 \ldots x^n\right)$, заданих на многовиді.

Знайдемо частинну похідну поля $\vec{a}\left(x\right)$ за змінною x^k:

$$\frac{\partial \vec{a}\left(x\right)}{\partial x^k} = \frac{\partial a^i\left(x\right)}{\partial x^k}e_i\left(x\right) + a^i\left(x\right)\frac{\partial e_i\left(x\right)}{\partial x^k}.$$

Центральним пунктом цих міркувань є те, що у загальному випадку криволінійних координат вектори базис у $e_1\left(x\right), e_2\left(x\right), \ldots e_n\left(x\right)$ не є незмінними: від точки до точки вони повертаються і змінюють свої абсолютні величини. Очевидно $\frac{\partial e_i\left(x\right)}{\partial x^k}$ є вектором і він може бути розкладений за векторами базису:

$$\frac{\partial e_i\left(x\right)}{\partial x^k} = \Gamma^j_{ik}\left(x\right)e_j\left(x\right).$$

Таким чином

$$\frac{\partial \vec{a}\left(x\right)}{\partial x^k} = \frac{\partial a^i\left(x\right)}{\partial x^k}e_i\left(x\right) + a^i\left(x\right)\Gamma^j_{ik}\left(x\right)e_j\left(x\right).$$

Перейменовуючи німі індекси, маємо:

$$\frac{\partial \vec{a}\left(x\right)}{\partial x^k} = \frac{\partial a^i\left(x\right)}{\partial x^k}e_i\left(x\right) + a^j\left(x\right)\Gamma^i_{jk}\left(x\right)e_i\left(x\right) = \left(\frac{\partial a^i\left(x\right)}{\partial x^k} + a^j\left(x\right)\Gamma^i_{jk}\left(x\right)\right)e_i\left(x\right).$$

Зауважимо, що коли ми говоримо про вектор, скажімо $\vec{a}(x)$, то дуже часто посилаємось на нього, як сукупність його контраваріантних координат $a^i(x)$, оскільки вважаємо систему локальних коодинат а з нею і базиси дотичного простору в кожній точці многовиду M наперед визначеними. Таке диференціювання координат вектору ми називаємо *коваріантним диференціюванням* вектору, і позначаємо таким чином:

$$\nabla_k a^i(x) = \frac{\partial a^i(x)}{\partial x^k} + a^j(x)\Gamma^i_{jk}(x).$$

Наведений знак коваріантного диференціювання має назву «набла». Часто для компактного запису замість набла використовують знак коми або крапки з комою:

$$a^i_{,k}(x) = \frac{\partial a^i(x)}{\partial x^k} + a^j(x)\Gamma^i_{jk}(x).$$

Тривимірна матриця $\Gamma^i_{jk}(x)$ має назву *об'єкта афінної зв'язності*. Об'єкт афінної зв'язності не є тензором, оскільки при переході від системи координат $x^1, x^2 \ldots x^n$ до нової $x^1, x^2 \ldots x^n$ відбувається за формулами:

$$\Gamma'^{\,i'}_{j'k'}(x') = \Gamma^{\,i}_{jk}(x)\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^k}{\partial x^{k'}}\frac{\partial x^{i'}}{\partial x^i} + \frac{\partial^2 x^i}{\partial x^{j'}\partial x^{k'}}\frac{\partial x^i}{\partial x^{i}}. \qquad (1.3)$$

Такий закон перетворення об'єкту афінної зв'язності пов'язаний з тим, що коваріантна похідна вектору має перетворюватися при переході до іншої системи координат як тензор, тобто за формулою (1.2):

$$\nabla_{k'} a^{i'}(x') = \nabla_k a^i(x)\frac{\partial x^k}{\partial x^{k'}}\frac{\partial x^{i'}}{\partial x^i}.$$

Враховуючи формулу (1.2) можна також показати, що для того, щоб коваріантна похідна ковектору b_i також була тензором, потрібно щоб вона була обчислена за формулою:

$$\nabla_k b_i(x) = \frac{\partial b_i(x)}{\partial x^k} - b_j(x)\Gamma^j_{ik}(x).$$

Узагальнюючи, можна отримати формулу знаходження коваріантної похідної довільного тензора:

$$\nabla_k T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} = T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q, k} = \frac{\partial T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}}{\partial x^k} + T^{\alpha i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \Gamma^{i_1}_{\alpha k} + T^{i_1 \alpha \ldots i_p}_{j_1 j_2 \ldots j_q} \Gamma^{i_2}_{\alpha k} + \ldots$$

$$+ T^{i_1 i_2 \ldots \alpha}_{j_1 j_2 \ldots j_q} \Gamma^{i_p}_{\alpha k} - T^{i_1 i_2 \ldots i_p}_{\alpha j_2 \ldots j_q} \Gamma^{\alpha}_{j_1 k} - T^{i_1 i_2 \ldots i_p}_{j_1 \alpha \ldots j_q} \Gamma^{\alpha}_{j_2 k} - \ldots - T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots \alpha} \Gamma^{\alpha}_{j_q k}. \qquad (1.4)$$

Тут і надалі в усіх випадках, коли не відбувається заміна координат – ми не будемо вказувати x як аргумент тензорних функцій і об'єкта афінної зв'язності, тобто писатимемо $T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$ замість $T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}(x)$, Γ^j_{ik} замість $\Gamma^j_{ik}(x)$ тощо.

Ми називаємо тензор коваріантно сталим, якщо

$$\nabla_k T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} = 0.$$

Поняття коваріантної похідної дозволяє розвинути так званий апарат абсолютного диференціювання.

Абсолютним диференціалом тензора ми називаємо вираз: $DT^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} = \nabla_k T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} dx^k$.

Ми кажемо що тензор $T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$, заданий в кожній точці деякого шляху *паралельно переноситься уздовж цього шляху*, якщо його абсолютний диференціал при будь-якому нескінченно малому зміщені уздовж цього шляху дорівнює нулю.

Зокрема при паралельному перенесенні вектора ξ^i за напрямом $\{dx^k\}$ маємо:

$$D\xi^i = \nabla_k \xi^i dx^k = \frac{\partial \xi^i}{\partial x^k} dx^k + \Gamma^i_{jk} \xi^j dx^k = d\xi^i + \Gamma^i_{jk} \xi^j dx^k = 0.$$

Отже, отримуємо формулу паралельного перенесення вектора:

$$d\xi^i = -\Gamma^i_{jk} \xi^j dx^k.$$

Запис $\nabla_X Y$ означає коваріантну похідну вектору $Y = Y^i e_i$ за напрямом вектору $X = X^i e_i$. Якщо розписати докладно в координатах, отримуємо:

$$\nabla_X Y = \left(X^k \nabla_k Y^i \right) e_i = X^k \left(\frac{\partial Y^i}{\partial x^k} + \Gamma^i_{jk} Y^j \right) e_i.$$

Варто відзначити ще таку деталь, наприклад, запис

$$\nabla_X Y = 0$$

може бути витлумачений по-різному, залежно від контексту. Якщо X є довільним векторним полем, то це означатиме що $\nabla_k Y^i = 0$, тобто коваріантні похідні по усім змінним є нульовими, отже певне поле Y є коваріантно сталим. Якщо не тільки Y, але і X є також певним векторним полем, то відповідний запис в координатах означатиме $X^k \nabla_k Y^i = 0$, і певне поле Y є коваріантно сталим лише уздовж інтегральних кривих певного поля X.

Якщо виконується рівність

$$\nabla_k Y^i = \rho_k Y^i$$

(ρ_k – деякий ковектор), то векторне поле Y називають *рекурентним*. У безкоординатній формі це виглядатиме так:

$$\nabla_X Y = \rho(X) Y,$$

тут X є довільним векторним полем, тобто $\forall X\,(\nabla_X Y = \rho(X) Y)$.

1.5. Геодезичні лінії

Нехай на многовиді M задано криву $\gamma(t)$ як відображення $\gamma : \bullet \to M$, де $\bullet \subset \mathrm{R}$, тобто якийсь відрізок $a \le t \le b$. Оскільки ми обрали систему координат, то криву можна описати системою параметричних рівнянь.

$$x^i = x^i(t), (i = 1, 2, \ldots n).$$

Тоді дотичний до такої кривої вектор виглядатиме так:

$$\dot{x}^i = \frac{dx^i}{dt}, (i = 1, 2, \ldots n).$$

Тензорне поле $T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$ називають *рекурентним уздовж кривої* $\gamma(t)$, якщо

$$\nabla_{\dot{x}} T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} = \dot{x}^k \nabla_{\mathrm{k}} T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} = \mathrm{f}(t) T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q},$$

де $f(t)$ – деякий скаляр. Зокрема, для вектору ξ^i рекурентність уздовж кривої $\gamma(t)$ виглядатиме так:

$$\nabla_{\dot{x}}\xi^i = \dot{x}^k \nabla_k \xi^i = \frac{\partial \xi^i}{\partial x^k}\dot{x}^k + \Gamma^i_{jk}\xi^j \dot{x}^k = \mathrm{f}(t)\xi^i.$$

Оскільки $\frac{\partial \xi^i}{\partial x^k}\dot{x}^k = \frac{\partial \xi^i}{\partial x^k}\frac{dx^i}{dt} = \frac{d\xi^i}{dt}$ (ланцюгове правило диференціювання), то вимогу рекурентності уздовж кривої запишемо у вигляді:

$$\frac{d\xi^i}{dt} + \Gamma^i_{jk}\xi^j \dot{x}^k = \mathrm{f}(t)\xi^i.$$

За допомогою властивості рекурентності ми можемо ввести нове, дуже важливе в геометрії, поняття.

Крива $\gamma(t)$, що задана рівняннями $x^i = x^i(t), (i = 1,2,\ldots n)$ має назву *геодезичної многовиду M у зв'язності* Γ^i_{jk}, якщо дотичний до цієї кривої вектор є рекурентним уздовж неї:

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = f(t)\frac{dx^i}{dt}. \tag{1.5}$$

Зауважимо, що можна обрати таку параметризацію $s = s(t)$ кривої γ, що рівняння геодезичної прийме вигляд:

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0, \tag{1.6}$$

Тобто дотичний вектор тепер переноситься паралельно. Такий параметр має назву *канонічного*.

Поняття геодезичної є дуже важливим. Справа в тому, що простори з якими працює ріманова геометрія не містять прямих ліній в тому сенсі до яких ми звикли у евклідовій геометрії. Геодезичні лінії є фактично є аналогами прямих у неевклідовому просторі, «найпрямішими» серед кривих.

1.6. Плоска зв'язність, тензор кривини

Взагалі кажучи, об'єкт афінної зв'язності Γ^i_{jk} є досить довільною структурою, головною вимогою є лише перетворення його при заміні координат згідно з формулою (1.3). Також, у загальному випадку не

є обов'язковим, щоб зв'язність була симетричною, тобто можливі випадки, коли

$$\Gamma^{i}_{jk} \neq \Gamma^{i}_{kj}.$$

В такому випадку кажуть, що зв'язність має ненульове *кручення*.

Об'єкт, який описує кручення має вигляд:

$$T^{i}_{jk} = \Gamma^{i}_{jk} - \Gamma^{i}_{kj}$$

і крім того є тензором. Це – так званий *тензор кручення*. Кол зв'язність є симетричною, то тензор кручення дорівнює нулю.

Дослідимо випадок, коли можна перейти до такої системи координат $y^1, y^2 \ldots y^n$, в якій

$$\acute{\Gamma}^{i}_{jk}(y) \equiv 0. \tag{1.7}$$

Запишемо (1.3) у вигляді:

$$\acute{\Gamma}^{i}_{jk}(x) = \Gamma^{\gamma}_{\alpha\beta}(y)\frac{\partial y^{\alpha}}{\partial x^{j}}\frac{\partial y^{\beta}}{\partial x^{k}}\frac{\partial x^{i}}{\partial y^{\gamma}} + \frac{\partial^2 y^{\gamma}}{\partial x^{j}\partial x^{k}}\frac{\partial x^{i}}{\partial y^{\gamma}}.$$

«Німі» індекси, тобто ті, за якими відбувається згортка позначено грецькими літерами. Враховуючи (1.7) отримуємо звідси

$$\Gamma^{i}_{jk}(x) = \frac{\partial^2 y^{\gamma}}{\partial x^{j}\partial x^{k}}\frac{\partial x^{i}}{\partial y^{\gamma}}.$$

Остаточно це можна записати у вигляді системи диференціальних рівнянь в частинних похідних відносно *n* невідомих функцій:

$$y^{i} = y^{i}\left(x^1, x^2 \ldots x^n\right),$$

$$\frac{\partial^2 y^{i}}{\partial x^{j}\partial x^{k}} = \Gamma^{m}_{jk}(x)\frac{\partial y^{i}}{\partial x^{m}}\left(i = 1,2,\ldots n\right). \tag{1.8}$$

Така система рівнянь, розв'язаних відносно старших похідних називають *системою типу Коші*. Якщо вона має розв'язок, то частинні похідній його мають згідно за теоремою Шварца задовільняти рівності (рівність мішаних похідних незалежно від порядку диференціювання):

$$\frac{\partial}{\partial x^{l}}\left(\frac{\partial^2 y^{i}}{\partial x^{j}\partial x^{k}}\right) - \frac{\partial}{\partial x^{k}}\left(\frac{\partial^2 y^{i}}{\partial x^{j}\partial x^{l}}\right) = 0, \left(i = 1,2,\ldots n\right). \tag{1.9}$$

Зазвичай, (1.9) називають *умовами інтегровності* (1.8)

Враховуючи (1.8), з (1.9) отримуємо:

$$\frac{\partial}{\partial x^l}\left(\Gamma^m_{jk}(x)\frac{\partial y^i}{\partial x^m}\right)-\frac{\partial}{\partial x^k}\left(\Gamma^m_{jl}(x)\frac{\partial y^i}{\partial x^m}\right)=0,$$

$$\frac{\partial y^i}{\partial x^m}\frac{\partial\Gamma^m_{jk}(x)}{\partial x^l}+\Gamma^m_{jk}(x)\frac{\partial^2 y^i}{\partial x^l\partial x^m}-\frac{\partial y^i}{\partial x^m}\frac{\partial\Gamma^m_{jl}(x)}{\partial x^k}+\Gamma^m_{jl}(x)\frac{\partial^2 y^i}{\partial x^k\partial x^m}=0,$$

$$\frac{\partial y^i}{\partial x^m}\frac{\partial\Gamma^m_{jk}(x)}{\partial x^l}+\Gamma^m_{jk}(x)\Gamma^n_{lm}(x)\frac{\partial y^i}{\partial x^n}-\frac{\partial y^i}{\partial x^m}\frac{\partial\Gamma^m_{jl}(x)}{\partial x^k}+\Gamma^m_{jl}(x)\Gamma^n_{km}(x)\frac{\partial y^i}{\partial x^n}=0.$$

Перейменовуючи «німі» індекси, маємо:

$$\frac{\partial y^i}{\partial x^m}\frac{\partial\Gamma^m_{jk}(x)}{\partial x^l}+\Gamma^n_{jk}(x)\Gamma^m_{ln}(x)\frac{\partial y^i}{\partial x^m}-\frac{\partial y^i}{\partial x^m}\frac{\partial\Gamma^m_{jl}(x)}{\partial x^k}+\Gamma^n_{jl}(x)\Gamma^m_{kn}(x)\frac{\partial y^i}{\partial x^m}=0.$$

Виносимо спільний множник за дужки:

$$\frac{\partial y^i}{\partial x^m}\left(\frac{\partial\Gamma^m_{jk}(x)}{\partial x^l}+\Gamma^n_{jk}(x)\Gamma^m_{ln}(x)-\frac{\partial\Gamma^m_{jl}(x)}{\partial x^k}-\Gamma^n_{jl}(x)\Gamma^m_{kn}(x)\right)=0.$$

Остаточно (1.9) можна записати у вигляді:

$$\frac{\partial y^i}{\partial x^m}R^m_{jlk}(x)=0,\left(i=1,2,\ldots n\right), \tag{1.10}$$

де

$$R^m_{jlk}(x)=\frac{\partial\Gamma^m_{jk}(x)}{\partial x^l}+\Gamma^n_{jk}(x)\Gamma^m_{ln}(x)-\frac{\partial\Gamma^m_{jl}(x)}{\partial x^k}-\Gamma^n_{jl}(x)\Gamma^m_{kn}(x). \tag{1.11}$$

Символом $R^m_{jlk}(x)$ ми позначаємо один раз контраваріантний, 3 рази коваріантний тензор, що має назву *тензора кривини, а також тензора Рімана, або символа Рімана другого роду*.

Оскільки матриця Якобі $\left(\frac{\partial y^i}{\partial x^m}\right)$ є невиродженою, тобто $det\left(\frac{\partial y^i}{\partial x^m}\right)\neq 0$, то умова інтегровності (10) буде виконуватися, якщо тензор кривини буде тотожно дорівнювати нулю:

$$R^m_{jlk}(x)\equiv 0.$$

Зв'язність, коефіцієнти можна перетворенням координат зробити тотожно рівними нулю $\left(\Gamma'^{\,i}_{jk}\left(y\right)\equiv 0\right)$ має назву *плоскої*. Ми бачимо, що плоскою може бути лише така зв'язність, тензор кривини якої тотожно дорівнює нулю. Додамо також, що з рівняння (1.8) є очевидним, що плоскою зв'язністю може бути лише симетрична зв'язність, тобто для якої $\Gamma^i_{jk}=\Gamma^i_{kj}$.

1.7. Ріманова метрика і зв'язність Леві – Чивіта

Нехай заданий многовид M із симетричною зв'язністю $\Gamma^i_{jk}\left(x\right)$ та метричним тензором $g_{ij}\left(x\right)$.

Як ми вже бачили, зв'язність на многовиді визначає зокрема як мають змінюватися координати тензора (зокрема вектору) при його паралельному перенесенні, враховуючи як змінюється від однієї точки до іншої дотичний простір, а з ним і базис цього дотичного простору. Метричний тензор задає формулу для обчислення скалярного добутку у дотичному просторі. Дослідимо, як можна узгодити зв'язність та метрику.

Розглянемо задану на многовиді M деяку криву $\gamma\left(t\right)$ (не обов'язково геодезичну) що описується системою параметричних рівнянь.

$$x^i=x^i\left(t\right),\left(i=1,2,\ldots n\right).$$

Нехай уздовж цієї кривої паралельно переносяться два вектори: $\vec{a}\left(x\right)=a^i\left(x\right)e_i\left(x\right)$ *та* $\vec{b}\left(x\right)=b^i\left(x\right)e_i\left(x\right)$.

Це означає що їх координати уздовж кривої $\gamma\left(t\right)$ задовільнятимуть рівнянням:

$$\nabla_{\dot{x}}a^i=\dot{x}^k\nabla_k a^i=\frac{\partial a^i}{\partial x^k}\dot{x}^k+\Gamma^i_{jk}a^j\dot{x}^k=0,$$
$$\nabla_{\dot{x}}b^i=\dot{x}^k\nabla_k b^i=\frac{\partial b^i}{\partial x^k}\dot{x}^k+\Gamma^i_{jk}b^j\dot{x}^k=0. \qquad (1.12)$$

Тут $\dot{x}^i=\dfrac{dx^i}{dt}$ – дотичний до кривої $\gamma\left(t\right)$ вектор.

Скалярний добуток цих векторів обчислюємо за формулою:

$$\vec{a} \cdot \vec{b} = a^i g_{ij} b^j.$$

Варто очікувати, що оскільки вектори переносяться паралельно уздовж кривої, то їх скалярний добуток має бути незмінний:

$$\frac{d\left(\vec{a}\left(x(t)\right) \cdot \vec{b}\left(x(t)\right)\right)}{dt} = 0.$$

Тому маємо:

$$\frac{d\left(\vec{a}\left(x(t)\right) \cdot \vec{b}\left(x(t)\right)\right)}{dt} = \nabla_{\dot{x}} a^i g_{ij} b^j + a^i \nabla_{\dot{x}} g_{ij} b^j + a^i g_{ij} \nabla_{\dot{x}} b^j = 0.$$

Оскільки вектори переносяться паралельно, внаслідок (1.12) в останньому рівнянні перший та третій доданки дорівнюють нулю. Отже

$$a^i \nabla_{\dot{x}} g_{ij} b^j = 0.$$

Оскільки вектори a^i та b^i – довільні[1], то

$$\nabla_{\dot{x}} g_{ij} = \frac{dx^k}{dt} \nabla_{\dot{k}} g_{ij} = 0.$$

Крива $\gamma(t)$ є також довільною, тому маємо:

$$\nabla_{\dot{k}} g_{ij} \equiv 0. \qquad (1.13)$$

Коли виконується умова (1.13), то кажуть що *зв'язність є узгодженою з метрикою*.

Згідно із загальною формулою (1.4) для коваріантного диференціювання довільного тензору, (1.13) можна записати таким чином:

$$\frac{\partial g_{ij}}{\partial x^k} - g_{\alpha j} \Gamma^{\alpha}_{ik} - g_{i\alpha} \Gamma^{\alpha}_{jk} = 0. \qquad (1.14)$$

Опускаючи індекси у коефіцієнтах зв'язності за формулою:

$$\Gamma_{ik,j} = g_{\alpha j} \Gamma^{\alpha}_{ik}, \qquad (1.15)$$

[1] Строго кажучи, насправді довільними є початкові значення координат векторів $a^i\left(x(t_0)\right)$ та $b^i\left(x(t_0)\right)$ для t_0. Значення координат векторів на шляху паралельного перенесення $\gamma(t)$ однозначно визначаються з рівнянь (1.12). Однозначність випливає з припущення локальної аналітичності функцій з якими ми оперуємо, зокрема $x^i(t)$ та $\Gamma^i_{jk}(x)$.

З (1.14) маємо:

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ik,j} + \Gamma_{jk,i}. \tag{1.16}$$

Перейменуємо індекси у (1.15):

$$\frac{\partial g_{ik}}{\partial x^j} = \Gamma_{ij,k} + \Gamma_{kj,i}, \tag{1.17}$$

$$\frac{\partial g_{kj}}{\partial x^i} = \Gamma_{ki,j} + \Gamma_{ji,k}. \tag{1.18}$$

Додаючи та віднімаючи за схемою $(1.17)+(1.17)-(1.18)$, отримуємо:

$$\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ij,k} + \Gamma_{kj,i} + \Gamma_{ki,j} + \Gamma_{ji,k} - \Gamma_{ik,j} - \Gamma_{jk,i}.$$

Враховуючи симетричність символу Крістофеля першого роду за першими двома нижніми індексами, маємо

$$\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ij,k} + \Gamma_{ji,k} = 2\Gamma_{ij,k}.$$

Остаточно

$$\Gamma_{ij,k} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} \right). \tag{1.19}$$

Отримана таким чином величина (1.19) має назву *символа Крістофеля першого роду*.

Враховуючи (1.15) піднімаємо у (1.19) індекс k:

$$\Gamma_{ij}^k = \frac{g^{k\alpha}}{2}\left(\frac{\partial g_{i\alpha}}{\partial x^j} + \frac{\partial g_{\alpha j}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^\alpha} \right). \tag{1.20}$$

Величину, отриману за допомогою формули (1.20) називають *символом Крістофеля другого роду*. Можна легко показати, що обидві величини, визначені в (1.19) і (1.20) не є тензорами.

Така симетрична зв'язність, що є узгодженою з метрикою називають *зв'язністю Леві – Чівіта*. Серед симетричних зв'язностей лише вона одна є узгодженою з метрикою, тобто є коваріантно сталою, тому має місце тотожність (1.14).

Але маючи зв'язність Леві – Чівіта ми можемо побудувати і інші, неузгоджені з даною нам метрикою зв'язності.

Формула проста:

$$\tilde{\Gamma}_{ij}^{k} = \Gamma_{ij}^{k} + T_{ij}^{k}. \tag{1.21}$$

Тут T_{ij}^{k} – довільний 1 раз контраваріантний і 2-коваріантний тензор, Γ_{ij}^{k} – символ Крістофеля другого роду. Якщо ми хочемо, щоб нова зв'язність $\tilde{\Gamma}_{ij}^{k}$ була симетричною, то треба вимагати, щоб T_{ij}^{k} був симетричним по коваріантним індексам. Зауважимо також, що $\tilde{\Gamma}_{ij}^{k}$ ми не можемо називати символом Крістофеля другого роду, оскільки ці коефіцієнти не є узгодженими формулою (1.20) з даною нам метрикою $g_{ij}(x)$. Поняття символу Крістофеля першого роду для $\tilde{\Gamma}_{ij}^{k}$ – взагалі позбавлене сенсу, оскільки нам для того, щоб опустити верхній індекс потрібна нова метрика $\tilde{g}_{ij}(x)$, з якою узгоджена нова зв'язність. Тим не менш, її можливо отримати, якщо розв'язати систему лінійних рівнянь (1.14), де нам відомі лише коефіцієнти нової зв'язності $\tilde{\Gamma}_{ij}^{k}$.

$$\frac{\partial \tilde{g}_{ij}}{\partial x^{k}} = \tilde{g}_{\alpha j}\tilde{\Gamma}_{ik}^{\alpha} + \tilde{g}_{i\alpha}\tilde{\Gamma}_{jk}^{\alpha}. \tag{1.22}$$

Лише отримавши розв'язок системи рівнянь (1.22), ми можемо називати $\tilde{\Gamma}_{ij}^{k}(x)$ символом Крістофеля другого роду, і за допомогою знайденої метрики $\tilde{g}_{ij}(x)$ отримати відповідний символ Крістофеля першого роду. Втім, може виявитися, що система (1.22) не має розв'язку. Тоді кажуть, що зв'язність $\tilde{\Gamma}_{ij}^{k}(x)$ не є метризовною.

Але повернемося до зв'язності Леві – Чівіта. Якщо вона є плоскою, то це означає, що існує система координат, в якій

$$\Gamma_{ij}^{k}(x) \equiv 0. \tag{1.23}$$

З рівняння (1.14) випливає, що

$$\frac{\partial g_{ij}}{\partial x^{k}} = 0. \tag{1.24}$$

Умови інтегрованості для (1.24) задовільняються тотожно. Отже, рівняння (1.24) є цілком інтегровними, що забезпечує локально в досліджуваній області многовиду задати евклідову метрику:

$$g_{ij} = \delta_{ij} = \begin{cases} 1, \text{ якщо } i = j, \\ 0, \text{ якщо } i \neq j. \end{cases}$$

1.7. Геодезичні у зв'язності Леві – Чивіта

Згадаємо відому із загального курсу задачу. Нехай У тривимірному евклідовому просторі R^3 задана крива параметричним рівнянням:

$$\begin{cases} x = x(t), \\ y = y(t), \\ z = z(t). \end{cases}$$

Потрібно знайти довжину відрізка кривої, що з'єднує точки $\left(x(t_1), y(t_1), z(t_1)\right)$ та $\left(x(t_2), y(t_2), z(t_2)\right)$.

Ми користувались формулою:

$$l = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt.$$

Підкореневий вираз є скалярним квадратом дотичного до кривої вектора:

$$\vec{r}\left(\dot{x}, \dot{y}, \dot{z}\right),$$

де $\dot{x} = \dfrac{dx(t)}{dt}$, $\dot{y} = \dfrac{dy(t)}{dt}$, $\dot{z} = \dfrac{dz(t)}{dt}$:

$\vec{r} \cdot \vec{r} = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$.

Скалярний квадрат обчислено для евклідової метрики $g_{ij} = \delta_{ij}$.

Якщо деяку криву $\gamma(t)$ задано системою параметричних рівнянь.

$$x^i = x^i(t), (i = 1, 2, \ldots n).$$

на неевклідовому многовиді M з метрикою g_{ij}, то скалярний квадрат дотичного вектору $\vec{\dot{x}} \cdot \vec{\dot{x}} = g_{ij}\dot{x}^i\dot{x}^j$.

Внаслідок цього, для обчислення довжини відрізка кривої на неевклідовому многовиді M матимемо:

$$l = \int_{t_1}^{t_2} \sqrt{g_{ij}\dot{x}^i\dot{x}^j}\, dt. \qquad (1.25)$$

Виникає питання: якими мають бути функції $x^i = x^i(t)$, щоб інтеграл (1.25) мав найменше значення з можливих. Тобто, потрібно знайти криву, що з'єднуючі задані точки матиме найменшу довжину.

У випадку евклідового простору, зрозуміло, що це будуть прямі лінії.

Якщо ми проаналізуємо ситуацію для довільної метрики g_{ij}, то можна довести, що значення інтегралу (1.25) буде мати екстремальне значення саме у випадку, коли криві є геодезичними, а саме, відповідають рівнянню (1.6).

1.8. Підмноговиди. Автопаралельні та цілком геодезичні підмноговиди

Розглянемо многовид M_m розмірності m що вкладений як поверхня у многовид N_n. Нехай у M_m задана координатна система $x^1, x^2 \dots x^m$, а у N_n координатна система $y^1, y^2 \dots y^n$ [14]. Рівняння такої поверхні задаються так:

$$\begin{cases} y^1 = y^1\left(x^1, x^2 \dots x^m\right) \\ y^2 = y^2\left(x^1, x^2 \dots x^m\right) \\ \dots\dots\dots\dots\dots\dots\dots \\ y^n = y^n\left(x^1, x^2 \dots x^m\right) \end{cases}. \tag{1.26}$$

Тоді є i-ю координатою у просторі, дотичному до N_n вектора, що є дотичним до координатної лінії x^α на поверхні M_m. Зрозуміло, що вектори $\frac{\partial \boldsymbol{y}}{\partial x^1}, \frac{\partial \boldsymbol{y}}{\partial x^2}, \dots, \frac{\partial \boldsymbol{y}}{\partial x^m}$ мають утворювати базис дотичного простору[2] до поверхні M_m у довільній точці цієї поверхні $\boldsymbol{y}_0\left(x_0^1, x_0^2, \dots, x_0^m\right)$. Тому ці вектори мають бути лінійно незалежними. Звідси випливає, що ранг матриці Якобі

$$\left(\frac{\partial y^i}{\partial x^\alpha}\right) = \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \dots & \frac{\partial y^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y^n}{\partial x^1} & \dots & \frac{\partial y^n}{\partial x^m} \end{pmatrix}.$$

[2] Такий дотичний простір часто позначається як $T_{x_0}\left(M_m\right)$. Очевидно, що $T_{x_0}\left(M_m\right)$ є підпростором до $T_{y(x_0)}\left(N_n\right)$ – простору, дотичного до N_n в точці $\boldsymbol{y}_0\left(x_0^1, x_0^2, \dots, x_0^m\right)$.

має дорівнювати m. Для подальшого викладення домовимося позначати:

$$B_{\alpha}^{i} = \frac{\partial y^{i}}{\partial x^{\alpha}}. \tag{1.27}$$

Якщо на N_n задано метрику, g_{ij} то на M_m буде індуковано метрику $\overset{\prime}{g}_{\alpha\beta}$:

$$\overset{\prime}{g}_{\alpha\beta} = B_{\alpha}^{i} g_{ij} B_{\beta}^{j}, \tag{1.28}$$

І відповідно зв'язність Леві – Чівіта:

$$\overset{\prime}{\Gamma}_{\alpha\beta}^{\gamma}(x) = B_{i}^{\gamma}\left(\Gamma_{jk}^{i}\left(y(x)\right)B_{\alpha}^{j}B_{\beta}^{k} + \frac{B_{\alpha}^{i}}{\partial x^{\beta}}\right). \tag{1.29}$$

Тут B_{i}^{γ} обчислюється за допомогою матриці $\overset{\prime}{g}^{\alpha\gamma}$, що є оберненою до матриці індукованої метрики (1.28):

$$B_{i}^{\gamma} = B_{\alpha}^{j} g_{ij} \overset{\prime}{g}^{\alpha\gamma}.$$

Об'єкт (1.27) є так званим «мішаним» тензором. Це пов'язано з тим, що верхній індекс у B_{α}^{i} відноситься до структури, пов'язної з N_n, нижній відповідно до структури M_m. Тому при коваріантному диференціюванні треба враховувати як зв'язність охоплюючого простору N_n так і індуковану зв'язність (1.29) вкладеного простору M_m.

$$\overset{\prime}{\nabla}_{\gamma} B_{\alpha}^{i} = \frac{\partial B_{\alpha}^{i}}{\partial x^{\gamma}} + B_{\alpha}^{j} B_{\gamma}^{k} \Gamma_{jk}^{i} - B_{\beta}^{i} \overset{\prime}{\Gamma}_{\alpha\gamma}^{\beta}. \tag{1.30}$$

Коваріантна похідна $\overset{\prime}{\nabla}$ мішаного тензора має назву *коваріантної похідної Ван дер Вардена-Бортолотті.*

Коваріантно диференціюючи (1.28) використовуючи (1.30) можна впевнитися, що

$$\overset{\prime}{\nabla}_{\gamma} B_{\alpha}^{i} g_{ij} B_{\beta}^{j} = 0. \tag{1.31}$$

З (1.31) випливає, що є ортогональним до всіх векторів, що утворюють дотичний простір $T_{x_0}\left(M_m\right)$.

Тому ми можемо в кожній точці M_m побудувати *n-m* взаємно перпендикулярних одиничних векторів:

$$C_{|m+1}^{i}, C_{|m+2}^{i}, \ldots, C_{|n}^{i}$$

так щоб разом з векторами $B_1^i, B_2^i, \ldots, B_m^i$ утворювали базис всього охоплюючого дотичного простору $T_{y(x_0)}\left(N_n\right)$.

Враховуючи (1.31), можна записати розкладання по векторах утвореного базису:

$$\overset{\prime}{\nabla}_{\gamma} B^i_{\alpha} = \sum_{\sigma=m+1}^{n} H_{\gamma\alpha|\sigma} C^i_{|\sigma}. \tag{1.32}$$

Коефіцієнти розкладання $H_{\gamma\alpha|\sigma}$ вектора $\overset{\prime}{\nabla}_{\gamma} B^i_{\alpha}$ по векторах $C^i_{|\sigma}$, $\sigma = m+1, m+2, \ldots, n$ є *другими фундаментальними тензорами* по відношенню до векторів $C^i_{|\sigma}$ [14]. Першим *фундаментальним тензором* вважається індукована на многовиді метрикою $\overset{\prime}{g}_{\alpha\beta}$.

Якщо ми коваріантно продиференціюємо нормальний вектор, то виявляється, що його похідна має вигляд:

$$\overset{\prime}{\nabla}_{\gamma} C^i_{|\sigma} = -\sum_{\alpha=1}^{m} H^{\alpha}_{\gamma|\sigma} B^i_{\alpha} + \sum_{\tau=m+1}^{n} L_{\gamma|\sigma\tau} C^i_{|\tau}. \tag{1.33}$$

Формула (1.32) має назву *формули Гауса* а (1.33) *формули Вейнгартена.* Також вони мають спільну назву – *деріваційні рівняння* M_m вкладеного у N_n.

Щоб виписати умови інтегрованості деріваційних рівнянь, нам знадобиться ще одне поняття. В (1.11) ми ввели тензор Рімана другого роду. *Тензор Рімана першого роду* можна отримати, опускаючи у тензора Рімана другого роду контраваріантний індекс за допомогою метрики:

$$R_{ijkl} \overset{\text{def}}{=} g_{ih} R^h_{jkl}. \tag{1.34}$$

Зауважимо, що тензор Рімана другого роду можна отримати для будь-якої афінної зв'язності, наявність метрики не є обов'язковою. Тензор Рімана першого роду (1.34) можна отримати для метричного простору, причому зв'язність, для якої розраховано R^h_{jlk} має бути узгоджена з метрикою g_{ij}.

Якщо дослідити умови інтегрованості (1.32) та (1.33), то ми отримаємо такі рівняння:

$$\overset{`}{R}_{\alpha\beta\gamma\delta} = \sum_{\sigma=m+1}^{n} \left(H_{\alpha\gamma|\sigma} H_{\beta\delta|\sigma} - H_{\alpha\gamma|\sigma} H_{\beta\gamma|\sigma} \right) + R_{ijkl} B^i_{\alpha} B^j_{\beta} B^k_{\gamma} B^l_{\delta}, \tag{1.35}$$

$$H_{\alpha\beta,\gamma|\sigma} - H_{\alpha\gamma,\beta|\sigma} = \sum_{\tau=m+1}^{n} \left(L_{\gamma|\tau\sigma} H_{\alpha\beta|\tau} - L_{\beta|\tau\sigma} H_{\alpha\gamma|\sigma} \right) + R_{ijkl} B^i_{\alpha} B^k_{\beta} B^l_{\gamma} C^j_{|\sigma}, \tag{1.36}$$

$$\sigma = m+1, m+2, \ldots, n$$

Рівняння (1.35) називають *формулою Гауса,* а (1.36) – *формулою Кодаццi.*

Важливим частинним випадком підмноговидів є поняття автопаралельних та цілком геодезичних многовидів.

Многовид M_m вкладений у N_n називають *геодезичним у точці* $p \in M_m$, якщо будь-яка геодезична многовиду M_m, що проходить через точку p є також геодезичною для простору N_n. Якщо M_m є геодезичним у будь-якій свої точці, то тоді його називають *цілком геодезичним підмноговидом* простору N_n [3; 12]. Підмноговид $M_m \subset N_n$ називають *автопаралельним*, якщо для кожного вектора, дотичного $T_{x_0}\left(M_m\right)$ і кожної кривої γ, що виходить з x_0, паралельний перенос цього вектору вздовж γ (відносно аффінної зв'язності охоплюючого простору N_n) призводить до того, що перенесений вектор лишається дотичним до M_m. Наприклад, сфера – не є автопаралельним у нашому тривимірному просторі. Якщо ми візьмемо будь-яку точку на мередіані глобуса, уявімо собі дотичний до мередіана вектор і почнемо його паралельно переносити уздовж меридіану – то побачимо, що він у процесі перенесення перестає бути дотичним як до меридіана, так і до кола в цілому (рис. 3).

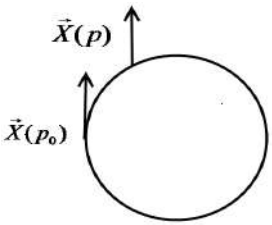


Рисунок 3 – Паралельне перенесення вектора відносно афінної зв'язності охоплюючого простору $\mathbb{R}^3$ уздовж мередіана сфери – підмноговида тривимірного простору

До речі, саме меридіани є геодезичними лініями сфери. А от серед паралелей – геодезичною є лише екватор.

Існує таке твердження: кожний автопаралельний підмноговид M_m, вкладений у N_n буде в N_n цілком геодезичним, якщо афінна зв'язність в N_n є симетричною. І навпаки, симетричність афінної зв'язності N_n

гарантує, що будь-який цілком геодезичний його підмноговид M_m буде автопаралельним.

Оскільки ми будемо в наших задачах стикатися лише із симетричною зв'язністю, то для нас автопаралельний та цілком геодезичні многовиди будуть синонімічними поняттями.

Дуже важливою ознакою того що многовид M_m, вкладений у N_n буде цілком геодезичним є той факт, що коефіцієнти другої квадратичною форми тотожно дорівнюють нулю:

$$H_{\gamma\alpha|\sigma} \equiv 0, \gamma, \alpha = 1, 2, \ldots, m; \sigma = m+1, m+2, \ldots, n. \qquad (1.37)$$

І навпаки, якщо підмноговид є цілком геодезичним (автопаралельним), то має виконуватися (1.35).

У такому випадку деривaційні формули приймають вигляд:

$$\overset{\prime}{\nabla}_\gamma B^i_\alpha = 0;$$

$$\overset{\prime}{\nabla}_\gamma C^i_{|\sigma} = \sum_{\tau=m+1}^{n} L_{\gamma|\sigma\tau} C^i_{|\tau}.$$

Стосовно формул Гауса та Кодацці, то вони відповідно виглядатимуть для цілком геодезичних підмноговидів так:

$$\overset{\prime}{R}_{\alpha\beta\gamma\delta} = R_{ijkl} B^i_\alpha B^j_\beta B^k_\gamma B^l_\delta;$$

$$R_{ijkl} B^i_\alpha B^k_\beta B^l_\gamma C^j_{|\sigma} = 0.$$

З викладеного видно, що перший основний тензор $\overset{\prime}{g}_{\alpha\beta}$ цілком визначає внутрішню геометрію вкладеного многовиду, другий основний тензор $H_{\gamma\alpha|\sigma}$ – показує як саме підноговид M_m є розташованим у охоплюючому просторі N_n.

Наостанок зауважимо ще один дуже важливий момент. Ми дуже часто казали і будемо казати, що підмноговид M_m є **вкладений** у N_n. Термінологічно – це не зовсім точно, оскільки весь наведений вище математичний апарат працює лише локально а не в цілому.

Термін **вкладання** – гарантує що мноногвид M_m взаємно-однозначно відобразиться на деяку поверхню у N_n, без розривів, самоперетинів, склеювання. Тобто, як кажуть, із повним збереженням топології. При локальному підході – ми можемо гарантувати взаємно-однозначність лише в межах деякої локальної карти

многовида. Для такого випадку існує термін: *занурення*, або *імерсія*. Але оскільки дослідження топології статистичних многовидів лежить поза межами даної роботи, то поняття «вкладання» і «занурення» ми будемо використовувати як синонімічні.

2. Основи інформаційної геометрії

2.1. Метод максимальної правдоподібності (Maxitum likelihood estimation, MLE)

Нехай у нас є результати *n* спостережень деякої випадкової величини $x_1, x_2, \ldots x_n$, нам відома загальна формула її розподілу $p(x,\xi)$, але параметри $\xi = (\xi^1, \xi^2, \ldots, \xi^m)$ цього розподілу невідомі.

Зауваження 2.1. *Уточнимо, що без втрати загальності ми можемо вважати випадкову величину одновимірною. Для k-вимірної величини – кожному результату випробування відповідає k-вимірний вектор, а інтегральна функція розподілу виражається k-кратним інтегралом.*

Підставляємо результати у функцію розподілу і отримані *n* виразів перемножуємо:

$$L(x_1, x_2, \ldots x_n; \xi) = p(x_1, \xi)\, p(x_2, \xi) \ldots p(x_n, \xi).$$

Отримана функція має назву *функції правдоподібності [3].*

Зручніше користуватися *логарифмічною функцією правдоподібності:*

$$l(x_1, x_2, \ldots x_n; \xi) = \ln L(x_1, x_2, \ldots x_n; \xi) = \sum_{i=1}^{n} \ln p(x_i, \xi).$$

Вона є функцією від *m* змінних $\xi^1, \xi^2, \ldots, \xi^m$.

Сутність методу максимальної правдоподібності полягає в тому, що значення невідомих параметрів мають бути такими, що при даних емпіричних результатах функція правдоподібності набувала б найбільшого значення, тобто ми маємо задачу знаходження безумовного екстремуму функції.

Для цього знайдемо частині похідні:

$$\frac{\partial l}{\partial \xi^j} = \sum_{i=1}^{n} \frac{1}{p(x_i, \xi)} \frac{\partial p(x_i, \xi)}{\partial \xi^j}.$$

Отриманий вектор градієнта має назву *функція внеску (score function).*

Для визначення оцінок параметрів прирівнюємо функцію внеску до нуля:

$$\frac{\partial l}{\partial \xi^j} = \sum_{i=1}^{n} \frac{1}{p(x_i, \xi)} \frac{\partial p(x_i, \xi)}{\partial \xi^j} = 0 \, (i = 1, 2, \ldots m) . \quad (2.1)$$

Ми отримали систему m рівнянь з m невідомими. Взагалі кажучи, система є нелінійною, і тому її розв'язування може бути пов'язаним з обчислювальними труднощами, знайдена критична точка $\hat{\xi} = \left(\hat{\xi}^1, \hat{\xi}^2, \ldots, \hat{\xi}^m \right)$ може бути точкою лише локального екстремуму, тощо. В цьому полягає один з недоліків методу.

Але в будь-якому випадку, якщо вдалося знайти критичну точку, її потрібно перевірити на те, що вона є точкою хоча б локального максимуму.

Для цього складаємо матрицю Гессе:

$$H\left(\hat{\xi}\right) = \begin{pmatrix} \frac{\partial^2 l}{\partial \xi^{12}} & \cdots & \frac{\partial^2 l}{\partial \xi^1 \partial \xi^k} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 l}{\partial \xi^k \partial \xi^1} & \cdots & \frac{\partial^2 l}{\partial \xi^{k2}} \end{pmatrix}. \quad (2.2)$$

Якщо матриця (2.2) відповідатиме негативно визначеній квадратичній формі,

$$\sum_{i,j} \frac{\partial^2 l}{\partial \xi^i \partial \xi^j} d\xi^i d\xi^j < 0,$$

це гарантує, що в критичній точці $\hat{\xi} = \left(\hat{\xi}^1, \hat{\xi}^2, \ldots, \hat{\xi}^m \right)$ функція правдоподібності має локальний максимум. Таким чином, ми знайшли оцінки параметрів розподілу. Такі оцінки є *слушними (конзистентними)*, але на жаль не завжди є незміщеними. Якщо для параметра набору параметрів $\hat{\xi}$ існує ефективна оцінка, то система (2.1) матиме один і тільки один розв'язок.

Якщо виконується остання нерівність, то зрозуміло, що виконується протилежна:

$$-\sum_{i,j}\frac{\partial^2 l}{\partial\xi^i\partial\xi^j}d\xi^i d\xi^j > 0.$$

А це означає що матриця $\left(-\frac{\partial^2 l}{\partial\xi^i\partial\xi^j}\right)$ задає позитивно визначену квадратичну форму, і тому може відігравати роль метрики на деякому многовиді. Розглянемо такий многовид. Що стосується метода максимальної правдоподібності, то ми до нього повернемось у розділі 4.3.

2.2. Статистичні многовиди

Розглянемо регулярну статистичну модель

$$M = \{p(x,\xi)\}.$$

Вона представляє собою множину функцій що є щільностями розподілу деякої випадкової величини x, параметризовану m-вимірним вектором параметрів $\xi = (\xi^1, \xi^2, \ldots, \xi^m)$. Таку модель можна розглядати як m-вимірний многовид, причому $\xi = (\xi^1, \xi^2, \ldots, \xi^m)$ є локальною системою координат, що задана на цьому многовиді.

В даному випадку *функцією внеску (score function)* ми також називаємо m-вимірний вектор $s(x,\xi)$ але цього разу компоненти його визначаються як частинні похідні натурального логарифму щільності розподілу $l_\xi = \ln p(x,\xi)$:

$$s_i(x,\xi) = \partial_i l_\xi = \frac{\partial_i p(x,\xi)}{p(x,\xi)}, \quad i = 1,\ldots,m. \tag{2.3}$$

Тут ми позначили $\partial_i = \frac{\partial}{\partial\xi^i}$.

Очевидно, що виконується така рівність, що свідчить про нульове математичне сподівання функції внеску (аналог формули (2.1)):

$$E_\xi\left[s_i(x,\xi)\right] = 0, \quad i = 1,\ldots,m. \tag{2.4}$$

Тут величина E_ξ – математичне сподівання $p(x,\xi)$.

Фактично маємо [15; 3]:

$$E_\xi\left[s_i(x,\xi)\right]=\int\partial_i l_\xi(x,\xi)\,p(x,\xi)\,dx=\int\partial_i\left(\ln\left(p(x,\xi)\right)\right)p(x,\xi)\,dx=$$
$$=\int\frac{\partial_i p(x,\xi)}{p(x,\xi)}\,p(x,\xi)\,dx=\int\partial_i p(x,\xi)\,dx=\partial_i\int p(x,\xi)\,dx=\partial_i 1=0. \qquad (2.4')$$

Зауваження 2.2. *У формулах* (2.4') *мається на увазі визначений інтеграл, хоча границь інтегрування не вказано. Інтегрування відбувається по всій області що є множиною усіх можливих значень випадкової величин, або як кажуть у математичному аналізі, по всьому «носію» функції* $p(x,\xi)$ *стосовно аргументу x. Носієм функції* $f(x)$ *ми називаємо множину значень аргументу, для яких f* $(x)\neq 0$. *Також, якщо випадкова величина є k-вимірною, то у (2.4) матиметься на увазі k-кратний інтеграл по k-вимірній області. Але оскільки обчислення у (2.4') мають суто загальний характер – можна задовільнитися спрощеною символікою.*

Компонент $s_i(x,\xi)=\partial_i l_\xi$ можна розглядати як дотичний вектор до многовиду *M* в напрямі координатної кривої ξ^i (Рис. 4). Звідси базисний вектор:

$$e_i=\frac{\partial}{\partial\xi^i}$$

є представленим внеском $s_i(x,\xi)$. Очевидно, він є випадковою змінною. Тому ми маємо випадкове представлення базисних векторів $e_i=\partial_i l_\xi$.

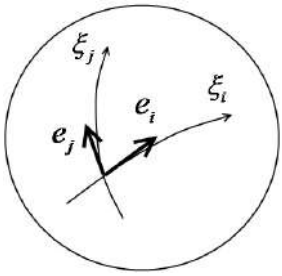


Рисунок 4 – Вектори базису дотичного простору до статистичного многовиду

Далі ми можемо визначити скалярний добуток двох базисних векторів:

$$e_i, e_j = E_\xi\left[\partial_i l_\xi \partial_j l_\xi\right] = \int \frac{\partial \ln\left(p(x,\xi)\right)}{\partial \xi^i} \frac{\partial \ln\left(p(x,\xi)\right)}{\partial \xi^j} p(x,\xi)\,dx = \\ = \int \frac{\partial p(x,\xi)}{\partial \xi^i} \frac{\partial p(x,\xi)}{\partial \xi^j} \frac{1}{p(x,\xi)}\,dx. \tag{2.5}$$

Таким чином ми отримуємо матрицю $G(\xi) = g_{ij}$ яка також має назву *інформаційна матриця Фішера:*

$$g_{ij} \overset{def}{=} e_i, e_j = E\left[\partial_i l_\xi \partial_j l_\xi\right]. \tag{2.5'}$$

Розглянемо $E\left[\partial_j \partial_i l_\xi\right]$:

$$E\left[\partial_j \partial_i l_\xi\right] = \int \frac{\partial}{\partial \xi^j}\left(\frac{\partial \ln\left(p(x,\xi)\right)}{\partial \xi^i}\right) p(x,\xi)\,dx = \\ = \int \frac{\partial}{\partial \xi^j}\left(\frac{\partial_i p(x,\xi)}{p(x,\xi)}\right) p(x,\xi)\,dx = \\ = \int \frac{\partial_j \partial_i p(x,\xi)\, p(x,\xi) - \partial_i p(x,\xi) \partial_j p(x,\xi)}{\left(p(x,\xi)\right)^2} p(x,\xi)\,dx = \\ = \int \frac{\partial_j \partial_i p(x,\xi)\, p(x,\xi) - \partial_i p(x,\xi) \partial_j p(x,\xi)}{p(x,\xi)}\,dx = \\ = \int \frac{\partial_j \partial_i p(x,\xi)\, p(x,\xi)}{p(x,\xi)}\,dx - \int \frac{\partial_i p(x,\xi) \partial_j p(x,\xi)}{p(x,\xi)}\,dx = \\ = \int \partial_j \partial_i p(x,\xi)\,dx - E\left[\partial_i l_\xi \partial_j l_\xi\right].$$

Внаслідок (2.4'):

$$\int \partial_j \partial_i p(x,\xi)\,dx = \partial_j \int \partial_i p(x,\xi)\,dx = 0.$$

Отже остаточно:

$$E\left[\partial_j \partial_i l_\xi\right] = -E\left[\partial_i l_\xi \partial_j l_\xi\right].$$

Отримаємо:

$$g_{ij} = -E\left[\partial_j \partial_i l_\xi\right]. \tag{2.6}$$

Також можна довести, що для g_{ij} існує і таке представлення:

$$g_{ij} = 4\int \partial_i \sqrt{p(x,\xi)}\,\partial_j \sqrt{p(x,\xi)}\,dx. \tag{2.7}$$

Очевидно що матриця $G(\xi) = g_{ij}$ симетричною і для будь-якого m-вимірного вектору $\boldsymbol{a} = \left(a^1, a^2, \ldots, a^m\right)^T$ (символ T означає транспонування) маємо:

$$aG(\xi)a^T = g_{ij}a^i a^j = \int\left(\sum_{i=1}^{m} a^i \partial_i l_\xi(x,\xi)\right)^2 p(x,\xi)\,dx \geq 0,$$

тобто квадратична форма, що задається матрицею $G(\xi)$ є позитивно напіввизначеною. Але оскільки ми вважаємо вектори $\left\{\partial_1 l_\xi, \partial_2 l_\xi, \ldots, \partial_m l_\xi\right\}$ лінійно-незалежними, то побудована нами за допомогою них матриця задаватиме позитивно визначену квадратичну форму. Отже, g_{ij} задає метрику на многовиді M.

Очевидно, оскільки на многовиді задана метрика, то і існуватиме і узгоджена з нею зв'язність. Її компоненти можуть бути визначені за формулою (1.20).

$${}^0\Gamma_{ij}^k = \frac{g^{k\alpha}}{2}\left(\frac{\partial g_{i\alpha}}{\partial \xi^j} + \frac{\partial g_{\alpha j}}{\partial \xi^i} - \frac{\partial g_{ij}}{\partial \xi^\alpha}\right), \tag{2.8}$$

де $g^{k\alpha}$ означає матрицю, обернену до метричного тензору $g_{i\alpha}$. Нулем, що стоїть зліва від гамми, позначається об'єкт зв'язності Леві – Чівіти.

Статистичним многовидом ми називаємо трійку (M, g, T), де M є зв'язним та гладкий многовид, g – ріманова метрика і нарешті T є гладке коваріантне цілком симетричне тензорне поле валентності три, що має назву *асиметрії (skewness)* [3; 15].

Асиметрія визначається за формулою:

$$T_{ijk} \overset{\text{def}}{=} E\left[\partial_i l_\xi \partial_j l_\xi \partial_k l_\xi\right], \tag{2.9}$$

або більш докладно:

$$T_{ijk} = \int \frac{\partial \ln\left(p(x,\xi)\right)}{\partial \xi^i} \frac{\partial \ln\left(p(x,\xi)\right)}{\partial \xi^j} \frac{\partial \ln\left(p(x,\xi)\right)}{\partial \xi^k} p(x,\xi)\,dx. \tag{2.9'}$$

Використовуючи тензор (2.9) ми можемо задати так звану α-зв'язність $\nabla^{\pm}$ і її коефіцієнти визначаються таким чином:

$$ {}^{\alpha}\Gamma_{ij,k} \overset{\text{def}}{=} {}^{0}\Gamma_{ij,k} - \frac{\alpha}{2} T_{ijk}. \tag{2.10} $$

Загалом, α є деяким дійсним числом. Але для статистичних досліджень вважають $\alpha = \pm 1$. Це буде обговорено нижче.

Приймаючи до уваги (2.5), (2.6), (2.8)–(2.10) в кожній точці $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$ ми отримуємо:

$$ {}^{\alpha}\Gamma_{ij,k} = \int \left(\partial_j \left(\frac{\partial_i p(x,\xi)}{p(x,\xi)} \right) + \frac{1-\alpha}{2} \frac{\partial_i p(x,\xi)}{p(x,\xi)} \frac{\partial_j p(x,\xi)}{p(x,\xi)} \right) \times $$

$$ \times \frac{\partial_k p(x,\xi)}{p(x,\xi)} p(x,\xi) dx = $$

$$ = \int \left(\begin{array}{l} \dfrac{\partial_j \partial_i p(x,\xi) \partial_k p(x,\xi)}{\left(p(x,\xi)\right)^2} - \dfrac{\partial_i p(x,\xi)}{p(x,\xi)} \dfrac{\partial_j p(x,\xi)}{p(x,\xi)} \dfrac{\partial_k p(x,\xi)}{p(x,\xi)} + \\ + \dfrac{1-\alpha}{2} \dfrac{\partial_i p(x,\xi)}{p(x,\xi)} \dfrac{\partial_j p(x,\xi)}{p(x,\xi)} \dfrac{\partial_k p(x,\xi)}{p(x,\xi)} \end{array} \right) p(x,\xi) dx. $$

Остаточно маємо:

$$ {}^{\alpha}\Gamma_{ij,k} = $$

$$ = \int \left(\frac{\partial_j \partial_i p(x,\xi) \partial_k p(x,\xi)}{p(x,\xi)} - \frac{1+\alpha}{2} \frac{\partial_i p(x,\xi) \partial_j p(x,\xi) \partial_k p(x,\xi)}{\left(p(x,\xi)\right)^2} \right) dx. \tag{2.11} $$

або

$$ {}^{\alpha}\Gamma_{ij,k} = E\left[\left(\partial_j \partial_i l_\xi(x,\xi) + \frac{1-\alpha}{2} \partial_i l_\xi(x,\xi) \partial_j l_\xi(x,\xi) \right) \partial_k l_{\xi(x,\xi)} \right]. \tag{2.12} $$

Також, диференціюючи (2.5) по ξ^k маємо:

$$ \begin{aligned} \frac{\partial g_{ij}}{\partial \xi^k} &= E\left[\left(\partial_k \partial_i l_\xi \right) \left(\partial_j l_\xi \right) \right] + E\left[\left(\partial_i l_\xi \right) \left(\partial_k \partial_j l_\xi \right) \right] + \\ &+ E\left[\left(\partial_i l_\xi \right) \left(\partial_j l_\xi \right) \left(\partial_k l_\xi \right) \right] = {}^{0}\Gamma_{ki,j} + {}^{0}\Gamma_{kj,i}. \end{aligned} \tag{2.13} $$

Розглядаючи (2.12) та (2.13), порівнюючи з (2.8), ми приходимо до висновку, що коли $\alpha = 0$, ми отримуємо 0-зв'язність ∇^0, яка є ріманновою зв'язністю, узгодженою з метрикою Фішера.

В заключення цього розділу варто зазначити, що матриця, обернена до інформаційної матриці Фішера використовується для оцінки середньої квадратичної помилки незміщеної оцінки а α-зв'язність $\nabla^{\pm}$ застосовується для обчислення дивергенції між різними статистичними розподілами.

2.3. Приклади статистичних многовидів

Розглянемо деякі приклади [3]. Позначимо $\mathcal{X}$ множину значень, яку може приймати випадкова величина.

Відповідно, $\mathcal{P}(\mathcal{X})$ – множину можливих розподілів $p(x)$ випадкової величини, літерою Ξ позначимо простір параметрів розподілу.

Формально це виглядатиме так:

$$\mathcal{P}(\mathcal{X}) \overset{\text{def}}{=} \left\{ p : \mathcal{X} \to \mathbb{R} \,\middle|\, \forall x \in \mathcal{X}\, p(x) \rangle 0, \int p(x)\,dx = 1 \right\}.$$

1. Нормальний розподіл. $\mathcal{X} = \mathbb{R}$, розмірність множини параметрів $\dim \Xi = 2$,

$$\Xi = \left\{ [\mu, \sigma] \,|\, -\infty < \mu < \infty, 0 < \sigma < \infty \right\};$$

$$p(x, \xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right).$$

2. Багатовимірний нормальний розподіл. $\mathcal{X} = \mathbb{R}^k$, розмірність множини параметрів $\dim \Xi = k + \dfrac{k(k+1)}{2}$,

$$\Xi = \left\{ [\mu, \Sigma] \,|\, \mu \in \mathbb{R}^k, \Sigma \in \mathbb{R}^{k \times k} \right\},$$

$$p(x, \xi) = \frac{1}{\left(\sqrt{2\pi}\right)^k \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu) \right).$$

Зазначимо, що матриця Σ утворює позитивно визначену квадратичну форму.

3. Розподіл Пуассона. $\mathcal{X}=\{0,1,2,\ldots.\}$, розмірність множини параметрів $\dim\Xi=1$,

$$\Xi=\{\xi|\xi\rangle 0\},$$

$$p(x,\xi)=e^{-\xi}\frac{\xi^x}{x!}.$$

4. Розподіл дискретної випадкової величини із скінченною кількістю можливих значень. $\mathcal{X}=\{x_0,x_1,x_2,\ldots x_n\}$, розмірність множини параметрів $\dim\Xi=n$

$$\Xi=\left\{\left[\xi^0,\xi^1,\ldots\xi^n\right]|\ \forall i\ \xi^i>0,\sum_{i=1}^{n}\xi^i<1\right\},$$

$$p(x,\xi)=\begin{cases}\xi^i\ (1\le i\le n)\\ 1-\sum_{i=1}^{n}\xi^i\ (i=0).\end{cases}$$

3. Геометричні властивості статистичних многовидів

3.1. Статистичні многовиди і деякі дифеоморфізми

Нехай афінна зв'язність ∇ допускає взаємно однозначне диференційоване відображення f на іншу афінну зв'язність $\bar{\nabla}$ таке, що обернене відображення є також диференційованим. Локально обидві зв'язності є симетричними та віднесені до спільної координатної системи $\xi=\left(\xi^1,\xi^2,\ldots,\xi^m\right)$.

Тензор деформації відносно відображення f виглядає так:

$$P_{ij}^h(\xi)=\bar{\Gamma}_{ij}^h(\xi)-\Gamma_{ij}^h(\xi),\qquad(3.1)$$

де $\Gamma_{ij}^h(\xi)$ та $\bar{\Gamma}_{ij}^h(\xi)$ – компоненти афінних зв'язностей відповідно ∇ та $\bar{\nabla}$ відповідно.

Дифеоморфізм $f:\nabla\to\bar{\nabla}$ називається *геодезичним* відображенням, якщо будь-яка геодезична лінія відносно зв'язності ∇ відображається під дією f на геодезичну відносно $\bar{\nabla}$.

Відомо, що для того, щоб відображення f простору з афінною зв'язністю ∇ на інший простір з афінною зв'язністю $\bar{\nabla}$ було

геодезичним, необхідно та достатньо, щоб у спільній системі координат тензор деформації можна було представити у вигляді:

$$P_{ij}^{h}\left(\xi\right)=\psi_{i}\left(\xi\right)\delta_{j}^{h}+\psi_{j}\left(\xi\right)\delta_{i}^{h}, \tag{3.2}$$

де δ_{i}^{h} є вже знайомим символом Кронекера ψ_{i} – деякий ковектор. Геодезичне вдбраження називають *нетривіальним*, якщо $\psi\not\equiv 0$.

Виникає питання, чи є можливим, щоб усі геодезичні відносно рімановоі зв'язності ∇^{0} співпадали з геодезичними α-зв'язності $\nabla^{\pm}$ якщо $\alpha\neq 0$. Якщо ми маємо таку відповідність, то

$$P_{ijk}=-\frac{\alpha}{2}T_{ijk}. \tag{3.3}$$

Прийнявши до уваги (3.2) і опускаючи верхній індекс за допомогою g_{hk}, маємо:

$$-\frac{\alpha}{2}T_{ijk}=\psi_{i}g_{jk}+\psi_{j}g_{ik}. \tag{3.4}$$

Міняючи у (3.4) місцями два останні індекси, ми отримуємо:

$$-\frac{\alpha}{2}T_{ikj}=\psi_{i}g_{kj}+\psi_{k}g_{ij}.$$

Оскільки тензор $r\ T_{ijk}$ є цілком симетричним, можна записати:

$$\psi_{i}g_{jk}+\psi_{j}g_{ik}=\psi_{i}g_{kj}+\psi_{k}g_{ij},$$

або

$$\psi_{j}g_{ik}=\psi_{k}g_{ij}. \tag{3.5}$$

Згортаючи (3.5) з $\psi^{j}=\psi_{h}g^{hj}$, ми отримуємо:

$$\left\|\psi\right\|^{2}g_{ik}=\psi_{k}\psi_{i}. \tag{3.6}$$

Тут $\left\|\psi\right\|^{2}=\psi^{i}\psi^{j}g_{ij}=\psi^{j}\psi_{j}$. Але відомо, що $rang\left(g_{ij}\right)\geq 2$. (3.5) містить протиріччя. Звідси ми отримуємо теорему.

Теорема 3.1. *Статистичний многовид* $\left(M,g,T\right)$ *не дозволяє нетривіальних геодезичних відображень між рімановою зв'язністю* ∇^{0} *та будь-якою* α*-зв'язністю* ∇^{α} *якщо* $\alpha\neq 0$.

Тепер розглянемо випадок зв'язності Вейля, тобто існування такої зв'язності Вейля по відношенню до метричного тензору g_{ij},

яка співпадатиме з α-зв'язністю $\nabla^{\pm}$ якщо $\alpha \neq 0$. В цьому випадку (3.1) та (3.3) також виконуються, але замість (3.2) ми маємо:

$$P_{ij}^{h}(\xi) = \varphi_i(\xi)\delta_j^h + \varphi_j(\xi)\delta_i^h - \varphi^h g_{ij},$$

де φ_i – деякий ковектор. Відповідно, замість (3.4) запишемо:

$$-\frac{\alpha}{2}T_{ijk} = \varphi_i g_{jk} + \varphi_j g_{ik} - \varphi_k g_{ij}.$$

Застосовуючи техніку подібну до того, як ми робили вище, отримуємо:

$$2\varphi_j g_{ik} = 2\varphi_k g_{ij}. \tag{3.7}$$

Згорнемо (3.7) з $\psi^j = \psi_h g^{hj}$ та поділимо на 2. Це нам дає

$$\|\varphi\|^2 g_{ik} = \varphi_k \varphi_i,$$

де $\|\varphi\|^2 = \varphi^i \varphi^j g_{ij} = \varphi^j \varphi_j$. Як і вище, ми отримали протиріччя, і тому можемо сформулювати теорему.

Теорема 3.2. *Статистичний многовид* (M, g, T) *не дозволяє зв'язності Вейля по відношенню до метричного тензору* g_{ij}*, яка співпадатиме з* α*-зв'язністю* ∇^{α} *якщо* $\alpha \neq 0$.

3.2. Експоненціальні і мішані сімейства як структури Гессе

Експоненціальне сімейство розподілів ймовірності може бути представленим записом щільності у так званій канонічній формі:

$$p(x,\theta) = \exp\left(C(x) + \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right).$$

Для взаємно-однозначної параметризації є необхідним та достатнім, щоб $n+1$ функцій $\{F_1, \ldots, F_n, 1\}$, де 1 означає сталу фукцію, що тотожно дорівнює одиниці. Параметри θ^i мають назву *канонічних*. Очевидно, $\psi(\theta)$ не є довільною функцією. Вона однозначно визначається формулою:

$$\psi(\theta) = \ln \int exp\left(C(x) + \sum_{i=1}^{n} \theta^i F_i(x) \right) dx.$$

Відповідно до формул (2.5) та (2.6) для експоненціального сімейства:

$$g_{ij} = \frac{\partial^2 \psi(\theta)}{\partial \theta^i \partial \theta^j}. \tag{3.8}$$

З (2.8), (2.12) та (3.8) випливає, що

$${}^0\Gamma_{ij,k} = \frac{1}{2} \frac{\partial^3 \psi(\theta)}{\partial \theta^i \partial \theta^j \partial \theta^k}.$$

Такі структури мають назву *структур Гессе* [3; 17]. Існує така теорема [17].

Теорема 3.3. *Нехай* (M, g) *є рімановим многовидом. Тоді наступні умови є еквівалентними.*

1. g_{ij} є метрика Гессе, тобто локально існує потенціальна функція $\psi(\theta)$ така, що $g_{ij} = \partial_i \partial_j \psi(\theta)$.

2. $\dfrac{\partial g_{ij}}{\partial \theta^k} = \dfrac{\partial g_{ik}}{\partial \theta^j}$.

3. Символ Кристофеля першого роду є цілком симетричним: $\Gamma_{ij,k} = \Gamma_{ik,j}$.

Це з очевидністю випливає леми Пуанкаре.

Також з (2.9), впливає, що

$$T_{ijk} = \frac{\partial^3 \psi(\theta)}{\partial \theta^i \partial \theta^j \partial \theta^k}. \tag{3.9}$$

Беручі до уваги (2.10), маємо

$${}^\alpha\Gamma_{ij,k} = \frac{1-\alpha}{2} \frac{\partial^3 \psi(\theta)}{\partial \theta^i \partial \theta^j \partial \theta^k}. \tag{3.10}$$

Ми бачимо, що ${}^1\Gamma_{ij,k} = 0$, тобто $\left[\theta^i\right]$ – афінна система координат і наш статистичний многовид є 1-плоский. Використовуючи (1.11), обчислимо ріманів тензор кривини:

$${}^\alpha R^h_{ijk} = \frac{1-\alpha^2}{2}\left(g^{ht} T_{tsk} g^{sp} T_{ijp} - g^{ht} T_{tsj} g^{sp} T_{ikp}\right). \tag{3.11}$$

Тут g^{ht}-матриця, обернена до g_{ij}, що розрахована за формулою (3.8), T_{ijk} отриманий згідно (3.9). Ми можемо бачити, що тензор кривини стає

нульовим не тільки, якщо $\alpha = 1$, але і у випадку коли $\alpha = -1$. Це означає, що також існує координатна система в якій ${}^{-1}\Gamma_{ij,k}$ тотожно дорівнює нулю. Щоб отримати таку координатну систему ми здійснимо так зване перетворення Лежандра:

$$\eta_i = \frac{\partial \psi(\theta)}{\partial \theta^i}.$$

Воно дає нам дуальну опуклу функцію:

$$\varphi(\eta) = \theta^i(\eta)\eta_i - \psi\left(\theta(\eta)\right)$$

де

$$\theta^i = \frac{\partial \varphi(\eta)}{\partial \eta_i}$$

є оберненим перетворенням.

Ми отримуємо, що [3]

$${}^{\alpha}\Gamma^{ij,k} = \frac{1+\alpha}{2}\frac{\partial^3 \varphi(\eta)}{\partial \eta_i \partial \eta_j \partial \eta_k}.$$

Можна побачити, що коефіцієнти зв'язності обертаються в нуль при $\alpha = -1$.

Зауваження 3.1. *Порівнюючи формули (3.9) та (3.10) можна записати:*

$${}^{\alpha}\Gamma_{ij,k} = \frac{1-\alpha}{2}T_{ijk}.$$

Тим не менш, ми цього намагалися уникати. Справа в тому, що воно є вірним лише для даної системи координат. Якщо ми перейдемо до іншої, то взагалі кажучи, ця рівність перестане бути рівною, оскільки об'єкт зв'язності ${}^{\alpha}\Gamma_{ij,k}$ *перетворюється за законом (1.3), а тензор* T_{ijk} *за формулами (1.2).*

Тепер розглянемо *n*-вимірну модель, виражену у термінах лінійно незалежних функцій $\{C, F_1, \ldots, F_n\}$ як

$$p(x,\theta) = C(x) + \sum_{i=1}^{n} \theta^i F_i(x). \tag{3.12}$$

Формула (3.12) визначає так зване *мішане сімейство* з *мішаними параметрами* θ^i. Для мішаного сімейства, застосувавши (2.3) маємо

$$\partial_i l_\theta(x,\theta) = \frac{F_i(x)}{p(x,\theta)}, \quad \partial_i \partial_j l_\theta(x,\theta) = -\frac{F_i(x)F_j(x)}{(p(x,\theta))^2}.$$

Ми бачимо, що

$$\partial_i \partial_j l_\theta(x,\theta) + \partial_i l_\theta(x,\theta)\partial_j l_\theta(x,\theta) = 0. \tag{3.13}$$

З (3.12) та (2.11) випливає, що ${}^{-1}\Gamma_{ij,k} = 0$.

Для довільного α маємо

$${}^{\alpha}\Gamma_{ij,k}(\theta) = -\frac{1+\alpha}{2} E\left[\partial_i l_\theta(x,\theta)\partial_j l_\theta(x,\theta)\partial_k l_\theta(x,\theta)\right]. \tag{3.14}$$

З (3.14) випливає, що при $\alpha = 0$ ми отримуємо цілком симетричний символ Крістофеля першого роду для метрики Фішера. Згідно із теоремою 3.3 це означає, що на многовиді існує потенціальна функція $\psi(\theta)$, така що $g_{ij} = \partial_i \partial_j \psi(\theta)$. Таким чином, ми довели теорему.

Теорема 3.4. *Інформаційна матриця Фішера обчислена для моделі мішаного сімейства є метрикою Гессе.*

Розглянемо довільний статистичний многовид (M, g, T). Нехай для деякого значення α усі компоненти ${}^{\alpha}\Gamma_{ij,k}$ перетворюються в нуль. Згідно з (2.10).

$${}^{0}\Gamma_{ij,k} = \frac{\alpha}{2} T_{ijk}. \tag{3.15}$$

Ми знаємо, що згідно з означенням асиметрії (skewness) T_{ijk} є цілком симетричний. Тоді внаслідок (3.15) символ Крістофеля першого роду ${}^{0}\Gamma_{ij,k}$ є цілком симетричним. Ми маємо теорему.

Теорема 3.5. *Нехай статистичний многовид (M, g, T) і для деякого значення компоненти α-зв'язності тотожно обертаються в нуль. Тоді його інформаційна матриця Фішера g є метрикою Гессе.*

Стосовно структур Гессе варто додати ще одне важливе міркування. Справа в тому, що «гессевість» многовида не є інваріантною властивістю для перетворень координат, виключаючи лінійні [3; 17]. Деякий многовид може бути многовидом Гессе для однієї координатної системи, і не бути многовидом Гессе для іншої. Розглянемо випадок, коли метрика многовиду, взагалі кажучи, не є метрикою

Гессе по відношенню до старої координатної системи $(x_1,\ldots,x_n)$, але допускає таку трансформацію координат, що в новій координатній системі $(y_1,\ldots,y_n)$, многовид має метрику Гессе.

Відомо, що символ Крістофеля першого роду у новій координатній системі є представленим в такому вигляді [14]:

$$ {}^0\overline{\Gamma}_{ij,k}(y) = {}^0\Gamma_{\alpha\beta,\gamma}(x)\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j}\frac{\partial x^\gamma}{\partial y^k} + \frac{\partial^2 x^\alpha}{\partial y^i \partial y^j} g_{\alpha\beta}\frac{\partial x^\beta}{\partial y^k}. \quad (3.16) $$

Згідно з теоремою 3.3 рівняння:

$$ {}^0\overline{\Gamma}_{ij,k}(y) = {}^0\overline{\Gamma}_{ik,j}(y) \quad (3.17) $$

має виконуватись тотожно в нових координатах. Беручі до уваги (3.16) та (3.17) отримуємо:

$$ \begin{aligned} &{}^0\Gamma_{\alpha\beta,\gamma}(x)\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j}\frac{\partial x^\gamma}{\partial y^k} - {}^0\Gamma_{\alpha\beta,\gamma}(x)\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^k}\frac{\partial x^\gamma}{\partial y^j} + \\ &+\frac{\partial^2 x^\alpha}{\partial y^i \partial y^j} g_{\alpha\beta}\frac{\partial x^\beta}{\partial y^k} - \frac{\partial^2 x^\alpha}{\partial y^i \partial y^k} g_{\alpha\beta}\frac{\partial x^\beta}{\partial y^j} = 0. \end{aligned} \quad (3.18) $$

Очевидно, якщо система диференціальних рівнянь (3.18) має розв'язок відносно функцій $x^i = x^i(y_1,\ldots,y_n)$, $i=\overline{1,n}$, тоді розглянутий многовид локально допускає можливість введення метрики Гессе.

3.3. Приклади експоненціальних сімейств

1. Нормальний розподіл.

$$ C(x)=0,\ F_1(x)=x,\ F_2(x)=x^2, \theta^1=\frac{\mu}{\sigma^2},\ \theta^2=-\frac{1}{2\sigma^2}, $$

$$ \psi(\theta)=\frac{\mu^2}{2\sigma^2}+\ln\left(\sigma\sqrt{2\pi}\right)=-\frac{\left(\theta^1\right)^2}{4\theta^2}+\frac{1}{2}\ln\left(-\frac{\pi}{\theta^2}\right). $$

2. Багатовимірний нормальний розподіл. Тут можна покласти

$$ C(x)=0,\ F_i(x)=x_i,\ F_{ij}(x)=x_i x_j,\ (i\le j), $$

$$ \theta^i=\sum_j\left(\Sigma^{-1}\right)^{ij}\mu_j,\ \theta^{ii}=-\frac{1}{2}\left(\Sigma^{-1}\right)^{ii},\ \theta^{ij}=-\left(\Sigma^{-1}\right)^{ij}, $$

і також

$$F_A(x) = x\ ,\ F_B(x) = xx^T\ ,\ \theta^A = \Sigma^{-1}\mu\ ,\ \theta^B = -\frac{1}{2}\Sigma^{-1}.$$

В результаті маємо

$$p(x,\theta) = \exp\left(\sum_{1\le i\le k}\theta^i F_i(x) + \sum_{1\le i\le j\le k}\theta^{ij}F_{ij}(x) - \psi(\theta)\right) =$$

$$= \exp\left(\left(\theta^A\right)^T F_A(x) + tr\left(\theta^B F_B(x)\right) - \psi(\theta)\right),$$

де

$$\psi(\theta) = \frac{1}{2}\mu^T\Sigma^{-1}\mu + \frac{1}{2}\ln\left(\det(2\pi\Sigma)\right) =$$

$$= -\frac{1}{4}\left(\theta^A\right)^T\left(\theta^B\right)^{-1}\theta^A + \frac{1}{2}\ln\left(-\pi\left(\theta^B\right)^{-1}\right).$$

3. Розподіл Пуассона.

$$C(x) = -\ln(x!),\ F(x) = x,\ \theta = ln\xi,$$

$$\psi(\theta) = \xi = e^{\xi}.$$

4. Розподіл дискретної випадкової величини із скінченною кількістю можливих значень.

$$C(x) = 0,\ F_i(x) = \begin{cases} 1(x = x_i) \\ 0(x \ne x_i). \end{cases}$$

$$\theta^i = \ln\frac{p(x_i)}{p(x_0)} = \ln\frac{p(x_i)}{1 - \sum_{i=1}^{n}\xi^i}\ ,\ (i = 1,2\ldots n),$$

$$\psi(\theta) = -\ln\left(p(x_0)\right) = -\ln\left(1 - \sum_{i=1}^{n}\xi^i\right) = \ln\left(1 + \sum_{i=1}^{n}\exp(\theta^i)\right).$$

3.4. *e*-зв'язність та *m*-зв'язність на статистичному многовиді

Порівняємо вирази для коефіцієнтів α-звязності експоненціального та мішаного сімейств, тобто вирази (3.10) та (3.14) відповідно:

Експоненціальне сімейство:

$${}^{\alpha}\Gamma_{ij,k} = \frac{1-\alpha}{2}\frac{\partial^3\psi(\theta)}{\partial\theta^i\partial\theta^j\partial\theta^k}.$$

Мішане сімейство:

$$^{\alpha}\Gamma_{ij,k}(\theta) = -\frac{1+\alpha}{2} E\left[\partial_i l_\theta(x,\theta)\partial_j l_\theta(x,\theta)\partial_k l_\theta(x,\theta)\right].$$

Ми бачимо, що у випадку експоненціального сімейства коефіцієнти зв'язності стають нульовими, якщо $\alpha = 1$, а для мішаного сімейства це ж саме відбувається коли $\alpha = -1$. У зв'язку з цим, систему координат, для якої $^{1}\Gamma_{ij,k} = 0$ ми називаємо *е-координатною системою,* а в якій таке «обнулення» відбувається при $\alpha = -1$, тобто $^{-1}\Gamma_{ij,k} = 0$ – називаємо відповідно *m-координатною системою.*

4. Статистичний висновок і геометрія

4.1. Достатня статистика і нерівність Рао – Крамера

Перш ніж застосовувати апарат інформаційної геометрії розглянемо деякі фундаментальні поняття статистики.

Нехай задана регулярна статистична модель:

$$M = \{p(x,\xi)\}, \tag{4.1}$$

і також відображення $F : \mathcal{X} \to \mathcal{Y}$, що перетворюя значення випадкової величини $X \epsilon \mathcal{X}$ у $Y = F(X)$. Тоді даний розподіл $p(x,\xi)$ перетворюється у $q(y,\xi)$, що задає поведінку іншої випадкової величини, а саме *Y*. Ми кажемо, що $F(x)$ є **достатньою статистикою** для моделі *M* якщо $p(x,\xi)$ можна представити наступним чином:

$$p(x,\xi) = q(F(x),\xi) r(x). \tag{4.2}$$

Якщо ми запишемо (4.2) у вигляді $r(x) = \dfrac{p(x,\xi)}{q(F(x),\xi)}$,

то стає зрозумілим, що $F(x)$ приймає такі значення, що є узгодженими з певним набором значень параметрів ξ. Це означає, що та частина розподілу $p(x,\xi)$, що залежить від ξ цілком зосереджена в розподілі $q(y,\xi)$ величини $Y = F(X)$ і тому для того, щоб оцінити невідомі параметри ξ (невідомий розподіл $p(x,\xi)$) достатньо знати лише значення *Y* [3]. Отже, $F(x)$ можна застосовувати для оцінювання значень параметрів ξ. Рівність (4.2) ще

має назву теореми про факторізацію. З неї ми маємо дуже просте і водночас потужне знаряддя для статистичного дослідження. Працює воно наступним чином.

Нехай задана модель (4.1) і запропоновано функцію $y = F(x)$, що ми розраховуємо представити розподіл $p(x,\xi)$ у вигляді добутку:

$$p(x,\xi) = q(F(x),\xi)r(x).$$

Тоді нам потрібно розглянути індуковану модель:

$$M_F \overset{\text{def}}{=} \{q(F(x),\xi)\}. \tag{4.3}$$

І порівняти її з даною, тобто (4.1).

Для цього ми маємо знайти за формулою (2.5) метрики Фішера $G(\xi) = (g_{ij}(\xi))$ для моделі (4.1) і для індукованої моделі (4.3), яку позначимо $G^F(\xi) = (g^F{}_{ij}(\xi))$. З цього приводу існує теорема [3].

Теорема 4.1 *Інформаційна матриця Фішера* $G^F(\xi) = (g^F{}_{ij}(\xi))$ *обчислена для індукованої моделі* $M_F = \{q(F(x),\xi)\}$ *задовільняє нерівність*

$$G^F(\xi) \le G(\xi),$$

в тому сенсі, що відповідна квадратична форма для матриці $\Delta G(\xi) = G(\xi) - G^F(\xi)$ є позитивно напіввизначена. Необхідна та достатня умова для того, щоб статистика $F(x)$ була достатньою є рівність $G(\xi) = G^F(\xi)$.

Кажуть, що різниця:

$$\Delta G(\xi) = G(\xi) - G^F(\xi) \tag{4.4}$$

визначає *втрату інформацій (information loss).*

Узагальнюючи (4.4) ми отримуємо ще три такі властивості інформаційної матриці Фішера, що є дуже широко застосовними у статистичних дослідженнях.

1. Нехай поряд з моделлю (4.1) матриця Фішера якої є $G(\xi)$ задано умовний розподіл $\kappa = \{\kappa(y|x) \ge 0; x \in \mathcal{X}, y \in \mathcal{Y}\}$, такий, що для будь-якого x:

$$\int \kappa(y|x)\,dy = 1,$$

і $G_{\kappa}(\xi)$ є матрицею Фішера для індукованої моделі:

$$q(y,\xi)=\int \kappa(y|x)\,p(x,\xi)\,dx.$$

Тоді,

$$G(\xi)=G_{\kappa}(\xi)+\Delta G(\xi). \tag{4.5}$$

З (4.1) випливає, що $G(\xi)\geq G_{\kappa}(\xi)$.

2. Нехай задано дві незалежних моделі виду (4.1):

$$M_1=\{p_1(x_1,\xi)\}\, та\, M_2=\{p_2(x_2,\xi)\},$$

що мають той самий простір елементарних подій (sample set) $\mathcal{X}$ і той самий простір параметрів Ξ. Тоді, матриця Фішера $G_{12}(\xi)$ для моделі-добутку, із щільністю розподілу $p_{12}(x_1,x_2,\xi)=p_1(x_1,\xi)\,p_2(x_2,\xi)$ матиме вигляд:

$$G_{12}(\xi)=G_1(\xi)+G_2(\xi), \tag{4.6}$$

де $G_1(\xi)$ та $G_2(\xi)$ є відповідно матрицями Фішера моделей M_1 та M_2.

Властивість (4.6) має назву *адитивності*.

3. Нехай задано дві незалежних моделі виду (4.1):

$$M_1=\{p_1(x_1,\xi)\}\, та\, M_2=\{p_2(x_2,\xi)\},$$

що мають той самий простір елементарних подій $\mathcal{X}$ і той самий простір параметрів Ξ і відповідно мають матриці Фішера $G_1(\xi)$ та $G_2(\xi)$.

Тоді для моделі із щільністю розподілу:

$$p_{\lambda}(x,\xi)=\lambda p_1(x,\xi)+(1-\lambda)\,p_1(x,\xi), 1\leq\lambda\leq 1$$

матриця Фішера задовільнятиме нерівності:

$$G_{\lambda}(\xi)\leq\lambda G_1(\xi)+(1-\lambda)G_2(\xi). \tag{4.7}$$

Рівняння (4.7) відображує властивість **опуклості** матриці Фішера.

Тепер розглянемо відому нерівність *Рао – Крамера*. Нехай дані *x* випадково генеруються згідно з розподілом, що є невідомим, але вважатимемо, що цей розподіл належить *n*-вимірній моделі (4.1):

$$M=\{p(x,\xi)\,|\,\xi=(\xi^1,\xi^2,\ldots,\xi^m)\in\Xi\}.$$

Нам потрібно оцінити невідомий параметр $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$ за допомогою функції $\hat{\xi}(x)$, аргументом якої є статистичні дані.

Відображення

$$\hat{\xi} = \left(\hat{\xi}^1, \hat{\xi}^2, \ldots, \hat{\xi}^m\right) : \mathcal{X} \to \mathbb{R}^m \tag{4.8}$$

має назву *оцінки (estimator)*. Ми кажемо, що (4.8) є *незміщеною оцінкою*, якщо математичне сподівання відносно простору елементарних подій відповідає рівності:

$$E_\xi\left[\hat{\xi}(x)\right] = \xi. \tag{4.9}$$

Середня квадратична похибка незміщеної оцінки $\hat{\xi}$ може бути виражено як коваріаційна матриця $V_\xi\left[\hat{\xi}\right] = \left(v_\xi^{ij}\right)$ обчислена як [3]:

$$v_\xi^{ij} = E_\xi\left[\left(\hat{\xi}^i(x) - \xi^i\right)\left(\hat{\xi}^j(x) - \xi^j\right)\right]. \tag{4.10}$$

Теорема 4.2. (*Рао – Крамера*) *Коваріаційна матриця* $V_\xi\left[\hat{\xi}\right]$ *для незміщеної оцінки* $\hat{\xi}$ *задовільняє нерівність:*

$$V_\xi\left[\hat{\xi}\right] \geq G^{-1}(\xi). \tag{4.11}$$

В тому сенсі що $V_\xi\left[\hat{\xi}\right] - G^{-1}(\xi)$ *визначає позитивно напіввизначену квадратичну форму.*

$G^{-1}(\xi)$ є матрицею, оберненою до матриці Фішера $G(\xi)$.

Якщо для незміщеної оцінка $\hat{\xi}$ виконується рівність:

$$V_\xi\left[\hat{\xi}\right] = G^{-1}(\xi), \tag{4.12}$$

то кажуть, що $\hat{\xi}$ є *ефективною оцінкою*. Є очевидним, що ефективна оцінка є найкращою незміщеною оцінкою в тому сенсі, що її дисперсія є мінімальною серед дисперсій усіх незміщених оцінок. Обернене твердження, загалом, не є вірним. Справа у тому, що найкраща незміщена оцінка не завжди досягає межи Рао – Крамера, тобто $V_\xi\left[\hat{\xi}\right] = G^{-1}(\xi)$. Більше того, відомий випадок, коли зміщена оцінка виявилася такою, що має меншу середню квадратичну похибку, ніж ефективна оцінка.

Зазначимо, що може і не існувати ефективна оцінка моделі (4.1). Але слід наголосити, що завжди існує послідовність оцінок $\left\{\hat{\xi}_N = \hat{\xi}_N\left(x_1, x_2, \ldots x_N\right)\right\}_{N=1}^{\infty}$, що асимптотично досягається рівність (4.12) при $N \to \infty$. Така послідовність оцінок має назву *асимптотично ефективної оцінки*, або *ефективної оцінки першого порядку*. Таким чином, матриця $G^{-1}(\xi)$ відображує, наскільки близько асимтотично оптимальна оцінка попадає у окіл істинного значення параметра $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$. Інакше кажучи, чим менші значення компонент $G^{-1}(\xi)$, тим кращою є оцінка. Зрозуміло, що відповідно значення компонент $G(\xi)$ мають бути якнайбільшими.

4.2. Оцінювання базоване на незалежних дослідженнях

Розглянемо сімейство розподілів ймовірностей $M = \{p(x, \xi)\}$ що залежить від *m* параметрів $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$. Згідно прийнятним умовам регулярності[3], ми розглядаємо *M* як *m*-вимірний многовид, в якому $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$ виступає в якості локальної системи координат. Нехай у нас є *N* незалежних спостережень $x_1, x_2, \ldots x_N$ випадкової величини ***x***, яка в загальному випадку є *n*-вимірною, тобто кожне спостереження виступає у вигляді деякого *n*-вимірного вектора. Ця випадкова величина розподілена згідно щільністю розподілу $p(x, \xi)$. Позначимо результати наших спостережень $x^N = \left(x_1, x_2, \ldots x_N\right)$.

Мета статистичного висновку – вивести розподіл ймовірності, базуючись на даних, що є вибіркою *N* значень. *Оцінювання* – це такий тип задач, де метою є знайти оцінку $\hat{\xi} = \left(\hat{\xi}^1, \hat{\xi}^2, \ldots, \hat{\xi}^m\right)$ для параметру (параметрів) $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$. З іншого боку, існує тип задач перевірки істиності гіпотез, де метою є визначення що саме треба зробити – чи прийняти «нульову» гіпотезу $H_0 : \xi = \xi_0$ всупереч альтернативної гіпотези $H_1 : \xi \neq \xi_0$, чи навпаки, відкинути.

[3] Ми маємо на увазі, що усі функції, використані в розрахунках локально мають неперервні похідні достатньо високого порядка.

Очевидно, що розподіл ймовірностей випадкової величини $x^N = (x_1, x_2, \ldots x_N)$ може бути записаний з використанням функції розподілу результату одиничного випробування $p(x, \xi)$ як

$$p_N(x^N, \xi) = \prod_{t=1}^{N} p(x_t, \xi).$$

Нагадаємо, що у загальному випадку $\boldsymbol{x}_t$ є n-вимірним вектором $x_t = (x_{1t}, x_{2t}, \ldots x_{mt})$. Отже ми маємо:

$$\ln\left(p_N(x^N, \xi)\right) = \sum_{t=1}^{N} \ln\left(p(x_t, \xi)\right). \tag{4.13}$$

Розглядаючи $\boldsymbol{x}^N$ як випадкову змінну, ми отримуємо модель $M_N = \left\{p(x^N, \xi)\right\}$, яка подібно $M = M_1$ є многовидом з $\xi = (\xi^1, \xi^2, \ldots, \xi^m)$ як координатної системи. Тому, з рівняння (4.13) і відповідних означень, ми знайдемо об'єкти геометричної структури на многовиді M_N:

$$^{N}g_{ij}(\xi) = N g_{ij}(\xi), \tag{4.14}$$

$$^{N(\alpha)}\overline{\Gamma}_{ij,k}(\xi) = {}^{(\alpha)}\overline{\Gamma}_{ij,k}(\xi). \tag{4.15}$$

Тут $g_{ij}(\xi)$ та ${}^{(\alpha)}\overline{\Gamma}_{ij,k}(\xi)$ – відповідно метрика Фішера та об'єкт α-зв'язність моделі (4.1) розраховані за формулами (2.5) та (2.11). Як ми бачимо, геометрична структура, яку ми ввели на M_N фактично є такою, як і на M, тільки відрізняється масштабуючим множником N.

Отже, для дослідження геометрії моделі M_N цілком достатньо використовувати геометрію моделі (4.1). Згідно із загальною теорією оцінювання, оцінка є визначеною як функція на N точках, які власне і складають вибірку $x^N = (x_1, x_2, \ldots x_N)$:

$$\hat{\xi}_N = \hat{\xi}(x^N) = \hat{\xi}(x_1, x_2, \ldots x_N). \tag{4.16}$$

Якщо ми розглядаємо $\boldsymbol{x}^N$ як випадкову змінну, то маємо розглядати $\hat{\xi}_N$ теж як випадкову змінну. Крім того, нам потрібна умова, якій має задовільняти $\hat{\xi}$, щоб гарантувати близькість оцінки $\hat{\xi}$ до реального значення параметра ξ, що входить до аналітичного виразу розподілу $p(x, \xi)$. Такою умовою є незміщеність. Грубо кажучи, вона полягає в тому, випадкова величина $\hat{\xi}$ є розподіленою навколо ξ. Цей факт

був вже записаний як (4.9). для випадку $N=1$. Ми можемо розширити цю формулу для $N>1$.

$$E_{\xi}\left[\hat{\xi}\right]=\xi. \tag{4.17}$$

Ми бачимо, що формули (4.9) та (4.17) фактично співпадають. Але якщо у випадку $N=1$ математичне сподівання функції $\hat{\xi}\left(x_1,x_2,\ldots x_n\right)$ n-вимірної випадкової величини (4.9) обчислюється як n-кратний інтеграл, то коли для $N>1$ формула (4.17) виглядатиме як інтеграл кратості $n\cdot N$ в якому замість $p\left(x,\xi\right)$ фігуруватиме роподіл $p_N\left(x^N,\xi\right)$.

Середня квадратична похибка використовується як засіб вимірювання якості оцінки $\hat{\xi}$ і може бути виражено як коваріаційна матриця $V_{\xi}\left[\hat{\xi}\right]=\left(v_{\xi}^{ij}\right)$ обчислена як [3]:

$$v_{\xi}^{ij}=E_{\xi}\left[\left(\hat{\xi}^{i}\left(x\right)-\xi^{i}\right)\left(\hat{\xi}^{j}\left(x\right)-\xi^{j}\right)\right]. \tag{4.18}$$

Як ми бачимо, (4.18) повністю співпадає з (4.10), але як у випадку (4.17) – інтегрування проводиться не згідно розподілу $p\left(x,\xi\right)$, а $p_N\left(x^N,\xi\right)$.

Варто зазначити, що як незміщеність, так і середня квадратична похибка не є геометрично інваріантними. Тобто вони суттєво залежать від вибору координатнної системи. Нижня межа середньої квадратичної похибки незміщеної оцінки $\hat{\xi}$ дається нерівністю Рао – Крамера (Теорема 4.2), яке зараз приймає форму:

$$\left(v_{\xi}^{ij}\right)\geq\frac{1}{N}\left(g^{ij}\left(\xi\right)\right). \tag{4.19}$$

Тут $g^{ij}\left(\xi\right)$ означає матрицю, обернену до інформаційної матриці Фішера $\left(g_{ij}\left(\xi\right)\right)$ моделі M. Множник $\frac{1}{N}$ на відміну від (4.11) виникає внаслідок того, що $\left(g_{ij}\left(\xi\right)\right)$ розраховано для моделі з розподілом $p\left(x,\xi\right)$, а $\left(v_{\xi}^{ij}\right)$ вже обчислено для $p_N\left(x^N,\xi\right)$.

4.3. Достатня статистика у статистичних дослідженнях

У секції 4.1 ми ввели поняття достатньої статистики слідуючи [3]. Однак для більш глибокого розуміння варто розібрати застосування поняття достатньої статистики на прикладах, пов'язаних

із статистичним дослідженням. Перш за все наголосимо, що статистичне дослідження – це в першу чергу отримання репрезентативної вибірки $X_1,X_2,\ldots X_N$. В контексті цієї секції *статистикою* називатимемо будь-яку функцію вигляду:

$$Y = F\left(X_1, X_2, \ldots, X_N\right).$$

Зрозуміло, що дослідника цікавитимуть саме такі функції від значень, що прийняли елементи випадкової вибірки, що можна використовувати як оцінки значень невідомих йому параметрів. Питання полягає в тому, як знайти таку функцію, щоб здобувала з вибірки саме ту інформацію, що характеризує генеральну сукупність. В зв'язку з цим ми можемо дати таке евристичне означення *достатньої статистики* [19].

Ми кажемо, що $Y = F\left(X_1, X_2, \ldots, X_N\right)$ є *достатньою статистикою,* якщо для оцінювання параметру ξ знання лише одного числового значення Y є цілком рівноцінним знанню всієї вибірки. Тобто, функція, яку ми називаємо *достатньою статистикою* екстрагує з вибірки з N значень – лише одне, але таке, яке характеризує вибірку в цілому.

Для знаходження такої функції застосовують вже згадану у секції 4.1 теорему про факторизацію. Але перш ніж її сформулювати у вигляді, застосовному у статистиці, зробимо деякі пояснення. Як ми вже писали, що розподіл ймовірностей випадкової величини $x^N = \left(x_1, x_2, \ldots x_N\right)$ може бути записаний з використанням функції розподілу результату одиничного випробування $p\left(x, \xi\right)$ як

$$p_N\left(x^N, \xi\right) = \prod_{t=1}^{N} p\left(x_t, \xi\right).$$

Тобто, в таких дослідженнях тип розподілу вважається вже відомим. Дослідник має знайти оцінки для невідомих параметрів.

Вибірка $X_1, X_2, \ldots X_N$ – фактично є результатом розігрування випадкових величин $x_1, x_2, \ldots x_N$, а функція

$$p\left(x_1, x_2, \ldots x_N | \xi\right) = p_N\left(x^N, \xi\right) = \prod_{t=1}^{N} p\left(x_t, \xi\right)$$

є їх спільним розподілом.

Таким чином, теорема про факторизацію виглядає так.

Теорема 4.3 *Нехай* $X_1, X_2, \ldots X_N$ *утворює випадкову вибірку із спільним розподілом* $p(x_1, x_2, \ldots x_N|\xi)$. *Статистика* $Y = F(X_1, X_2, \ldots, X_N)$ *є достатньою для параметра* ξ, *якщо і тільки якщо цей спільний розподіл може бути розкладеним на множники наступним чином:*

$$p(x_1, x_2, \ldots x_N|\xi) = q(F(x_1, x_2, \ldots x_N), \xi) r(x_1, x_2, \ldots x_N), \quad (4.20)$$

де $q(F(x_1, x_2, \ldots x_N), \xi)$ *та* $r(x_1, x_2, \ldots x_N)$, *є функціями, що приймають лише невід'ємні значення. Функція* $r(x_1, x_2, \ldots x_N)$ *залежить від усіх значень повного об'єму вибірки, але не залежить від оцінюванного невідомого параметру* ξ. *Функція* $q(F(x_1, x_2, \ldots x_N), \xi)$ *може залежати від параметру* ξ, *але може залежати від вибірки лише через значення функції* $F(x_1, x_2, \ldots x_N)$.

Зауважимо, що $r(x_1, x_2, \ldots x_N)$ може залежати від інших параметрів, окрім ξ. Головне, щоб $r(x_1, x_2, \ldots x_N)$ не залежав саме від того параметру, для якого ми шукаємо достатню статистику. Додамо, що (4.20) – це є нічим іншим, як детальним представленням (4.2) у вигляді, що є застосовним у статистичних дослідженнях.

Розберемо деякі приклади

1. *Нормальний розподіл, невідоме математичне сподівання, дисперсія відома*. Отже, ми розглядаємо нормально розподілену генеральну сукупність, для якої треба оцінити математичне сподівання μ, дисперсію σ^2 вважаємо відомою. Спільний розподіл має вигляд:

$$p(x_1, x_2, \ldots x_N \mid \xi) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right) =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{n\mu^2}{2\sigma^2}\right) =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2\right)\exp\left(-\frac{n\mu^2}{2\sigma^2} + \frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i\right) =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2\right)\exp\left(-\frac{n\mu^2}{2\sigma^2} + \frac{\mu}{\sigma^2}F(x_1, x_2, \ldots x_N)\right).$$

Оскільки дисперсія σ^2 є відомою, можна вважати

$$r\left(x_1, x_2, \ldots x_N\right) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right)$$

та

$$q\left(F\left(x_1, x_2, \ldots x_N\right), \xi\right) = \exp\left(-\frac{n\mu^2}{2\sigma^2} + \frac{\mu}{\sigma^2} F\left(x_1, x_2, \ldots x_N\right)\right),$$

де

$$F\left(x_1, x_2, \ldots x_N\right) = \sum_{i=1}^{n} x_i.$$

Згідно із теоремою про факторизацію (4.20) з цього випливає, що $\sum_{i=1}^{n} X_i$ є достатньою статистикою. З цього випливає, що середня вибіркова

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{n} X_i$$

також є достатньою статистикою.

2. *Розподіл Пуасона, невідоме математичне сподівання [20].*

Для цієї дискретно розподіленої випадкової величини потрібно знайти достатню статистику, виходячи з вибірки $X_1, X_2, \ldots X_N$.

Факторизація згідно із (4.20) дає:

$$p\left(x_1, x_2, \ldots x_N | \xi\right) = \prod_{t=1}^{N} p\left(x_t, \xi\right) = \prod_{t=1}^{N} e^{-\xi}\frac{\xi^{x_t}}{x_t!} = \frac{e^{-n\xi}\xi^{\sum_{t=1}^{N} x_t}}{\prod_{t=1}^{N} x_i!}.$$

де

$$q\left(F\left(x_1, x_2, \ldots x_N\right), \xi\right) = e^{-n\xi}\xi^{\sum_{t=1}^{N} x_t},$$

та

$$r\left(x_1, x_2, \ldots x_N\right) = \frac{1}{\prod_{t=1}^{N} x_i!}.$$

Тут також статистика

$$F\left(x_1, x_2, \ldots x_N\right) = \sum_{i=1}^{n} x_i$$

є достатньою.

3. *Однорідний розподіл [19].* X_i розподілена однорідно на проміжку $[0,b]$, параметр b невідомий. Тоді спільна щільність розподілу

$$p(x_1,x_2,\ldots x_N|\xi)=b^{-n}1(x_i\le b,i=1.2,\ldots,N),$$

де $1(E)$ – так звана індикаторна функція. Вона дорівнює 1, якщо подія E відбулась, і дорівнює 0 в протилежному випадку. Ми бачимо, що $x_i\le b,i=1.2,\ldots,N$ тоді і тільки тоді, якщо $max\{x_1,x_2,\ldots x_N\}\le b$.

Таким чином отримуємо:

$$p(x_1,x_2,\ldots x_N|\xi)=b^{-n}1(max\{x_1,x_2,\ldots x_N\}\le b).$$

Згідно з теоремою про факторизацію,

$$F(x_1,x_2,\ldots x_N)=max\{x_1,x_2,\ldots x_N\}$$

і є ефективною статистикою.

4.4. Асимптотична теорія оцінювання

В асимптотичній теорії оцінювання засобом є представлення оцінки шляхом граничного переходу $N\to\infty$. В цьому випадку незміщеність не є настільки значущою, як тоді, коли N є фіксованим. Натомість ми зазвичай вимагаємо, щоб оцінка $\hat{\xi}_N=\hat{\xi}(x^N)$, або якщо бути точнішим, послідовність оцінок $\{\hat{\xi}_N;N=1,2,\ldots\}$ була *слушною (конзистентною)* в тому сенсі, що для будь-якого параметру ξ його оцінка $\hat{\xi}(x^N)$ збігається за ймовірністю до ξ по мірі того, як $N\to\infty$, коли x^N є розподіленою за законом $p_N(x^N,\xi)$. Інакше кажучи, для будь-якого $\varepsilon>0$

$$\lim_{N\to\infty}P\left(\left|\hat{\xi}_N-\xi\right|>\varepsilon\right)=0,$$

тобто ймовірність того, що значення оцінки буде суттєво відрізнятися від справжнього значення параметру, прямує до нуля. Ймовірносний розподіл слушної оцінки $\hat{\xi}_N$ є усе більше і більше зконцентрованим навколо цього реального значення ξ разом з тим, що $N\to\infty$. Згідно з вимогами регулярності математичне сподівання випадкової величини, якою є $\hat{\xi}_N$, рівномірно збігається до ξ, отже:

$$\lim_{N\to\infty} E_{\xi}\left[\hat{\xi}_N\right] = \xi$$

та

$$\lim_{N\to\infty} \frac{\partial E_{\xi}\left[\hat{\xi}^i_N\right]}{\partial \xi^j} = \partial_j \xi^i = \delta^i_j.$$

Середня квадратична похибка такої асмптотично незміщеної оцінки задовільняє *асимптотичну нерівність Рао – Крамера:*

$$\lim_{N\to\infty} N\left(v^{ij}_{\xi}\left(\hat{\xi}^i_N\right)\right) \geq \left(g^{ij}\left(\xi\right)\right). \tag{4.21}$$

Слушна оцінка, що утворює знак рівності у (4.21) має *назву асимптотично ефективної оцінки*, або *ефективної оцінки першого порядку.* Така оцінка є оптимальною по відношенню до середньої квадратичною похибки з точністю порядку $\frac{1}{N}$. Виходячи з цього виникає питання, чи завжди існуватиме для довільної моделі асимптотично ефективна оцінка. На відміну від випадку ефективної оцінки для скінченного значення *N*, відповідь на це питання є позитивною і в якості прикладу можна навести метод максимальної правдоподібності (Maximum likelihood estimation, MLE), з яким ми познайомилися у розділі 2.1. Зараз з метою обгрунтування опишемо його більш формально, враховуючи викладений теоретичний апарат.

Маючи емпіричні дані x^N, розглянемо функцію $p_N\left(x^N, \xi\right)$ як функцію від $\xi = \left(\xi^1, \xi^2, \ldots, \xi^m\right)$, і назвемо її *функцією правдоподібності.* Назвемо $\hat{\xi}_N$ *оцінкою максимальної правдоподібності (maximum likelihood estimator, MLE),* позначивши як $\hat{\xi}_{MLE}$, якщо вона задовільняє умові:

$$\max_{\xi} p_N\left(x^N, \xi\right) = p_N\left(x^N, \hat{\xi}_N\right), \tag{4.22}$$

тобто приймає значення для ξ, яке максимізує правдоподібність. Наступна теорема лежить в основі методу максимальної правдоподібності [3].

Теорема 4.3 О*цінка максимальної правдоподібності* $\hat{\xi}_{MLE}$ *є асимптотично ефективною. Це виражається наступною рівністю:*

$$\lim_{N\to\infty} N v^{ij}_{\xi}\left(\hat{\xi}_{MLE}\right) = g^{ij}\left(\xi\right). \tag{4.23}$$

Фактично, є можливим довести більш сильне твердження, а саме, що оцінка $\hat{\xi}$, як випадкова величина має багатовимірний нормальний розподіл із математичним сподіванням ξ і коваріаційною матрицею $\frac{1}{N}\left(g^{ij}(\xi)\right)$.

Середня квадратична похибка асмптотично ефективної оцінки $\hat{\xi}$ може бути записаною так:

$$v_{\xi}^{ij}\left(\hat{\xi}\right)=\frac{1}{N}g^{ij}(\xi)+\mathrm{O}\left(\frac{1}{N^2}\right).$$

Аналіз доданку другого порядку є предметом *асимптотичної теорії вищого порядку*, і вона має розглядатись окремо.

5. Дуальні зв'язності

5.1. α-афінні многовиди і α-сімейства

В цій секції ми розширимо і узагльнемо деякі результати отримані стосовно α-зв'язності, отримані в розділах 2 та 3.

Ми трохи послабимо наші фундаментальні припущення, яких ми досі трималися. Досі ми розглядали статистичну модель $M=\left\{p(x,\xi)\right\}$, кожний елемент якої є розподілом ймовірностей $p(x,\xi)\in\mathcal{P}(\mathcal{X})$. В цій секції ми розглянемо випадки, коли

$$\int p(x,\xi)dx$$

може приймати довільні позитивні значення. Рівняння (2.5) та (2.11), що визачають відповідно метрику Фішера та α-зв'язність лишаються в силі, в той самий час (2.4) та (2.6) взагалі кажучи, не є такими, що обов'язково мають виконуватись.

Для кожного значення $\alpha\in\mathbb{R}$ задано:

$$L^{(\alpha)}(u)\overset{\text{def}}{=}\begin{cases}\dfrac{2}{1-\alpha}u^{\frac{1-\alpha}{2}} & (\alpha\neq 1)\\ \ln u & (\alpha=1)\end{cases}, \tag{5.1}$$

$$l^{(\alpha)}(x,\xi)\overset{\text{def}}{=}L^{(\alpha)}\left(p(x,\xi)\right). \tag{5.2}$$

Зазначимо, що зокрема $l^{(\alpha)}(x,\xi)=l(x,\xi)$ що $l^{(-1)}(x,\xi)=p(x,\xi)$. Для дотичного вектора $X\in T_{\xi}(M)$ функцію від змінної *x*:

$$X^{(\alpha)}(x)\overset{\text{def}}{=}Xl^{(\alpha)}(x,\xi)=X^{i}\frac{\partial l^{(\alpha)}(x,\xi)}{\partial\xi^{i}}. \tag{5.3}$$

ми вважатимемо його α-представленням. Відомі нам *e*-, *m*- та 0-представлення, що ми згадували в секції 3.4 є нічим іншим, як випадками, що відповідають $\alpha=1;-1;0$. Оскільки:

$$\partial_i l^{(\alpha)}=p^{\frac{1-\alpha}{2}}\partial_i l \text{ та } \partial_i\partial_j l^{(\alpha)}=p^{\frac{1-\alpha}{2}}\left(\partial_j\partial_i l+\frac{1-\alpha}{2}\partial_i l\partial_j l\right),$$

Ми можемо переписати (2.5) та (2.11) у вигляді:

$$g_{ij}(\xi)=\int\partial_i l^{(\alpha)}(x,\xi)\partial_j l^{(-\alpha)}(x,\xi)dx, \tag{5.4}$$

$${}^{\alpha}\Gamma_{ij,k}(\xi)=\int\partial_i\partial_j l^{(\alpha)}(x,\xi)\partial_i l^{(-\alpha)}(x,\xi)dx. \tag{5.5}$$

Ці рівняння можуть бути прочитаними в тому сенсі, що α-зв'язність на *M* індукована афінною структурою простору $\mathbb{R}^{\mathcal{X}}$ функцій на $\mathcal{X}$ через вкладання $\xi\to l^{(\alpha)}(x,\xi)$. Зауважимо, що у випадку коли $M=\{p(x,\xi)\}$ складається з розподілів ймовірності, відповідно для (2.4) та (2.6), ми маємо:

$$\int p^{\frac{1+\alpha}{2}}(x,\xi)\partial_i l^{(\alpha)}(x,\xi)dx=0, \tag{5.6}$$

Зафіксуємо певне значення α. Якщо в деякій системі координат $\left[\theta^{i}\right]$:

$$\frac{\partial^2 l^{(\alpha)}(x,\theta)}{\partial\theta^{i}\partial\theta^{j}}=0, \tag{5.7}$$

то з рівняння (5.5) ми бачимо, що $\left[\theta^{i}\right]$ є α-афінна система координат, отже $M=\{p(x,\xi)\}$ є α-плоским. Ми називатимемо такий *M* α-*афінним многовидом*. Умова (5.7) є еквівалентною існуванню таких функцій $\{F_1,\ldots,F_n,C\}$, на $\mathcal{X}$, що:

$$l^{(\alpha)}(x,\xi)=C(x)+\sum_{i=1}^{n}\theta^{i}F_i(x). \tag{5.8}$$

5.2. Дуальність зв'язностей

Нехай M є мновидом із рімановою метрикою g_{ij} та афінними зв'язностями ∇ та ∇^*. В довільній точці p задано два дотичних до многовиду вектори $X(p)$, $Y(p) \in T_p(M)$. Крім того, точка p і інша точка $q(p \neq q)$ з'єднані кривою $\gamma(t)$. Вектори X та Y паралельно переносяться уздовж кривої $\gamma(t)$. Це означає, що їх координати уздовж $\gamma(t)$ мають змінюватися згідно рівнянню (1.12). Але від класичного випадку ситуація відрізняється тим, що вектор X переноситься паралельно у зв'язності ∇, а вектор Y – відповідно у зв'язності ∇^*.

Тобто,

$$\frac{\partial X^i}{\partial \xi^k}\dot{\gamma}^k + \Gamma^i_{jk} X^j(t)\dot{\gamma}^k(t) = 0, \tag{5.9}$$

$$\frac{\partial Y^i}{\partial \xi^k}\dot{\gamma}^k + \Gamma^{*i}_{jk} Y^j(t)\dot{\gamma}^k(t) = 0. \tag{5.10}$$

Тут $\dot{\gamma}^k(t)$ – дотичний до кривої $\gamma(t)$ вектор, а Γ^i_{jk} та Γ^{*i}_{jk} – коефіцієнти зв'язностей ∇ та ∇^* відповідно.

Вимагаємо, щоб скалярний добуток паралельно перенесених векторів зберігався, тобто

$$X^i(p) g_{ij} Y^i(p) = X^i(q) g_{ij} Y^i(q).$$

Це означає, що

$$\frac{d}{dt}\left(X^i(t) g_{ij} Y^i(t)\right) = 0, \tag{5.11}$$

або

$$\nabla_{\dot{\gamma}(t)} X^i(t) g_{ij} Y^i(t) + X^i(t) \tilde{\nabla}_{\dot{\gamma}(t)} g_{ij} Y^i(t) + X^i(t) g_{ij} \nabla^*_{\dot{\gamma}(t)} Y^i(t) = 0.$$

Для коваріантних похідних $\nabla_{\dot{\gamma}(t)} X^i$ та $\nabla^*_{\dot{\gamma}(t)} Y^i(t)$ зрозуміло, що вони мають вираховуватись за формулами (5.9) та (5.10), отже перший та останній доданок обертаються в нуль. Звідси випливає, що

$$X^i(t) \tilde{\nabla}_{\dot{\gamma}(t)} g_{ij} Y^i(t) = 0.$$

Поки що не зрозуміло, як має виглядати коваріантна похідна метричного тензору. Але оскільки крива $\gamma(t)$ є довільною, то очевидно що

$$\tilde{\nabla}_k g_{ij} = 0. \tag{5.12}$$

З іншого боку, мають задовільнятись (5.09) та (5.10).

Враховуючи (5.12), маємо

$$\tilde{\nabla}_k g_{ij} = \frac{\partial g_{ij}}{\partial x^k} - \Gamma^t_{ik} g_{tj} - \Gamma^{*t}_{jk} g_{it} = 0.$$

Звідси випливає, що для будь-якого векторних полів (цього разу довільних)

$$\frac{\partial}{\partial \xi^k}\left(X^i g_{ij} Y^i\right) = \nabla_k X^i g_{ij} Y^i + X^i g_{ij} \nabla^*_k Y^i. \tag{5.13}$$

Пара зв'язностей, що відповідають рівнянню (5.13) мають назву *дуальних,* або *спряжених одна до одної по відношенню до метрики Фішера.*

Очевидно, що $\left(\nabla^*\right)^* = \nabla$. Також можна легко переконатись, що $\frac{\nabla + \nabla^*}{2}$ є метричною, тобто зв'язністю Леві – Чивіта. З іншого боку, навпаки, якщо нам відомо що ∇ та ∇' є симетричними зв'язностями і $\frac{\nabla + \nabla'}{2}$ є метричною, то можна стверджувати, що ∇' і ∇ є взаємно дуальними, тобто $\nabla = \nabla^*$.

Проаналізувавши (5.4) та (5.5) з очевидністю отримуємо теорему.

Теорема 5.1. *Для будь-якої статистичної моделі* α*-зв'язність і* (−α)*-зв'язність є взаємно дуальними по відношенню до метрикі Фішера.*

Для двох взаємно дуальних зв'язностей ∇ та ∇^* за формулами (1.11) можна розрахувати відповідні тензори кривини: R^h_{ijk} та R^{*h}_{ijk}. З очевидністю бачимо, що

$$g_{lh} R^h_{ijk} = -g_{ih} R^{*h}_{ljk}, \tag{5.14}$$

де g_{ij}-метрика Фішера. Враховуючи невиродженість цієї матриці, отримуємо наступну теорему.

Теорема 5.2 *Якщо тензор кривини однієї з двох взаємно дуальних зв'язностей* ∇ *та* ∇^* *дорівнює нулю, то нулю дорівнюватиме і тензор кривини іншої.*

5.3. Дивергенції

Нехай M є многовидом і на ньому задана гладка функція $D(\cdot|\cdot): M \times M \to \mathbb{R}$ Позначати будемо $D\big(p(x,\xi)|p(x,\xi')\big)$, або $D(\xi|\xi')$ чи $D(p|p')$ в залежності від контекста задачі. Для будь-якої пари розподілів $p = p(x,\xi) \in M$ та $p' = p(x,\xi') \in M$ має виконуватись:

$$\begin{aligned} &1.\ D(p|p') \geq 0, \\ &2.\ D(p|p') = 0, \text{ тоді і тільки тоді } p = p'. \end{aligned} \tag{5.15}$$

Ця функція є у деякому сенсі мірою віддаленням однією точки від іншої і в цьому схожа на функцію відстані, але загалом, не задовольняє аксіомам симетричності та нерівності трикутника.

Введемо деякі позначення

$$D\Big(\big(\partial_i\big)_p|p'\Big) \overset{\text{def}}{=} \partial_i D(p|p'),$$

$$D\Big(\big(\partial_i\big)_p|\big(\partial_j\big)_{p'}\Big) \overset{\text{def}}{=} \partial_i \partial'_j D(p|p'),$$

$$D\Big(\big(\partial_i\partial_j\big)_p|\big(\partial_k\big)_{p'}\Big) \overset{\text{def}}{=} \partial_i \partial_j \partial'_k D(p|p'), \text{ тощо,}$$

де вказані в правих частинах оператори $\partial_i = \dfrac{\partial}{\partial \xi^i}$ застосовуються до відповідної змінної у $p = p(x,\xi)$, а $\partial'_i = \dfrac{\partial}{\partial \xi'^i}$ – до відповідної змінної у $p' = p(x,\xi')$. Визначені таким чином операції можуть бути розширені до таких:

$$\begin{aligned} &D\Big(\big(X_1,\ldots,X_l\big)_p|p'\Big),\ D\Big(p|\big(Y_1,\ldots,Y_m\big)_{p'}\Big) \\ &\text{та } D\Big(\big(X_1,\ldots,X_l\big)_p|\big(Y_1,\ldots,Y_m\big)_{p'}\Big), \end{aligned} \tag{5.16}$$

для будь-яких векторних полів, визначених на M.

Зауваження 5.1 *В щойно наведених формулах векторні поля – це диференціальні оператори вигляду:*

$$X = X^1 \frac{\partial}{\partial \xi^1} + X^2 \frac{\partial}{\partial \xi^2} + \ldots + X^m \frac{\partial}{\partial \xi^m}.$$

Взагалі, якщо на деякому многовиді задані $X = X^1 e_1 + X^2 e_2 + \ldots + X^m e_m$ – векторне поле, $f(x_1, x_2, \ldots, x_m)$ – функція, в диференціальній геометрії вважається, що запис Xf означає:

$$Xf = X^1 \frac{\partial f}{\partial x^1} + X^2 \frac{\partial f}{\partial x^2} + \ldots + X^m \frac{\partial f}{\partial x^m}.$$

Саме таким чином треба розуміти формули (5.16).

Тепер розглянемо обмеження цих формул на діагональ $\{(p,p) | p \in M\} \subset M \times M$ і позначимо відображення, індуковані на M таким чином:

$$D(X_1, \ldots, X_l | \cdot) : p \to D\left((X_1, \ldots, X_l)_p | p\right),$$

$$D(\cdot | Y_1, \ldots, Y_m) : p \to D\left(p | (Y_1, \ldots, Y_m)_p\right),$$

$$D(X_1, \ldots, X_l | Y_1, \ldots, Y_m) : p \to D\left((X_1, \ldots, X_l)_p | (Y_1, \ldots, Y_m)_p\right).$$

З рівнянь (5.14) випливає:

$$D(\partial_i | \cdot) = D(\cdot | \partial_i) = 0, \tag{5.17}$$

$$D(\partial_i \partial_j | \cdot) = D(\cdot | \partial_i \partial_j) = -D(\partial_i | \partial_j) \left(\stackrel{\text{def}}{=} g_{ij}^{(D)} \right). \tag{5.18}$$

Крім того, матриця $\left(g_{ij}^{(D)}\right)$ породжує позитивно напіввизначену квадратичну форму. Якщо породжена форма строго позитивно визначена в усіх точках M, то ми кажемо, що D є *дивергенцією*, або *контрастною функцією* на M. Дивергенція однозначно задає ріманову метрику як:

$$\partial_i, \partial_j^{(D)} = g_{ij}^{(D)},$$

або у інваріантній формі:

$$X, Y^{(D)} = -D(X | Y).$$

Ця метрика з'являється, якщо ми випишемо для дивергенції D апроксимацію до другого порядку включно:

$$D(p|q) = \frac{1}{2} g_{ij}^{(D)}(q) \Delta\xi^i \Delta\xi^j + \mathrm{o}\left(\Delta\xi^2\right), \tag{5.19}$$

де $\Delta\xi^i = \xi^i(p) - \xi^i(q)$ і $\mathrm{o}\left(\Delta\xi^2\right)$ означає доданки, що прямують до нуля швидше, ніж $\Delta\xi^2$ по мірі наближення p до q. Додамо, що доданків першого порядку у (5.19) немає внаслідок (5.17).

Маючи дивергенцію D можна також задати афінну зв'язність $\nabla^{(D)}$ з коефіцієнтами:

$$\Gamma_{ij,k}^{(D)} = -D\left(\partial_i \partial_j | \partial_k\right), \tag{5.20}$$

або

$$\nabla_X^{(D)} Y, Z^{(D)} = -D(XY|Z).$$

Зауважимо, що отримана зв'язність – симетрична: $\Gamma_{ij,k}^{(D)} = \Gamma_{ji,k}^{(D)}$. Використовуючи метрику і коефіцієнти зв'язності можна отримати для дивергенції D апроксимацію третього порядку:

$$D(p|q) = \frac{1}{2} g_{ij}^{(D)}(q) \Delta\xi^i \Delta\xi^j + \frac{1}{6} h_{ijk}^{(D)}(q) \Delta\xi^i \Delta\xi^j \Delta\xi^k + \mathrm{o}\left(\Delta\xi^3\right), \tag{5.21}$$

де

$$h_{ijk}^{(D)} = D\left(\partial_i \partial_j \partial_k | \cdot\right).$$

Таким чином коефіцієнти $h_{ijk}^{(D)}$ пов'язані з $g_{ij}^{(D)}$ та $\Gamma_{ij,k}^{(D)}$ таким чином

$$h_{ijk}^{(D)} = \partial_i g_{jk}^{(D)} + \Gamma_{jk,i}^{(D)}. \tag{5.22}$$

І навпаки, як метрика так і коефіцієнти зв'язності визначаються через розкладання (5.21) та співвідношення (5.22).

Тепер розглянемо дивергенцію $D^*(p|q) = D(q|p)$. Її називають *дуальною* до $D(p|q)$. Розрахувавши метрику, використовуючи дуальну дивергенцію отримуємо:

$$g_{ij}^{(D^*)} = g_{ij}^{(D)}$$

і

$$\Gamma_{ij,k}^{(D^*)} = -D\left(\partial_k | \partial_i \partial_j\right). \tag{5.23}$$

Є очевидною теорема.

Теорема 5.2 *Зв'язності* $\nabla^{(D)}$ *та* $\nabla^{(D^*)}$ *є взаємно дуальними по відношенню до метрики* $g_{ij}^{(D)}$.

Використовуючи дуальні величини, отримуємо рівняння, що є аналогічними (5.21) і (5.22).

$$D(p|q) = D^*(q|p) = \frac{1}{2} g_{ij}^{(D)}(p) \Delta\xi^i \Delta\xi^j - \\ -\frac{1}{6} h_{ijk}^{(D^*)}(p) \Delta\xi^i \Delta\xi^j \Delta\xi^k + \mathrm{o}\left(\Delta\xi^3\right), \tag{5.24}$$

$$h_{ijk}^{(D^*)} = D\left(\cdot | \partial_i \partial_j \partial_k\right) = \partial_i g_{jk}^{(D)} + \Gamma_{jk,i}^{(D^*)}. \tag{5.25}$$

Таким чином, ми бачимо, що будь-яка дивергенція породжує дуальну структуру з нульовим крученням. Навпаки, будь-який набір $\left(g, \nabla, \nabla^*\right)$, тобто метрика і пара узгоджених з нею взаємно дуальних зв'язностей може породжувати дивергенцію. Маючи таку трійку в координатній системі $\xi^1, \xi^2, \ldots, \xi^m$ можна локально задати дивергенцію:

$$D(p|q) \overset{\text{def}}{=} \frac{1}{2} g_{ij}(q) \Delta\xi^i \Delta\xi^j + \frac{1}{6} h_{ijk}(q) \Delta\xi^i \Delta\xi^j \Delta\xi^k, \tag{5.26}$$

де $\Delta\xi^i = \xi^i(p) - \xi^i(q)$ і

$$h_{ijk}(q) \overset{\text{def}}{=} \partial_i g_{jk} + \Gamma_{jk,i} = \Gamma_{ij,k} + \Gamma_{ik,j}^* + \Gamma_{jk,i}. \tag{5.27}$$

Оскільки $\Gamma_{ik,j}^*$ відрізняється від $\Gamma_{ik,j}$ на величину, пропорційну симетричному по всіх трьох індексах коваріантному тензору (див. (2.10), (2.11)), то ми бачимо, що h_{ijk} має бути також абсолютно симметричною величиною. Тому з (5.26) та (5.27) ми отримаємо $D\left(\partial_i \partial_j | \cdot\right) = g_{ij}$ та $D\left(\partial_i \partial_j \partial_k | \cdot\right) = h_{ijk}$, з чого випливає, що трійка $\left(g, \nabla, \nabla^*\right)$ породжує дивергенцію, яка в свою чергу однозначно породжує таку ж саме трійку. Додамо, що виконання умови невід'ємності у (5.15) забезпечується саме позитивною визначеністю метричного тензора g_{ij}. Тим не менш, дивергенцію, що відповідає цій трійці $\left(g, \nabla, \nabla^*\right)$ – можна побудувати і іншим шляхом. Зокрема так:

$$D(p|q) \overset{\text{def}}{=} \frac{1}{2} g_{ij}(p) \Delta\xi^i \Delta\xi^j - \frac{1}{6} h^*_{ijk}(p) \Delta\xi^i \Delta\xi^j \Delta\xi^k,$$

де $\Delta\xi^i = \xi^i(p) - \xi^i(q)$ і

$$h^*_{ijk} \overset{\text{def}}{=} \partial_i g_{jk} + \Gamma^*_{jk,i} = \Gamma^*_{ij,k} + \Gamma_{ik,j} + \Gamma^*_{jk,i}.$$

Як ми вже казали, що запропоновані вирази для обчислення дивергенції отримані нами з локальних міркувань, але у [18] доведено, що для будь-якої дуальної структури існує глобально визначена дивергенція, що її породжує.

Тепер розглянемо дуже важливий клас дивергенцій, що є застосовними для дослідження статистичних моделей [19]. Нехай $f(u)$ є опукою функцією. Для будь-якої пари розподілів $p(x,\xi)$ та $p(x,\xi')$ задаємо дивергенцію наступним чином:

$$D_f\left(p(x,\xi)|p(x,\xi')\right) \overset{\text{def}}{=} \int p(x,\xi) f\left(\frac{p(x,\xi')}{p(x,\xi)}\right) dx.$$

Визначена таким контрастна функція має назву *f-дивергенції*. Познайомимося з її властивостями. По-перше, згадаємо так звану нерівність Йєнсена. Вона полягає в тому, що для будь-якої опуклої функції $f(x)$ і будь-якого розподілу $p(x,\xi)$ завжди виконується

$$f\left(\int p(x,\xi) g(x) dx\right) \le \int p(x,\xi) f\left(g(x)\right) dx.$$

Застосовуючи його для *f*-дивергенції, маємо:

$$\begin{aligned} D_f\left(p(x,\xi)|p(x,\xi')\right) &= \int p(x,\xi) f\left(\frac{p(x,\xi')}{p(x,\xi)}\right) dx \ge \\ &\ge f\left(\int p(x,\xi) \frac{p(x,\xi')}{p(x,\xi)} dx\right) = f\left(\int p(x,\xi') dx\right) = f(1). \end{aligned} \tag{5.28}$$

З іншого боку, якщо $\xi = \xi'$, то згідно з (5.14) $D_f\left(p(x,\xi)|p(x,\xi)\right) = 0$, отже враховуючи (5.28) маємо, що $f(1) = 0$.

Важливим прикладом гладкої *f*-дивергенції є α-дивергенція, яка позначається $D^{(\alpha)} = D_{f(\alpha)}$ для дійсного числа α. Вона породжується функцією:

$$f^{(\alpha)}(u) \overset{\text{def}}{=} \begin{cases} \dfrac{4}{1-\alpha^2}\left(1-u^{\frac{1+\alpha}{2}}\right) (\alpha \neq \pm 1) \\ u \ln u \ (\alpha = 1) \\ -\ln u \ (\alpha = -1) \end{cases}. \quad (5.29)$$

Для $\alpha \neq \pm 1$ ми маємо:

$$D^{(\alpha)}\left(p(x,\xi)|p(x,\xi)\right) = \frac{4}{1-\alpha^2}\left(1-\int\left(p(x,\xi)\right)^{\frac{1-\alpha}{2}}\left(p(x,\xi)\right)^{\frac{1+\alpha}{2}} dx\right). \quad (5.30)$$

У випадку $\alpha = \pm 1$:

$$D^{(-1)}\left(p(x,\xi)|p(x,\xi)\right) = D^{(1)}\left(p(x,\xi)|p(x,\xi)\right) = \int p(x,\xi) ln\left(\frac{p(x,\xi)}{p(x,\xi)}\right) dx. \quad (5.31)$$

Ми бачимо, що α-дивергенція $D^{(\alpha)}$ індукує структуру $\left(g, \nabla^{(\alpha)}\right)$. Варто зазначити, що рівність

$$D^{(\alpha)}\left(p(x,\xi)|p(x,\xi)\right) = D^{(-\alpha)}\left(p(x,\xi)|p(x,\xi)\right).$$

Виконується для будь-якого α. Зокрема, $D^{(0)}\left(p(x,\xi)|p(x,\xi)\right)$ є симетричною, більше того, $\sqrt{D^{(0)}\left(p(x,\xi)|p(x,\xi)\right)}$ задовільняє усім аксіомам відстані. Дійсно, з (5.31) при $\alpha = 0$ маємо

$$D^{(0)}\left(p(x,\xi)|p(x,\xi)\right) = 4\left(1-\int\left(p(x,\xi)\right)^{\frac{1}{2}}\left(p(x,\xi)\right)^{\frac{1}{2}} dx\right) =$$

$$= 2\left(2-2\int\left(p(x,\xi)\right)^{\frac{1}{2}}\left(p(x,\xi)\right)^{\frac{1}{2}} dx\right) =$$

$$= 2\left(\int p(x,\xi)dx + \int p(x,\xi)dx - 2\int\left(p(x,\xi)\right)^{\frac{1}{2}}\left(p(x,\xi)\right)^{\frac{1}{2}} dx\right) =$$

$$= 2\int\left(p(x,\xi) + p(x,\xi) - 2\left(p(x,\xi)\right)^{\frac{1}{2}}\left(p(x,\xi)\right)^{\frac{1}{2}}\right) dx =$$

$$= 2\int\left(\left(p\left(x,\xi\right)\right)^{\frac{1}{2}} - \left(p\left(x,\xi\ \right)\right)^{\frac{1}{2}}\right)^2 dx.$$

Ми тут суттєво використовували той факт, що

$$\int p\left(x,\xi\right)dx = \int p\left(x,\xi\ \right)dx = 1.$$

$\sqrt{D^{(0)}\left(p\left(x,\xi\right)|p\left(x,\xi\ \right)\right)}$ має назву *відстані Гелінгера.*

Випадок $\alpha = \pm 1$, так звана ± 1 – дивергенція має декілька назв, зокрема *дивергенція Каллбека*, *інформація Кулбека-Ляйблера*, *відносна ентропія*, або просто дивергенція.

Знайдемо явні вирази для метрики та коефіцієнтів для $\nabla^{(\alpha)}$ зв'язності.

Щоб знайти метрику

$$g_{ij}^{(D)} = D\left(\partial_i\partial_j|\cdot\right),$$

знайдемо спочатку

$$\frac{\partial}{\partial\xi^i}\frac{\partial}{\partial\xi^j}\left(\frac{4}{1-\alpha^2}\left(1-\int\left(p\left(x,\xi\right)\right)^{\frac{1-\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{1+\alpha}{2}}dx\right)\right) =$$

$$= \frac{4}{1-\alpha^2}\frac{1-\alpha}{2}\frac{\partial}{\partial\xi^j}\left(-\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\left(p\left(x,\xi\right)\right)^{-\frac{1+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{1+\alpha}{2}}dx\right) =$$

$$= -\frac{}{+\alpha}\left(\int\left(\begin{array}{c}\frac{\partial\ p\left(x\ \xi\right)}{\partial\xi^i\partial\xi^j}\left(p\left(x\ \xi\right)\right)^{\frac{}{+\alpha}} - \\ -\frac{}{+\alpha}\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\left(p\left(x\ \xi\right)\right)^{\frac{}{+\alpha}}\end{array}\right)\left(p\left(x\ \xi'\right)\right)^{\frac{}{+\alpha}}dx\right) =$$

$$= -\frac{2}{1+\alpha}\int\frac{\partial^2 p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}\left(p\left(x,\xi\right)\right)^{-\frac{1+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{1+\alpha}{2}}dx +$$

$$+\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\left(p\left(x,\xi\right)\right)^{-\frac{3+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{1+\alpha}{2}}dx.$$

Враховуючи, що

$$D\left(\partial_i\partial_j|\cdot\right)=\frac{\partial}{\partial\xi^i}\frac{\partial}{\partial\xi^j}D\left(p\left(x,\xi\right)|\,p\left(x,\xi'\right)\right)_{|\xi=\xi'},$$

Отримуємо:

$$g_{ij}^{(D)}=D\left(\partial_i\partial_j|\cdot\right)=-\frac{2}{1+\alpha}\int\frac{\partial^2p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}dx+\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\frac{1}{p\left(x,\xi\right)}dx.$$

Як ми бачимо, перший доданок внаслідок (2.4) обертається в нуль, а другий – власне і є інформаційною метрикою Фішера згідно з (2.5):

$$g_{ij}^{(D)}=D\left(\partial_i\partial_j|\cdot\right)=\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\frac{1}{p\left(x,\xi\right)}dx=E\left[\partial_i l_\xi\partial_j l_\xi\right].$$

Тепер знайдемо коефіцієнти зв'язності:

$$\Gamma_{ij,k}^{(D)}=-D\left(\partial_i\partial_j|\partial_k\right).$$

Спочатку знайдемо

$$\frac{\partial}{\partial\xi'^k}\frac{\partial}{\partial\xi^i}\frac{\partial}{\partial\xi^j}D\left(p\left(x,\xi\right)|\,p\left(x,\xi'\right)\right)=$$

$$=\frac{\partial}{\partial\xi'^k}\left(\begin{array}{l}-\frac{2}{1+\alpha}\int\frac{\partial^2p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}\left(p\left(x,\xi\right)\right)^{-\frac{1+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{1+\alpha}{2}}dx+\\+\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\left(p\left(x,\xi\right)\right)^{-\frac{3+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{1+\alpha}{2}}dx\end{array}\right)=$$

$$=-\int\frac{\partial^2p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}\left(p\left(x,\xi\right)\right)^{-\frac{1+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{\alpha-1}{2}}\frac{\partial p\left(x,\xi\right)}{\partial\xi^k}dx+$$

$$+\frac{1+\alpha}{2}\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\left(p\left(x,\xi\right)\right)^{-\frac{3+\alpha}{2}}\left(p\left(x,\xi'\right)\right)^{\frac{\alpha-1}{2}}\frac{\partial p\left(x,\xi'\right)}{\partial\xi'^k}dx.$$

Остаточно маємо:

$$\Gamma_{ij,k}^{(\alpha)} = -D\left(\partial_i\partial_j|\partial_k\right) = -\frac{\partial}{\partial\xi'^k}\frac{\partial}{\partial\xi^i}\frac{\partial}{\partial\xi^j}D\left(p\left(x,\xi\right)|p\left(x,\xi'\right)\right)_{|\xi=\xi} =$$

$$= \int\frac{\partial^2 p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}\frac{\partial p\left(x,\xi\right)}{\partial\xi^k}\frac{1}{p\left(x,\xi\right)}dx - \tag{5.32}$$

$$-\frac{1+\alpha}{2}\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\frac{\partial p\left(x,\xi\right)}{\partial\xi^k}\frac{1}{\left(p\left(x,\xi\right)\right)^2}dx.$$

Це цілком збігається із формулою (2.12). Більше того, якщо ми повторимо наші обчислення метрики та коефіцієнтів зв'язності для дивергенції Кулбека – Ляйблера (5.31), то отримаємо:

$$g_{ij}^{(KL)} = D\left(\partial_i\partial_j|\cdot\right) = \int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\frac{1}{p\left(x,\xi\right)}dx,$$

$$“_{ij,k}^{(KL)} = \Gamma_{ij,k}^{(-1)} = \int\frac{\partial^2 p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}\frac{\partial p\left(x,\xi\right)}{\partial\xi^k}\frac{1}{p\left(x,\xi\right)}dx.$$

$$“_{ij,k}^{(KL^*)} = \Gamma_{ij,k}^{(1)} = \int\frac{\partial^2 p\left(x,\xi\right)}{\partial\xi^i\partial\xi^j}\frac{\partial p\left(x,\xi\right)}{\partial\xi^k}\frac{1}{p\left(x,\xi\right)}dx -$$

$$-\int\frac{\partial p\left(x,\xi\right)}{\partial\xi^i}\frac{\partial p\left(x,\xi\right)}{\partial\xi^j}\frac{\partial p\left(x,\xi\right)}{\partial\xi^k}\frac{1}{\left(p\left(x,\xi\right)\right)^2}dx.$$

Це цілком збігатиметься з тим, якщо ми підставимо $\alpha=-1$ та $\alpha=1$ у (5.22) відповідно.

5.4. Дуально плоскі простори

Нехай $\left(g,\nabla,\nabla^*\right)$ – деяка дуальна структура на статистичному многовиді M. З теореми 5.2 випливає, що з того, що ∇ є плоскою – випливає, що плоскою також є і ∇^*. Зокрема, ми бачили, що експоненціальні структури є 1-плоскими, а мішані структури в свою чергу є -1-плоскими (секція 3.2). Як виявилося, вони обидві є ±1-плоскими. З іншого боку, це вже було видно по виразу тензора кривини експоненціального простору (3.11):

$$ {}^{\alpha}R^{h}_{ijk} = \frac{1-\alpha^2}{2}\left(g^{ht}T_{tsk}g^{sp}T_{ijp} - g^{ht}T_{tsj}g^{sp}T_{ikp} \right). $$

Ми бачили, що він обертається в нуль не тільки коли $\alpha = \pm 1$.

В загалі, многовид M із заданою на ньому дуальною структурою $\left(g, \nabla, \nabla^*\right)$ – позначатимемо $\left(M, g, \nabla, \nabla^*\right)$ – будемо називати *дуально плоским простором,* якщо обидві здуальні зв'язності ∇ та ∇^* є плоскими.

При певних обставинах така властивість може розповсюджуватися і на підмноговиди. Нагадаємо (секція 1.8), що підмноговид $S \subset M$ називають *автопаралельним*, якщо для кожного вектора, дотичного $T_{x_0}\left(S\right)$ і кожної кривої γ, що виходить з x_0, паралельний перенос цього вектору вздовж γ (відносно аффінної зв'язності охоплюючого простору M) призводить до того, що перенесений вектор лишається дотичним до S. Звісно, треба завжди пояснювати в контексті якої зв'язності многовид є автопаралельним. Зараз, звичайно мова йтиме про α-зв'язності. Зрозуміло, що при одному значенні α вкладений підмноговид може бути автопаралельним, при іншому – ні. Тому, в таких випадках коректно говорити по α-автопаралельні многовиди.

Існує така теорема [3].

Теорема 5.3 *Нехай* $\left(M, g, \nabla, \nabla^*\right)$ *є дуально плоским простором. Якщо підмноговид* $S \subset M$ *є автопаралельним по відношенню до* ∇ *або* ∇^**, тоді S є дуально плоским многовидом по відношенню до дуальної структури* $\left(g_M, \nabla_M, \nabla^*_M\right)$*, індукованою на S дуальною структурою* $\left(g, \nabla, \nabla^*\right)$.

Розглянемо загальну структуру дуально плоского простору $\left(M, g, \nabla, \nabla^*\right)$. Безпосередньо з означення випливає існування на многовиді M ∇-афінної координатної системи, яку ми позначимо $\left[\theta^i\right]$, і ∇^*-афінної координатної системи, яку ми позначимо відповідно $\left[\eta_j\right]$. Відповідно, задаємо диференціальні оператори $\partial_i = \dfrac{\partial}{\partial \theta^i}$, які визначають дотичні вектори до координатних кривих $\left[\theta^i\right]$. Аналогічно задаємо диференціальні оператори $\partial^j = \dfrac{\partial}{\partial \eta_j}$.

Зауваження 5.2 *Ми вже писали, зокрема в Зауваженні 5.1, що дуже часто диференціальні оператори інтепретують як вектори. Зокрема, диференціальний оператор* $\partial_i = \dfrac{\partial}{\partial x^i}$ *прямо називають базисним вектором, що є дотичним до координатної кривої* x^i. *Це пов'язано в першу чергу з тим, що при заміні координат диференціальні оператори перетворюються так само, як і базисні вектори. Порівняйте:*

Перетворення базисних векторів $e'_i = \dfrac{\partial x^i}{\partial x^{i'}} e_i$,

і перетворення диференціальних операторів $\dfrac{\partial}{\partial x^{i'}} = \dfrac{\partial x^i}{\partial x^{i'}} \dfrac{\partial}{\partial x^i}$.

Це має прояв, наприклад, коли ми перераховуємо градієнт у новій системі координат, користуючись ланцюговим правилом:

$$\frac{\partial f}{\partial x^{i'}} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial f}{\partial x^i}.$$

згідно з ланцюговим правилом. Формула (1.2)*, згідно з якою відбувається перерахунок компонент тензора загального вигляду – якраз є узагальненням щойно згаданого правила перерахунку градієнта.*

Отже, надалі ми будемо сприймати $\partial_i = \dfrac{\partial}{\partial x^i}$ *як базисний вектор* e_i, *що є дотичним до координатної кривої* x^i *або як векторне поле у якого усі контраваріантні координати дорівнюють нулю, крім і-ї компоненти, яка дорівнює одиниці.*

Зауваження 5.3 *Дуже важливо відрізняти загальну ситуацію, коли ми позначали систему координат статистичного многовиду як* $\left[\xi^i\right]$ *від нинішньої, коли у нас є дві спеціальні системи координат, а саме,* $\left[\theta^i\right]$ *та* $\left[\eta_j\right]$. *В першому, загальному випадку системи координат* $\left[\xi^i\right]$ *верхній і нижній індекси є відповідно контра- і коваріантними координатами тензорів, що враховується при перерахунку за формулою* (1.2)*.*

Коли ми працюємо в системах координат $\left[\theta^i\right]$ *та* $\left[\eta_j\right]$ *– то ми стикаємося з такими властивостями, які проявляються лише конкретно в даних спеціальних системах координат. Тому*

врахування контра- і коваріантності при заміні систем координат – просто втрачає сенс. У зв'язку з тим – індекси розташовують, головним чином, згідно з тим, наскільки це буде зручно при застосування правила Ейнштейна для сумації за одноіменними індексами. Наприклад, вектор для базисного вектору $\partial^j = \dfrac{\partial}{\partial \eta_j}$ *в системі координат* $\left[\eta_j\right]$ *ми будемо використовувати в тому числі позначення* e^j*. Це ні в якому разі не означає що такі вектори змінюватимуться в загальному випадку до чогось «контраваріантно».*

Але, узагальнюючи, можна рекомендувати таке правило. У системі координат $\left[\eta_j\right]$*, індекси що зазвичай контраваріантні – записуються як коваріантні і навпаки.*

Очевидно, що векторні $\partial_i = \dfrac{\partial}{\partial \theta^i}$, будуть плоскими у зв'язності ∇ в системі координат $\left[\theta^i\right]$, а $\partial^j = \dfrac{\partial}{\partial \eta_j}$ – плоскими у зв'язності ∇^* в системі координат $\left[\eta_j\right]$:

$$\nabla_{\frac{\partial}{\partial \theta^j}} \frac{\partial}{\partial \theta^i} = \sum_{t=1}^{m} \Gamma_{ij}^t(\theta) \frac{\partial}{\partial \theta^t} = 0, \nabla_{\frac{\partial}{\partial \eta_j}} \frac{\partial}{\partial \eta_i} = \sum_{t=1}^{m} \Gamma_t^{ij*}(\eta) \frac{\partial}{\partial \eta_t} = \mathbf{0}. \quad (5.33)$$

Очевидно, що (5.33) є інваріантними відносно лінійного перетворення систем координат. Тому, для дуально плоского простору $\left(M, g, \nabla, \nabla^*\right)$, маючи ∇-афінну коодинатну систему $\left[\theta^i\right]$, підібрати ∇^*-афінну коодинатну систему $\left[\eta_j\right]$ так, щоб для скалярного добутку двох векторів $\partial_i = \dfrac{\partial}{\partial \theta^i}$ та $\partial^j = \dfrac{\partial}{\partial \eta_j}$ виконувалось [3]:

$$\partial_i, \partial^j = \delta_i^j, \quad (5.34)$$

де δ_i^h є вже знайомим символом Кронекера.

Якщо координатні системи $\left[\theta^i\right]$ та $\left[\eta_j\right]$, задані на римановому многовиді задовільняють (5.33), то їх називають *взаємно дуальними*. Для дуально плоскої структури $\left(M, g, \nabla, \nabla^*\right)$ така пара координатних систем завжди існує. З іншого боку, можна стверджувати, існування для деякого ріманового многовиду $\left(M, g\right)$ таких систем $\left[\theta^i\right]$ та $\left[\eta_j\right]$,

що відносно них зв'язності ∇, ∇^* є відповідно афінними, і $\frac{\nabla+\nabla^*}{2}$ буде зв'язністю Леві – Чівіта – забезпечує виконання того факту, що $\left(M, g, \nabla, \nabla^*\right)$ буде дуально плоскою структурою.

Компоненти рімановой метрики для такого простору визначаються за формулами:

$$\partial_i, \partial_j = g_{ij}, \partial^i, \partial^j = g^{ij}. \tag{5.35}$$

Базисні вектори в обох системах координат пов'язані наступним чином:

$$\partial^j = \left(\partial^j \theta^i\right)\partial_i = g^{ji}\partial_i, \partial_i = \left(\partial_i \eta_j\right)\partial^j = g_{ij}\partial^j.$$

Отже, очевидним є

$$\frac{\partial \eta_j}{\partial \theta^i} = g_{ij}, \frac{\partial \theta^i}{\partial \eta_j} = g^{ij}. \tag{5.36}$$

Звідси ми бачимо, що $g_{ik}g^{kj} = \delta_i^j$.

Оскільки матриця g_{ij} є симетричною, це означає, що

$$\frac{\partial \eta_j}{\partial \theta^i} - \frac{\partial \eta_i}{\partial \theta^j} = 0,$$

Тобто диференціальна форма $\eta_j d\theta^j$ э замкненою. Оскільки за лемою Пуанкаре принаймні локально існує така функція (скаляр) $\psi(\theta): M \to \mathbb{R}$, що

$$\eta_i = \frac{\partial \psi}{\partial \theta^i}. \tag{5.37}$$

З (5.36) та (5.37) випливає, що метрика в системі координат $\left[\theta^i\right]$ може бути представленою у вигляді:

$$g_{ij}(\theta) = \frac{\partial^2 \psi(\theta)}{\partial \theta^i \partial \theta^j}, \tag{5.38}$$

як це було зроблено у секції 3.2, рівняння (3.8). Нагадаємо, що многовиди з такою метрикою, утворюють *структури Гессе* (див. Теорему 3.3). Оскільки g_{ij}, то відповідна квадратична форма є позитивно визначеною, звідси $\psi(\theta)$ є строго опуклою функцією змінних $\theta^1, \theta^2, \ldots, \theta^m$.

Подібним чином ми можемо проаналізувати рівняння,

$$\theta^i = \frac{\partial \varphi}{\partial \eta_i}. \tag{5.39}$$

яке виникає з огляду того, що

$$\frac{\partial \theta^i}{\partial \eta_j} - \frac{\partial \theta^j}{\partial \eta_i} = 0.$$

Отже,

$$g^{ij}(\eta) = \frac{\partial^2 \varphi(\eta)}{\partial \eta_i \partial \eta_j}, \tag{5.40}$$

де $\varphi(\eta)$ є строго опуклою функцією змінних $\eta_1, \eta_2, \ldots, \eta_m$.

Якщо виписати зв'язок між $\varphi(\eta)$ та $\psi(\theta)$, отримаємо

$$\varphi(\eta) = \theta^i \eta_i - \psi(\theta). \tag{5.41}$$

Однак, в (5.41) не відзеркалено зв'язок між змінними $\theta^1, \theta^2, \ldots, \theta^m$ та $\eta_1, \eta_2, \ldots, \eta_m$ що описується у (5.37) та (5.39). Якщо ми хочемо в (5.41) врахувати зв'язок (5.37) та строгу опуклість $\psi(\theta)$, то це потрібно записати так

$$\varphi(\eta) = \max_{\theta} \left(\theta^i \eta_i - \psi(\theta) \right). \tag{5.42}$$

Аналогічно, для врахування опуклості $\varphi(\eta)$ та (5.39) – маємо

$$\psi(\theta) = \max_{\eta} \left(\theta^i \eta_i - \varphi(\eta) \right). \tag{5.43}$$

Перетворення систем координат $\left[\theta^i\right] \leftrightarrow \left[\eta_j\right]$, що представлені (5.42) та (5.43) мають назву перетворень Лежандра. Таким чином, функції $\psi(\theta)$ та $\varphi(\eta)$ є потеціалами, які породжують взаємно дуальні структури Гессе.

Також

$$\Gamma^*_{ij,k}(\theta) = \nabla^*_{\partial_i} \partial_j, \partial_k = \partial_i \partial_j \partial_k \psi(\theta), \tag{5.44}$$

$$\Gamma^{ij,k}(\eta) = \nabla_{\partial^i} \partial^j, \partial^k = \partial^i \partial^j \partial^k \varphi(\eta). \tag{5.45}$$

Зважимо, що

$$\Gamma^{*ij,k}(\eta) = \Gamma_{ij,k}(\theta) = 0, \tag{5.46}$$

оскільки згідно з (5.12) та (5.38)

$$\partial_i\partial_j\partial_k\psi(\theta)=\frac{\partial g_{ij}}{\partial x^k}=\Gamma^t_{ik}g_{tj}+\Gamma^{*t}_{jk}g_{it}=\Gamma_{ik,j}+\Gamma^*_{jk,i}.$$
$$\partial^i\partial^j\partial^k\varphi(\eta)=\Gamma^{ik}_{,t}(\eta)g^{tj}(\eta)+\Gamma^{*jk}_{,t}(\eta)g^{ti}(\eta)=\Gamma^{ik,j}(\eta)+\Gamma^{*jk,i}(\eta).$$

Усі наведені міркування ми можемо підсумувати у вигляді теореми.

Теорема 5.4 *Якщо для дуально плоского многовиду* (M,g,∇,∇^*) *існує* ∇*-афінна координатна система* $\left[\theta^i\right]$*, то також існує дуальна до* $\left[\theta^i\right]$ ∇^**-афінна координатна система* $\left[\eta_j\right]$*. Ці координатні системи пов'язані одна з одною через перетворення Лежандра* (5.42), (5.43) *потенціалів* $\psi(\theta)$ *та* $\varphi(\eta)$*. Для безпосереднього перерахунку координат використовують формули* (5.37) *та* (5.39)*. Компоненти метричного тензора в цих системах координат обчислюються відповідно через формули* (5.38) *та* (5.40)*.*

Коенфіцієнти об'єктів ∇*- та* ∇^**-зв'язностей обчислюються за допомогою* (5.44)–(5.46)*.*

Зауваження 5.3 *Варто нагадати, що коли ми для короткості називали об'єкти* $\Gamma_{ik,j}$, $\Gamma^*_{jk,i}$, $\Gamma^{ik,j}$ *та* $\Gamma^{*jk,i}$ *– коефіцієнтами зв'язності (зокрема, щойно посилаючись на* (5.44)–(5.46)*) – то це було не зовсім вірно. Перш ніж підставити отримані «гамми» у рівняння для коваріантного диференціювання – потрібно метричним тензором змінити локацію останнього індексу. А саме, у координатах* $\left[\theta^i\right]$ *коефіцієнти зв'язності мають вигляд* $\Gamma^k_{ij}(\theta)=\Gamma_{ij,t}(\theta)g^{tk}(\theta)$ *та* $\Gamma^{*k}_{ij}(\theta)=\Gamma^*_{ij,t}(\theta)g^{tk}(\theta)$.

Відповідно, у координатах $\left[\eta_j\right]$:
$\Gamma^{ij}_{,k}(\eta)=\Gamma^{ij,t}(\eta)g_{tk}(\eta)$ *та* $\Gamma^{*ij}_{,k}(\eta)=\Gamma^{*ij,t}(\eta)g_{tk}(\eta)$.

Наприклад, у координатах $\left[\theta^i\right]$ *рівняння e-геодезичної виглядатиме так:*

$$\frac{d^2\theta^i}{ds^2}+\Gamma_{jk,t}(\theta)g^{ti}(\theta)\frac{d\theta^j}{ds}\frac{d\theta^k}{ds}=0. \tag{5.47}$$

У тих саме координатах рівняння m-геодезичної запишемо таким чином:

$$\frac{d^2\theta^i}{ds^2}+\Gamma^*_{ij,t}(\theta)g^{tk}(\theta)\frac{d\theta^j}{ds}\frac{d\theta^k}{ds}=0. \tag{5.48}$$

У координатах $\left[\eta_j\right]$ *рівняння т-геодезичної:*

$$\frac{d^2\eta_i}{ds^2}+\Gamma^{*jk,t}\left(\eta\right)g_{ti}\left(\eta\right)\frac{d\eta_j}{ds}\frac{d\eta_k}{ds}=0. \tag{5.49}$$

І нарешті, е-геодезична у координатах $\left[\eta_j\right]$*:*

$$\frac{d^2\eta_i}{ds^2}+\Gamma^{jk,t}\left(\eta\right)g_{ti}\left(\eta\right)\frac{d\eta_j}{ds}\frac{d\eta_k}{ds}=0. \tag{5.50}$$

Враховуючи (5.44)–(5.46)*, маємо:*

е-геодезична в системі $\left[\theta^i\right]$

$$\frac{d^2\theta^i}{ds^2}=0. \tag{5.47'}$$

т-геодезична в системі $\left[\theta^i\right]$

$$\frac{d^2\theta^i}{ds^2}+\partial_i\partial_j\partial_t\psi\left(\theta\right)g^{tk}\left(\theta\right)\frac{d\theta^j}{ds}\frac{d\theta^k}{ds}=0. \tag{5.48'}$$

т-геодезична в системі $\left[\eta_j\right]$

$$\frac{d^2i}{ds^2}=0. \tag{5.49'}$$

е-геодезична в системі $\left[\eta_j\right]$

$$\frac{d^2\eta_i}{ds^2}+\partial^j\partial^k\partial^t\varphi\left(\eta\right)g_{ti}\left(\eta\right)\frac{d\eta_j}{ds}\frac{d\eta_k}{ds}=0. \tag{5.50'}$$

Саме простота рівнянь (5.48') *та* (5.50') *демонструє, яким потужним знаряддям у інформаційній гометрії є апарат взаємно дуальних зв'язностей.*

5.5. Канонічна дивергенція

У секції 5.2 ми розглядали дивергенції, і побачили що деяка коректно задана дивергенція однозначно породжує метрику і дві дуальних зв'язності, пов'язаних із цією метрикою. З іншого боку, одна і та сама дуальна структура може породжуватись різними дивергенціями. Але існує так звана *канонічна дивергенція*, яка на дуально плоскому просторі задається однозначно.

Нехай заданий дуально плоский многовид $\left(M,g,\nabla,\nabla^*\right)$ на якому існують взаємно дуальні координатні системи $\left[\theta^i\right]$ та $\left[\eta_j\right]$ і відповідні потенціали $\psi(\theta)$ та $\varphi(\eta)$. На цьому многовиді виберемо довільним чином два різних розподіла $p=p(x,\xi)$ та $p'=p(x,\xi')$, таких що $p,p'\in M$. Літерою ξ ми позначаємо параметри розподілів що задані в якійсь довільній координатній системі $\xi=\left(\xi^1,\xi^2,\ldots,\xi^m\right)$, що була задана на M з самого початку. Звісно, оперуючи з нашими спеціальними координатними системами нам доведеться із довільної переходити до $\theta=\left(\theta^1,\theta^2,\ldots,\theta^m\right)$ або до $\eta=(\eta_1,\eta_2,\ldots,\eta_m)$. Тобто, ми матимемо функції $\theta=\theta(\xi)$, $\eta=\eta(\xi)$. Оскільки розподіли $p=p(x,\xi)$ та $p'=p(x,\xi')$ є різними, то різними будуть і їх представлення у нових координатних системах: для $p=p(x,\xi)\theta=\theta(\xi)$ та $\eta=\eta(\xi)$, а для $p'=p(x,\xi')$ відповідно $\theta'=\theta(\xi')$ та $\eta'=\eta(\xi')$.

Враховуючи викладене, ми можемо ввести на многовиді $\left(M,g,\nabla,\nabla^*\right)$ дивергенцію наступним чином:

$$D\left(p(x,\xi)|p\left(x,\xi\right)\right)\overset{\text{def}}{=}\psi\left(\theta(\xi)\right)+\varphi\left(\eta\left(\xi\right)\right)-\theta^i(\xi)\eta_i\left(\xi\right). \quad (5.51)$$

З іншого боку, якщо для нас не буде суттєвою якась третя система координат, і будуть задані два різних розподіла – скажімо p та q з їхніми координатами у відповідних системах, а саме $\theta(p)$ та $\theta(q)$ у системі $\left[\theta^i\right]$ і $\eta(p)$ та $\eta(q)$ у системі $\left[\eta_j\right]$ – тоді дивергенція двох розподілів виглядатиме так:

$$D(p|q)\overset{\text{def}}{=}\psi\left(\theta(p)\right)+\varphi\left(\eta(q)\right)-\theta^i(p)\eta_i(q). \quad (5.52)$$

Величину, що обчислюється за формулою (5.51) або (5.52) – називатимемо *канонічною дивергенцією для структури* $\left(M,g,\nabla,\nabla^*\right)$, або задля короткості (g,∇)-*дивергенцією*.

Ми бачимо, що згідно (5.42) та (5.43) виконуються властивості $D(p|q) \geq 0$ та $D(p|q) = 0$ якщо і тільки якщо $p = q$. Є також справедливим

$$D\left(\left(\partial_i \partial_j\right)_p | q\right) = g_{ij}(p) ma D\left(p | \left(\partial_i \partial_j\right)_q\right) = g^{ij}(q). \qquad (5.53)$$

З виразу (5.53) ми бачимо, що з міркувань симетрії, можна ввести спряжену $\left(g, \nabla^*\right)$-*дивергенцією*

$$D^*(p|q) = D(q|p). \qquad (5.54)$$

Якщо існує підмноговид $S \subset M$, причому він є автопаралельним по відношенню як для зв'язності ∇ так і для ∇^* то на ньому очевидно буде індуковано дуальну плоску структуру $\left(S, g_S, \nabla_S, \nabla_S^*\right)$. Можна довести, що на ньому індукована $\left(g_S, \nabla_S\right)$-*дивергенція,* така, що для будь-яких $p, q \in S$

$$D_S(p \mid q) = D(q \mid p). \qquad (5.55)$$

Аналогічне твердження є справедливим для спряженої дивергенції:

$$D_S(q|p) = D(q|p) = D^*(p|q) = D_S^*(p|q).$$

Якщо ∇ є зв'язністю Леві – Чівіта, узгодженою із метрикою Фішера, то маємо $\nabla = \nabla^*$. Для дуально-плоского простору це просто означає, що в ньому можна задати систему координат $\left[\theta^i\right]$, що є евклідовою. Така координатна система буде «автодуальною», тобто $\theta^i = \eta_i$. Очевидно, що потенціал бути мати вигляд:

$$\psi = \varphi = \frac{1}{2} \sum_{i=1}^{m} \left(\theta^i\right)^2.$$

Підставивши це у (5.54), отримуємо:

$$\begin{aligned} D(p|q) &= \frac{1}{2} \sum_{i=1}^{m} \left(\left(\theta^i(p)\right)^2 + \left(\theta^i(q)\right)^2 - 2\theta^i(p)\theta^i(q) \right) = \\ &= \frac{1}{2} \sum_{i=1}^{m} \left(\theta^i(p) - \theta^i(q)\right)^2 = \frac{1}{2} \left(d(p,q)\right)^2. \end{aligned} \qquad (5.56)$$

Тут через $d(p,q)$ позначено евклідову відстань

$$d(p,q)=\sqrt{\sum_{i=1}^{m}\left(\theta^i(p)-\theta^i(q)\right)^2}.$$

Загалом, канонічна дивергенція $D(p|q)$ на дуально плоскому просторі лише приблизно дорівнює $\frac{1}{2}\left(d(p,q)\right)^2$ в сенсі формули (5.19), тобто з точністю до доданків $o\left(\Delta\xi^2\right)$.

5.6. Найважливіші властивості канонічної дивергенція

Сформулюємо декілька теорем, що відзеркалюють властивості канонічної дивергенції [3].

Теорема 5.5 *Нехай на дуально плоскому многовиді* $\left(M,g,\nabla,\nabla^*\right)$ *задані взаємно дуальні координатні системи* $\left[\theta^i\right]$ *та* $\left[\eta_j\right]$, *і D є дивергенцією на M. Для того, щоб D була канонічною дивергенцією, необхідно і достатньо, щоб для будь-яких розподілів* $p,q,r\in M$ *виконувалось наступне триангуляційне співвідношення*:

$$D(p|q)+D(q|r)-D(p|r)=\sum_{i=1}^{m}\left(\theta^i(p)-\theta^i(q)\right)\left(\eta_i(r)-\eta_i(q)\right). \quad (5.57)$$

Також, ми можемо узагальнити теорему Піфагора, яку було сформульовано для евклідових відстаней на випадок канонічної дивергенції що є заданою на довільному дуально плоскому просторі $\left(M,g,\nabla,\nabla^*\right)$.

Теорема 5.6 *Нехай на дуально плоскому многовиді* $\left(M,g,\nabla,\nabla^*\right)$ *задані три точки (розподіли)* $p,q,r\in M$. *При цьому,* ∇*-геодезична* γ_1 *з'єднує точки p та q, а* ∇^**-геодезична* γ *з'єднує точки q та r так, що в точці q обидві геодезичні є ортогональними згідно із заданою метрикою g. Тоді має місце рівність Піфагора (рис. 5).*

$$D(p|r)=D(p|q)+D(q|r). \quad (5.58)$$

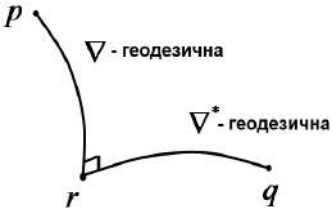


Рисунок 5 – Рівність Піфагора для (g,∇) – дивергенції

Наступне твердження є наслідком Теореми 5.6.

Наслідок 5.7 *Нехай на дуально плоскому многовиді* (M,g,∇,∇^*) *задано точку* $p\in M$ *і нехай S є* ∇^**-автопаралельним підмноговидом M, причому точка* $q\in S$. *Тоді необхідною та достатньою умовою* $D(p|q)=\min\limits_{r\in S}D(p|r)$ *є ортогональність до S тієї* ∇*-геодезичної, що з'єднує точки p та q (рис. 6).*

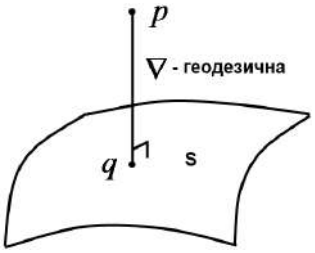


Рисунок 6 – Теорема про проєкцію для (g,∇) – дивергенції

Точка *q* про яку йде мова у Наслідку 5.7 має *назву* ∇ *-проекції точки p на підмноговид S.*

Наступне твердження є справедливим для будь-якого підмноговида *S*.

Теорема 5.8 *Нехай на дуально плоскому многовиді* (M,g,∇,∇^*) *задано точку* $p\in M$ *і нехай S є підмноговидом M, причому точка* $q\in S$. *Тоді необхідною та достатньою умовою, щоб q була стаціонарною точкою функції* $D(p|\cdot):r\to D(p|r)$ *обмеженої на підмноговид S (в тому сенсі, що усі частинні похідні відносно координат на S мають бути рівні нулю) є ортогональність до* $S\nabla$*-геодезичної, що з'єднує точки p та q*

Наслідок 5.9 *Нехай на дуально плоскому многовиді* $\left(M, g, \nabla, \nabla^*\right)$ *задано точку* $p \in M$ *і нехай* S *є гіперповерхнею у* M*, так, що* $S = \left\{q \in M | D\left(p|q\right) = c\right\}$*, тобто* S *є «D-сферою» радіуса* $c = const$*. Тоді кожна* ∇*-геодезична, що походить через центр, яким є точка* p *отогонально перетинає* S*.*

5.7. Взаємно дуальні шарування

Нехай заданий дуально плоский m-вимірний многовид $\left(M_m, g, \nabla, \nabla^*\right)$ на якому існують взаємно дуальні координатні системи $\left[\theta^i\right]$ та $\left[\eta_i\right]$. Розіб'ємо множину значень індексу $i = 1, 2, \ldots, m$ на дві частини: $i = 1, 2, \ldots, k$ та $i = k+1, k+2, \ldots, m$. Назвемо ці частини секцією I та секцією II відповідно. Нехай $S\left(c_I\right)$ буде множиною точок, чиї координати $\left[\eta_i\right]$ в секції I є фіксованими наборами констант $c_I = \left(c_{I,i}\right)$ для $i = 1, 2, \ldots, k$

$$S\left(c_I\right) = \left\{p \in M_m | \eta_1\left(p\right) = c_{I,1}, \eta_2\left(p\right) = c_{I,2}, \ldots, \eta_k\left(p\right) = c_{I,k}\right\}.$$

Таким чином утворений $\left(m-k\right)$-вимірний ∇^* – автопаралельний підмноговид. Множина $S\left(c_I\right)$ визначається конкретними значеннями набору c_I, і якщо $c_I \neq c_I$, то

$$S\left(c_I\right) \cap S\left(c_I\right) = \varnothing,$$

і є очевидним, що мом можемо представити весь многовид M_m, як об'єднання усіх можливих

$$\bigcup_{c_I} S\left(c_I\right) = M_m.$$

Тобто, M_m є розбитим на підмноговиди. Таке розбиття називають *шаруванням* M_m.

Подібним чином розглянемо множину $A\left(d_{II}\right)$ точок, чиї координати $\theta^{k+1}, \theta^{k+2}, \ldots, \theta^m$ у θ-координатній системі є фіксованими набором $d_{II}^{k+1}, \ldots, d_{II}^m$:

$$A\left(d_{II}\right) = \left\{p \in M_m | \theta^{k+1}\left(p\right) = d_{II}^{k+1}, \theta^{k+2}\left(p\right) = d_{II}^{k+2}, \ldots, \theta^m\left(p\right) = d_{II}^m\right\}.$$

Ми отримали ∇-автопаралельний підмноговид, і $\{A(d_{II})\}$ формує ще одне шарування многовиду M_m.

Будь-яка точка p однозначно визначає певні $S(c_I)$ та $A(d_{II})$, що містить цю точку. Таким чином, точка p є перетином цих $S(c_I)$ та $A(d_{II})$. Дотичним простором $T_p(S)$ до $S(c_I)$ в цій точці є $(m-k)-$вимірний простір, натягнутий на базисні вектори $\{\partial^{k+1},\partial^{k+2},\ldots,\partial^m\}$ а дотичним простором $T_p(A)$ до $A(d_{II})$ є відповідно k-вимірний простір, натягнутий на базисні вектори $\{\partial_1,\partial_2,\ldots,\partial_k\}$. Дуже важливим фактом є те, що

$$\partial_i,\partial^j=0, \text{ якщо } i\neq j.$$

Тому $T_p(S)$ та $T_p(A)$ є взаємно ортогональними. Такі шарування, що є ортогональними одне до одного називають *взаємно дуальними*.

Можна розділити координати точки p на перші $(m-k)$ вимірів і решту k вимірів: $\eta=(\eta_I,\eta_{II})$. Або подібним чином можна розбити θ-координати: $\theta=(\theta^I,\theta^{II})$. Але, оскільки точка p є перетином $S(\eta_I)$ та $A(\theta^{II})$, то дуже корисним є вибрати координати наступним чином: $\xi=(\eta_I,\theta^{II})$.

Такий вибір має назву *мішаної системи координат* (рис. 7).

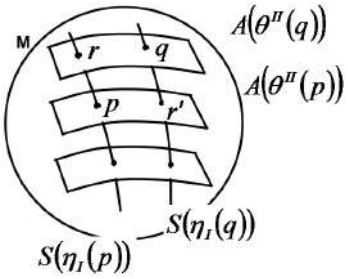


Рисунок 7 – Взаємно дуальні шарування многовиду M

Нехай p та q є точками, чиїми мішаними координатами є $(\eta_I(p),\theta^{II}(p))$ та $(\eta_I(q),\theta^{II}(q))$ відповідно. Далі, розглянемо точки з такими мішаними координатами r: $(\eta_I(p),\theta^{II}(q))$ та r': $(\eta_I(q),\theta^{II}(p))$. Тоді r є ∇-проекцією q на $S(\eta_I(p))$ та r' є ∇^*-проекцією q на $A(\theta^{II}(p))$.

Звідси ми бачимо, що

$$D(p|q) = D(p|r) + D(r|q) \text{ та}$$

$$D(q|p) = D(q|r\ ) + D(r\ |p),$$

де символом D позначено (g, ∇^*)-дивергенцією.

6. Висновки

Виконане дослідження доводить, що інформаційна геометрія є актуальною та потужною математичною основою для сучасної кібербезпеки, дозволяючи моделювати ймовірнісні розподіли, аналізувати інформаційні відстані та виявляти аномалії у великих потоках даних. Геометричні методи, які ґрунтуються на статистичних многовидах, метриці Фішера та дивергенції, показують переваги у виявленні складних атак і оптимізації захисту даних, перевищуючи класичні підходи, особливо для аналізу гібридних загроз та в умовах невизначеності. Практична цінність полягає у поєднанні інформаційної геометрії з алгоритмами штучного інтелекту, що дозволяє підвищувати ефективність систем раннього попередження і знижувати обчислювальні витрати. Водночас існують обмеження: висока обчислювальна складність, нестача стандартів впровадження та потреба адаптації до українського законодавства. Перспективи подальших досліджень охоплюють інтеграцію з квантовими технологіями та блокчейном, оптимізацію алгоритмів, емпіричне тестування на реальних даних, застосування до нових сфер (наприклад, соціальних мереж чи біометрії), а також розробку стандартів та освітніх програм у співпраці з українськими установами. Інформаційна геометрія має всі шанси стати основою для більш стійких і адаптивних систем кіберзахисту. Ці перспективи дозволять не лише теоретично вдосконалити методи, але й забезпечити їх практичне впровадження для підвищення рівня кібербезпеки в Україні та світі.

Список використаних джерел

1. National Institute of Standards and Technology. (2023). Cybersecurity Framework. Gaithersbaurg: NIST.

2. Про основні засади забезпечення кібербезпеки України : Закон України від 05.10.2017 № 2163-VIII : станом на 20 квітня 2025 року. URL: https://zakon.rada.gov.ua/laws/show/2163-19#Text

3. Amari S. Methods of information geometry. Providence, RI : American Mathematical Society, 2000. 206 p.

4. Dodson, C. T. J. (2017). On Some Information Geometric Approaches to Cyber Security. DOI: 10.1007/978-3-319-51500-7_9

5. Cherevko Y., Chepurna O., Kuleshova Y. On information geometry methods for data analysis. *Geometry, Integrability and Quantization*. 2024. Issue 29. P. 11–22. DOI: 10.7546/giq-29-2024-11-22

6. Chuang I. L., Nielsen M. A. Prescription for experimental determination of the dynamics of a quantum black box. Journal of Modern Optics. 1997. Vol. 44, No. 11–12. P. 2455–2467. DOI: https://doi.org/10.1080/09500349708231894

7. Wang X., Li Y., Zhang Z. Data Privacy and Security Risks in AI-Driven Organizations. *Journal of Information Privacy and Security.* 2019. Vol. 37, No. 187–200. P. 112–130.

8. Watanabe S. Mathematical Theory of Bayesian Statistics. Boca Raton: CRC Press. 2018.

9. Chen Y., Liu H., Zhang Q. Information geometry for big data security analytics. *IEEE Transactions on Information Forensics and Security.* 2020. Issue 15. P. 2100–2115.

10. Семенов П. В. Нейронні мережі та їх застосування. Київ : Техніка, 2023. 230 с.

11. Semenova O. O., Semenov A. O., Voitsekhovska O. O. Using a neural network for the vertical handover procedure. *Information technology and computer engineering.* 2020. Vol. 49, No. 3. P. 14–21. DOI: https://doi.org/10.31649/1999-9941-2020-49-3-14-21 (date of access: 28.07.2025).

12. Mikeš J. et al. Differential Geometry of Special Mappings. Palacky Univ. Press. Olomouc. 2019.

13. Eisenhart L. Riemannian Geometry. Princeton Univ. Press. 1926.

14. Amari S. Differential Geometry of Curved Exponential Families. Curvature and Information Loss. *Ann. Stat.* 1982. Issue 10. P. 357–385.

15. Cartan, Elie Leçons sur la géométrie des espaces de Riemann (2ème. ed. rev. et aug. ed.). Paris: Gauthier-Villars. p. VIII, 378. ISBN 287647008X. Zbl 0060.38101.

16. Shima H. The Geometry of Hessian Structures. World Sci. Publ. 2007.

17. Matumoto T. Any statistical manifold has a cotrast function -on the C3-functions taking the minimum at the diagonal of the product manifold. *Hiroshima Mathematical Journal*. 1993. Issue 23. P. 327–332.

18. Csiszar I. Information type measures of difference of probability distributions and indirect observations. S*tudia Scientiarum Mathematicarum Hungarica.* 1967. Issue 2. P. 299–318.

19. Kennedy, T. G. (Spring 2006). Some notes on sufficient statistics. Math 466/566 lecture notes, University of Arizona.

20. Lauritzen, S. L. Sufficiency and Unbiased Estimation. (2004, 19 October). Lecture 3 of BS2 Statistical Inference. University of Oxford. URL: https://www.stats.ox.ac.uk/~steffen/teaching/bs2siMT04/si3c.pdf

21. Lauritzen, S. L. Extremal Families and Systems of Sufficient Statistics Lecture Notes in Statistics. Vol. 49. (1988). Springer.

РОЗДІЛ 3. КРИПТОГРАФІЧНІ ТА СТЕГАНОГРАФІЧНІ МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

Соколов Артем Вікторович
д.т.н., проф., професор кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

Евдокимов Денис Віталійович
аспірант кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

КОНЦЕПЦІЯ КОДОВОГО УПРАВЛІННЯ У ЦИФРОВІЙ СТЕГАНОГРАФІЇ: ТЕОРЕТИЧНІ ОСНОВИ, РЕАЛІЗАЦІЇ ТА ПЕРСПЕКТИВИ

DOI https://doi.org/10.36059/978-966-397-537-5-6

1.1. Вступ і постановка задачі

Однією з ключових складових сучасних систем захисту інформації, зокрема при її передаванні за допомогою сучасних телекомунікаційних систем, є стеганографічні методи, що дозволяють приховати сам факт передачі секретної інформації [1]. Розвиток кіберпростору у бік багаторазового збільшення обсягів переданої мультимедійної інформації, а також поява пристроїв IoT і мобільних пристроїв [2–4], що володіють високими вимогами до обчислювальної ефективності застосовуваних методів, дозволяють сформулювати ключові вимоги до стеганографічних методів: висока надійність сприйняття стеганоповідомлення; стійкість стеганоповідомлення до атак проти вбудованого повідомлення, в першу чергу, атак стисненням із втратами; висока швидкодія та простота алгоритмічної реалізації.

Для забезпечення високої стійкості стеганографічного методу до атак стисненням, вбудовування інформації має відбуватися в низькочастотні та середньочастотні складові контейнера [5], що найбільш просто зробити із застосуванням областей перетворення контейнера:

дискретного косинусного перетворення (ДКП), вейвлет-перетворення, спектрального/сингулярного розкладання матриць блоків контейнера. Наслідком цієї обставини є те, що більшість відомих сьогодні стеганографічних методів, що позиціонуються як стійкі до атак проти вбудованого повідомлення [6–15] виконують операції вбудовування та вилучення інформації у зазначених областях перетворень. Тим не менш, використання областей перетворення контейнера для стеганоперетворення призводить до ускладнення алгоритмічної реалізації стеганографічних методів, до збільшення обчислювальної складності за рахунок наявності переходів з просторової області в область перетворення та назад, а також, часто, до неконтрольованих змін яскравості пікселів зображення у просторовій області, що може привести до погіршення надійності сприйняття стеганоповідомлення [1].

Принципова перспективність використання просторової області для вбудовування інформації була обґрунтована в роботі [16], тим не менш, до розробки концепції кодового управління, реалізація відповідності стеганографічного метода основним критеріям ефективності була утруднена.

Проривом стало створення концепції кодового управління вбудовуванням додаткової інформації [17], яка отримала свій подальший розвиток у роботах [18–20]. Зазначена концепція дозволяє здійснювати контроль за вбудовуванням додаткової інформації у ту або іншу частотну складову без необхідності здійснення затратного з обчислювальної точки зору переходу в області перетворень.

Метою дослідження є аналітичне узагальнення результатів, отриманих авторами у рамках розробки концепції кодового управління в цифровій стеганографії, із подальшим обґрунтуванням ефективності відповідних методів в умовах практичних атак і обмежень сучасних інформаційних систем.

Зважаючи на поставлену мету сформулюємо задачі дослідження:

1. Проаналізувати сучасні підходи до вбудовування додаткової інформації в цифрові контейнери з позицій стійкості до атак, надійності сприйняття та обчислювальної ефективності.

2. Систематизувати основні положення концепції кодового управління в цифровій стеганографії, зокрема у контексті використання досконалих алгебраїчних конструкцій.

3. Оцінити переваги та обмеження використання просторової області зображення для вбудовування інформації без необхідності переходу до частотної області.

4. Узагальнити та формалізувати результати, отримані в рамках класичного стеганографічного методу з кодовим управлінням.

5. Узагальнити та формалізувати результати стосовно стеганографічного методу з кодовим управлінням та можливістю сліпого декодування.

6. Узагальнити та формалізувати підходи до побудови стеганографічного методу з кодовим управлінням на основі досконалих двійкових решіток.

7. Визначити перспективні напрямки подальшого розвитку концепції кодового управління в умовах сучасних викликів інформаційної безпеки.

Об'єкт дослідження – процеси стеганографічного перетворення в цифрових контейнерах.

Предмет дослідження – методи стеганографічного вбудовування з використанням концепції кодового управління, зокрема їх теоретичні засади, кодові конструкції, властивості стійкості, надійності та можливість сліпого декодування.

Розглянемо далі основні наявні сьогодні теоретичні та практичні здобутки, отримані під час розвитку цієї концепції, а також основні проблеми, які необхідно вирішити для подальшого розвитку та практичного впровадження концепції кодового управління.

1.2. Основні визначення

Коротко розглянемо основні визначення, які застосовуються в роботі.

У сучасних методах цифрової стеганографії важливу роль відіграють перетворення зображень у частотну область, які дозволяють досягти високої непомітності та стійкості прихованої інформації. Одним із найпоширеніших таких перетворень є дискретне косинусне перетворення (ДКП) [21], що широко використовується в алгоритмах стиснення зображень (наприклад, JPEG) і є природним середовищем для приховування даних. Перехід до частотного представлення зображення дозволяє обирати зони з меншою чутливістю зорової системи людини для модифікації, знижуючи ризик візуального або статистичного виявлення. Матриця трансформант S ДКП визначається за допомогою наступного співвідношення:

$$S = C_N X C_N^T, C_N = \left\| c_{i,j} \right\|, i, j = 0,1,..., N-1, c_{i,j} =$$
$$= \begin{cases} \dfrac{1}{\sqrt{N}}, & \text{при } i = 0; \\ \sqrt{\dfrac{2}{N}} \cos\left(\dfrac{(2j+1) \cdot i \cdot \pi}{2N} \right), & \text{при } i > 0, \end{cases} \quad (1)$$

де X – фрагмент вихідного зображення розміру $N \times N$.

Перетворенням, яке покладено в основу стеганографічного методу з кодовим управлінням вбудовуванням інформації є двовимірне перетворення Уолша – Адамара. Окрім класичних частотних методів, у стеганографії набувають популярності альтернативні ортогональні перетворення, зокрема перетворення Уолша – Адамара [22]. Це перетворення базується на використанні функцій, що приймають лише два значення (±1), і не потребує операцій множення, що забезпечує його високу обчислювальну ефективність. Завдяки своїй простоті, перетворення Уолша – Адамара є привабливим для реалізації у системах з обмеженими ресурсами, а також дає можливість створювати ефективні алгоритми вбудовування інформації зі збереженням високої швидкодії. Матриця трансформант W для перетворення Уолша – Адамара визначаються за допомогою наступного співвідношення:

$$W = H'_N X H'^T_N, H'_N = \frac{1}{\sqrt{N}} H_N, H_N = \begin{bmatrix} H_{N/2} & H_{N/2} \\ H_{N/2} & -H_{N/2} \end{bmatrix}, H_1 = [1], \quad (2)$$

де H_N і H'_N – ненормована та нормована матриці Уолша – Адамара порядку $N \times N$, відповідно.

Вектор трансформант V одновимірного перетворення Уолша – Адамара вектора Y довжини N визначається за допомогою наступного співвідношення:

$$V = YH_N. \quad (3)$$

Одним з теоретичних здобутків, який лежить в основі концепції кодового управління вбудовуванням інформації є встановлений в роботі [23] взаємозв'язок між трансформантами двовимірного та одновимірного перетворення Уолша – Адамара, який може бути

записаний (з точністю до коефіцієнта $1/N$) за допомогою оператора $\tilde{A}$, який визначає запис матриці A порядку $N \times N$ у вигляді вектору-рядка довжини N^2 шляхом послідовної конкатенації рядків вихідної матриці A:

$$\tilde{W} = \tilde{X} H_{N^2}, \tag{4}$$

при цьому вираз (4) є справедливим і для інших видів перетворень, насамперед, ДКП, однак, замість матриці H_{N^2} в цьому випадку має застосовуватися матриця $A1_{N^2}$, що побудована відповідно до [23].

Реалізація основної ідеї кодового управління вбудовуванням інформації, яка була запропонована у роботі [17] зводиться до додавання матриць блоків контейнера і кодових слів, елементи яких належать алфавіту $\{\pm 1\}$, тоді блок стеганоповідомлення, а також, із врахуванням (4), його трансформанти перетворення Уолша – Адамара визначається як

$$\begin{gathered} M = X + T, \\ \tilde{M} H_{N^2} = (\tilde{X} + \tilde{T}) H_{N^2} = \tilde{X} H_{N^2} + \tilde{T} H_{N^2}, \end{gathered} \tag{5}$$

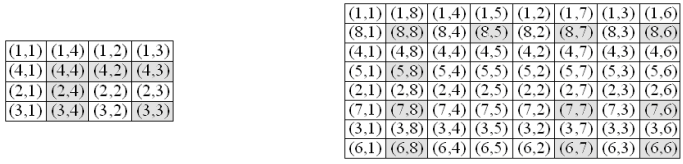
де M – блок стеганоповідомлення розміру $\mu \times \mu$, T – кодове слово розміру $\mu \times \mu$, елементи якого $t_{i,j} \in \{\pm 1\}, i, j = 1, 2, ..., \mu$.

З виразу (5) безпосередньо витікає той факт, що підбираючи такі кодові слова T, які характеризуються впливом на ті чи інші трансформанти перетворення Уолша – Адамара можна забезпечити строго заданий вплив на необхідні трансформанти перетворення Уолша – Адамара контейнера, таким чином, отримуючи задані властивості стеганоповідомлення M. В якості таких кодових слів T було запропоновано застосування $N \times N$-матричного подання рядків матриці Уолша – Адамара порядку N^2, тобто коду Ріда – Маллера першого порядку.

Беручи до уваги теоретичні результати роботи [23], можемо сформулювати дві достатні умови, які є основою для побудови кодових слів, що застосовуються у стеганографічних методах, заснованих на концепції кодового управління.

Достатня умова забезпечення надійності сприйняття стеганоповідомлення. Для забезпечення надійності сприйняття стеганоповідомлення достатньо проводити вбудовування додаткової інформації

таким чином, щоб в області перетворення Уолша – Адамара його результатом було збурення елементів, локалізація яких наведена на рис. 1. для $l \times l$-блоків розміру $l \in \{4,8,16\}$, при цьому саме вбудовування ДІ може здійснюватися не тільки безпосередньо в області Уолша – Адамара, а й у будь-якій іншій області контейнера (просторовій, перетворення). При необхідності використання блоків іншого розміру рекомендується проводити вбудовування ДІ таким чином, щоб результатом його було збурення елементів в області перетворення Уолша – Адамара в межах другого стовпця та другого рядка перетвореної матриці.



(1,1)	(1,4)	(1,2)	(1,3)
(4,1)	(4,4)	(4,2)	(4,3)
(2,1)	(2,4)	(2,2)	(2,3)
(3,1)	(3,4)	(3,2)	(3,3)

а

(1,1)	(1,8)	(1,4)	(1,5)	(1,2)	(1,7)	(1,3)	(1,6)
(8,1)	(8,8)	(8,4)	(8,5)	(8,2)	(8,7)	(8,3)	(8,6)
(4,1)	(4,8)	(4,4)	(4,5)	(4,2)	(4,7)	(4,3)	(4,6)
(5,1)	(5,8)	(5,4)	(5,5)	(5,2)	(5,7)	(5,3)	(5,6)
(2,1)	(2,8)	(2,4)	(2,5)	(2,2)	(2,7)	(2,3)	(2,6)
(7,1)	(7,8)	(7,4)	(7,5)	(7,2)	(7,7)	(7,3)	(7,6)
(3,1)	(3,8)	(3,4)	(3,5)	(3,2)	(3,7)	(3,3)	(3,6)
(6,1)	(6,8)	(6,4)	(6,5)	(6,2)	(6,7)	(6,3)	(6,6)

б

(1,1)	(1,16)	(1,8)	(1,9)	(1,4)	(1,13)	(1,5)	(1,12)	(1,2)	(1,15)	(1,7)	(1,10)	(1,3)	(1,14)	(1,6)	(1,11)
(16,1)	(16,16)	(16,8)	(16,9)	(16,4)	(16,13)	(16,5)	(16,12)	(16,2)	(16,15)	(16,7)	(16,10)	(16,3)	(16,14)	(16,6)	(16,11)
(8,1)	(8,16)	(8,8)	(8,9)	(8,4)	(8,13)	(8,5)	(8,12)	(8,2)	(8,15)	(8,7)	(8,10)	(8,3)	(8,14)	(8,6)	(8,11)
(9,1)	(9,16)	(9,8)	(9,9)	(9,4)	(9,13)	(9,5)	(9,12)	(9,2)	(9,15)	(9,7)	(9,10)	(9,3)	(9,14)	(9,6)	(9,11)
(4,1)	(4,16)	(4,8)	(4,9)	(4,4)	(4,13)	(4,5)	(4,12)	(4,2)	(4,15)	(4,7)	(4,10)	(4,3)	(4,14)	(4,6)	(4,11)
(13,1)	(13,16)	(13,8)	(13,9)	(13,4)	(13,13)	(13,5)	(13,12)	(13,2)	(13,15)	(13,7)	(13,10)	(13,3)	(13,14)	(13,6)	(13,11)
(5,1)	(5,16)	(5,8)	(5,9)	(5,4)	(5,13)	(5,5)	(5,12)	(5,2)	(5,15)	(5,7)	(5,10)	(5,3)	(5,14)	(5,6)	(5,11)
(12,1)	(12,16)	(12,8)	(12,9)	(12,4)	(12,13)	(12,5)	(12,12)	(12,2)	(12,15)	(12,7)	(12,10)	(12,3)	(12,14)	(12,6)	(12,11)
(2,1)	(2,16)	(2,8)	(2,9)	(2,4)	(2,13)	(2,5)	(2,12)	(2,2)	(2,15)	(2,7)	(2,10)	(2,3)	(2,14)	(2,6)	(2,11)
(15,1)	(15,16)	(15,8)	(15,9)	(15,4)	(15,13)	(15,5)	(15,12)	(15,2)	(15,15)	(15,7)	(15,10)	(15,3)	(15,14)	(15,6)	(15,11)
(7,1)	(7,16)	(7,8)	(7,9)	(7,4)	(7,13)	(7,5)	(7,12)	(7,2)	(7,15)	(7,7)	(7,10)	(7,3)	(7,14)	(7,6)	(7,11)
(10,1)	(10,16)	(10,8)	(10,9)	(10,4)	(10,13)	(10,5)	(10,12)	(10,2)	(10,15)	(10,7)	(10,10)	(10,3)	(10,14)	(10,6)	(10,11)
(3,1)	(3,16)	(3,8)	(3,9)	(3,4)	(3,13)	(3,5)	(3,12)	(3,2)	(3,15)	(3,7)	(3,10)	(3,3)	(3,14)	(3,6)	(3,11)
(14,1)	(14,16)	(14,8)	(14,9)	(14,4)	(14,13)	(14,5)	(14,12)	(14,2)	(14,15)	(14,7)	(14,10)	(14,3)	(14,14)	(14,6)	(14,11)
(6,1)	(6,16)	(6,8)	(6,9)	(6,4)	(6,13)	(6,5)	(6,12)	(6,2)	(6,15)	(6,7)	(6,10)	(6,3)	(6,14)	(6,6)	(6,11)
(11,1)	(11,16)	(11,8)	(11,9)	(11,4)	(11,13)	(11,5)	(11,12)	(11,2)	(11,15)	(11,7)	(11,10)	(11,3)	(11,14)	(11,6)	(11,11)

в

Рисунок 1 – Локалізація області можливого збурення в результаті стеганоперетворення в області перетворення Уолша – Адамара для $l \times l$-блоків цифрового зображення: а – $l = 4$; б – $l = 8$; в – $l = 16$

Достатня умова забезпечення нечутливості стеганоповідомлення до збурних дій. Для забезпечення нечутливості стеганоповідомлення до збурних дій достатньо проводити вбудовування

додаткової інформації таким чином, щоб в області перетворення Уолша – Адамара його результатом було збурення елементів, локалізованих як

ДКП	Перетворення Уолша-Адамара
$(1,1);\ (1,2);\ (2,1);$ $(3,1);\ (2,2);\ (1,3)...$	$4\times4:(1,1);\ (1,3);\ (3,1);\ (4,1);\ (3,3);\ (1,4)...$ $8\times8:(1,1);\ (1,5);\ (5,1);\ (7,1);\ (5,5);\ (1,7)...$ $16\times16:(1,1);\ (1,9);\ (9,1);\ (13,1);\ (9,9);\ (1,13)...$

, (6)

для $l\times l$-блоків розміру $l\in\{4,8,16\}$, при цьому саме вбудовування додаткової інформації може здійснюватися не тільки безпосередньо в області Уолша – Адамара, а й у будь-якій іншій області контейнера (просторовій, перетворення).

На практиці стеганографічні методи, побудовані на принципах кодового управління, характеризуються тим, що вплив на кожен окремий піксель контейнера обмежується зміною не більше ніж на ±1. Такий рівень модифікації є практично непомітним для людського зору, що забезпечує високу надійність візуального сприйняття модифікованого зображення. Відтак, основна увага при конструюванні кодових слів у межах таких методів зосереджується не на зменшенні помітності, а на максимізації стійкості до атак, спрямованих на виявлення або спотворення вбудованого повідомлення. Це визначає специфіку побудови кодів – вони мають бути не лише компактними, а й адаптованими до типових деструктивних впливів (наприклад, стиснення, фільтрації або статистичного аналізу).

З практичної точки зору інтерес являє побудова кодових слів різної довжини, насамперед практично важливих довжин 4×4, 8×8, що змінюють лише ті складові контейнеру, які були отримані у достатній умові забезпечення надійності стеганоповідомлення до збурних впливів.

Відповідно до встановленої відповідності між двовимірним і одновимірним перетворенням Уолша – Адамара в якості таких кодових слів візьмемо 5, 33, 37 і 1-ший рядки матриці Уолша – Адамара порядку $N^2=64$ (у вигляді відповідних їм матриць порядку $N=8$), для кожної з яких наведемо відповідну матрицю трансформант Уолша – Адамара (з точністю до коефіцієнта $1/N$).

$$T_1^+ = \begin{bmatrix} 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \end{bmatrix}, \quad W_1^+ = \begin{bmatrix} 0&0&0&0&64&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{bmatrix};$$

$$T_2^+ = \begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ -1&-1&-1&-1&-1&-1&-1&-1 \\ -1&-1&-1&-1&-1&-1&-1&-1 \\ -1&-1&-1&-1&-1&-1&-1&-1 \\ -1&-1&-1&-1&-1&-1&-1&-1 \end{bmatrix}, W_2^+ = \begin{bmatrix} 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 64&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{bmatrix}; \quad (7)$$

$$T_3^+ = \begin{bmatrix} 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ 1&1&1&1&-1&-1&-1&-1 \\ -1&-1&-1&-1&1&1&1&1 \\ -1&-1&-1&-1&1&1&1&1 \\ -1&-1&-1&-1&1&1&1&1 \\ -1&-1&-1&-1&1&1&1&1 \end{bmatrix}, W_3^+ = \begin{bmatrix} 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&64&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{bmatrix};$$

$$T_4^+ = \begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \end{bmatrix}, \quad W_4^+ = \begin{bmatrix} 64&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{bmatrix}.$$

У сучасних стеганографічних методах на основі кодового управління переважна більшість підходів використовує кодові слова, які впливають вибірково лише на одну трансформанту перетворення Уолша – Адамара, наприклад, кодові слова (7). Така стратегія дозволяє імплементувати досить прості та зрозумілі методи синтезу кодових слів, забезпечуючи як високу надійність сприйняття так і стійкість до атак проти вбудованого повідомлення. Проте вона має і суттєве обмеження – низьку інформаційну щільність, оскільки для

вбудовування одного біта інформації зазвичай використовується один блок або навіть одна трансформанта. Водночас можливість паралельного впливу кодових слів на кілька трансформант перетворення Уолша–Адамара відкриває потенціал для значного підвищення пропускної спроможності та стійкості до атак стеганоаналізу, особливо якщо вплив реалізується з урахуванням спектральних або кореляційних властивостей сигналу.

Однак синтез таких складних багатоканальних кодових конструкцій залишається майже недослідженою сферою. Вона потребує нових підходів до проєктування кодів із заданими спектральними характеристиками, здатних одночасно керувати групою трансформант без істотного порушення статистичних або візуальних параметрів контейнера. Це, зокрема, передбачає створення моделей залежностей між трансформантами, розробку метрик взаємного впливу та оптимізаційних критеріїв для балансування між непомітністю, стійкістю та пропускною спроможністю. Розвиток цього напряму відкриває перспективи для побудови високоефективних адаптивних методів стеганографії, здатних протистояти сучасним стеганоаналітичним атакам.

У наступних підрозділах буде детально проаналізовано три найбільш ефективні та практично значущі стеганографічні методи, що реалізують концепцію кодового управління: класичний метод, метод із сліпим декодуванням, а також перспективний метод, побудовану на основі досконалих двійкових решіток.

1.3. Класичний стеганографічний метод з кодовим управлінням

Класичний стеганографічний метод з кодовим управлінням [17] ґрунтується на надзвичайно елегантному та водночас ефективному принципі: приховане повідомлення вбудовується в контейнер шляхом адитивного накладення кодового слова на блок даних. У цьому підході кожен інформаційний біт визначає вибір конкретного кодового слова з заздалегідь заданого коду, яке потім додається до блоку контейнера (наприклад, трансформант або просторових значень). Завдяки цьому забезпечується керованість модифікацій,

передбачуваність структури спотворень і висока стеганографічна непомітність. Метод є не лише концептуально прозорим, а й надзвичайно гнучким: він легко адаптується до різних доменів перетворень, дозволяє реалізувати рівномірний розподіл змін, а також створює підґрунтя для подальших вдосконалень, зокрема в напрямку адаптивних схем і побудови спеціалізованих кодів з потрібними властивостями.

Вбудовування додаткової інформації

Крок 1. Виконуємо сегментацію вихідного зображення P розміру $m \times n$ на блоки $\mu \times \mu$.

Крок 2. Кожному блоку вихідного зображення розміру $\mu \times \mu$, задіяному в процесі стеганоперетворення, ставимо у відповідність λ біт додаткової інформації, таким чином отримуємо матрицю додаткової інформації D розміру $\frac{m}{\mu} \times \frac{n}{\mu}$, кожен елемент якої містить λ біт інформації. При цьому $\lambda = \log_2 J$, де J – кількість кодових слів, що використовуються.

Крок 3. Будуємо таблицю відповідності половини комбінацій із λ біт додаткової інформації кодовим словам $\{T_i^+\}$, тоді як друга половина комбінацій із λ біт додаткової інформації кодується за допомогою інверсій кодових слів $\{T_i^-\}$. Відзначимо, що така таблиця може бути частиною ключа. Отримуємо закодовану матрицю додаткової інформації шляхом подання кожного її елемента з λ біт додаткової інформації за допомогою кодових слів $\{T_i^+\}$ та $\{T_i^-\}$.

Крок 4. Виконуємо вбудовування інформації шляхом підсумовування матриці контейнера P з кодовою матрицею, що була отримана на *Кроці 3*, в результаті чого отримуємо стеганоповідомлення M, тобто

$$M = P + D. \tag{8}$$

Декодування додаткової інформації

Крок 1. Вилучення інформації відбувається шляхом віднімання з матриці можливо збуреного стеганоповідомлення $\overline{M}$ матриці контейнера P, що є частиною секретного ключа. В результаті вилучення, для кожного i-го блоку розміру $\mu \times \mu$ отримуємо матрицю Δ_i:

$$\{\Delta_i\} = \overline{M} - P,\ i = 0,1,\ldots mn-1. \quad (9)$$

Крок 2. Для кожної матриці T_i^+ робимо поелементне множення кожної отриманої на *Кроці 1* матриці Δ_i на матрицю T_i^+, після чого знаходимо суму всіх елементів результуючої матриці кожного кодового слова T_i^+, тобто розраховуємо значення:

$$\sigma_j = \sum_{l=0}^{\mu-1}\sum_{k=0}^{\mu-1}\Delta_i(l,k)T^+(l,k),\ j = 0,1,\ldots,J/2-1.$$

Крок 3. Серед отриманої для кожного блоку множини значень σ_i знаходимо максимальне за модулем значення. При цьому індекс знайденого значення буде відповідати індексу декодованого кодового слова $T_i^?$, в той час як знак знайденого максимуму буде відповідати знаку, з яким вказане кодове слово було вбудовано (прямий або інверсний вигляд).

Зазначимо, що *Крок 2* та *Крок 3* у частині декодування додаткової інформації в представленому методі, по суті, реалізують алгоритм оптимального прийому [24].

Зауваження. Зважаючи на те, що більшість використовуваних зображень сьогодні представлені з використанням моделі RGB, де для кодування кожного кольору відводиться 1 байт (кожна складова кольору представляється числами в діапазоні $[0,...,255]$), у разі наявності в блоці граничних для даного діапазону значень (0 або 255), вказаний блок не використовується в процесі стеганоперетворення у разі застосування розробленого стеганографічного методу.

На рис. 2 представлені графіки залежності кількості помилок, що виникають (у відсотках від загальної кількості біт додаткової інформації) від степеню стиснення *QF* зображення алгоритмом JPEG для кожної з розглянутих трансформант Уолша – Адамара (кодові слова (7) розміру 8×8, а також кодові слова, що одночасно впливають на всі складові блоків контейнера), що дозволяє оцінити ефективність їх використання для протистояння атакам стисненням на стеганоповідомлення. При цьому, через те, що всі кодові слова, які використовують одночасно всі частотні складові показують практично еквівалентні результати, на графіку (рис. 2) вони показані однією кривою.

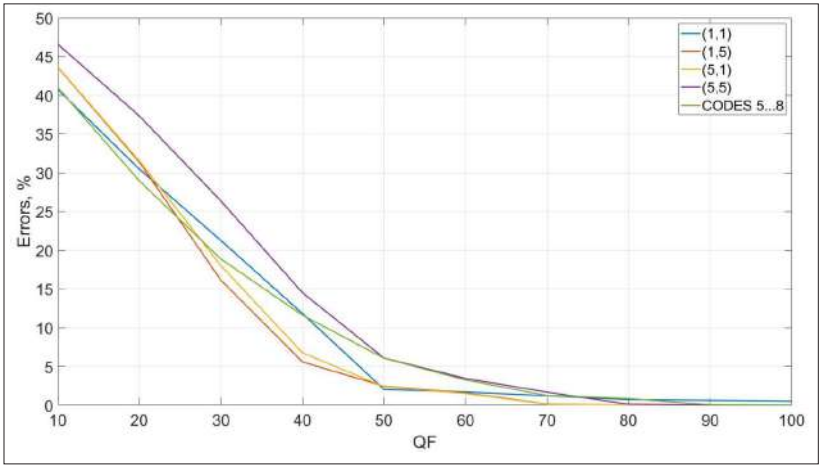


Рисунок 2 – Графік залежності кількості помилок, що виникають, від степеня *QF* стиснення стеганоповідомлення алгоритмом JPEG для різних трансформант Уолша – Адамара, в які здійснюється вбудовування

Аналіз даних, представлених на рис. 2 показує, що вбудовування інформації як з використанням кодових слів (7), так і кодових слів, які використовують одночасно всі частотні складові, формує стеганоповідомлення, нечутливе до атаки стисненням. При цьому, при стисненні з коефіцієнтом $QF = 60\,\%$ (на практиці вкрай нечасто використовуються більші степені стиснення), відсоток помилок, що виникають, не перевищує 5 %, що є показником високої стійкості стеганографічного методу.

На рис. 3 представлено графік залежності кількості помилок при декодуванні вбудованого повідомлення від степеню стиснення стеганоповідомлення алгоритмом JPEG під час проведення атаки для кодових слів розміру 16×16.

Аналіз даних рис. 3 дозволяє дійти висновку, що використання кодових слів розміру 16×16 забезпечує високий рівень стійкості розробленого стеганографічного методу. Так, навіть при атаці з рівнем стиску $QF = 30$ для будь-якого кодового слова кількість помилок при декодуванні не перевищує 5 %. Однак, така стійкість досягається за рахунок зниження пропускної спроможності стеганографічного методу до значення 1/256.

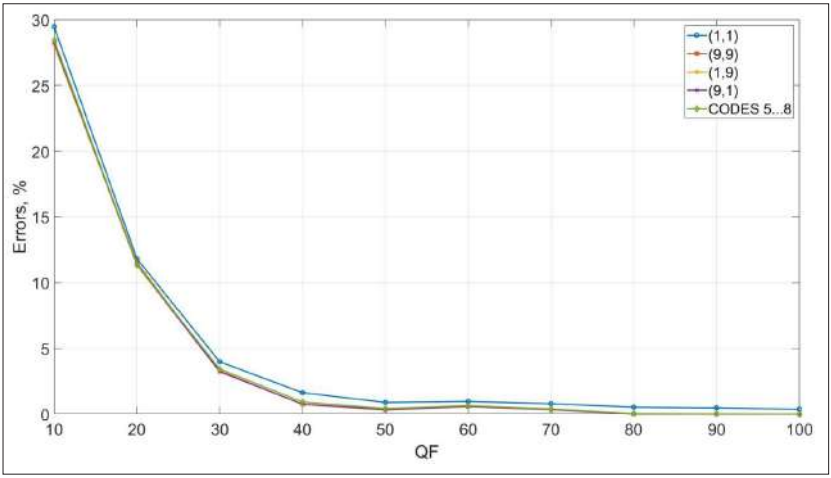
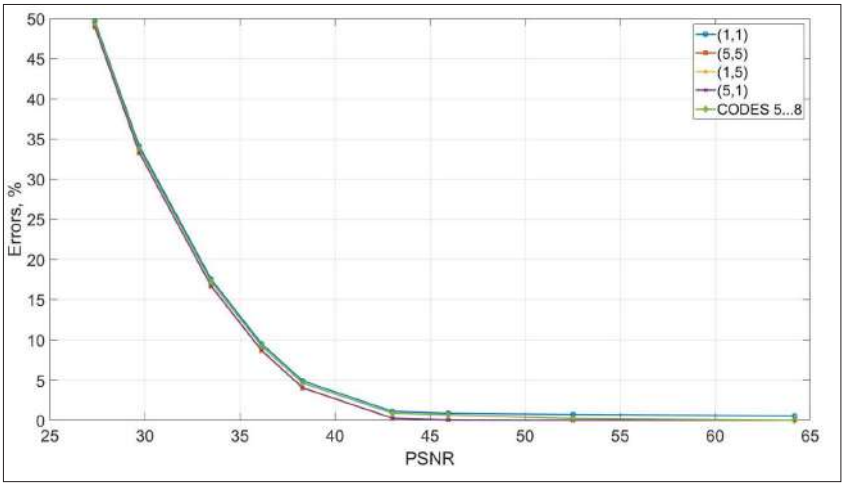


Рисунок 3 – Графік залежності кількості помилок, що виникають, від степеня *QF* стиснення зображеного алгоритмом JPEG для $\mu = 16$

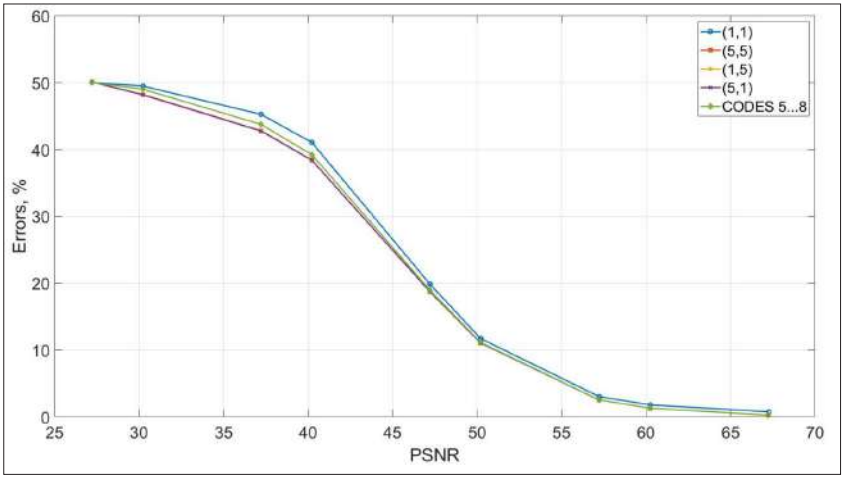
На рис. 4 показано графік залежності кількості помилок при декодуванні додаткової інформації при її вилученні від степеню зашумленості стеганоповідомлення гаусівським шумом (а) та шумом типу “salt&pepper” (б), який виражено у PSNR для блоків розміру 8×8.

Аналіз даних, представлених на рис. 4 показує, що розроблений стеганографічний метод з кодовим управлінням забезпечує рівень стійкості до виникнення помилок у стеганоповідомленні на рівні менше за 5 % при зашумленні гаусівським шумом з PSNR = 43 дБ або зашумленні на рівні PSNR = 57 дБ для випадку шуму “salt&pepper”.

На рис. 5 представлений приклад вбудовування додаткової інформації в контейнер за допомогою розробленого стеганографічного методу, при цьому використані параметри $\lambda = 1$, $\mu = 8$, а також для вбудовування інформації застосовано кодове слово T_2^+. Розмір контейнера становить 2592×3872, при цьому вбудовування інформації відбувалося в кожну колірну складову. Таким чином, загальний обсяг вбудованої інформації для даного контейнера склав 470 448 біт.

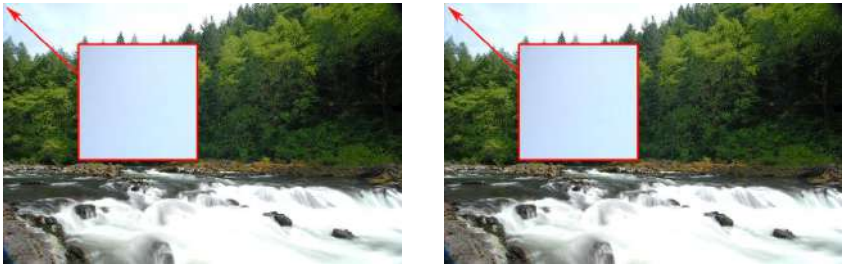


а



б

Рисунок 4 – Графік залежності кількості помилок при декодуванні зашумленого стеганоповідомлення від PSNR при використанні кодових слів розміру 8 × 8: а – зашумлення гаусівським шумом, б – зашумлення шумом типу “salt&pepper”



а б

Рисунок 5 – Приклад використання розробленого стеганографічного методу: а – контейнер; б – стеганоповідомлення

Суб'єктивне ранжування зображень, що представлені на рис. 5 не дозволяє виявити артефакти, або будь-які інші відмінності стеганоповідомлення від вихідного контейнера. Даний факт є очікуваним, оскільки результатом стеганоперетворення є збурення значень яскравості пікселів контейнера на ±1.

Зауваження. Обчислювальна складність алгоритмічної реалізації розробленого методу визначається кількістю блоків, що використовуються в процесі стеганоперетворення, і в найгіршому випадку складе $\underline{O}(nm)$ операцій. З урахуванням блоковості методу очевидним є його внутрішній паралелізм, що разом із використанням просторової області для стеганоперетворення забезпечує потенційну можливість його використання в режимі реального часу для потокового контейнера.

1.4. Стеганографічний метод з кодовим управлінням та сліпим декодуванням

Серед варіацій стеганографічного методу з кодовим управлінням особливе місце посідає модифікація, що забезпечує сліпе декодування [19]. На відміну від класичного методу, який передбачає наявність оригінального контейнера для витягування повідомлення, цей варіант дозволяє відновлювати приховану інформацію без доступу до незміненого зображення, що значно підвищує

практичну гнучкість методу та його придатність для реальних умов застосування.

У стеганографічному методі з кодовим управлінням вбудовуванням додаткової інформації із сліпим декодуванням блок контейнера X застосовується для вбудовування одного біта додаткової інформації d, тоді як зазначений біт розподіляється по підблокам χ_j, $j = 1,2,...,4$. Вбудовування додаткової інформації відбувається у адитивний спосіб, тоді як блок стеганоповідомлення визначатиметься у наступний спосіб:

$$M = X + T_{\mu}, \tag{10}$$

де T_{μ} – кодове слово розміру $\mu \times \mu$, яке формується наступним чином:

$$T_{\mu} = \left[\begin{array}{c|c} t_1 & t_1 \\ \hline -t_1 & -t_1 \end{array}\right] + \left[\begin{array}{c|c} t_2 & t_2 \\ \hline -t_2 & -t_2 \end{array}\right] + ... + \left[\begin{array}{c|c} t_n & t_n \\ \hline -t_n & -t_n \end{array}\right], \tag{11}$$

де $t_i, i = 1,2,...,n$ – кодові слова розміру $\mu/2 \times \mu/2$, що мають елементарну структуру [23] перетворення Уолша – Адамара $\left\{(\mu/2)^2(1), 0((\mu/2)^2 - 1)\right\}$, тобто кожне з кодових слів $\left[\begin{array}{c|c} t_i & t_i \\ \hline -t_i & -t_i \end{array}\right]$ при адитивному вбудовуванні (10) здатне забезпечити вибірковий вплив на задану трансформанту перетворення Уолша – Адамара підблоків χ_j, $j = 1,2,...,4$.

Відповідно до [23], для того, щоб кодові слова t_i мали елементарну структуру $\left\{(\mu/2)^2(1), 0((\mu/2)^2 - 1)\right\}$, вони мають бути побудовані шляхом послідовного заповнення відповідної матриці t_i розміру $\mu/2 \times \mu/2$ елементами відповідного рядка матриці Уолша – Адамара $H_{(\mu/2)^2}$.

Твердження 1. Сумарний вплив на трансформанти перетворення Уолша – Адамара підблоків χ_j, $j = 1,2,...,4$ при адитивному вбудовуванні додаткової інформації (10) із застосуванням кодового слова T_{μ} буде визначатися як сума впливів від кожного кодового слова $t_i, i = 1,2,...,n$.

Доказ. Оскільки спосіб розбиття кодового слова T_{μ} на кодові слова $t_i, i = 1,2,...,n$ розміру $\mu/2 \times \mu/2$ відповідає способу розбиття

блоку контейнеру X на підблоки, то можемо представити блок стеганоповідомлення M розміру $\mu \times \mu$ у вигляді наступної конструкції:

$$M = \left[\begin{array}{c|c} m_1 & m_2 \\ \hline m_3 & m_4 \end{array}\right] = \left[\begin{array}{c|c} \chi_1 & \chi_2 \\ \hline \chi_3 & \chi_4 \end{array}\right] + \left[\begin{array}{c|c} t_1 & t_1 \\ \hline -t_1 & -t_1 \end{array}\right] + \left[\begin{array}{c|c} t_2 & t_2 \\ \hline -t_2 & -t_2 \end{array}\right] + \ldots + \left[\begin{array}{c|c} t_n & t_n \\ \hline -t_n & -t_n \end{array}\right], \quad (12)$$

де

$$\begin{cases} m_1 = \chi_1 + t_1 + t_2 + \ldots + t_n; \\ m_2 = \chi_2 + t_1 + t_2 + \ldots + t_n; \\ m_3 = \chi_3 - t_1 - t_2 - \ldots - t_n; \\ m_4 = \chi_4 - t_1 - t_2 - \ldots - t_n, \end{cases} \quad (13)$$

тоді враховуючі (4), трансформанти перетворення Уолша – Адамара підблоків m_1, m_2, m_3, m_4 будуть визначатися як:

$$\begin{aligned} \tilde{W}_{m_1} &= \tilde{W}_{\chi_1} + \tilde{W}_{t_1} + \tilde{W}_{t_2} + \ldots + \tilde{W}_{t_n}; \\ \tilde{W}_{m_2} &= \tilde{W}_{\chi_2} + \tilde{W}_{t_1} + \tilde{W}_{t_2} + \ldots + \tilde{W}_{t_n}; \\ \tilde{W}_{m_3} &= \tilde{W}_{\chi_3} - \tilde{W}_{t_1} - \tilde{W}_{t_2} - \ldots - \tilde{W}_{t_n}; \\ \tilde{W}_{m_4} &= \tilde{W}_{\chi_4} - \tilde{W}_{t_1} - \tilde{W}_{t_2} - \ldots - \tilde{W}_{t_n}, \end{aligned} \quad (14)$$

що доводить Твердження 1.

При цьому, якщо в якості матриць t_i розміру $\mu/2 \times \mu/2$ застосовуються кодові слова, що мають елементарну структуру перетворення Уолша – Адамара $\left\{(\mu/2)^2(1), 0((\mu/2)^2 - 1)\right\}$, результуючий вплив на підблоки χ_j, $j = 1,2,\ldots,4$ розміру $\mu/2 \times \mu/2$ буде складатися із впливів на трансформанти перетворення Уолша – Адамара, що забезпечують застосовані кодові слова t_i.

Твердження 2. Найбільше (за модулем) значення амплітуди впливу кодового слова T_μ на елемент блоку зображення X при застосуванні кодових слів, що засновані на рядках матриці Уолша – Адамара $H_{(\mu/2)^2}$ буде дорівнювати n.

Доказ. Оскільки за побудовою (2) перший стовбець матриці $H_{(\mu/2)^2}$ буде складатися з елементів +1, елементи всіх кодових

слів t_i з індексом (1,1), що побудовані на основі відповідного рядка матриці Уолша – Адамара будуть містити значення +1. Таким чином, у виразі (13), амплітуда максимального значення суми $t_1 + t_2 + \ldots + t_n$, або амплітуда мінімального значення різниці $-t_1 - t_2 - \ldots - t_n$ у позиції (1,1) дорівнюватиме n за модулем, тобто кількості застосованих кодових слів t_i.

Твердження 3. Структура кодового слова T_μ, яка обумовлена впливами кодових слів t_i на трансформанти підблоків $\tilde{W}_{\chi_1}, \tilde{W}_{\chi_2}, \tilde{W}_{\chi_3}, \tilde{W}_{\chi_4}$ блоку X може бути відновлена по блокам стеганоповідомлення m_1, m_2, m_3, m_4 за умови достатньо низького відхилення трансформант перетворення Уолша – Адамара, на які впливають кодові слова t_i підблоків χ_j, $j = 1,2,...,4$ від середнього арифметичного значення відповідних трансформант даних підблоків.

Доказ. Розглянемо вектор, що містить середні арифметичні значення трансформант перетворення Уолша – Адамара:

$$\begin{aligned}\tilde{W}_{\overline{m}} = \frac{1}{4}(\tilde{W}_{m_1} + \tilde{W}_{m_2} + \tilde{W}_{m_3} + \tilde{W}_{m_4}) = \frac{1}{4}(\tilde{W}_{\chi_1} + \tilde{W}_{t_1} + \tilde{W}_{t_2} + \ldots + \tilde{W}_{t_n} + \\ + \tilde{W}_{\chi_2} + \tilde{W}_{t_1} + \tilde{W}_{t_2} + \ldots + \tilde{W}_{t_n} + \tilde{W}_{\chi_3} - \tilde{W}_{t_1} - \tilde{W}_{t_2} - \ldots - \tilde{W}_{t_n} + \\ \tilde{W}_{\chi_4} - \tilde{W}_{t_1} - \tilde{W}_{t_2} - \ldots - \tilde{W}_{t_n}) = \frac{1}{4}(\tilde{W}_{\chi_1} + \tilde{W}_{\chi_2} + \tilde{W}_{\chi_3} + \tilde{W}_{\chi_4}),\end{aligned} \quad (15)$$

які визначаються виключно трансформантами перетворення Уолша – Адамара підблоків χ_j, $j = 1,2,...,4$.

Розглянемо матрицю різниць трансформант перетворення Уолша – Адамара підблоків стеганоповідомлення m_1, m_2, m_3, m_4 і середнього значення $\tilde{W}_{\overline{m}}$.

$$\left[\begin{array}{c|c} \Delta_1 & \Delta_2 \\ \hline \Delta_3 & \Delta_4 \end{array}\right] =$$

$$= \left[\begin{array}{c|c} \tilde{W}_{\chi_1} - \tilde{W}_{\overline{m}} + \tilde{W}_{t_1} + \tilde{W}_{t_2} + \ldots + \tilde{W}_{t_n} & \tilde{W}_{\chi_2} - \tilde{W}_{\overline{m}} + \tilde{W}_{t_1} + \tilde{W}_{t_2} + \ldots + \tilde{W}_{t_n} \\ \hline \tilde{W}_{\chi_3} - \tilde{W}_{\overline{m}} - \tilde{W}_{t_1} - \tilde{W}_{t_2} - \ldots - \tilde{W}_{t_n} & \tilde{W}_{\chi_4} - \tilde{W}_{\overline{m}} - \tilde{W}_{t_1} - \tilde{W}_{t_2} - \ldots - \tilde{W}_{t_n} \end{array}\right]. \quad (16)$$

Неважко побачити, що за виконання умов:

$$\begin{cases} \tilde{W}_{\chi_1} - \tilde{W}_{\overline{m}} \to 0; \\ \tilde{W}_{\chi_2} - \tilde{W}_{\overline{m}} \to 0; \\ \tilde{W}_{\chi_3} - \tilde{W}_{\overline{m}} \to 0; \\ \tilde{W}_{\chi_4} - \tilde{W}_{\overline{m}} \to 0, \end{cases} \quad (17)$$

де під 0 розуміється нульовий вектор порядку $(\mu/2)^2$, матриця $\left[\begin{array}{c|c} \Delta_1 & \Delta_2 \\ \hline \Delta_3 & \Delta_4 \end{array}\right]$ відтворюватиме структуру кодового слова T_μ, що обумовлює принципову можливість його декодування.

Відзначимо, що на практиці для контейнерів, у які відбувається вбудовування додаткової інформації, умова (17) не виконується, що потенційно веде до виникнення помилок при декодуванні кодового слова T_μ. Такі помилки будемо називати обумовленими варіацією підблоків.

Із застосуванням результатів Твердження 1 і Твердження 2 запишемо алгоритм вбудовування додаткової інформації:

Крок 1. Оригінальне зображення розбивається на блоки X_i розміру $\mu \times \mu$ стандартним способом.

Крок 2. У кожен із блоків відбувається вбудовування біта додаткової інформації d_i за допомогою кодового слова T_μ розміру $\mu \times \mu$, тобто i-й блок стеганоповідомлення визначається як

$$M_i = X_i + d_i T_\mu,\ T_\mu = \left[\begin{array}{c|c} t_1 & t_1 \\ \hline -t_1 & -t_1 \end{array}\right] + \left[\begin{array}{c|c} t_2 & t_2 \\ \hline -t_2 & -t_2 \end{array}\right] + \ldots + \left[\begin{array}{c|c} t_n & t_n \\ \hline -t_n & -t_n \end{array}\right], \quad (18)$$

де $t_i, i = 1,2,\ldots,n$ – кодові слова розміру $\mu/2 \times \mu/2$, що впливають на обрані трансформанти Уолша – Адамара.

Із застосуванням результатів Твердження 3 запишемо алгоритм вилучення додаткової інформації:

Крок 1. Стеганоповідомлення M' розбити на блоки M'_i розміру $\mu \times \mu$.

Крок 2. Для чергового блоку M'_i розміру $\mu \times \mu$ виконати його розбиття ще на 4 блоки розміру $\mu/2 \times \mu/2$ відповідно до наступної конструкції $M'_i = \left[\begin{array}{c|c} m_1 & m_2 \\ \hline m_3 & m_4 \end{array}\right]$.

Крок 3. Для кожного блоку M_i' розрахувати n матриць розміру 2×2:

$$u_i = \left[\begin{array}{c|c} \sum_{a=1}^{4}\sum_{b=1}^{4} m_1(a,b)t_i(a,b) & \sum_{a=1}^{4}\sum_{b=1}^{4} m_2(a,b)t_i(a,b) \\ \hline \sum_{a=1}^{4}\sum_{b=1}^{4} m_3(a,b)t_i(a,b) & \sum_{a=1}^{4}\sum_{b=1}^{4} m_4(a,b)t_i(a,b) \end{array} \right], i = 1,2,...,n, \quad (19)$$

де запис $m(a,b)$ означає елемент матриці m з індексом (a,b).

Крок 4. Знайти середні значення $\overline{u_i} = \sum_{l=1}^{2}\sum_{m=1}^{2} u_i(l,m), i = 1,2,...,n$.

Крок 5. Знайти значення вилученого біту додаткової інформації для даного блоку M_i' як

$$d_i' = \text{sign}\left(\sum_{i=1}^{n} \left((u_i(1,1) - \overline{u_i}) + (u_i(1,2) - \overline{u_i}) - (u_i(2,1) - \overline{u_i}) - (u_i(2,2) - \overline{u_i}) \right) \right). \quad (20)$$

Зважаючи на зміст Твердження 1 і Твердження 2, а також на практичні результати, які були отримані у [14], доцільним вбачається використання на практиці значення розміру блоку $\mu = 8$ і значення $n = 3$, які розглядаються далі в цій роботі. Тим не менш, питання оптимізації даних параметрів, на думку авторів, становлять окремий інтерес що виходить за рамки представлених у статті результатів.

У стеганографічному методі з кодовим управлінням вбудовуванням додаткової інформації викладеному в [17] в якості кодових слів обґрунтовано використання матричного представлення рядків матриці Уолша – Адамара порядку N^2 і показано, що тільки такі кодові слова забезпечують вибірковий вплив на ту, або іншу трансформанту перетворення Уолша – Адамара.

Зважаючи на те, що кодові слова t_1 і t_2 мають порядок $N = 4$, вони можуть бути побудовані на основі матриці Уолша – Адамара порядку $N^2 = 16$ шляхом представлення її рядків у вигляді матриць розміру 4×4. Покажемо зазначені кодові слова у табл. 1, де для кожного кодового слова зверху зазначено на яку трансформанту Уолша – Адамара воно впливає.

Таблиця 1 – Можливі варіанти кодових слів t_1 і t_2

(1,1)	(1,2)	(1,3)	(1,4)
$\xi_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	$\xi_2 = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$	$\xi_3 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$	$\xi_4 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$
(2,1)	(2,2)	(2,3)	(2,4)
$\xi_5 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$	$\xi_6 = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$	$\xi_7 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$	$\xi_8 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$
(3,1)	(3,2)	(3,3)	(3,4)
$\xi_9 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$	$\xi_{10} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$	$\xi_{11} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$	$\xi_{12} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$
(4,1)	(4,2)	(4,3)	(4,4)
$\xi_{13} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	$\xi_{14} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$	$\xi_{15} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$	$\xi_{16} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Для повноти викладу матеріалу розглянемо конкретний приклад. Нехай визначено кодове слово ξ_3, що здійснює зосереджений вплив на трансформанту Уолша – Адамара (1,3). Знайдемо для нього трансформанти перетворення Уолша – Адамара (2) (з точністю до нормувального коефіцієнта $1/N$):

$$W_{\xi_3} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 16 & 0 \\ 0 & \mathbf{0} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (21)$$

Як бачимо з (21), кодове слово ξ_3 дійсно має матрицю трансформант перетворення Уолша – Адамара, що характеризується лише однією ненульовою трансформантою перетворення Уолша – Адамара.

Проведені дослідження дозволили встановити, що при виборі того чи іншого набору кодових слів $\xi_i, \xi_j, i, j = 1,2,...,16$ в якості t_1 і t_2, на результуючу стійкість стеганографічного методу будуть впливати два основних фактори, які розглянемо докладно.

1. Помилки обумовлені варіацією підблоків.

Згідно до алгоритму вилучення додаткової інформації у досліджуваному стеганографічному методі, стеганоповідомлення сегментується на блоки $\mu \times \mu$ після чого кожний отриманий блок додатково сегментується на підблоки $\mu/2 \times \mu/2$, після чого визначення значень елементів вбудованого кодового слова відбувається відповідно до «нульової точки», в якості якої слугують середні значення $\overline{u_i}, i = 1,2,...,n$.

Зважаючи на те, що використовувані у стеганографічному методі кодові слова вибірково впливають на задану трансформанту перетворення Уолша – Адамара, середні значення $\overline{u_i}, i = 1,2,...,n$ по суті являють собою середні значення трансформант Уолша – Адамара, на які впливають кодові слова $t_i, i = 1,2,...,n$. На практиці можливою є ситуація, коли розкид значень даних трансформант Уолша – Адамара у зазначених підблоках розміру $\mu \times \mu$ блоку $\mu/2 \times \mu/2$ перевищує сумарну амплітуду впливу на дану трансформанту кодового слова, що веде до неможливості вилучення додаткової інформації із даного блоку.

Розглянемо конкретний приклад такої ситуації. Нехай задано вихідний блок зображення розміру 8×8:

$$X = \begin{bmatrix} 150 & 143 & 146 & 145 & 142 & 142 & 145 & 144 \\ 129 & 126 & 129 & 132 & 134 & 135 & 132 & 131 \\ 119 & 118 & 120 & 126 & 129 & 129 & 129 & 128 \\ 117 & 121 & 123 & 126 & 129 & 136 & 137 & 132 \\ 121 & 128 & 130 & 132 & 139 & 137 & 139 & 139 \\ 124 & 128 & 133 & 140 & 143 & 137 & 138 & 139 \\ 128 & 128 & 137 & 141 & 138 & 142 & 140 & 138 \\ 129 & 132 & 139 & 138 & 137 & 145 & 142 & 133 \end{bmatrix}, \quad (22)$$

у який необхідно здійснити вбудовування біта додаткової інформації $d = 1$ із застосуванням $n = 2$ однакових кодових слів:

$$t_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}, t_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}, \quad (23)$$

тобто для вбудовування інформації буде застосовуватися кодове слово:

$$(-1)T_8 = \begin{bmatrix} -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 \\ -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 \\ -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 \end{bmatrix}, \quad (24)$$

тоді як стеганоповідомлення матиме вигляд:

$$M = X + (-1)T_8 = \begin{bmatrix} 148 & 141 & 144 & 143 & 140 & 140 & 143 & 142 \\ 127 & 124 & 127 & 130 & 132 & 133 & 130 & 129 \\ 121 & 120 & 122 & 128 & 131 & 131 & 131 & 130 \\ 119 & 123 & 125 & 128 & 131 & 138 & 139 & 134 \\ 123 & 130 & 132 & 134 & 141 & 139 & 141 & 141 \\ 126 & 130 & 135 & 142 & 145 & 139 & 140 & 141 \\ 126 & 126 & 135 & 139 & 136 & 140 & 138 & 136 \\ 127 & 130 & 137 & 136 & 135 & 143 & 140 & 131 \end{bmatrix}. \quad (25)$$

Спробуємо вилучити додаткову інформацію з блоку (25), знаходимо, що

$$u_1 = u_2 = \begin{bmatrix} 98 & 24 \\ -4 & 28 \end{bmatrix}, \quad (26)$$

тоді як середні значення $\overline{u_1} = 36{,}5$ і $\overline{u_2} = 36{,}5$. Таким чином, відповідно до (20) маємо

$$\begin{aligned} d_i' = \operatorname{sign}((98-36{,}5)+(24-36{,}5)-(-4-36{,}5)-(28-36{,}5)+ \\ +(98-36{,}5)+(24-36{,}5)-(-4-36{,}5)-(28-36{,}5))= \\ = \operatorname{sign}(61{,}5-12{,}5+40{,}5+8{,}5+ \\ +61{,}5-12{,}5+40{,}5+8{,}5) = 196. \end{aligned} \quad (27)$$

Як бачимо, суттєва відмінність у значеннях елементів матриць u_1 і u_2 від середніх значень $\overline{u_1}$ і $\overline{u_2}$, яка обумовлена особливостями структури блоку зображення, та перевищує вплив амплітуди кожного з застосованих кодових слів $\left[\begin{array}{c|c} t_i & t_i \\ \hline -t_i & -t_i \end{array}\right]$, призвела до того, що значення результуючої суми під знаком sign стало додатним, що призвело до виникнення помилки при вилученні додаткової інформації.

Таким чином, ризик виникнення помилки існує у тих блоків зображення розміру 8×8, у яких відмінність значень застосованих для вбудовування додаткової інформації трансформант Уолша – Адамара

від середнього значення серед підблоків розміру 4×4 перевищує значення амплітуди впливу кожного з кодових слів $\left[\begin{array}{c|c} t_i & t_i \\ \hline -t_i & -t_i \end{array}\right]$.

Іншими словами, для уникнення ризику помилки при вбудовуванні інформації, середньоквадратичне відхилення δ трансформант Уолша – Адамара, у які відбувається вбудовування інформації, має бути меншим від значення амплітуди впливу кожного з кодових слів $\left[\begin{array}{c|c} t_i & t_i \\ \hline -t_i & -t_i \end{array}\right]$, застосованих для вбудовування інформації.

Для нашого прикладу, вбудовування інформації відбувається із застосуванням кодових слів (23) у трансформанту (3,1). В нашому випадку для блоку (22) маємо наступні значення трансформант (3,1) Уолша – Адамара $[130,56,-36,-4]$, а отже значення середньоквадратичного відхилення складає $\sigma=73{,}073$, що значно перевищує значення 16. Таким чином, блок (22) знаходиться у групі ризику щодо можливого виникнення помилок, пов'язаних із структурними особливостями зображення.

Проведені дослідження дозволили дійти висновку, що для реальних зображень різні трансформанти перетворення Уолша – Адамара характеризуються різними очікуваними значеннями σ. Для оцінки значень σ для різних трансформант Уолша – Адамара був проведений наступний експеримент. Для 500 зображень з бази NRCS [25] у форматі JPEG була виконана сегментація на блоки X_i розміру 4×4, для кожного з яких було знайдено двовимірне перетворення Уолша – Адамара відповідно до (2).

Далі, серед отриманих блоків, була знайдена кількість таких, для яких середньоквадратичне відхилення відповідних трансформант Уолша – Адамара підблоків 4×4 перевищує значення 16, після чого була побудована відповідна стовпчикова діаграма, що показана на рис. 6.

Аналіз даних, представлених на рис. 6 дозволяє дійти висновку про те, що найменшими значеннями кількості блоків із $\sigma\geq16$ володіють трансформанти перетворення Уолша – Адамара із номерами (2,2), (2,4), (4,2), (4,4).

Таким чином, застосування саме даних трансформант Уолша – Адамара для вбудовування додаткової інформації за допомогою стеганографічного методу з кодовим управлінням із сліпим декодуванням дозволить мінімізувати кількість помилок, що обумовлені варіацією підблоків.

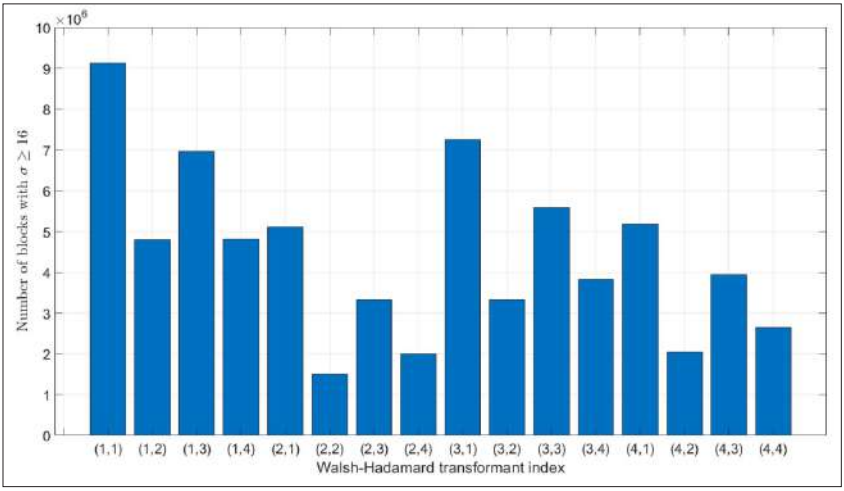


Рисунок 6 – Стовпчикова діаграма, що ілюструє кількості блоків із значенням $\sigma \geq 16$

2. Помилки, що обумовлені атаками проти вбудованого повідомлення. На сьогодні, відомо чимало можливих атак проти вбудованого повідомлення, однак, найбільш розповсюдженою є атака стисненням. Це пояснюється тим, що на практиці зображення нечасто передаються по телекомунікаційним системам або зберігаються у нестисненому вигляді, найчастіше до них застосовуються алгоритми стиску із втратами, найбільш поширеним серед яких є алгоритм JPEG.

Незважаючи на те, що такі алгоритми стиску із втратами, побудовані таким чином, щоб зробити спотворення, що вносяться у зображення найменш помітними людському оку, їх впливу достатньою для знищення додаткової інформації, що вбудована із застосуванням більшості стеганографічних методів, окрім тих, які розроблювалися як стійкі до атак проти вбудованого повідомлення.

Відзначимо, що у відповідності до структури формування кодового слова згідно до (11), отримуємо результуючі кодові слова T_8 при $n=2$, що впливають на відповідні трансформанти перетворення Уолша – Адамара блоку розміру 8×8, при цьому конкретні трансформанти перетворення Уолша – Адамара, на які здійснюють вплив ті чи інші комбінації кодових слів $\xi_i, \xi_i, i = 1,2,...,16$, показані у табл. 2.

Таблиця 2 – Трансформанти Уолша – Адамара, на які здійснюють вплив кодові слова T_8 на основі кодових слів t_1 і t_2

Кодове слово	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8
Трансформанта	(5,1)	(5,2)	(5,3)	(5,4)	(6,1)	(6,2)	(6,3)	(6,4)
Кодове слово	ξ_9	ξ_{10}	ξ_{11}	ξ_{12}	ξ_{13}	ξ_{14}	ξ_{15}	ξ_{16}
Трансформанта	(7,1)	(7,2)	(7,3)	(7,4)	(8,1)	(8,2)	(8,3)	(8,4)

Розглянемо дані, представлені у табл. 2 на конкретному прикладі. Сформуємо із застосуванням формули (11) кодове слово T_8 на основі ξ_3:

$$T_8 = \begin{bmatrix} \xi_3 & \xi_3 \\ -\xi_3 & -\xi_3 \end{bmatrix} + \begin{bmatrix} \xi_3 & \xi_3 \\ -\xi_3 & -\xi_3 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \end{bmatrix}. \quad (28)$$

Знайдемо двовимірне перетворення Уолша – Адамара матриці (28) із застосуванням виразу (2):

$$W_{T_8} = H_8 T_8 H_8^T =$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \\ -2 & -2 & 2 & 2 & -2 & -2 & 2 & 2 \end{bmatrix} \times$$

$$\times\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 128 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (29)$$

Як бачимо з виразу (29), кодове слово T_8, яке побудоване на основі ξ_3, що впливають на трансформанту (3,1) дійсно забезпечує вплив на трансформанту (5,3), як це показано у табл. 2.

При цьому необхідно зазначити, що одночасне застосування різних кодових слів t_1 і t_2 веде до вбудовування додаткової інформації у відповідну комбінацію трансформант Уолша – Адамара. Наприклад, в разі застосування в якості кодових слів t_1 і t_2 матриць ξ_1 і ξ_9, результуючі кодові слова T_8 будуть виконувати вбудовування додаткової інформації у частотні складові (5,1) і (7,1).

Таким чином, з аналізу даних табл. 2, а також враховуючи достатню умову забезпечення нечутливості стеганоповідомлення до збурних дій можемо прийти до висновку, що найкращими кодовими словами з точки зору стійкості стеганоповідомлення до атаки стиском є, насамперед, кодові слова ξ_1 і ξ_9.

Тим не менш, відповідно до даних табл. 1, кодові слова ξ_1 і ξ_9 впливають на трансформанти Уолша – Адамара (1,1) і (3,1), відповідно, які, згідно до даних рис. 6 мають найбільшу кількість блоків із значенням $\sigma \geq 16$, що ведуть до помилок, обумовлених варіацією підблоків.

І, навпаки, застосування в якості t_1 і t_2 кодових слів ξ_6, ξ_8, ξ_{14} і ξ_{16}, які впливають на трансформанти, що характеризуються найнижчою кількістю блоків із значенням $\sigma \geq 16$ веде до таких результуючих кодових слів T_8, що забезпечують вбудовування додаткової інформації у досить високочастотні трансформанти Уолша – Адамара контейнера (6,2), (6,4), (8,2) і (8,4), відповідно, що не дозволяє забезпечити нечутливість стеганоповідомлення до збурних дій.

Задля практичного підтвердження викладених міркувань, проведемо ряд експериментів.

Для дослідження роботи стеганографічного методу з кодовим управлінням вбудовуванням додаткової інформації із сліпим декодуванням під впливом атаки стисненням проти вбудованого повідомлення вбудовування додаткової інформації у 150 випадково обраних зображень з бази NRCS [25] здійснювалося із застосуванням кодових слів T_8^+ і T_8^-, що побудовані на основі двох однакових кодових слів $\xi_i, \xi_i, i = 1,2,...,16$, які використовуються в якості t_1 і t_2. Після вбудовування додаткової інформації зображення піддавалися стисненню алгоритмом JPEG із значеннями $QF = \{100,90,80,70,60,50,40,30,20,10\}$. Далі з атакованого стисненням стеганоповідомлення відбувалося вилучення додаткової інформації із вимірюванням кількості помилок, що відбулися. Результати щодо кількості помилок, що відбулися, для кожного з кодових слів $\xi_i, \xi_i, i = 1,2,...,16$, а також для кожного значення *QF* наведені у табл. 3.

Таблиця 3 – Результати щодо кількості помилок при вилученні додаткової інформації в умовах атаки стисненням алгоритмом JPEG

ξ_i, ξ_j / *QF*	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}	ξ_{11}	ξ_{12}	ξ_{13}	ξ_{14}	ξ_{15}	ξ_{16}
100	9,5	0,5	2,6	1,1	2,5	0	0,3	0,1	5,9	0,2	1,7	0,8	2,5	0	0,4	0,2
90	9,6	0,9	2,8	1,2	2,5	0,4	0,8	0,5	5,7	0,4	2,0	1,1	2,5	0,5	0,9	0,6
80	9,8	1,1	3,6	1,6	2,3	24,2	9,4	20,2	6,3	0,8	2,7	2,0	3,0	22,8	4,6	17,5
70	9,8	17,1	4,0	2,5	3,9	27,7	19,3	29,5	6,8	18,9	3,4	3,2	3,1	27,3	16,8	29,6
60	9,9	24,1	4,7	3,3	13,9	29,9	24,7	31,8	7,1	24,7	4,1	10,0	3,6	30,1	25,3	32,0
50	10,3	27,0	5,5	4,9	19,3	31,2	28,3	33,3	7,4	28,0	7,3	29,1	4,6	31,6	28,4	33,4
40	11,2	30,0	9,7	15,4	25,2	33,2	31,1	34,9	8,3	30,8	19,7	33,4	10,7	33,3	31,3	34,7
30	12,6	33,0	24,9	33,0	29,4	35,1	33,9	36,3	9,7	33,9	29,3	35,7	23,0	35,2	34,0	36,4
20	18,7	36,1	33,5	37,9	34,1	37,4	37,0	38,5	21,5	36,8	35,3	38,4	31,5	37,7	37,0	38,5
10	36,0	41,3	40,8	42,5	39,3	41,8	42,1	42,3	37,7	41,6	41,7	42,4	38,7	42,0	41,9	42,5

Аналіз даних, представлених у табл. 3 підтверджує викладені раніше міркування. Так, бачимо, що кодові слова, засновані на матрицях ξ_6, ξ_8 та ξ_{14} показують найкращі результати щодо кількості помилок при вилученні додаткової інформації при значеннях $QF = \{100,90\}$, тобто фактично за умови відсутності атаки проти вбудованої інформації, що пояснюється невеликою кількістю блоків із значенням $\sigma \geq 16$.

З іншого боку, застосування кодових слів ξ_1 та ξ_9, що забезпечують вбудовування додаткової інформації у найбільш низькочастотні для досліджуваного стеганографічного методу компоненти, як бачимо з результатів представлених у табл. 3, найменше збільшують кількість помилок при вилученні додаткової інформації із зменшенням значення *QF*. Тим не менш, велика кількість помилок при декодуванні, що обумовлена варіацією підблоків, не дозволяє рекомендувати їх до практичного застосування. Таким чином, з практичної точки зору, для забезпечення роботи стеганографічного методу з кодовим управлінням вбудовуванням додаткової інформації із сліпим декодуванням необхідним є компроміс між мінімізацією кількості помилок, що обумовлені варіацією підблоків (для чого вбудовування додаткової інформації має відбуватися у високочастотні складові) та максимізацією стійкості до збурних дій (для чого вбудовування додаткової інформації має відбуватися у низькочастотні складові блоку). Завдання пошуку такого компромісу вирішується в рамках наступних експериментів.

Наведемо результати дослідження кодового слова, яке дає найкращу стійкість до атаки стисненням, для чого вбудовування додаткової інформації у 150 випадковим чином обраних зображень у форматі JPEG з бази NRCS [25] здійснювалося із застосуванням кодових слів T_8. Кодові слова T_8 складаються з кодових слів t_1, t_2 відповідно до формули $T_8 = \left[\begin{array}{c|c} t_1 & t_1 \\ \hline -t_1 & -t_1 \end{array}\right] + \left[\begin{array}{c|c} t_2 & t_2 \\ \hline -t_2 & -t_2 \end{array}\right]$, де кодові слова t_1, t_2 розміром $\mu/2 \times \mu/2$ впливають на дві обрані трансформанти Уолша – Адамара (табл. 1). У табл. 4 показані відсотки помилок усіх комбінацій кодових слів t_1 та t_2 при $QF = 70$.

Таблиця 4 – Відсотки помилок усіх комбінацій кодових слів t_1 та t_2

$\xi_i \backslash \xi_j$	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}	ξ_{11}	ξ_{12}	ξ_{13}	ξ_{14}	ξ_{15}	ξ_{16}
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ξ_1	9,8	13,9	11,0	10,6	13,2	14,2	13,9	14,1	10,7	13,9	11,8	13,3	10,3	14,2	13,8	14,2
ξ_2	13,8	17,1	7,4	11,8	19,6	25,4	23,2	25,2	10,9	23,5	13,0	23,2	9,0	26,4	25,0	25,4
ξ_3	11,1	7,4	4,1	6,2	9,4	10,1	9,6	10,0	8,2	9,7	6,9	8,9	6,2	10,9	11,6	10,1
ξ_4	10,6	11,8	6,2	2,5	11,0	13,2	11,7	13,9	8,3	11,8	7,4	12,7	6,2	13,2	11,8	18,8
ξ_5	13,2	19,6	9,4	11,1	3,9	21,6	20,1	21,9	9,0	20,2	12,3	19,6	8,4	21,7	20,1	21,9
ξ_6	14,2	25,4	10,1	13,2	21,6	27,7	24,6	29,5	11,6	24,7	10,5	27,0	9,4	29,2	25,7	29,8

Продовження таблиці 4

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ξ_7	13,9	23,2	9,6	11,7	20,1	24,6	19,3	26,0	11,1	21,1	5,8	23,8	8,8	25,9	23,5	26,3
ξ_8	14,1	25,1	10,0	13,9	21,9	29,5	26,1	29,5	11,7	26,0	14,2	15,4	9,3	29,4	25,9	30,9
ξ_9	10,7	10,9	8,2	8,3	9,0	11,6	11,1	11,7	6,8	11,1	8,7	10,5	7,7	11,6	10,9	11,7
ξ_{10}	13,9	23,5	9,7	11,9	20,2	24,6	21,1	26,0	11,1	18,9	7,1	23,9	8,9	26,1	23,4	26,4
ξ_{11}	11,7	13,0	6,9	7,4	12,3	10,5	5,9	14,2	8,7	7,0	3,4	12,9	6,4	14,2	13,1	14,3
ξ_{12}	13,3	23,2	8,8	12,7	19,6	27,0	23,8	15,5	10,5	23,9	12,9	3,2	8,3	26,8	23,8	28,3
ξ_{13}	10,3	9,0	6,2	6,2	8,4	9,4	8,8	9,3	7,7	8,9	6,4	8,3	3,1	9,3	8,6	9,3
ξ_{14}	14,2	26,4	10,8	13,3	21,7	29,2	25,9	29,4	11,6	26,2	14,1	26,9	9,3	27,3	23,8	29,6
ξ_{15}	13,7	25,0	11,6	11,8	20,1	25,7	23,5	25,9	10,9	23,4	13,1	23,8	8,7	23,8	16,8	26,2
ξ_{16}	14,2	25,4	10,1	18,8	21,9	29,8	26,3	30,9	11,7	26,4	14,3	28,2	9,3	29,6	26,2	29,6

Для оцінки рівня стійкості стеганографічного методу з кодовим управлінням вбудовуванням додаткової інформації та сліпим декодуванням з бази даних NRCS [25] було випадковим чином обрано 150 зображень у форматі JPEG та проведено подальші експерименти з вилучення додаткової інформації під впливом атаки стисненням за алгоритмом JPEG:

– оригінальний стеганографічний метод з кодовим управлінням вбудовуванням додаткової інформації за допомогою кодового слова розміру $\mu = 8$, яке впливає на трансформанту Уолша – Адамара (5,4);

– оригінальний стеганографічний метод з кодовим управлінням вбудовуванням додаткової інформації з кодовим словом розміру $\mu = 8$, яке впливає на трансформанту (5,1);

– стеганографічний метод з кодовим управлінням вбудовуванням додаткової інформації зі сліпим декодуванням на основі кодових слів ξ_9, ξ_6;

– стеганографічний метод з кодовим управлінням вбудовуванням додаткової інформації за допомогою кодових слів, для вибору яких запропоновано обґрунтування: ξ_4, ξ_4, які впливають на трансформанту (5,4), а також на основі кодових слів ξ_7, ξ_{11}, які впливають на трансформанти (6,3) та (7,3).

На рис. 7 наведено графіки залежності кількості помилок, що виникають від степеню стиснення зображення алгоритмом JPEG для всіх розглянутих вище кодових слів, що дозволяє оцінити ефективність їх використання в протистоянні атакам стисненням проти вбудованого повідомлення.

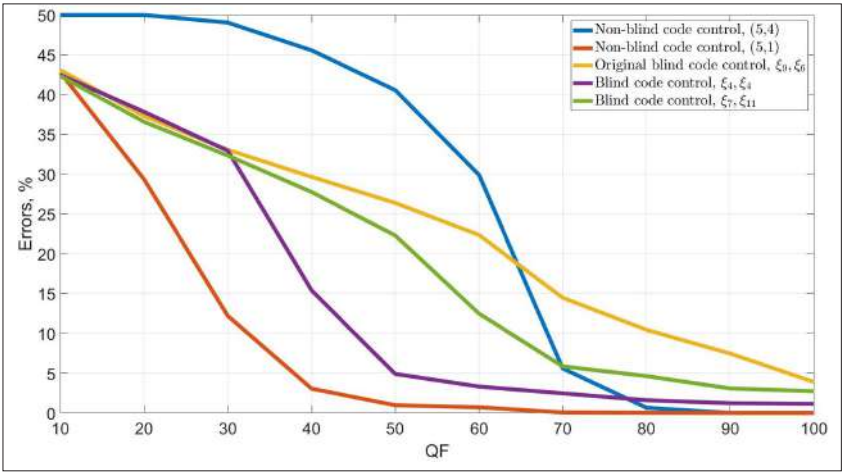


Рисунок 7 – Графік залежності кількості помилок при вилученні додаткової інформації від значення *QF* для стеганографічного методу з кодовим управлінням у різних його варіаціях

Аналіз даних, представлених на рис. 7, показує, що використання кодових слів ξ_4, ξ_4 та ξ_7, ξ_{11} у стеганографічному методі з кодовим управлінням вбудовуванням додаткової інформації та сліпим декодуванням під впливом атаки стисненням JPEG дозволяє зменшити кількість помилок у декодованих значеннях під впливом атаки стисненням JPEG у порівнянні з версією цього методу [26], яка базується на кодових словах ξ_9, ξ_6.

Так, згідно з даними, представленими на рис. 7, при використанні кодових слів ξ_4, ξ_4 у стеганографічному методі з кодовим управлінням вбудовуванням додаткової інформації та сліпим декодуванням під впливом атаки стисненням спостерігається лише на 2,4 % більше помилок у порівнянні з оригінальним стеганографічним методом з кодовим управлінням без сліпого декодування [17] на основі одного з найкращих кодових слів розміру $\mu = 8$, що впливає на трансформанту Уолша – Адамара (5,1).

Стеганографічний метод з кодовим управлінням із сліпим декодуванням характеризується тими самими перевагами, що і класичний варіант стеганографічного методу з кодовим управлінням

вбудовуванням додаткової інформації [17], зокрема, забезпеченням високої надійності сприйняття стеганоповідомлення. Задля оцінки надійності сприйняття стеганоповідомлення загальноприйнятним є показник PSNR (хоча він і не є таким, що вичерпно оцінює надійність сприйняття), який визначається наступним чином:

$$PSNR = 20\lg\left(\frac{255}{\sqrt{MSE}}\right), \tag{30}$$

де показник MSE визначається як

$$MSE = \frac{1}{nm}\sum_i\sum_j \left|X(i,j) - M(i,j)\right|^2. \tag{31}$$

де X – матриця оригінального зображення, M – матриця стеганоповідомлення.

Зважаючи на структуру стеганографічного методу з кодовим управлінням вбудовуванням додаткової інформації показник PSNR є стабільним, не залежить від розміру контейнера, залежить лише від застосованого набору кодових слів. Так, встановлено, що у разі застосування в якості t_1 і t_2 однакових кодових слів $\xi_i, \xi_i, i = 1,2,...,16$ всі елементи результуючих кодових слів T_8 є рівними за амплітудою, яка дорівнює 2. При цьому, значення PSNR стеганоповідомлення дорівнює 36,1315 дБ.

У разі, якщо в якості t_1 і t_2 обрано різні кодові слова $\xi_i, \xi_j, i, j = 1,2,...,16, i \neq j$, елементи результуючих кодових слів T_8 є різними за амплітудою, і при цьому максимальна амплітуда даних елементів не перевищує значення 2. При цьому PSNR результуючого стеганоповідомлення дорівнює 42,1524 дБ.

У будь-якому випадку, при застосуванні стеганографічного методу з кодовим управлінням із сліпим декодуванням, вплив на інтенсивність пікселів контейнеру є незначним і суворо контрольованим, що дозволяє домогтися як високих показників PSNR, так і фактично виключити виникнення будь-яких артефактів у результуючих стеганоповідомленнях. Наприклад, на рис. 8 показаний приклад оригінального зображення і стеганоповідомлення, що отримано із застосуванням стеганографічного методу з кодовим управлінням вбудовуванням додаткової інформації із сліпим декодуванням

із застосуванням в якості t_1 і t_2 кодових слів ξ_4, ξ_4. При цьому розмір контейнера у форматі JPEG, в який відбувалося вбудовування додаткової інформації склав 1200×1200 пікселів, вбудовування відбувалося у просторі YCbCr у компоненту Y, таким чином обсяг додаткової інформації склав 22 500 біт.

а б

Рисунок 8 – Приклад оригінального зображення (а) і стеганоповідомлення (б), що отримано із застосуванням стеганографічного методу з кодовим управлінням вбудовуванням додаткової інформації із сліпим декодуванням

Аналіз даних рис. 8 шляхом суб'єктивного ранжування дозволяє дійти висновку про відсутність видимих спотворень або артефактів на стеганоповідомленні.

1.5. Стеганографічний метод з кодовим управлінням на основі досконалих двійкових решіток

Подальший розвиток концепції стеганографії з кодовим управлінням пов'язаний із необхідністю підвищення пропускної спроможності методів без істотної втрати непомітності та стійкості. У цьому контексті особливий інтерес викликає підхід, який базується на використанні досконалих двійкових решіток – спеціальних

кодових структур, що дозволяють ефективно представляти множину кодових слів із заданими структурними властивостями. На відміну від класичних схем, даний метод орієнтований на вбудовування декількох бітів у межах одного блоку, що істотно розширює пропускну спроможність. При цьому структура досконалих двійкових решіток забезпечує контроль над відстанями між кодовими словами, що сприяє збереженню стійкості до атак і зниженню ймовірності помилкового декодування.

Введемо необхідні визначення.

Визначення 1 [27]. Досконала двійкова решітка – це двовимірна послідовність (матриця):

$$H(N) = \left\| h_{i,j} \right\|, \quad i, j = 0,1,...,N-1, \\ h_{i,j} \in \{-1,1\}, \tag{32}$$

що має ідеальну двовимірну періодичну автокореляційну функцію (2ПАКФ), елементи якої:

$$R(m,n) = PACF\left(m,n\right) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} h_{i,j} h_{i+m,j+n} = \\ = \begin{cases} N^2, \text{ для } m = n = 0; \\ 0, \text{ для будь-яких інших } m \text{ і } n, \end{cases} \tag{33}$$

де $m,n = 0,1,...,N-1$, а всі індекси елементів $h_{i+m,j+\tau}$ зведені за модулем N. Наведемо як приклад досконалу двійкову решітку порядку $N = 4$, а також її двовимірну періодичну автокореляційну функцію:

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix},$$

$$R = \begin{bmatrix} 64 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (34)$$

Твердження 4 [24]. Кожна досконала двійкова решітка порядку N генерує E(N)-клас еквівалентних матриць – досконалих двійкових решіток за допомогою операцій циклічного зсуву по рядкам і стовпцям та інверсій, при цьому потужність класу еквівалентних матриць дорівнює:

$$J_{E(N)} = 2N^2. \quad (35)$$

З дослідження [24] відомо наступне. Якщо задано довільну породжуючу досконалу двійкову решітку $A_0(N)$ порядку N, то всі її циклічні зсуви визначаються як $L_{k_1} A Q_{k_2}$, $k_1, k_2 = 0,1,...,N-1$ і нехай двовимірна періодична взаємнокореляційна функція (2ПВКФ) між $A_0(N)$ та її циклічно зсунутою решіткою визначається співвідношенням:

$$B(m,n) = A_0(N) ** L_{k_1} A_0(N) Q_{k_2} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} h_{i,j} h_{i+m,j+\tau}, \quad (36)$$

де символ ** позначає двовимірну кореляцію (згортку); $m, n = 0,1,...,N-1$.

Відоме наступне твердження:

Твердження 5 [24]. 2ПВКФ $B(m,n)$ решітки $A_0(N)$ та решітки $L_{k_1} A_0(N) Q_{k_2}$ порядку N, $k_1, k_2 = 0,1,...,N-1$, є 2ПАКФ $\|R(m,n)\|$ решітки $A_0(N)$, зсунута на k_1 рядків та k_2 стовпців, тобто

$$B(m,n) = R(m+k_1, n+k_2), \quad m,n = 0,1,...,N-1. \quad (37)$$

Запропонований стеганографічний метод базується на принципі модуляції інформації, який використовує ключову властивість досконалих двійкових решіток: 2ПВКФ між досконалою двійковою решіткою та її циклічно зсунутою версією (по рядкам і стовбцям) структурно еквівалентна відповідному циклічному зсуву 2ПАКФ вихідної решітки. Ця алгебраїчна симетрія дозволяє точно та передбачувано маніпулювати піками кореляції. Ми пропонуємо кодувати інформацію, вибираючи один з можливих станів, за допомогою контрольованих циклічних зсувів (або по рядкам, або по стовпцям) у поєднанні з додатковою двійковою інверсією. Таким чином, процес вбудовування відповідає певному перетворенню структури масиву в блоці зображення. Під час вилучення метод спирається на несліпе декодування, типове для схем з кодовим управлінням: реконструюється вбудований компонент, обчислюється 2ПВКФ, і положення її глобального максимуму однозначно визначає біти вбудованого повідомлення.

На відміну від традиційних підходів на основі застосування областей перетворень, запропонований метод працює повністю в просторовій області, використовуючи досконалі двійкові решітки як кодові носії інформації для вбудовування додаткової інформації, що дозволяє здійснювати кодове управління вбудовуванням у частотні складові. Ключова ідея полягає у використанні унікальних характеристик автокореляції та взаємної кореляції досконалих двійкових решіток, зокрема їхньої поведінки при циклічних зсувах та інверсії. Ці алгебраїчні симетрії забезпечують надійне та високопродуктивне кодування даних, зберігаючи при цьому низьку обчислювальну складність та сумісність із ресурсообмеженими середовищами. Далі ми описуємо процедури вбудовування та вилучення, а також схему модуляції, що використовується для представлення інформації за допомогою структурних перетворень досконалих двійкових решіток.

Наведемо основні кроки запропонованого стеганографічного методу на основі [17].

Вбудовування додаткової інформації.

Крок 1. Виконайте стандартне розбиття зображення вихідного контейнера на неперекриваючі блоки розміру $N \times N$.

Крок 2. Виберіть опорну досконалу двійкову решітку $A_0(N)$ розміру N.

Крок 3. Нехай X – наступний блок контейнера, що бере участь у стеганографічному перетворенні. Виберіть вектор $D = \{d_1 \quad d_2 \quad ... \quad d_{\log_2(2N^2)}\}$, який містить наступні $\log_2(2N^2)$ бітів інформації, які будуть вбудовані в цей блок контейнера.

Крок 4. Визначте значення біта d_1 як знак кодування досконалою двійкової решітки, десятковий еквівалент бітів $k_1 = \{d_2 \quad d_3 \quad ... \quad d_{\log_2(N)+1}\}_{10}$ як значення зсуву по рядкам, а десяткових еквівалент бітів $k_2 = \{d_{\log_2(N)+2} \quad d_{\log_2(N)+3} \quad ... \quad d_{\log_2(2N^2)}\}_{10}$ як зсув по стовпцям.

Крок 5. Побудуйте решітку $L_{k_1} A_0(N) Q_{k_2}$ з $E(N)$-класу для вбудовування додаткової інформації та виконайте вбудовування, тоді наступний блок стеганографічного повідомлення буде визначено як

$$M = X + L_{k_1} A_0(N) Q_{k_2}. \tag{38}$$

Зауважте, що при вбудовуванні значення +1 у значення пікселя контейнера 255, а також при вбудовуванні значення −1 у значення пікселя контейнера 0, операція вбудовування для цих пікселів не виконується.

Вилучення додаткової інформації.

Крок 1. Виконайте стандартне розбиття стеганографічного повідомлення на блоки, що не перекриваються розміру $N \times N$.

Крок 2. Нехай $\overline{M}$ – наступний блок можливо збуреного стеганографічного повідомлення, що відповідає блоку X контейнера.

2.1. Побудуйте матрицю $\Delta = \overline{M} - X$ з елементами $\Delta(i, j)$, $i, j = 0,1,...,\mu - 1$, для якого потрібно побудувати матрицю 2ПВКФ:

$$B(m,n) = A_0(N) ** \Delta. \tag{39}$$

2.2. Знайдіть індекси x по рядкам і y по стовпцям максимального (по модулю) значення матриці $B(m,n)$.

2.3. Відновіть вбудовані $\log_2(2N^2)$ біт інформації як

$$\begin{cases} \tilde{d}_1 = sign(B(x,y)) - 1 (\mathrm{mod}\, 3); \\ \{\tilde{d}_2 \quad \tilde{d}_3 \quad ... \quad \tilde{d}_{\log_2(N)+1}\}_{10} = x; \\ \{\tilde{d}_{\log_2(N)+2} \quad \tilde{d}_{\log_2(N)+3} \quad ... \quad \tilde{d}_{\log_2(2N^2)}\}_{10} = y. \end{cases} \tag{40}$$

Ефективність методу оцінюється відповідно до ключових стеганографічних показників, включаючи пропускну спроможність, надійність сприйняття та стійкість до атак проти вбудованого повідомлення. Щоб оцінити ефективність запропонованого методу з точки зору корисного навантаження даних, ми порівнюємо його пропускну спроможність з пропускною спроможністю класичного стеганографічного методу з кодовим управлінням, на основі функцій Уолша – Адамара. Порівняння проводиться для різних розмірів блоків. Тоді як класичний метод зазвичай кодує один біт на блок незалежно від розміру блоку, запропонована схема використовує структурні властивості досконалих двійкових решіток для кодування до $\log_2(2N^2)$ бітів на блок. У табл. 5 підсумовано отримані пропускні спроможності з точки зору кількості бітів на блок та кількості бітів на піксель.

Таблиця 5 – Пропускна спроможність стеганографічних методів

Розмір блоку *N*	**Класичний метод з кодовим управлінням**		**Метод на основі досконалих двійкових решіток**	
	Біт на блок	**Біт на піксель**	**Біт на блок**	**Біт на піксель**
4	1	0,0625	5	0,3125
8	1	0,0156	7	0,1093
16	1	0,0039	9	0,0351

Дані, представлені в табл. 5, демонструють значну перевагу запропонованого методу над класичним методом з кодовим управлінням з точки зору пропускної спроможності. Хоча класичний метод послідовно вбудовує лише 1 біт на блок незалежно від розміру блоку, запропонований метод використовує структурні особливості досконалих двійкових решіток для досягнення значно більшого корисного навантаження. Примітно, що для малих розмірів блоків (наприклад, $N=4$) стеганографічний метод із застосуванням досконалих двійкових решіток має п'ятикратне збільшення пропускної спроможності у вимірі біт на піксель. Навіть для більших блоків, де щільність вбудовування зазвичай зменшується, метод зберігає значну перевагу як у кількості бітів на блок, так і в кількості бітів на піксель. Це підкреслює потенціал методу для застосувань, що потребують високої пропускної спроможності, особливо коли важливо підтримувати візуальну якість. Оцінка точності

сприйняття є складним завданням через суб'єктивну природу людського зорового сприйняття. Зорова система людини демонструє різну чутливість до просторових, частотних та кольорових спотворень, що робить формальну оцінку обмеженою за своєю суттю. Тим не менш, у запропонованому методі модифікація елементів контейнера суворо обмежена: зміни не перевищують за величиною ±1. Це обмеження забезпечує мінімальні спотворення, які, як очікується, будуть непомітними за нормальних умов перегляду. Для підтвердження цього твердження ми надаємо кількісний аналіз з використанням метрики пікового співвідношення сигнал/шум (PSNR) для різних щільностей вбудовування, що відображає частку модифікованих блоків у зображенні.

PSNR оцінюється як

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right), \tag{41}$$

де MSE – це середньоквадратична помилка між оригінальним та зміненим зображенням. Якщо пікселі змінюються на ±1, то

$$MSE = p\frac{1^2}{100} = \frac{p}{100}. \tag{42}$$

У табл. 6 ми наводимо значення PSNR для різних відсотків блоків, що задіяні у стеганоперетвореннi.

Таблиця 6 – Значення PSNR для різних відсотків блоків, що задіяні у стеганоперетворенні

Відсоток блоків, що задіяні у стеганоперетворенні	**MSE**	**PSNR (dB)**
25 %	0,25	54,15
50 %	0,50	51,14
75 %	0,75	49,38
100 %	1,00	48,13

Результати, представлені в табл. 6, підтверджують, що запропонований метод зберігає значну надійність сприйняття за різних відсотків блоків, що приймають участь у стеганоперетворенні. Навіть при 100 % вбудовуванні, де кожен блок бере участь у приховуванні даних, PSNR залишається вище 48 дБ – у межах діапазону, який

вважається візуально непомітним. Ця стійкість випливає з фундаментальної конструкції методу, успадкованої від концепції кодового управління, де модифікації пікселів обмежені значеннями ±1. Завдяки ретельному обмеженню амплітуди змін, запропонований підхід гарантує, що вбудована інформація не вносить помітних візуальних артефактів, таким чином зберігаючи візуальну цілісність зображення.

У табл. 7 ми показуємо отриману залежність коефіцієнта помилок декодування від коефіцієнта стиснення *QF*.

Таблиця 7 – Залежність коефіцієнта помилок при декодуванні від *QF* в умовах атаки стисненням JPEG

QF	100	90	80	70	60
Кількість помилок, %	0,02	26,53	42,88	46,41	47,77
QF	50	40	30	20	10
Кількість помилок, %	48,40	48,88	49,23	49,53	49,82

Експериментальні результати показують, що навіть за відносно невеликого розміру блоку та вбудовування 7 бітів на блок, що представляє семикратне збільшення корисного навантаження порівняно з класичними методами з кодовим управлінням, запропонований метод підтримує прийнятну стійкість за низьких рівнів стиснення JPEG, що є вражаючим результатом, враховуючи поєднання агресивного збільшення пропускної спроможності та стиснення з втратами. Ці результати піднімають важливе та малодосліджене питання в стеганографічному проєктуванні: чи ефективніше використовувати невеликі блоки з обмеженою пропускною спроможністю для забезпечення стійкості, чи використовувати більші блоки, які вміщують більше даних на одиницю?

Як зазначив Шеннон у своїй фундаментальній роботі [28], більші коди часто забезпечують кращу ефективність та стійкість. Екстраполюючи цей принцип, більші розміри блоків потенційно можуть забезпечити підвищену стійкість не лише до артефактів стиснення, але й до стеганалітичних атак через підвищену структурну складність та варіабельність вбудовування. Ця гіпотеза припускає, що традиційна перевага малих одиниць вбудовування в просторовій та трансформаційно-доменній стеганографії, можливо, потребує переоцінки. Отже, це відкриває переконливий напрямок для переосмислення самих

основ стеганографічного проєктування, що може призвести до появи нових принципів та вбудованих архітектур, заснованих на великоблочних алгебраїчних фреймворках.

1.6. Порявніння стеганографічних методів

У табл. 8 проводиться порівняння різних стеганографічних методів на основі концепції кодового управління за такими критеріями: стійкість до атаки стисненням, надійність сприйняття, що визначається показником PSNR, пропускна спроможність R, область вбудовування, можливість сліпого декодування.

Таблиця 8 – Порівняння стеганографічних методів на основі концепції кодового управління із сучасними аналогами

Алгоритм / Метод	% помилок при заданому рівні QF										PSNR, dB	R	Обл, вб,	Сліпе дек,
	10	20	30	40	50	60	70	80	90	100				
Стеганографічний метод з кодовим управлінням [13]														
$T_{(5,1)}$	42,78	29,35	12,21	3,05	0,97	0,72	0,07	0,03	0	0	48,1	1/16	S	–
$T_{(5,4)}$	50	50	49,04	45,56	40,56	29,90	5,60	0,70	0	0	48,1	1/64	S	–
Стеганографічний метод з кодовим управлінням із сліпим декодуванням														
Original [14]	43,09	37,29	33,06	29,66	26,40	22,36	14,48	10,45	7,47	3,89	42,2	1/64	S	+
ξ_4, ξ_4	42,47	37,83	32,96	15,40	4,91	3,32	2,47	1,60	1,23	1,14	36,1	1/64	S	+
ξ_7, ξ_{11}	42,33	36,58	32,35	27,76	22,30	12,48	5,86	4,63	3,09	2,73	42,2	1/64	S	+
Стеганографічний метод з кодовим управлінням на основі досконалих двійкових решіток														
$A_0(N)$	49,8	49,5	49,2	48,9	48,4	47,8	46,4	42,9	26,5	0,02	48,13	7/64	S	–
Сучасні аналоги														
Li, 2021, [6]	–	–	0,02	–	0,02	–	0,01	–	0,01	0,01	~27,1	0,01	NN	+
Wang, 2021, [7]	–	–	0	–	0	–	–	0	–	0	~38,5	0,01	DCT	+
Zhu, 2019, [8]	–	–	–	–	–	–	33,9	7,4	0,3	–	~45	<1/8	DCT	+
Chanu, 2014, [9]	–	–	–	–	23,88	14,1	2,76	0,08	0,08	–	~32,7	1/16	SVD	+
Мельник, 2012, [10]	13	7	5	4	2	2	2	2	2	–	~ 34,7	1/64	SVD	+
Chang, 2007, [11]	–	–	–	–	24,7	14,4	2,71	0,2	0,1	–	~32,7	1/64	SVD	+

У табл. 8 прийняті такі умовні позначення: S – просторова область, DCT – область трансформант ДКП блоків зображення, SVD – область сингулярного розкладання блоків зображення, NN – стеганоперетворення відбувається засобами нейронної мережі.

1.7. Висновки

1. Узагальнено концепцію кодового управління як ефективну основу для побудови сучасних стеганографічних методів, що забезпечують керований вплив на частотні або просторові складові цифрових контейнерів без переходу в частотну область.

2. Формалізовано класичний стеганографічний метод з кодовим управлінням, який характеризується високою надійністю сприйняття стеганоповідомлення, стійкістю до атак стисненням JPEG та незначним впливом на візуальні властивості контейнера.

3. Досліджено варіант методу зі сліпим декодуванням, що дозволяє здійснювати вилучення прихованої інформації без доступу до оригінального контейнера, забезпечуючи при цьому високу точність декодування в умовах атак.

4. Показано обмеження методу сліпого декодування, пов'язані з варіаціями підблоків, і визначено трансформанти Уолша – Адамара, що є оптимальними з погляду компромісу між стійкістю до атак і мінімізацією помилок.

5. Обґрунтовано доцільність використання досконалих двійкових решіток як основи для побудови ефективних кодових конструкцій, які дозволяють підвищити пропускну спроможність стеганографічних методів без втрати стійкості.

6. Проведено комплексне експериментальне оцінювання ефективності різних варіантів кодових слів, що підтвердило можливість адаптації методів до практичних умов – зокрема до атак стиснення, шумових впливів та інших деструктивних дій.

7. Визначено перспективні напрямки подальших досліджень, зокрема розробку багатоканальних кодових конструкцій, побудову метрик взаємного впливу трансформант та створення адаптивних схем вбудовування з урахуванням статистичних властивостей сигналу.

Список використаних джерел

1. Кобозева А. А., Хорошко В. А. Анализ информационной безопасности. Киев : Изд. ГУИКТ. 2009. 251 с.

2. Wang J. et al. The evolution of the Internet of Things (IoT) over the past 20 years. *Computers & Industrial Engineering.* 2021. Vol. 155. P. 107174.

3. Alaba F. A. The Evolution of the IoT. Internet of Things: A Case Study in Africa. Cham : Springer Nature Switzerland, 2024. P. 1–18.

4. Coiduras-Sanagustín A., Manchado-Pérez E., García-Hernández C. Understanding perspectives for product design on personal data privacy in internet of things (IoT): A systematic literature review (SLR). *Heliyon*. 2024. Vol. 10 (9). P. 1–17.

5. Lam E. Y., Goodman J. W. A mathematical analysis of the DCT coefficient distributions for images. *IEEE transactions on image processing.* 2000. Vol. 9 (10). P. 1661–1666.

6. Li Z., Zhang M., Liu J. Robust image steganography framework based on generative adversarial network. *Journal of Electronic Imaging.* 2021. Vol. 30. Issue 2. P. 023006. doi: 10.1117/1.JEI.30.2.023006

7. Wang S., Zheng N., Xu M. A Compression Resistant Steganography Based on Differential Manchester Code. *Symmetry*. 2021. Vol. 13 (2). P. 345. DOI: 10.3390/sym13020165

8. Zhu Z., Zheng N., Qiao T., Xu M. Robust Steganography by Modifying Sign of DCT Coefficients. *IEEE Access.* 2019. Vol. 7. P. 168613–168628. DOI: 10.1109/access.2019.2953504

9. Chanu Y. J., Singh Kh. M., Tuithung T. A Robust Steganographic Method based on Singular Value Decomposition. *International Journal of Information & Computation Technology.* 2014. Vol. 4 (7). P. 717–726. DOI: 10.1049/ic:20070706

10. Melnik M. A. Compression-resistant steganographic algorithm. *Information security.* 2012. Vol. 2 (8). P. 99–106.

11. Chang C. C., Lin C. C., Hu Y. S. An SVD oriented watermark embedding scheme with high qualities for the restored images. *International journal of innovative computing, information & control.* 2007. Vol. 3 (3). P. 609–620.

12. Evsutin O., Melman A., Meshcheryakov R. Digital Steganography and Watermarking for Digital Images: A Review of Current Research

Directions. *IEEE Access.* 2020. Vol. 8. P. 166589–166611. DOI: 10.1109/ACCESS.2020.3022779

13. Bao Z. et al. A robust image steganography based on the concatenated error correction encoder and discrete cosine transform coefficients. *Journal of Ambient Intelligence and Humanized Computing.* 2020. Vol. 11 (5). P. 1889–1901. DOI: 10.1007/s12652-019-01345-8

14. Bhattacharyya S., Mondal S., Sanyal G. A Robust Image Steganography using Hadamard Transform. International Conference on Information Technology in Signal and Image Processing, Mumbai, 2013. P. 416–426.

15. Tao J. et al. Towards robust image steganography. IEEE Transactions on Circuits and Systems for Video Technology. 2018. Vol. 29(2). P. 594–600. DOI: 10.1109/TCSVT.2018.2881118

16. Костирка О. В. Анализ преимуществ пространственной области цифрового изображения-контейнера для стеганопреобразования. *Информатика и математические методы в моделировании.* 2013. № 3. С. 275–282.

17. Kobozeva A. A., Sokolov A. V. Robust Steganographic Method with Code-Controlled Information Embedding. *Problemele energeticii regionale*. 2021. Vol. 4 (52). P. 115–130.

18. Kobozeva A. A., Sokolov A. V. Steganographic Method with Code Control of Information Embedding Based on Multi-level Code Words. *Radioelectronics and Communications Systems.* Vol. 66 (4). P. 173–189.

19. Sokolov A. V., Ihnatenko O. O., Balandina N. M. Increasing the Efficiency of Blind Decoding of the Steganographic Method with Code Control of Additional Information Embedding. *Problems of regional energetics.* 2024. Vol. 62 (2). P. 121–137. DOI: 10.52254/1857-0070.2024.2-62.11

20. Sokolov A. V. Multiple access steganographic method based on code control and frequency arrangements. *Informatics and Mathematical Methods in Simulation.* Vol. 2021. Vol. 11 (3). P. 147–161.

21. Strang G. The discrete cosine transform. SIAM review. 1999. Vol. 41 (1). P. 135–147.

22. Jayathilake A., Perera A. A. I., Chamikara M. A. P. Discrete Walsh-Hadamard transform in signal processing. *IJRIT Int. J. Res. Inf. Technol*. 2013. Vol. 1. P. 80–89.

23. Kobozeva A. A., Sokolov A. V. The Sufficient Condition for Ensuring the Reliability of Perception of the Steganographic Message in the Walsh-Hadamard Transform Domain. *Problemele Energeticii Regionale.* 2022. 54 (2). P. 84–100.

24. Мазурков М. И. Системы широкополосной радиосвязи. Одесса : Наука и Техника, 2010. 340 с.

25. Natural Resources Conservation Service (NRCS), United States Department of Agriculture. URL: https://www.nrcs.usda.gov.

26. Зігінова Ю. К. Модифікований стеганографічний метод з кодовим управлінням вбудовуванням додаткової інформації із сліпим декодуванням. *Сучасні аспекти діджиталізації та інформатизації в програмній та комп'ютерній інженерії* : міжнародна науково-практична конференція, 01–03 червня 2023 року. С. 68–70.

27. Mazurkov M. I., Chechelnytskyi V. Y., Murr P. Information security method based on perfect binary arrays. *Radioelectronics and Communications Systems.* Vol. 51 (11). P. 612–614. DOI: 10.3103/s0735272708110095

28. C. E. Shannon, A mathematical theory of com-munication. *The Bell system technical journal.* 1948. Vol. 27 (3). P. 379–423.

Ахмаметьєва Ганна Валеріївна
к.т.н., доцент, доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

СВІТОВІ ТЕНДЕНЦІЇ РОЗВИТКУ СТЕГАНОАНАЛІЗУ ЦИФРОВИХ ЗОБРАЖЕНЬ

DOI https://doi.org/10.36059/978-966-397-537-5-7

Широке застосування мультимедійних даних в мережі Internet призводить до швидкого розповсюдження спотворень цифрових зображень (ЦЗ), що полягають у ненавмисних пошкодженнях, таких як незначні похибки при передачі каналами зв'язку або результат фільтрації або стиснення при користуванні деякими додатками, або навмисних, зокрема цілеспрямоване пошкодження зображень, їх розмиття, клонування, фальсифікація (як авторства, так і самих даних), тощо. Через обмеження неконтрольованого застосування криптографічних засобів захисту в ряді країн світу набув розвиток стеганографії – наукового напряму, який займається розробкою та дослідженням методів організації прихованого каналу зв'язку через вбудовування конфіденційної інформації в цифровий контейнер. Найчастіше в якості контейнеру застосовуються ЦЗ через особливості їх структури, що містить достатній обсяг надлишкової інформації для приховування в них секретних даних.

На даний час існує величезна кількість стеганографічних методів, які дозволяють вбудувати повідомлення в просторову або частотну область ЦЗ. Просторові стеганографічні методи, такі як метод заміни найменшого значущого біту LSB (Least Significant Bit) та його варіації [1–4], численні модифікації методів PVD (Pixel Value Differencing) [5–7] забезпечують високу пропускну спроможність прихованого каналу зв'язку, тобто обсяг вбудованої інформації (payload) в біт/піксель (bpp), є простими в реалізації. Такі методи часто забезпечують високу візуальну якість заповненого контейнеру (стеганоповідомлення, СП, stego), однак є нестійкими до ряду атак, таких як просторова фільтрація, яскравісні перетворення, стиснення, які спричиняють повне або часткове руйнування вбудованих даних.

У сучасних дослідженнях, присвячених стеганоаналізу (основною задачею якого є виявлення і доведення факту існування прихованого повідомлення в цифровому контенті), серед просторових методів вбудовування широкого поширення набули методи HUGO [8], WOW [9], S-UNIWARD [10], MiPOD [11], Hill [12] через наявність їх програмних реалізації мовами програмування C++ (Windows, Linux) та MatLAB у відкритому доступі [13].

На відміну від просторових методів частотні застосовують різні перетворення, зокрема дискретне косинусне перетворення [14–16], Вейвлет-перетворення [17–18], перетворення Фур'є [19] та інші. Стеганографічні методи, побудовані на їх основі, більш складні в реалізації, в деякій мірі допускають можливість незначного спотворення прихованої інформації, проте забезпечують стійкість до атак стиснення та деяких інших, проти яких просторові методи стеганографії безсильні.

З метою протидії безконтрольного обміну злочинних та протизаконних повідомлень через приховані канали зв'язку необхідно розвивати та удосконалювати методи стеганоаналізу. Через те, що існує величезна кількість методів вбудовування секретної інформації та наявні різні параметри як самого алгоритму вбудовування/вилучення, також існує можливість варіювати пропускну спроможність або задавати стеганографічний ключ, який задає порядок вибору елементу контейнера, відновлення прихованого повідомлення – вкрай складна і на даний час нереалізована задача. Однак довести факт існування в цифровому контенті деякого секретного повідомлення, оцінити приблизну пропускну спроможність або визначити область або елементи, використані для вбудовування, на даний час не викликає певних труднощів.

Слід зазначити, що на ефективність виявлення вкладень в значній мірі впливає пропускна спроможність прихованого каналу зв'язку, адже незважаючи на можливість ЦЗ вмістити значний обсяг додаткової інформації, найчастіше її вбудовують з малим значенням payload: від 0,05 до 0,5 bpp, що забезпечує більш високу якість результуючого зображення та відсутність небажаних артефактів.

Серед стеганоаналітичних методів розрізняють спеціальні – спрямовані проти виявлення стеганоповідомлень, вбудованих конкретним

визначеним стеганографічним методом, та універсальні, що охоплюють широку групу методів вбудовування. Прийнято розрізняти універсальні методи, що здатні виявляти наявність секретної інформації, вбудованої в просторову область контейнера, та універсальні методи, спрямовані проти стеганографії в області перетворень. Останнім часом загально поширеною є тенденція застосування нейронних мереж для навчання на основі тестової вибірки та подальшого аналізу випадкових даних.

Метою дослідження є аналіз сучасних тенденцій в області стеганоаналізу з урахуванням ефективності виявлення вкладень в цифрових зображеннях, вбудованими поширеними стеганографічними алгоритмами при малих значеннях пропускної спроможності прихованого каналу зв'язку в просторову область та область перетворень контейнерів.

Для розв'язання поставленої мети необхідно:

– проаналізувати сучасні рішення стеганоаналізу для виявлення вкладень в ЦЗ, вбудованих в просторову область контейнера;

– навести порівняльний аналіз ефективності стеганоаналітичних методів, що виявляють вкладення на основі просторової області ЦЗ;

– порівняти ефективність раніше розробленого автором методу Color Triads з сучасними рішеннями стеганоаналізу;

– проаналізувати сучасні стеганоаналітичні методи для виявлення вкладень в ЦЗ, вбудованих в область перетворень контейнера;

– навести порівняльний аналіз ефективності стеганоаналітичних методів, що виявляють вкладення в області перетворень ЦЗ.

Об'єкт дослідження – процеси організації і виявлення стеганографічного каналу зв'язку.

Предмет дослідження – стеганоаналітичні методи для цифрових зображень.

Оцінювати ефективність виявлення наявності прихованих повідомлень будемо на основі показника точності детектування (Accuracy – *Acc*), який визначається на основі бінарної класифікації (табл. 1) за формулою:

$$Acc = \frac{TP + TN}{P + N}, \quad P = TP + FN, \quad N = FP + TN.$$

Таблиця 1 – Позначення бінарної класифікації в задачах стеганоаналізу

		Вірна гіпотеза	
		Контейнер	**Стегано-повідомлення**
Результат класифікації	Контейнер	*TN* – True Negative	*FP* – False Positive
	Стегано-повідомлення	*FN* – False Negative	*TP* – True Positive

Тобто, *TN* – деяке ЦЗ визначається як незаповнений контейнер, і це так насправді є; *FP* – незаповнений контейнер визначається як стеганоповідомлення (в літературі такі помилки називають помилки II роду); *FN* – стеганоповідомлення визначається як незаповнений контейнер (помилки I роду); *TP* – заповнене ЦЗ правильно визначається як стеганоповідомлення.

В сучасній обробці даних, зокрема цифрових зображень, поширено використання згорткових нейронних мереж (Convolutional Neural Networks, CNN), які побудовані на основі матричної структури та здатні автоматично і адаптивно виявляти та навчатися просторової ієрархії ознак, наприклад, просторові особливості перепадів яскравості, високочастотні, середньочастотні або низькочастотні характеристики на основі дискретного косинусного перетворення, перетворення Фур'є та інших. Часто CNN поєднують з технологією глибокого навчання (Deep Learning Networks, DNN), коли неможливо чітко визначити ті характеристики, які властиві, наприклад, лише стеганоповідомленням. В більшості випадків неможливо спрогнозувати, які саме характеристичні особливості будуть збурені тим чи іншим стеганографічним методом, оскільки велике значення в процесі стеганоперетворення має і обраний контейнер.

Розглянемо деякі сучасні методи стеганоаналізу, побудовані на основі нейронних мереж, та проаналізуємо, наскільки ефективно вони виявляють вкладення, вбудовані за умови малої пропускної здатності.

У роботі [20] автори пропонують стеганоаналіз зображень, сформованих з низькою пропускною спроможністю, побудований на основі згорткової нейронної мережі, що включає шар передобробки

(виділення характерних особливостей на основі 30 високочастотних фільтрів), транспонована згортка (Transposed convolutional layer – TC), звичайні згортки (Convolutional layer – Conv) та повнозв'язні згортки (Fully connected layer – FC). Дослідженні були проведені для ЦЗ в градаціях сірого при використанні стеганографічних методів S-UNIWARD [10], WOW [9] та HUGO [8] для пропускної спроможності прихованого каналу зв'язку 0,2, 0,1 та 0,05 bpp. В роботі наведено порівняння результатів детектування з [21–23].

Мережа, запропонована в роботі [21], відрізняється від вищезгаданої кількістю звичайних згорткових шарів при аналогічних інших параметрах, в роботах [22; 23] додатково використовується усереднююче об'єднання(Average Polling layer – AP), в [22] – чотири AP та три FC, в [23] – три AC та один FC при одному високочастотному фільтрі на шарі передоборобки. У всіх згорткових мережах [21–23] відсутній шар TC. Порівнюючи вищезгадані чотири нейронні мережі лише перші дві надають адекватні результати детектування: при пропускній здатності 0,2 bpp отримано середню точність детектування 89 % та 85 % відповідно для [20] і [21], точність детектування для [22; 23] для 0,2 bpp не перевищує 65 %.

В роботі [24] запропоновано стеганоаналітичну нейронну мережу, що містить п'ять звичайних згорткових шарів, п'ять FC шари та два шари підвищення частоти дискретизації. Особливістю підходу є аналіз всіх частотних складових вхідних зображень, не лише високочастотних, як в попередніх методах. В статті наведені результати виявлення стеганоповідомлень, сформованих алгоритмами S-UNIWARD [10], WOW [9] та MiPOD [11], для пропускної здатності 0,1, 0,2, 0,3, 0,4, 0,5 bpp, які є достатньо високими для ємності 0,4, 0,5 bpp, однак значно просідають при 0,1 bpp (від 66 до 78 % в залежності від стеганоалгоритму).

Стаття [25] пропонує складну нейронну мережу, яка складається зі звичайних згорток з різними згортковими ядрами, роздільних згорток (Separable convolution), які оперують з 30 та 60 фільтрами, функцій пакетної нормалізації, поканальних згорток (Deethwise convolution), AP шарів та кінцевої глобальної AP-згортки. Наведені результати детектування для різних стеганографічних методів, результати яких наведено нижче в таблицях порівняння. Незважаючи на складність

запропонованої нейронної мережі результати детектування навіть для пропускної спроможності 0,4 bpp не перевищують 90 %.

В [26] запропоновано мережу на основі компонентів, які містить мережа [25] з додаванням шару просторового відсіву (Spatial Dropout Layer). Наведена точність детектування стеганоповідомлень на основі алгоритмів S-UNIWARD [10], WOW [9] та MiPOD [11], найкращі результати для пропускної спроможності 0,2 bpp отримано при виявленні WOW-стеганоповідомлень: точність виявлення становить 92 %. При менших значеннях пропускної здатності результати не перевищують 76 %.

У роботі [27] автори пропонують відносно просту нейронну мережу, у складі якої є звичайні згортки, AP-згортки та кінцева FC-згортка. Додатковим елементом мережі є коваріаційний пул (Covariance Pool). Передобробка ЦЗ включає пошук характеристик на основі високочастотних фільтрів. Особливістю нейронної мережі є можливість виявляти наявність прихованої інформації, вкладеної як просторовими стеганографічними методами, так і методами, що використовують область перетворень.

Автори статті [28] пропонують ще один метод, спроможний виявляти вкладання, вбудовані як просторовими стеганографічними методами, так і вбудовуванням в область перетворень. Стеганоаналітичний метод базується на аналізі характеристичних особливостей пересічних блоків, причому розмір цих блоків може бути довільним. На основі виявлених характеристик проводиться тренування нейронної мережі та подальше виявлення наявності секретних повідомлень. Результати алгоритму наведені лише для стеганометодів S-UNIWARD [10] та PQ [29], найкраща точність детектування відповідно складає 88,64 % та 82,5 % при пропускної спроможності 0,4 bpp. Також наведено порівняння результатів детектування з роботами [30–33].

У статті [30] наведена нейронна мережа для відокремлення незаповнених контейнерів від стеганоповідомлень, сформованих стеганоалгоритмами S-UNIWARD [10], WOW [9] та Hill [12]. Однак за наведеними результатами мережа здатна виявляти вбудовані вкладення з точністю не вище 90 % при пропускній здатності 0,5 bpp.

У [31] для стеганографічних методів S-UNIWARD [10], WOW [9], MiPOD [11] та HUGO [8] наведені результати

стеганоаналізу нейронною мережею, точність виявлення вкладень не перевищує 90 % для пропускної спроможності 0,4 bpp.

В роботі [32] проводиться дослідження нейронної мережі для виявлення прихованих вкладень в залежності від комбінації складових згортки, наведено результати тестування для стеганографічних алгоритмів S-UNIWARD [10] та WOW [9]. Завдяки запропонованій мережі найкращий результат складає 90,2 % для пропускної здатності 0,2 bpp при виявленні стеганоповідомлень, вбудованих алгоритмом WOW.

Серед просторових стеганографічних методів зберігають актуальність методи заміни найменшого значущого біту (Least Significant Bit, LSB), з'являються його різновиди [34–37], найчастіше метод LSB застосовується при малих значеннях прихованої пропускної спроможності.

Серед стеганоаналітичних розробок, спрямованих на виявлення вкладень LSB-алгоритмами, можна відзначити роботи [38; 39]. В публікації [38] запропонована нейронна мережа для здійснення WS-стеганоаналізу, яка досліджувалась для трьох випадків: при застосуванні лінійного усереднюючого фільтру (AVG), при використанні вибірки KB-dropout, на основі мережевої структури U-Net, а також наведено порівняння з результатами стеганоаналізу B0 [40], ns-B0 [41] та ns/r-B0 [42].

Робота [39] пропонує стеганоаналіз п'яти варіантів методу LSB, зокрема LSB Replacement (LSBR) при заміні одного та двох молодших бітів, LSB Matching, LSBMR, що описується формулою:

$$LSBMR = LSBR\,(f\,(A, B)),$$

де $f(A, B)$ – функція над парою пікселів A, B; та LSBRmod5:

$$LSBR \bmod 5(X) = \arg\min_{Y \bmod 5 - M} |X - Y|,$$

де X, Y – значення пікселів, $X, Y \in [0, 255]$, M – секретне повідомлення в бітах.

Навчання та аналіз передбачають вилучення характерних наборів особливостей на основі кореляційних зв'язків та блокових пеатернів LDP, LFP (в статті наведені результати для трьох наборів, які

позначені як Set F1, Set F2 та SPAM). На основі проведеного огляду в табл. 2–7 зібрані результати виявлення вкладень, вбудованих стеганографічними методами S-UNIWARD, WOW, HUGO, MiPOD, Hill та LSB. В таблицях 2, 3, 5, 7 також наведені результати досліджень [43] розробленого автором стеганоаналітичного алгоритму Color Triads* [44], позначення * зазначає, що алгоритм не використовує в своїй основі нейронні мережі. Порівняння цього алгоритму з аналогами, що базуються на згорткових нейронних мережах або глибокому навчанні, обумовлено отриманими високими значеннями у якості виявлення вбудованих повідомлень. Жирним в табл. 2–7 позначені кращі результати детектування серед стеганоаналітичних методів, побудованих на основі нейронних мереж. Жирним з курсивом – результати стеганоаналізу Color Triads, якщо вони перевищують інші методи.

Таблиця 2 – Точність детектування стеганоповідомлень, сформованих алгоритмом S-UNIWARD, %

Стеганоаналітичний метод	**Payload, bpp**					
	0,5	**0,4**	**0,3**	**0,2**	**0,1**	**0,05**
HSDetect-Net (2025) [24]	94,22	89,36	83,32	81,87	66,78	–
TCSI-ECA-Transfer (2023) [20]	–	–	–	**91,60**	**84,00**	**74,20**
Net from [32] (2023)	–	93,10	–	79,30	–	–
Net from [26] (2023)	96,60	**94,80**	86,20	81,60	68,80	49,80
OBS (2023) [28]	–	88,64	79,50	74,25	66,70	–
Net from [27] (2022)	–	90,58	**86,85**	80,98	69,87	–
AG-Net (2022) [30]	89,27	85,49	80,66	–	–	–
Shen-Net (2021) [21]	–	–	–	88,17	78,90	69,17
GBRAS-net (2021) [25]	–	77,90	–	73,60	–	–
SFR-Net (2021) [31]	–	89,60	–	75,50	–	–
Yedroudj-Net (2018) [22]	–	–	–	56,01	–	–
SCA-TLU-CNN (2017) [33]	–	87,19	–	77,76	67,80	60,00
Xu-Net (2016) [23]	–	–	–	60,94	53,82	50,55
Color Triads* (2022) [43]	–	***99,67***	–	***99,43***	***97,70***	***92,53***

Якщо розглядати стеганоаналіз методу S-UNIWARD, то більшість нейронних мереж добре працюють при виявленні вкладень, вбудованих з пропускною спроможністю 0,3–0,5 bpp, серед них

можна визначити мережу [26; 27] та HSDetect-Net [24], які забезпечують точність детектування на рівні 83–87 % для 0,3 bpp, 92–98 % для 0,4 bpp та 99 % для 0,5 bpp. Для меншої пропускної спроможності серед нейронних мереж кращою є TCSI-ECA-Transfer [20], яка з високою точністю (91,2 %) виявляє стеганоповідомлення, сформовані з пропускною здатністю 0,2 bpp та 0,1 bpp (84 %).

Таблиця 3 – Точність детектування стеганоповідомлень, сформованих алгоритмом WOW, %

Стеганоаналітичний метод	**Payload, bpp**					
	0,5	**0,4**	**0,3**	**0,2**	**0,1**	**0,05**
HSDetect-Net (2025) [24]	99,07	**98,99**	88,44	82,09	76,92	–
TCSI-ECA-Transfer (2023) [20]	–	–	–	89,00	**79,84**	**69,73**
Net from [32] (2023)	–	94,40	–	90,20	–	–
Net from [26] (2023)	**99,60**	98,60	**94,00**	**92,00**	76,00	52,20
Net from [27] (2022)	–	92,09	89,22	84,61	75,70	–
AG-Net (2022) [30]	85,82	83,62	79,57	–	–	–
Shen-Net (2021) [21]	–	–	–	85,49	74,05	63,90
GBRAS-net (2021) [25]	–	89,80	–	80,30	–	–
SFR-Net (2021) [31]	–	87,90	–	76,80	–	–
Yedroudj-Net (2018) [22]	–	–	–	51,20	–	–
SCA-TLU-CNN (2017) [33]	–	90,41	–	83,09	75,58	65,50
Xu-Net (2016) [23]	–	–	–	65,43	55,86	50,76
Color Triads* (2022) [43]	–	***99,51***	–	***99,35***	***97,87***	***94,09***

Для виявлення стеганоповідомлень, сформованих стеганоаналітичним методом WOW з пропускною спроможністю 0,2–0,5 bpp, гарну точність детектування (більше 92 % для заданих значень payload) надає мережа [26], непогані результати спостерігаються також для мережі HSDetect-Net [24], особливо при значеннях payload 0,3–0,5 bpp: точність виявлення заповнених зображень і пустих контейнерів перевищує 88 %. Також можна відзначити мережу [32], що за умови пропускної здатності 0,2 bpp забезпечує 90,2 % точно виявлених стеганоповідомлень і незаповнених контейнерів. З малою пропускною спроможністю 0,05–0,1 bpp впорається мережа TCSI-ECA-Transfer [20], однак все ж таки відсоток помилкових детектувань залишається значним (21 % та 31 % відповідно для 0,1 bpp та 0,05 bpp).

Таблиця 4 – Точність детектування стеганоповідомлень, сформованих алгоритмом HUGO, %

Стеганоаналітичний метод	Payload, bpp			
	0,4	**0,2**	**0,1**	**0,05**
TCSI-ECA-Transfer (2023) [20]	–	**89,15**	**79,30**	**69,19**
Shen-Net (2021) [21]	–	51,32	72,93	62,78
GBRAS-net (2021) [25]	**84,50**	74,60	–	–
SFR-Net (2021) [31]	83,60	75,40	–	–
Yedroudj-Net (2018) [22]	–	84,66	–	–
Xu-Net (2016) [23]	–	66,79	59,49	52,83

Менш поширеним для вбудовування додаткової інформації є стеганографічний алгоритм HUGO, і найкраще визначає наявність вкладень мережа TCSI-ECA-Transfer [20], і хоч для неї не проводились дослідження щодо пропускної спроможності 0,4 bpp, з урахуванням 89 % вірних класифікацій при 0,2 bpp, можна очікувати точність детектування не менше 92–93 %.

Таблиця 5 – Точність детектування стеганоповідомлень, сформованих алгоритмом MiPOD, %

Стеганоаналітичний метод	Payload, bpp					
	0,5	**0,4**	**0,3**	**0,2**	**0,1**	**0,05**
HSDetect-Net (2025) [24]	95,68	89,99	87,80	84,77	**78,90**	–
Net from [26] (2023)	**97,20**	**96,40**	**89,00**	**86,80**	70,90	50,60
GBRAS-net (2021) [25]	–	81,40	–	68,30	–	–
SFR-Net (2021) [31]	–	84,10	–	75,20	–	–
Color Triads* (2022) [43]	–	***99,59***	**–**	***99,09***	***95,41***	***92,69***

Стеганографічні алгоритми MiPOD та Hill реалізовані лише мовою математичного середовища MatLAB, чим обумовлено їх менше поширення, тим не менш нейронні мережі [26] та HSDetect-Net [24] непогано визначають наявність вкладень, вбудованих з пропускною спроможністю 0,2–0,5 bpp.

Найкращі результати детектування стеганографічних вкладень, вбудованих методом Hill, визначає нейронна мережа [27], точність детектування складає 86 % для пропускної спроможності 0,4 bpp, очікуване значення *Acc* для 0,5 bpp має перевищувати 90 %.

Таблиця 6 – Точність детектування стеганоповідомлень, сформованих алгоритмом Hill, %

Стеганоаналітичний метод	Payload, bpp				
	0,5	**0,4**	**0,3**	**0,2**	**0,1**
Net from [27] (2022)	–	**86,23**	**82,75**	**77,50**	68,61
AG-Net (2022) [30]	84,60	81,95	77,96	–	–
GBRAS-net (2021) [25]	–	81,90	–	68,50	–

Таблиця 7 – Точність детектування стеганоповідомлень, сформованих алгоритмом LSB, %

Стеганоаналітичний метод		Payload, bpp							
		0,5	**0,4**	**0,3**	**0,25**	**0,2**	**0,1**	**0,05**	**0,01**
B0 [40] (2019)		–	–	–	–	83,10	81,60	78,50	62,90
ns-B0 [41] (2021)		–	–	–	–	88,60	88,60	86,60	**78,00**
ns/r-B0 [42] (2019)		–	–	–	–	91,80	89,80	86,80	73,40
[38] (2024)	AVG	–	–	–	–	91,50	81,60	76,40	57,10
	KB	–	–	–	–	98,90	95,30	87,00	61,90
	U-Net	–	–	–	–	**99,70**	**97,70**	**90,60**	64,20
[39] (2024)	Set F1	87,19	83,29	78,39	75,25	70,49	54,44	–	–
	Set F2	86,8	83,71	80,01	76,50	72,25	57,69	–	–
	SPAM	86,24	78,35	72,50	68,98	63,92	40,05	–	–
Color Triads* (2022) [43]		***100***	***100***	***100***	***100***	***100***	***99,9***	***95,9***	–

З таблиці 7 видно, що найкращий з нейромережевих стеганоаналітичних методів [38] у варіації U-Net забезпечує високу якість виявлення вбудованих повідомлень навіть за умови малих значень прихованої пропускної спроможності (для значень *payload* 0,2, 0,1 та 0,05 bpp досягається точність 99,7 %, 97,9 % та 90,6 % відповідно. Також статті [38, 40–42] приділяють увагу пропускній здатності 0,01 bpp, для якої найкращий результат детектування отриманий алгоритмом ns-B0 [41] і становить 78 % вірно класифікованих стеганоповідомлень та незаповнених контейнерів.

Щодо стеганоалгоритму Color Triads для значень *payload* 0,05–0,5 bpp досягаються максимальні значення точності детектування, які перевищують 95 % вірно класифікованих ЦЗ.

Щодо стеганоаналізу Color Triads слід зазначити важливу його особливість: він може приймати на вхід лише кольорові ЦЗ, представлені колірною схемою RGB. Алгоритм Color Triads побудований так, що для визначення прихованого повідомлення, яке може бути вбудовано в одну, дві або всі три колірні складові [45], аналізується кількість Red-, Green- та Blue-тріад, що утворюються послідовними групами колірних триплетів, тобто кольорів, які описуються схемою (*R*, *G*, *B*), де *R* – яскравість червоної колірної складової, *G* – яскравість зеленої колірної складової, *B* – яскравість синьої колірної складової. Наведені в таблицях 2, 3, 5 і 7 результати детектування ЦЗ властиві випадку заповнення однієї довільної колірної складової, притому досягається висока точність виявлення повідомлень навіть за малих значень пропускної здатності (значення *Acc* не менше 92 % навіть при *payload*=0,05 bpp). Методи, побудовані на основі нейронних мереж, аналізують зображення в градаціях сірого, а тому також здатні виявляти вкладення в кольорових контейнерах, аналізуючи кожну колірну складову незалежно від інших.

Серед стеганографічних методів, що вбудовують додаткову інформацію в область перетворень ЦЗ можна відзначити J-UNIWARD [46], nsF5 [47], UED [48], UERD [49] та PQ [29], зокрема вони використовуються дослідниками при розробці і тестуванні методів виявлення прихованих повідомлень.

При стеганоаналізі ЦЗ в форматі JPEG окрім пропускної спроможності також враховується якісь стиснення результуючих зображень (Quality Factor, QF), найчастіше розглядають значення QF, що дорівнюють 75 та 95.

В роботі [50] автори пропонують нейронну мережу, побудовану за структурою SRNet, причому до характеристичних ознак додаються додаткові, які базуються на стратегії редукції, включаючи max, min и mean, об'єднання яких може створювати додаткові ознаки, зокрема max_mean, min_mean, max_min та max_min_mean. Експериментальні дослідження проводились для чотирьох алгоритмів J-UNIWARD [46], nsF5 [47], UED [48], UERD [49] при значеннях *payload* 0,1–0,4 bpp та значеннях QF 75 та 95. Також в статті наводиться порівняння з мережею SRNet [51], яка стала основою для удосконалення поточного методу, та [52].

У статті [53] запропонований стеганоаналітичний метод, в основі якого відокремлення низькочастотної складової ЦЗ на основі вейвлет-перетворення, для якої за допомогою глибокої згорткової нейронної мережі вилучаються ознаки вмісту зображення. На попередньому етапі отримані характеристики кластеризуються для отримання міток відповідного класу, після чого стеганоаналітична мережа навчається окремо за допомогою зразків в кожному підкласі. В роботі наведені результати експериментів для стеганографічного методу J-UNIWARD [46] при значеннях *payload* 0,1–0,4 bpp та значенні $QF = 75$.

Робота [54] пропонує для стеганоаналізу схему автоматичного зіставлення розподілів, що базується на консенсусної кластеризації CADM. Наводяться результати досліджень при виявленні стеганографічних вкладень методами J-UNIWARD [46], nsF5 [47], UERD [49] та J-MiPOD [55] за умови пропускної здатності 0,2–0,3 bpp. Зокрема точність виявлення вкладень, сформованих алгоритмом J-MiPOD [55] при $QF = 75$ та 0,2 bpp складає 94,3 % та 87,2 % відповідно для 0,3 bpp та 0,2 bpp.

Стаття [56] пропонує стеганоаналітичний алгоритм виявлення вкладень, вбудованих стеганографічними методами J-UNIWARD [46] та UED [48] при $QF = 75$ та 95, а також пропускної спроможності 0,1–0,5 bpp. Стеганоаналіз базується на навчанні нейронної мережі з урахуванням графа, який попереджує глобальну втрату ознак через локальне навчання ознак згортковою нейронною мережею, та модуля покращення ознак, який запобігає ослабленню стеганографічної інформації через накладання згорткових шарів. Це дозволяє знизити кількість помилкових класифікацій, зокрема при пропускній здатності 0,5 bpp спостерігається лише 1,05 % та 6,23 % помилок виявлення вкладень стеганоалгоритмом UED та 4,42 % та 10,83 % помилок виявлення вкладень стеганоалгоритмом J-UNIWARD відповідно при $QF = 75$ та $QF = 95$.

У вищезгаданих роботах [27; 28] наводяться результати експериментів для запропонованих стеганоаналітичних мереж: в [27] на основі стеганографії J-UNIWARD [46] та UED [48], в [28] на основі стеганографії nsF5 [47] та PQ [29], дослідження яких були проведені для $QF = 75$, зокрема точність стеганоаналізу PQ-алгоритму

становить 55,45 %, 60,45 %, 67,04 % та 73,18 % для значень payload 0,1 bpp, 0,2 bpp, 0,3 bpp, 0,4 bpp відповідно. В статті [27] також наведено порівняння точності детектування з [51, 57, 58].

На основі описаних вище стеганоаналітичних методів та наведених в публікаціях порівняння з аналогами, сформовано узагальнені табл. 8–11, які містять точність детектування стеганоповідомлень, отриманих вбудовуванням додаткової інформації стеганографічними алгоритмами J-UNIWARD [46], nsF5 [47], UED [48], UERD [49]. Жирним в таблицях 8–11 позначені кращі результати детектування серед наведених стеганоаналітичних методів.

Таблиця 8 – Точність детектування стеганоповідомлень, сформованих алгоритмом J-UNIWARD, %

Стегано-аналітичний метод	Payload, bpp							
	0,4		0,3		0,2		0,1	
	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95
[51] (2018)	92,71	81,19	88,43	74,10	80,94	65,13	67,65	55,76
[52] (2019)	–	–	–	–	–	–	56,20	59,60
[57] (2019)	92,05	80,87	87,18	73,48	79,11	65,57	67,14	55,84
[58] (2019)	91,61	77,41	85,40	69,47	77,66	61,79	64,20	54,98
[27] (2022)	**93,18**	81,09	89,08	73,92	81,89	65,71	67,88	57,03
[50] (2023)	90,75	79,05	85,20	69,53	76,60	62,27	66,40	56,49
[53] (2023)	85,56	–	79,70	–	71,39	–	61,13	–
[54] (2023)	–	–	72,40	69,50	–	–	–	–
[56] (2023)	90,99	**84,05**	**89,51**	**76,13**	**85,08**	**69,50**	**69,86**	59,79

Порівнюючи результати класифікації і відокремлення стеганоповідомлень від незаповнених контейнерів, можна спостерігати, що незалежно від стеганографічного методу кращі значення точності детектування характерні для якості стиснення $QF = 75$, тобто чим вище якість збережених заповнених зображень, тим складнішим буде стеганоаналіз.

За результатами аналізу стеганографічного алгоритму J-UNIWARD [46] (табл. 8) можна відзначити нейронні мережі [27; 51; 57], які мають порівняний відсоток вірно класифікованих зображень.

Однак найкраща ефективність досягається методом [56], результати якого перевищують аналоги майже за всіма випадками, окрім $QF = 75$ при 0,4 bpp, де кращим виявився метод [27], який за цих умов досягає значення 93,18 %.

Таблиця 9 – Точність детектування стеганоповідомлень, сформованих алгоритмом UED, %

Стегано-аналітичний метод	Payload, bpp							
	0,4		0,3		0,2		0,1	
	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95
[51] (2018)	98,16	92,12	**97,14**	87,63	94,12	80,34	86,77	69,69
[52] (2019)	–	–	–	–	–	–	57,75	59,50
[57] (2019)	98,20	**92,22**	97,11	87,99	94,21	80,70	87,22	70,46
[58] (2019)	97,97	89,90	96,27	84,75	92,76	76,64	83,52	65,35
[27] (2022)	**98,34**	91,87	96,97	87,21	**94,71**	79,95	**88,16**	69,91
[50] (2023)	96,63	90,65	95,50	83,65	91,88	75,82	85,53	65,55
[56] (2023)	94,79	91,96	92,72	**88,78**	89,97	**81,85**	82,61	70,50

Застосування стеганоаналізу проти UED-стеганографії [48] (табл. 9) є більш ефективним при застосуванні нейромереж [51; 57; 27; 56], трохи відстають за показниками методи [50; 58], причому при збереженні стеганоповідомлень з якістю $QF = 75$ краще визначає наявність вбудованої інформації мережа [27], особливо при значеннях пропускної здатності 0,1 та 0,2 bpp (88,16 % та 94,71 % відповідно).

У випадку якості $QF = 95$ найбільш ефективним при значеннях *payload* 0,1, 0,2, 0,3 bpp є [56], де точність виявлення складає 70,5 %, 81,85 % та 88,78 % відповідно.

За умови більш високої пропускної спроможності прихованого каналу зв'язку можна відзначити методи [51] (97,14 % при $QF = 75$, 0,3 bpp), [57] (92,22 % при $QF = 95$, 0,4 bpp) та [27] (98,34 % при $QF = 75$, 0,4 bpp).

При виявленні стеганоповідомлень, сформованих методом nsF5 [47] (табл. 10) фаворитом є нейронна мережа [50], причому якість детектування зображень, збережених з QF=95 навіть трохи перевищує якість $QF = 75$, де різниця у точності детектування при однакових значеннях пропускної здатності не перевищує 1,25 %.

Таблиця 10 – Точність детектування стеганоповідомлень, сформованих алгоритмом nsF5, %

Стегано-аналітичний метод	Payload, bpp							
	0,4		0,3		0,2		0,1	
	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95
[51] (2018)	85,62	82,79	76,25	72,50	67,68	63,96	55,35	52,59
[52] (2019)	–	–	–	–	–	–	58	57,15
NOBS [28] (2023)	82,75	–	73,50	–	66,20	–	58,72	–
[50] (2023)	**97,12**	**98,37**	**92,63**	**93,55**	**82,98**	**83,40**	**68,72**	**68,95**
[54] (2023)	–	–	–	–	81,40	76,90	–	–

Таблиця 11 – Точність детектування стеганоповідомлень, сформованих алгоритмом UERD, %

Стегано-аналітичний метод	Payload, bpp							
	0,4		0,3		0,2		0,1	
	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95	QF 75	QF 95
[51] (2018)	80,87	75,15	79,19	71,94	75,28	66,71	62,63	57,76
[50] (2023)	**96,32**	**88,68**	**93,48**	83,55	**88,41**	75,82	**80,27**	**67,55**
[54] (2023)	–	–	92,90	**93,80**	79,20	**79,30**	–	–

Також нейронна мережа [50] добре зарекомендувала себе при виявленні вкладень на основі UERD-стеганографії [49] (табл. 11), однак за умови $QF = 95$ при значеннях *payload* 0,2 та 0,3 bpp точність методу [54] перевищує і складає 79,3 % та 93,8 % відповідно.

Висновки. Отже, проведений огляд сучасних методів стеганографічного аналізу показав, що все більшого поширення набувають нейронні мережі, побудовані на основі згорткової структури та глибокого навчання з виявленням численних ознак ЦЗ як на основі просторової області, так і в області перетворень на основі низьких, середніх та високих частот зображення.

Розглянуті мережі забезпечують високу ефективність виявлення вкладень за умови пропускної спроможності прихованого каналу зв'язку 0,4–0,5 bpp, однак за менших значеннях пропускної

здатності результати залишають бажати кращого. Особливо цей факт підтверджує порівняння алгоритму Color Triads [43] з аналогами, що працюють в просторовій області, який використовує в своїй основі лише три подібні характеристики, а саме кількість Red-, Green- та Blue-тріад, за якими визначається факт наявності вкладень. Отже точність детектування цього алгоритму свідчить про те, що точковий аналіз може дати набагато вищі результати, ніж навчання на основі великої кількості характеристичних ознак.

При цьому розвиток стеганоаналізу не припиняється, оскільки постійно з'являються нові методи стеганографії, спрямовані на ускладнення процесу виявлення вкладень, в сучасних дослідженнях автори охоплюють все більший спектр стеганографічних алгоритмів при урахуванні малих значень пропускної спроможності прихованого каналу зв'язку, що особливо актуально, оскільки саме такі вкладення набагато складніше.

Дослідники удосконалюють структуру нейронних мереж, комбінуючи різні функції та елементи аналізу і вилучення характеристичних ознак, що вносить значний вклад в посилення інформаційної безпеки та захисту інформації.

Список використаних джерел

1. J. Mielikainen. LSB matching revisited. *Signal Processing Letters, IEEE*. 2006. Vol. 13. P. 285–287.

2. Muhammad K., Sajjad M., Mehmood, I. A novel magic LSB substitution method (M-LSB-SM) using multi-level encryption and achromatic component of an image. *Multimed Tools Appl*. 2016. No. 75. P. 14867–14893. DOI: https://doi.org/10.1007/s11042-015-2671-9

3. Soukaena H. Hashem, Eman Taleb Zghaer. Review Study among Traditional LSB, Proposed Modified LSB and DWT Steganography Algorith. *Al-Mansour Journal*. 2017. Issue (27). P. 37–56.

4. Rahman S., Uddin J., Hussain H. et al. A novel and efficient digital image steganography technique using least significant bit substitution. *Scientific Reports*. 2025. Vol. 15, Is. 107. DOI: https://doi.org/10.1038/s41598-024-83147-3

5. Wu D. C., Tsai W. H. A Steganographic method for images by pixel-value differencing. *Pattern Recognition Letters*. 2003. Vol. 24. P. 1613–1626.

6. Da-Chun Wu, Zong-Nan Shih. Image Steganography by Pixel-Value Differencing Using General Quantization Ranges. *Computer Modeling in Engineering & Sciences*. August 2024. Vol. 141 (1). P. 1–10. DOI: 10.32604/cmes.2024.050813

7. Monalisa Sahu, Neelamadhab Padhy, Sasanko Sekhar Gantayat. Multi-directional PVD steganography avoiding PDH and boundary issue. *Journal of King Saud University – Computer and Information Sciences*. November 2022. Vol. 34 (10), Part A. P. 8838–8851. DOI: https://doi.org/10.1016/j.jksuci.2021.10.007

8. Filler T., Fridrich J. Gibbs Construction in Steganography. *IEEE Transactions on Information Forensics and Security*. 2010. Vol. 4 (5). P. 705–720.

9. Holub V., Fridrich J. Designing Steganographic Distortion Using Directional Filters. In: *2012 IEEE International Workshop on Information Forensics and Security* (WIFS 2012). IEEE, Tenerife, Canary Islands, 2012. P. 1–6.

10. Holub V., Fridrich J., Denemark T. Universal Distortion Function for Steganography in an Arbitrary Domain. *EURASIP Journal on Information Security* (Section: SI: Revised Selected Papers of ACM IH and MMS 2013). 2014. P. 1–13.

11. Sedighi V., Cogranne R., Fridrich J. Content-Adaptive Steganography by Minimizing Statistical Detectability. *IEEE Transactions on Information Forensics and Security*. 2016. Vol. 11, Is. 2. P. 221–234.

12. Shuliang Sun, Yongning Guo. A Novel Image Steganography Based on Contourlet Transform and Hill Cipher. *Journal of Information Hiding and Multimedia Signal Processing*. 2015. Vol. 6 (5). P. 889–897.

13. Steganographic Algorithms. DDE Lab at Binghamton University. [Electronic resource]. URL: https://dde.binghamton.edu/download/stego_algorithms/

14. Khalaf Ashraf A. M., Fouad, Osama, Hussein Aziza, Hamed, Hesham, Kelash Hamdy, Ali, Hanafy. Hiding data in images using DCT steganography techniques with compression algorithms. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*. 2019. Vol. 17 (3). P. 1168–1175. DOI: 10.12928/telkomnika.v17i3.

15. Ritonga Ramadani, Rosnelly Rika, Wanayumini, Wanayumini. Implementation of SLT-DCT Steganography Method on Images

to Improve the Quality of Steganography Images. *Proceeding of International Conference on Science and Technology UISU*. 2024. P. 264–271. DOI: 10.30743/qad1k130.

16. Tamer Rabie, Mohammed Baziyad, Ibrahim Kamel. High-Fidelity Steganography: A Covert Parity Bit Model-Based Approach. *Algorithms*. 2024. Vol. 17 (8). P. 328. DOI: https://doi.org/10.3390/a17080328

17. Tian Wu, Xuan Hu, Chunnian Liu, Yizhe Wang, Yiping Zhu. An efficient steganography scheme based on wavelet transformation for side-information estimation. Journal of King Saud University – Computer and Information Sciences. July 2024. Vol. 36, Is. 6, 102109. P. 1–12. https://doi.org/10.1016/j.jksuci.2024.102109

18. Alobaidi T., Mikhael W. An adaptive steganography insertion technique based on wavelet transform. *Journal of Engineering and Applied Science*. 2023. Vol. 70, P. 144. DOI: https://doi.org/10.1186/s44147-023-00300-x

19. Yujie Jiang, Jing Dong, Shutiao Luo. Robust Generative Image Steganography based on Frequency Domain using Fourier Transform. *BDICN '25: Proceedings of the 2025 4th International Conference on Big Data, Information and Computer Network*. 2025. P. 290–295. DOI: https://doi.org/10.1145/3727353.3727402

20. Shouyue Liu, Chunying Zhang, Liya Wang, Pengchao Yang, Shaona Hua, Tong Zhang. Image Steganalysis of Low Embedding Rate Based on the Attention Mechanism and Transfer Learning. *Electronics*. 2023. Vol. 12. P. 969. DOI: https://doi.org/10.3390/electronics12040969

21. Shen J., Liao X., Qin Z., Liu X.-C. Spatial Steganalysis of Low Embedding Rate Based on Convolutional Neural Network. *J. Softw.* 2021. Vol. 32. P. 2901–2915.

22. Yedroudj M., Comby F., Chaumont M. Yedroudj-net: An efficient CNN for spatial steganalysis. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018. P. 2092–2096.

23. Xu G., Wu H., Shi Y. Structural design of convolutional neural networks for steganalysis. IEEE Signal Process. *Lett*. 2016. Vol. 23. P. 708–712. DOI: https://doi.org/10.1109/LSP.2016.2548421

24. Ntivuguruzwa Jean De La Croix, Tohari Ahmad, Fengling Han, Royyana Muslim Ijtihadie. HSDetect-Net: A Fuzzy-Based Deep

Learning Steganalysis Framework to Detect Possible Hidden Data in Digital Images. *IEEE Access*. 2025. Vol. 13. P. 43013–43027. DOI: 10.1109/ACCESS.2025.3546510

25. Reinel T.-S., Brayan A. H., Alejandro B. M., Alejandro M.-R., Daniel A.-G., Alejandro A. J., B. A. Buenaventura, Simon O.-A., Gustavo I., Raúl R.-P. GBRAS-net: A convolutional neural network architecture for spatial image steganalysis. *IEEE Access.* 2021. Vol. 9. P. 14340–14350. DOI: 10.1109/ACCESS.2021.3052494

26. Croix N. J. D. L., Ahmad T., Han F. Enhancing secret data detection using convolutional neural networks with fuzzy edge detection. *IEEE Access.* 2023. Vol. 11. P. 131001–131016. DOI: 10.1109/ACCESS.2023.3334650

27. Deng X.-Q., Chen B.-L., Luo W.-Q., Luo D.. Universal image steganalysis based on convolutional neural network with global covariance pooling. *J. Comput. Sci. Technol.*, 2022. Vol. 37 (5). P. 1134–1145. DOI: 10.1007/s11390-021-0572-0

28. Sabeti V. An improved approach to blind image steganalysis using an overlapping blocks idea. *J. Electr. Comput. Eng. Innovations*. 2023. Vol. 11 (2). P. 563–276. DOI: 10.22061/jecei.2022.9241.592

29. Fridrich J., Goljan M., Soukal D. Perturbed quantization steganography. *ACM Multimedia System Journal*, 2005. Vol. 11 (2). P. 98–107.

30. Zhang H., Liu F., Song Z., Zhang X., Zhao Y. AG-Net: An advanced general CNN model for steganalysis. *IEEE Access*. 2022. Vol. 10. P. 44116–44122.

31. Xu G., Xu Y., Zhang S., Xie X. SFRNet: Feature extraction-fusion steganalysis network based on squeeze-and-excitation block and RepVgg Block. *Secur. Commun. Networks*. 2021. P. 1–10.

32. Jean De La Croix Ntivuguruzwa, Tohari Ahmad. A convolutional neural network to detect possible hidden data in spatial domain images. *Cybersecurity*. 2023. Vol. 6 (23). DOI: https://doi.org/10.1186/s42400-023-00156-x

33. Jian Ye, Jiangqun Ni, Yang Yi. Deep Learning Hierarchical Representations for Image Steganalysis. *IEEE Transactions on Information Forensics and Security*. 2017. Vol. 12 (11). P. 2545–2557.

34. Mielikainen J. LSB matching revisited. *Signal Processing Letters, IEEE*. 2006. Vol. 13. P. 285–287.

35. Muhammad K., Sajjad M., Mehmood I. A novel magic LSB substitution method (M-LSB-SM) using multi-level encryption and achromatic component of an image. *Multimed Tools Appl.* 2016. Vol. 75. P. 14867–14893. DOI: https://doi.org/10.1007/s11042-015-2671-9

36. Soukaena H. Hashem, Eman Taleb Zghaer. Review Study among Traditional LSB, Proposed Modified LSB and DWT Steganography Algorith. *Al-Mansour Journal.* 2017. Vol. 27. P. 37–56.

37. Marwa M. Emam, Abdelmgeid A. Aly, Fatma A. Omara. A Modified Image Steganography Method based on LSB Technique. International Journal of Computer Applications (0975 – 8887). September 2015. Vol. 125 (5). P. 12–17.

38. Martin Beneš, Rainer Böhme. Exploring Diffusion-Inspired Pixel Predictors for WS Steganalysis. In *Proceedings of the 2024 ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec '24)*, June 24–26, 2024, Baiona, Spain. P. 75–86. DOI: https://doi.org/10.1145/3658664.3659645

39. Veena Sivasamy Thanasekaran, Arivazhagan Selvaraj. Low Dimensional Multi Class Steganalysis of Spatial LSB based Stego Images Using Textural Features. *The International Arab Journal of Information Technology.* March 2024. Vol. 21 (2). P. 233–242. DOI: https://doi.org/10.34028/iajit/21/2/6

40. Mingxing Tan, Quoc Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *International Conference on Machine Learning (ICML)*. 2019. PMLR, P. 6105–6114.

41. Yassine Yousfi, Jan Butora, Jessica Fridrich, Clément Fuji Tsang. Improving Efficient-Net for JPEG Steganalysis. *In Workshop on Information Hiding and Multimedia Security (IH&MMSec)*. 2021. ACM, P. 149–157.

42. Mo Chen, Mehdi Boroumand, Jessica Fridrich. Reference Channels for Steganalysis of Images with Convolutional Neural Networks. *In Workshop onbInformation Hiding and Multimedia Security (IH&MMSec)*. 2019. ACM, P. 188–197.

43. Akhmametieva H. Universal approach to the identification of steganographic transformation of the spatial domain of digital images.

International Science Journal of Engineering & Agriculture. 2022. Vol. 1 (3). P. 133–142.

44. Anna Akhmametieva. Steganalysis of digital contents, based on the analysis of unique color triplets. *Annales Mathematicae et Informaticae*. 2017. Vol. 47. P. 3–18.

45. Akhmametieva A. V., Bwabwa M. C. Steganalysis of Digital Images under Different Content Fullness. *Telecommunications and Radio Engineering*. 2019. Vol. 78. P. 671–681. DOI: 10.1615/TelecomRadEng.v78.i8.30

46. Holub V., Fridrich J., Denemark T. Universal Distortion Function for Steganography in an Arbitrary Domain. *EURASIP Journal on Information Security* (Section:SI: Revised Selected Papers of ACM IH and MMS 2013). 2014 (1). P. 1–13.

47. Fridrich, J. Pevný T., Kodovský J. Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities. In J. Dittmann and J. Fridrich, editors, *Proceedings of the 9th ACM Multimedia & Security Workshop*, Dallas, TX, September 20–21, 2007. P. 3–14.

48. L. Guo, J. Ni, Y. Q. Shi. Uniform embedding for efficient JPEG steganography. *IEEE Transactions on Information Forensics and Security*. 2014. Vol. 9 (5). P. 814–825.

49. Guo L., Ni J., Su W., Tang C., Shi Y. Q. Using statistical image model for JPEG steganography: Uniform embedding revisited. *IEEE Transactions on Information Forensics and Security.* 2015. Vol. 10 (12). P. 2669–2680.

50. Yuchen Li, Baohong Ling, Donghui Hu, Shuli Zheng, Guoan Zhang. A Deep Learning Driven Feature Based Steganalysis Approach**.** *Intelligent Automation & Soft Computing*. 2023. Vol. 37 (2). P. 2213–2225. DOI: 10.32604/iasc.2023.029983

51. Boroumand M., Chen M., Fridrich J. Deep residual network for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security*. 2019. Vol. 14 (5). P. 1181–1193.

52. Xue Y., Yang L., Wen J., Niu S., Zhong P. A subspace learning-based method for JPEG mismatched steganalysis. *Multimedia Tools and Applications*. 2019. Vol. 78 (7). P. 8151–8166.

53. Mo Chengyu, Liu Fenlin, Zhu Ma, Yan Gengcong, Qi Baojun, Yang Chunfang. Image Steganalysis Based on Deep Content Features

Clustering. *Computers, Materials and Continua*. 2023. Vol. 76 (3). P. 2921–2936. DOI: https://doi.org/10.32604/cmc.2023.039540

54. Ju Jia, Meng Luo, Siqi Ma, Lina Wang, Yang Liu. Consensus-Clustering-Based Automatic Distribution Matching for Cross-Domain Image Steganalysis. *IEEE Transactions on Knowledge and Data Engineering*. June 2023. Vol. 35 (6). P. 5665–5679.

55. Cogranne R., Giboulot Q., Bas P. Steganography by minimizing statistical detectability: The cases of JPEG and color images. *In Proc. Workshop Informat. Hiding Multimedia Secur*. 2020. P. 161–167.

56. Qiyun Liu, Zhiguang Yang, Hanzhou Wu. JPEG Steganalysis Based on Steganographic Feature Enhancement and Graph Attention Learning. *Journal of Electronic Imaging.* 2023. P. 1–10. DOI: https://doi.org/10.48550/arXiv.2302.02276

57. Huang J., Ni J., Wan L., Yan J. A customized convolutional neural network with low model complexity for JPEG steganalysis. In Proc. the ACM Workshop on Information Hiding and Multimedia Security, July 2019. P. 198–203. DOI: 10.1145/3335203.3335734

58. Deng X., Chen B., Luo W., Luo D. Fast and effective global covariance pooling network for image steganalysis. In Proc. the ACM Workshop on Information Hiding and Multimedia Security, July 2019, P. 230–234. DOI: 10.1145/3335203.3335739

Овчинников Микола Юрійович
асистент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ЗАХИСТ ІНФОРМАЦІЇ В МЕДІАФАЙЛАХ

DOI https://doi.org/10.36059/978-966-397-537-5-8

Актуальність теми. У сучасному цифровому середовищі обсяги створення, обміну та зберігання медіафайлів (зображень, відео, аудіо) зростають експоненційно. Мультимедійний контент відіграє ключову роль у функціонуванні соціальних мереж, електронного управління, медіаіндустрії, освіти, реклами та навіть юриспруденції. Водночас медіафайли дедалі частіше стають об'єктами кіберзлочинів, включно з підробкою, модифікацією, перехопленням та несанкціонованим поширенням конфіденційної інформації. Це створює серйозні загрози як для окремих користувачів, так і для організацій, які зберігають або обробляють велику кількість цифрового контенту.

Особливо гостро проблема захисту медіафайлів постає у світлі стрімкого розвитку технологій штучного інтелекту, зокрема генеративних моделей, здатних створювати фальшиві зображення або відео, відомі як deepfake. Такі інструменти вже активно використовуються у дезінформаційних кампаніях, кібершантажі та спробах втручання в політичні процеси. У відповідь на ці виклики постає необхідність впровадження ефективних засобів захисту, здатних забезпечити автентичність, цілісність та конфіденційність мультимедійних даних.

Традиційні підходи до захисту інформації, зокрема криптографічні методи, залишаються актуальними, однак в умовах розвитку стеганографічних атак і змішаних загроз виникає потреба у використанні комбінованих технологій. Зокрема, зростає інтерес до цифрових водяних знаків, адаптивної стеганографії, верифікації метаданих, використання блокчейн-систем для підтвердження автентичності медіа та методів штучного інтелекту для виявлення змін у контенті. Усе це формує нову галузь у сфері інформаційної безпеки – захист інформації в медіафайлах, яка потребує системного дослідження, практичного впровадження та удосконалення з урахуванням динаміки сучасних кіберзагроз.

Захист інформації в медіафайлах формується як самостійний міждисциплінарний напрям, що перебуває на стику інформаційної безпеки, цифрової криміналістики, машинного навчання та авторського права. Цей напрям охоплює не лише технічні аспекти – як-от шифрування, водяні знаки або стеганографію, – а й правові, організаційні та етичні питання, пов'язані з автентичністю, доказовістю та непорушністю цифрового контенту.

Особливість цієї галузі полягає у необхідності глибокого розуміння формату та структури медіафайлів, властивостей цифрових слідів, взаємодії між метаданими та зображенням або аудіоінформацією, а також у вмінні виявляти штучно згенеровані або змінені фрагменти. Водночас, засоби захисту повинні бути малопомітними для користувача, не впливати на якість візуального або аудіовмісту, але бути надійними з точки зору стійкості до атак і здатності підтверджувати справжність навіть у ворожому середовищі.

Інтенсивний розвиток цієї сфери стимулюється практичними потребами: від підтвердження достовірності відеозаписів у судових процесах до захисту журналістських матеріалів, отриманих у зонах конфліктів. Державні органи, ЗМІ, приватні компанії й правозахисні організації все частіше шукають надійні методи ідентифікації автентичного контенту та виявлення маніпуляцій. У зв'язку з цим міжнародні ініціативи на кшталт Content Authenticity Initiative (Adobe, BBC, New York Times) або Coalition for Content Provenance and Authenticity (C2PA) розробляють стандарти, що регулюють доказовість цифрового вмісту.

Таким чином, захист інформації в медіафайлах стає критично важливою складовою кібергігієни та національної безпеки. Формування теоретичних основ і розробка практичних інструментів у цій галузі становить актуальну дослідницьку задачу, що потребує міжгалузевого підходу, високого рівня технічної компетентності та чіткого нормативного врегулювання.

Метою даного дослідження є систематичний аналіз сучасних методів і технологій захисту інформації в медіафайлах з урахуванням динаміки розвитку кіберзагроз, – включно зі стеганографією, цифровим водяним знаком, криптографічними гібридними методами та AI-підходами – з акцентом на практичні інструменти,

здатні протистояти поточним кіберзагрозам, та на розробку рекомендацій для їхнього застосування в локальному контексті. А також обґрунтування практичної ефективності інструментів, що забезпечують автентичність, цілісність та конфіденційність мультимедійного контенту. Дослідження має на меті виявити найбільш релевантні підходи до захисту зображень, відео та аудіо у цифровому середовищі та визначити їхню придатність до використання в різних прикладних сферах – від цифрової журналістики до військової справи та судової експертизи.

У центрі уваги знаходиться як технологічна складова (алгоритми шифрування, цифрові водяні знаки, стеганографічні методи, AI/ML-підходи), так і практичний аспект – зокрема, огляд реальних інструментів, що впроваджуються у відкритих або корпоративних інформаційних системах. Також передбачено аналіз поточних тенденцій щодо інтеграції таких інструментів у процеси цифрової автентифікації, верифікації джерел, захисту авторських прав і протидії інформаційним маніпуляціям.

Окремою задачею є визначення сильних і слабких сторін існуючих рішень, ступеня їх стійкості до актуальних типів атак, а також ступеня відповідності правовим і етичним стандартам. Це дозволяє не лише підвищити рівень теоретичного осмислення проблеми, а й закласти підґрунтя для подальших розробок у цій перспективній галузі кібербезпеки.

Досягнення цілей дослідження передбачає виконання низки взаємопов'язаних завдань. Серед них, визначити основні типи кіберзагроз на медіафайли, що включають як класичні атаки на цілісність і конфіденційність, так і складніші модифікації, зокрема deepfake-маніпуляції з застосуванням штучного інтелекту. Потрібно здійснити огляд і класифікацію сучасних методів захисту: гібридних стеганографічно-шифрувальних систем, зокрема тих, що застосовують DCT/DWT-класифікацію в поєднанні з криптографією (DCT+DWT+OTP), а також інноваційних підходів coverless стеганографії, заснованих на генеративних моделях – зокрема coverless-стегіографія з DiffStega 2024 і об'єктно-орієнтовані формати на основі YOLO. Проаналізувати інструментальні реалізації, що доступні й перспективні для практичного використання, включно

з цифровими водяними знаками, стеганографічними утилітами (OpenStego, Digimarc тощо) та блокчейн-інтегрованими методами. Оцінити ефективність захисних методів за такими критеріями: стійкість до змін (JPEG, переформатування, геометричні трансформації), прихованість кодованої інформації, здатність до верифікації джерела та захисту авторського права, а також відповідність технічним можливостям українських організацій. Сформувати рекомендації щодо впровадження комплексних технологій захисту медіаінформації в умовах атак національного рівня, гібридної агресії, дезінформації та кібервійськ.

Об'єктом дослідження є процеси забезпечення інформаційної безпеки в цифрових медіафайлах.

Предметом дослідження є методи, засоби та інструменти захисту мультимедійного контенту від несанкціонованого доступу, підробки та модифікації.

Сучасний стан досліджень. За останні роки наукова спільнота активно розвиває інструменти захисту медіафайлів, зокрема шляхом поєднання стеганографії, цифрових водяних знаків та криптографічних технологій із методами глибинного навчання. Глибокі нейронні мережі, зокрема автоенкодери, дедалі активніше використовуються у стеганографії, дозволяючи поєднувати класичні методи приховування інформації з інтелектуальними алгоритмами. Це відкриває нові можливості для підвищення надійності, пропускної здатності та стійкості систем до атак завдяки адаптивному кодуванню і здатності моделі до навчання на реальних даних [1]. В іншій роботі розроблено моделі та алгоритми створення цифрових водяних знаків (ЦВЗ) для аудіофайлів на основі сучасних методів інформаційної теорії, статистичного моделювання, цифрової обробки сигналів та нейронних мереж. Показано, що застосування двошарових штучних нейронних мереж прямого поширення для стеганографічного вбудовування дозволяє зменшити спотворення сигналу та підвищити стійкість до атак. Проведено статистичний аналіз виявлення і відновлення ЦВЗ стороннім спостерігачем, що дало змогу оцінити ефективність і безпеку маркування. Результати підтверджують доцільність використання таких методів для контролю авторських прав і діагностики цифрового аудіоконтенту [2].

В дослідженні Бондарь І. В. розроблено моделі та алгоритми створення цифрових водяних знаків (ЦВЗ) для аудіофайлів на основі сучасних методів інформаційної теорії, статистичного моделювання, цифрової обробки сигналів та нейронних мереж. Показано, що застосування двошарових штучних нейронних мереж прямого поширення для стеганографічного вбудовування дозволяє зменшити спотворення сигналу та підвищити стійкість до атак. Проведено статистичний аналіз виявлення і відновлення ЦВЗ стороннім спостерігачем, що дало змогу оцінити ефективність і безпеку маркування. Результати підтверджують доцільність використання таких методів для контролю авторських прав і діагностики цифрового аудіоконтенту [3]. Пропонується і комплексний огляд застосувань блокчейн-технологій для захисту мультимедійного контенту, розробляючи системну таксономію, що охоплює технічні аспекти блокчейну, методи захисту – такі як шифрування, цифрове управління правами, водяні знаки та відстеження транзакцій – а також критерії ефективності, й окреслюють ключові виклики та перспективи подальших досліджень у цій галузі [4]. У роботах різних авторів зустрічається систематизація підходів до прихованого вбудовування даних у зображення, аудіо та відео за допомогою нейромереж, зазначаючи вагомі переваги таких моделей у забезпеченні непомітності та надійності захисту. Подальший розвиток цих напрямків продовжився в гібридних методах, де використовуються перетворення DCT або DWT у поєднанні з GAN, що дозволяє вбудовувати дані в частотну область з підвищеною стійкістю до атак і мінімальним погіршенням візуальної якості. Алгоритми, що комбінують DCT, DWT, а також криптографічні підходи (наприклад, OTP чи Playfair), довели свою ефективність у задачах захисту авторських прав та забезпечення цілісності через вимірювання PSNR, SNR та MSE. Hiroki Tomosada та співавтори у своїй роботі пропонують високоякісний метод усунення розмиття зображень, який ґрунтується на використанні дискретного косинусного перетворення (DCT) і відзначається зниженою обчислювальною складністю. Модель була навчена на спеціально підготовленому наборі даних, що містить зображення з інтенсивними розмиттями, зумовленими рухом. У той час як традиційні підходи до відновлення зображень, засновані на згорткових нейронних мережах

(CNN) і генеративних змагальних мережах (GAN), часто використовують багатомасштабну або багатофрагментну архітектуру для зменшення артефактів і покращення чіткості, запропонований метод *DeblurDCTGAN* досягає подібних або кращих результатів без цих ускладнень. Ключовою інновацією є використання функції втрат на основі DCT, яка дозволяє порівнювати відновлене зображення з еталонним у частотній області. Завдяки цьому підходу вдається ефективно зменшити блокові шуми та кільцеві артефакти, зберігаючи при цьому якість розфокусованих областей. Результати експериментів демонструють, що *DeblurDCTGAN* перевершує традиційні методи за показниками PSNR, SSIM і часом обробки. Крім того, при тестуванні на реальних наборах даних (GoPro, DVD, NFS, HIDE) метод показав підвищену ефективність завдяки кастомізованому набору тренувальних зображень [5]. На Рисунку 1 представлено використання семантичної карти обличчя як апріорної для нейронної мережі для усунення розмиття [6]. Цей метод спочатку отримує семантичні мітки розмитих зображень обличчя за допомогою мережі розбору облич. Потім семантична карта об'єднується з розмитим зображенням у вимірі каналу як вхідними даними мережі для усунення розмиття. Відтворено з дозволу посилання © IEEE 2018.

Окрім того, у 2024 р. з'явилися приклади застосування гібридних підходів до відео – де поєднуються стеганографія та криптографія для вбудовування секретних даних у відеофрейми, що забезпечує подвійний рівень захисту: спершу шифрування, потім – приховане вбудовування у послідовність кадрів. У сфері стеганалітики демонструють ефективні результати невикористовувані CNN-моделі, що здатні виявляти навіть технологічно вдосконалені приховані зв'язки в медіафайлах. У дослідженні C. Ravichandran та співавторів розглянуто сучасні підходи до відеостеганографії як методу прихованого передавання інформації в умовах загроз з боку третіх осіб. Основна увага приділяється методу приховування даних на основі найменш значущого біта (LSB), модифікованого із застосуванням хеш-функцій для підвищення стійкості та безпеки. Автори провели ґрунтовний аналіз існуючих методів стеганографії, спрямованих на збереження якості відео при одночасному підвищенні захищеності прихованих даних, а також вказали на актуальність теми в умовах зростання загроз цифровій безпеці.

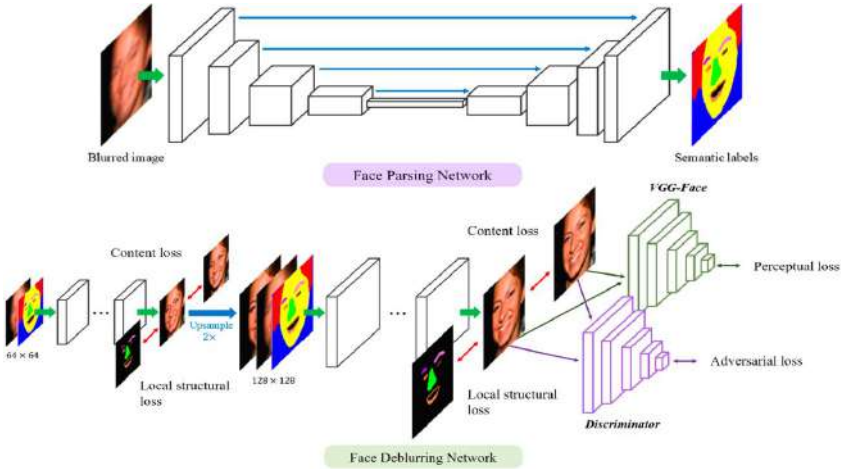


Рисунок 1 – Використання семантичної карти обличчя як апріорної для нейронної мережі для усунення розмиття [6]

Результати моделювання, проведеного у середовищі MATLAB, засвідчили перевагу запропонованого алгоритму над іншими сучасними рішеннями. Запропонований підхід забезпечує надійне приховування даних із мінімальними спотвореннями, що зберігає високу якість відеопотоку та ускладнює виявлення вбудованої інформації сторонніми спостерігачами [7].

Одним із ключових трендів стає розвиток так званої “coverless” стеганографії, при якій дані не вставляються безпосередньо до медіафайла, а реконструюються на основі моделі, що забезпечує високу непомітність для систем детекції. У роботі, присвяченій огляду алгоритмів безвбудовувальної (coverless) стеганографії, автори аналізують понад 90 наукових праць, що охоплюють основні етапи розвитку цієї технології у сфері зображень та відео. На відміну від традиційних методів стеганографії, які вносять незначні зміни в носій інформації (що може бути виявлено за допомогою алгоритмів стегоаналізу), безвбудовувальна стеганографія не модифікує сам носій, а встановлює відповідність між переданим повідомленням і характеристиками обраного об’єкта. Завдяки відсутності змін у структурі носія, така технологія значно ускладнює або навіть унеможливлює

виявлення прихованої інформації сторонніми засобами. У статті узагальнено внески існуючих методів, а також детально розглянуто актуальні проблеми обмеженої місткості, стійкості до атак та забезпечення конфіденційності. Особливу увагу приділено безпеці безвбудовувальної стеганографії – вперше виконано її аналіз як з теоретичної, так і з прикладної точки зору [8]. Таким чином, сучасні дослідження демонструють активний рух у напрямі мультидисциплінарних гібридних рішень, що поєднують крипто-, стего- та машинно-навчальні технології, створюючи новий рівень захисту інтелектуальної та конфіденційної інформації в медіафайлах. Стрімкий розвиток цифрових технологій та широке поширення мультимедійного контенту зумовлює необхідність постійного вдосконалення засобів захисту інформації в медіафайлах. На тлі зростання кількості кібератак, витоків персональних даних, атак типу "Man-in-the-Middle", а також масового використання deepfake-технологій, особливої актуальності набуває проблема забезпечення автентичності та цілісності аудіовізуального контенту. У статті, присвяченої атакам типу «людина посередині» (Man-In-The-Middle, MITM), які становлять одну з найсерйозніших загроз для комп'ютерної безпеки, основна увага приділяється аналізу способів реалізації MITM-атак у контексті моделі взаємодії відкритих систем (OSI), а також у популярних мережевих технологіях GSM та UMTS. Атаки класифікуються за такими параметрами, як розташування зловмисника в мережі, тип комунікаційного каналу та методи імітації (імперсонації). Для кожного з класів MITM-атак наводяться типові етапи їх виконання. Окремо розглянуто існуючі захисні засоби, проведено їх порівняльний аналіз, а також запропоновано систематизацію механізмів запобігання MITM-атакам. На завершення автори окреслюють перспективні напрямки подальших досліджень у цій сфері [9].

Медіафайли дедалі частіше використовуються як доказова база у судовій практиці, у сферах журналістики, правозастосування, розслідувань, військової розвідки, що вимагає надійних механізмів підтвердження їхнього джерела, захисту від маніпуляцій та несанкціонованого копіювання. На відміну від традиційного текстового контенту, медіа мають складну внутрішню структуру, що створює додаткові виклики в контексті прихованого вбудовування або

захисту від змін. Такі характеристики як колірні простори, частотні області, рівні стиснення та втрата якості під час передачі можуть істотно впливати на ефективність методів захисту. У зв'язку з цим все більшої уваги заслуговують алгоритми, які поєднують кілька методик: шифрування даних, цифрові водяні знаки, динамічну стеганографію, детекцію маніпуляцій із застосуванням машинного навчання. Методи захисту, засновані на мультимедійному шифруванні, стають дедалі актуальнішими, оскільки дозволяють безпечно обмінюватися контентом на цифрових платформах. У роботі, присвяченій шифруванню різних типів мультимедійних даних, розглянуто стан безпечних та конфіденційних схем шифрування, що застосовуються до цифрового мультимедіа, такого як цифрові зображення, цифрове відео та цифрове аудіо. Результати огляду сприятимуть кращому розумінню ефективності та надійності наявних рішень, а також стануть підґрунтям для розробки більш досконалих та безпечних схем мультимедійного шифрування в майбутньому [10].

Крім того, особливе значення ця тема має в умовах війни, коли медіаінформація використовується не лише для інформування, а й як інструмент психологічного впливу та дезінформації. Наявність технологій верифікації цифрових зображень та відео, захисту метаданих, а також можливості відстеження джерела походження медіа є критично важливими для забезпечення інформаційної безпеки держави, військових структур і цивільного суспільства.

Таким чином, вибір теми дослідження обумовлений не лише її високою науковою значущістю, але й чіткою практичною необхідністю розробки інструментів, здатних забезпечити цілісний захист інформації в медіафайлах в умовах сучасних гібридних і кіберзагроз.

У межах дослідження визначено проблему підвищення ефективності захисту інформації в медіафайлах, що включає необхідність обрання адаптивних та стійких до сучасних кіберзагроз методів. Актуальними є комбінації криптографічних протоколів, цифрових водяних знаків, стеганографії та машинного навчання з метою забезпечення цілісності, прихованості, незмінності джерела й атрибутації медіаконтенту. Головна задача – сформувати комплекс практичних рекомендацій щодо впровадження захисних технологій у системах, що обробляють цифрові зображення, аудіо- й відеофайли,

враховуючи наукову перспективу та реальні обмеження впровадження в Україні.

Захист інформації в медіафайлах реалізується через чотири основні напрями: криптографічні методи, цифрові водяні знаки, стеганографія та гібридні підходи, часто із застосуванням технологій штучного інтелекту.

Криптографічні технології зберігають конфіденційність даних, однак сам факт шифрування медіафайлу може привертати увагу потенційного зловмисника. Для подолання цих обмежень застосовують цифровий водяний знак – незначні модифікації медіа, що приховані від ока, але дозволяють перевірити походження й цілісність. Наприклад, M. Hamidi та співавтори описали комбіновану техніку на основі DWT, DCT та SIFT, яка забезпечує стійкість до геометричних трансформацій та компресії. У даній роботі запропоновано надійний гібридний метод водяного знакування зображень, що базується на дискретному вейвлет-перетворенні (DWT), дискретному косинусному перетворенні (DCT) та масштабно-інваріантному перетворенні ознак (SIFT). Основна мета дослідження – розробка стійкої до атак схеми маркування, здатної витримувати як обробку зображень, так і геометричні спотворення при збереженні високої непомітності. Для цього водяний знак вбудовується в середній діапазон коефіцієнтів DCT, отриманих з піддіапазону HL1 DWT, що забезпечує захист від стандартної обробки зображень. Алгоритм SIFT застосовується для виявлення ключових точок, які дозволяють компенсувати геометричні викривлення під час вилучення водяного знака. Результати численних експериментів підтверджують високу стійкість методу до атак та спотворень при збереженні високої якості зображення, а також демонструють його переваги порівняно з альтернативними підходами [11]. При цьому стеганографія – це приховане вбитовування повідомлення в сам файл. Класичні методи використовують LSB або частотні перетворення, а сучасні – нейронні мережі і “coverless” підходи, що не змінюють сам медіафайл. Серед них особливо перспективною є DiffStega, яка на основі дифузійної моделі створює stego-зображення без потреби оригінала, з високою стійкістю до атак та сильною прихованістю завдяки Noise-Flip і пароль-залежним запитом. DiffStega – це метод стеганографії зображень без покриття, який використовує

моделі дифузії для вбудовування секретних зображень. Стеганографія без покриття підвищує непомітність, усуваючи необхідність використання зображення-покриття. Замість того, щоб приховувати дані в існуючому зображенні, DiffStega генерує нове зображення, яке містить секретну інформацію. Цей підхід потенційно може запропонувати покращену безпеку та ємність порівняно з традиційними методами стеганографії [12; 13]. Інший напрям – DDIM-моделі для coverless стеганографії з реальним ключем, що дозволяє оновлювати stego-зображення без повторної передачі промтів. DDIM (Denoising Diffusion Implicit Models) – це варіант дифузійних моделей генерації зображень, який використовується для швидшого та керованого створення високоякісного контенту. У контексті coverless стеганографії – це сучасний підхід, коли секретна інформація не вбудовується у файл-носій, а відображається через генерацію або вибір певного медіаоб'єкта, що пов'язаний з цією інформацією. У дослідженні, проведеному авторами статті, присвяченій судово-медичному аналізу дипфейків, запропоновано нову форензичну методику ідентифікації навчального датасету (наприклад, *CelebA* або *FFHQ*), використаного для генерації зображень за допомогою Generative Adversarial Networks (GANs). Запропонована система ґрунтується на інтерпретованому аналізі ознак, який включає спектральні перетворення (Fourier/DCT), метрики колірного розподілу та дескриптори локальних ознак (SIFT). Застосування класифікаторів на основі машинного навчання (Random Forest, SVM, XGBoost) дало змогу досягти 98–99 % точності як у бінарній класифікації (реальні/синтетичні зображення), так і у багатокласовому визначенні джерела датасету. Дослідження підкреслює важливість частотних ознак у виявленні артефактів, характерних для конкретних датасетів, таких як патерни апсемплінгу та спектральні спотворення, тоді як колірні гістограми виявляють непрямі ознаки регуляризації, властиві тренуванню GAN. Окремо розглянуто правові та етичні аспекти, пов'язані з авторським правом, захистом персональних даних та дотриманням нормативних актів (наприклад, GDPR і каліфорнійського закону AB 602). Запропонований підхід сприяє підвищенню відповідальності та прозорості у сфері генеративного ШІ й може бути застосований у цифровій криміналістиці, модерації контенту та вирішенні спорів з інтелектуальної власності [14].

Інші приклади – системи, що командуються через YOLO і хвильові хеші. YOLO (You Only Look Once – Ви тільки дивитесь один раз) – це система виявлення об'єктів, а вейвлет-хеші пов'язані зі стеганографією зображень, яка може використовувати глибокі згорткові нейронні мережі. Виявлення об'єктів, для якого розроблена YOLO, включає ідентифікацію та визначення місцезнаходження об'єктів на зображенні або відео. Стеганографія, з іншого боку, зосереджена на приховуванні інформації в носії, такому як зображення, щоб запобігти виявленню. В одному із досліджень розглянуто сучасний підхід до стеганографії зображень на основі глибокого навчання. Зокрема, робота базується на розвитку методу **DeepWaveletFusion** шляхом інтеграції згорткових нейронних мереж (CNN) з дискретним вейвлет-перетворенням (DWT) для підвищення якості вбудовування та витягання прихованих даних. Запропонований метод, під назвою **DeepWaveletFusionToo**, вирізняється легкістю архітектури та ефективністю, оскільки використовує спеціалізований датасет зображень, трансформованих за допомогою DWT, і навчається з використанням функції втрат середньоквадратичної похибки (MSE), що суттєво знижує складність моделі та обчислювальні витрати. Результати експериментів підтверджують, що **DeepWaveletFusionToo** забезпечує вищу непомітність приховання інформації порівняно з попередником та демонструє конкурентну точність відновлення, що підкреслює простоту та ефективність підходу [15]. У другому дослідженні запропоновано новий підхід до відновлення стереозображень з використанням дифузійних моделей (DM), які раніше майже не застосовувалися в цій галузі через ряд обмежень. Зокрема, проблема подвійного відновлення зображень призводить до зростання обчислювальної складності, а більшість латентних дифузійних моделей орієнтовані на семантичний зміст і при цьому відкидають високочастотні компоненти, які є критично важливими для задач реставрації. Щоб подолати ці обмеження, було запропоновано **DiffStereo** – першу високочастотну обізнану дифузійну модель, спеціально розроблену для стереозображень. Метод передбачає навчання латентного представлення високочастотних компонентів (LHFR) високоякісних зображень. Дифузійна модель використовується для оцінки LHFR у парі стереозображень, після

чого ці представлення інтегруються у трансформерну мережу для відновлення зображень, забезпечуючи збагачення високочастотною інформацією. Важливо, що роздільна здатність LHFR зберігається на рівні з вхідними зображеннями, що дозволяє зберегти текстури, а зменшення кількості каналів знижує обчислювальні витрати. Також реалізовано спеціальний механізм позиційного кодування для забезпечення контекстної глибини під час відновлення. Результати експериментів демонструють, що поєднання дифузійної генеративної моделі з трансформерною архітектурою дозволяє **DiffStereo** перевершити сучасні підходи за точністю реконструкції та сприйняттям якості при вирішенні задач суперроздільності, деблурінгу та покращення зображень у слабкому освітленні [16] Zhiwei Kang описал метод, який за допомогою YOLO вибирає об'єкти для формування послідовності і ключів, що мають високий ступінь стійкості навіть при геометричних атаках. З огляду на переваги розпізнавання орієнтації у просторі скінченного ріджлет-перетворення було запропоновано метод вбудовування водяного знаку в зображення, стійкий до геометричних атак. На першому етапі оригінальне зображення піддається одномірному ріджлет-перетворенню. Далі виконується тривимірне вейвлет-декодування, після чого водяний знак вбудовується у коефіцієнти ріджлет-перетворення. Після цього проводиться зворотне відновлення зображення. Експериментальні результати демонструють високу стійкість запропонованого методу до геометричних атак, таких як обертання, масштабування, зсув і зсув по зрезу (shearing), а також до стандартних атак, включаючи JPEG-компресію, медіанну фільтрацію, додавання гаусового шуму та імпульсного шуму (сіль і перець) [17]. Ranjan та Kumar запропонували метод, що використовує DWT-хешування та Huffman-кодування для формування сховища, яке стійке до масштабування, шуму, JPEG-упаковки та інших атак. Стиснення даних, зокрема за допомогою дискретного вейвлет-перетворення (DWT) та кодування Хаффмана, має вирішальне значення для ефективного зберігання та передачі даних. Метою дослідження була розробка методу для покращення якості зображення та стиснення. Завдяки поєднанню PCA, DWT та канонічного кодування Хаффмана, запропонований метод перевершив існуючі методи з точки зору бітової швидкості, значень PSNR,

CR та SSIM. Об'єктивні та суб'єктивні тести доводять, що новий підхід є ефективнішим для стиснення зображень, що має потенціал для покращення якості зберігання та передачі даних зображень через цифрові мережі [18].

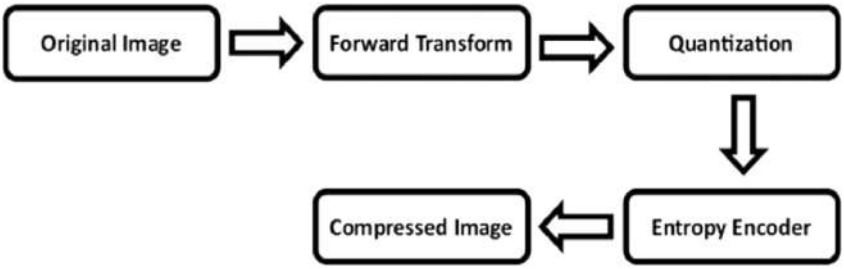


Рисунок 2 – Блок-схема стиснення зображень з втратами в загальному вигляді [18]

Обробка сигналів у зашифрованому середовищі – це перспективна технологія, що має на меті забезпечення захисту цінної інформації під час обробки у потенційно незахищених середовищах. У цьому дослідженні запропоновано метод реалізації дискретного вейвлет-перетворення (DWT) та багатороздільного аналізу (MRA) у гомоморфно зашифрованому середовищі. В одній з робіт спочатку представлено концептуальну модель для виконання DWT та оберненого DWT (IDWT) у зашифрованому вигляді, після чого проведено аналіз впливу розширення даних та квантувальних похибок у межах цієї моделі. З метою зменшення розширення даних – одного з критично важливих факторів для практичного застосування – автори запропонували спеціальний підхід до оптимізації виконання DWT та IDWT. Запропонований метод дозволяє здійснювати багаторівневе DWT/IDWT із мінімальним збільшенням об'єму даних у гомоморфно зашифрованому середовищі. Крім того, розроблено нову процедуру обробки сигналів, у якій на завершальному етапі застосовується метод оберненого множення для обмеження розширення даних. На прикладі 2D- вейвлет-перетворення Хаара було проведено низку експериментів, що продемонстрували ефективність запропонованого підходу для захищеної обробки зображень.

Також виконано аналіз обчислювальної складності та порівняльне дослідження. Наскільки відомо авторам, це перше повідомлення про реалізацію DWT та MRA у зашифрованому середовищі [19]. Значну наукову увагу також привертають гібридні стего-шифрувальні методи, де застосовують DCT/DWT + GAN. Наприклад, поєднання DCT із GAN для прихованого передавання даних у медіа з підвищеною стійкістю до атак та покращення співвідношення швидкість/якість.

Таким чином, сучасні дослідження демонструють поступ до мультидисциплінарних рішень, які комбінують криптографію, стеганографію, цифрові водяні знаки, метадані та машинне навчання. Ці гібриди дозволяють забезпечити ключові характеристики захисту: прихованість, стійкість до атак, перевірку походження контенту та збереження високої якості медіа. Для практики українських дослідників це означає не лише освоєння теоретичних алгоритмів, а й їхню адаптацію до локальних умов: обмеження по обчислювальних ресурсах, законодавчі рамки, інтеграцію в наявні IT-системи.

Класифікація медіафайлів та типів загроз.

Цифрові медіафайли умовно поділяють на зображення, аудіо та відео – кожен із них має унікальні характеристики, що впливають на уразливості. Зображення зазвичай містять дані в пікселях або частотних компонентах (DCT, DWT), аудіо – у тимчасових та спектральних доменах, відео – це послідовність кадрів із додатковими метаданими та руховими векторами. Загрози include стеганографії, де вміст файлу змінюється для прихованого передавання інформації, та stegomalware – випадки, коли зловмисний код вбудовується в мультимедіа й запускається після завантаження. Крім того, є атаки на цілісність: зміни кадрів, компресія, геометричні модифікації, deepfake-маніпуляції, що викликають серйозну загрозу автентичності.

У контексті інформаційної безпеки медіафайли – це цифрові об’єкти, які зберігають візуальну або аудіовізуальну інформацію. Їхня специфіка полягає в складній структурі (пікселі, кодеки, частотні компоненти, метадані), що обумовлює велику кількість векторів атак. Для ефективного захисту перш за все необхідно класифікувати як типи самих файлів, так і загрози, що їм притаманні.

Класифікація медіафайлів:

– Цифрові зображення (JPEG, PNG, BMP, TIFF та ін.). Мають структуру, засновану на матриці пікселів, із можливим використанням частотних перетворень (наприклад, JPEG-компресія використовує DCT). Сприйнятливі до змін у просторовому та частотному доменах.

– Аудіофайли (MP3, WAV, FLAC). Містять звукову інформацію, представлену у часовій або частотній формі. Часто піддаються втратній компресії, що ускладнює вбудовування прихованої інформації без спотворень.

– Відеофайли (MP4, AVI, MKV). Є поєднанням зображень (кадрів) та аудіо, іноді – субтитрів. Складаються з великої кількості компонентів (ключові кадри, рухові вектори, метадані), що відкриває численні вектори для атак і прихованого передавання даних.

– Гібридні медіа (інтерактивні, 3D-файли, AR/VR-контент). Новий тип медіа, що поєднує зображення, відео, аудіо, навігаційні дані, сенсорні метадані. Зокрема, файли формату USDZ або glTF мають складну структуру й передбачають нові ризики – як на рівні контенту, так і метаданих.

Тепер перейдемо до більш детального розгляду кожного типу медіафайлів. Систематизуємо ґрунтовний опис основних форматів мультимедійних даних, з урахуванням їхніх структурних особливостей, способів кодування інформації та специфіки використання в контексті інформаційної безпеки. Це дозволить краще зрозуміти, які характеристики визначають вразливість кожного формату до можливих загроз або маніпуляцій.

Цифрові зображення – це мультимедійні об’єкти, які зберігають інформацію у вигляді двовимірних масивів пікселів. Кожен піксель має колірні або яскравісні значення, що формують візуальне представлення зображення. Існує кілька поширених форматів цифрових зображень, кожен із яких має власні особливості структури та способи кодування даних:

BMP (Bitmap Image File) – формат із простим представленням зображення у вигляді незжатого масиву пікселів, що зберігає інформацію у растровому форматі. Підходить для високої якості, але займає багато пам’яті.

JPEG (Joint Photographic Experts Group) – широко використовуваний формат із стисненням із втратами, що застосовує дискретне косинусне перетворення (DCT) для ефективного стиснення. Забезпечує баланс між якістю та розміром файлу.

PNG (Portable Network Graphics) – формат із безвтратним стисненням, підтримує прозорість та зберігає якість зображення без втрат, використовуючи методи стиснення на основі алгоритмів DEFLATE.

TIFF (Tagged Image File Format) – гнучкий формат, який підтримує як стиснення без втрат, так і з втратами, широко застосовується у професійній графіці та медичних зображеннях.

GIF (Graphics Interchange Format) – формат з обмеженою палітрою кольорів (до 256), що підтримує анімацію, але використовується переважно для простих графічних елементів.

RAW – формат, що містить необроблені дані, отримані безпосередньо з сенсора камери, забезпечує максимальну якість і гнучкість для подальшої обробки.

HEIF/HEIC (High Efficiency Image Format) – сучасний формат із високоефективним стисненням, що дозволяє зберігати зображення з кращою якістю при меншому розмірі файлу, часто використовується на мобільних пристроях.

Ці формати можуть базуватись на різних типах кодування, зокрема:

Просторове кодування – безпосереднє збереження пікселів у матриці (BMP, RAW).

Частотне кодування – застосування перетворень, таких як дискретне косинусне перетворення (DCT, JPEG) або дискретне вейвлет-перетворення (DWT, TIFF), що дозволяє більш ефективно стиснути дані.

Структура і формат впливають на стійкість зображень до змін і на вибір методів захисту, таких як цифрове водякування, стеганографія, шифрування та інші. З позиції авторського права, вбудовані у медіафайли дані слугують надійним засобом запобігання претензіям третіх осіб на право власності. Водночас цифровий водяний знак має залишатися конфіденційним, доступним лише автору, та зберігати свою стійкість до різноманітних видів атак і спроб видалення [20]. Крім того, характеристики форматів визначають рівень

втрат при стисненні, можливість збереження метаданих (EXIF, XMP), підтримку прозорості, анімації або багатошаровості. Розуміння цих особливостей є критично важливим для розробки ефективних систем захисту інформації, оскільки різні формати зображень по-різному реагують на атаки і втручання у просторовому або частотному домені. Серед різноманітних підходів до захисту зображень за допомогою цифрових водяних знаків, особливу увагу приділяють методам, що базуються на частотних перетвореннях, таких як дискретне косинусне перетворення. В одній з робіт розглянуто підхід, у якому з метою забезпечення стійкості цифрових водяних знаків до різноманітних спотворень та стиснення зображень, при збереженні прийнятної візуальної якості, було запропоновано нові критерії вибору низькочастотних коефіцієнтів DCT, придатних для модифікації, а також умови, за якими здійснюється вбудовування цифрових водяних знаків [21].

Аудіофайли є важливим типом медіа, що містить звукову інформацію у цифровій формі, представлено у часовому або частотному домені. У контексті інформаційної безпеки аудіоформати відіграють особливу роль у задачах стеганографії, цифрового водякування, автентифікації й виявлення підробок. Залежно від формату, аудіофайли відрізняються типом стиснення, якістю відтворення, структурою та можливостями прихованого вбудовування даних [22].

WAV (Waveform Audio File Format) – це один із найпростіших та найбільш придатних для аналізу аудіоформатів, що зберігає сигнал у незжатому вигляді. Завдяки відсутності втрат під час збереження, WAV забезпечує повне збереження амплітудно-часової структури сигналу, що робить його ідеальним кандидатом для застосування цифрових підписів або вбудовування водяних знаків без спотворень.

MP3 (MPEG-1 Audio Layer III) є найпоширенішим форматом аудіо з втратами. Він застосовує складні методи перетворення, зокрема модифіковане дискретне косинусне перетворення (MDCT), щоб зменшити розмір файлу, зберігаючи прийнятну якість звуку. Проте під час такого стиснення втрачається частина звукової інформації, що ускладнює приховане вбудовування даних і робить такі файли менш надійними для задач захисту інформації.

FLAC (Free Lossless Audio Codec) – це формат зі стисненням без втрат, який зберігає високу якість сигналу при значно меншому обсязі, ніж WAV. FLAC підтримує додаткові метадані й є придатним для безпечного зберігання і обробки, що робить його ефективним для вбудовування захищеної інформації або використання в системах контролю автентичності.

AAC (Advanced Audio Coding) є сучасним форматом із втратами, який забезпечує кращу якість звуку порівняно з MP3 при однаковому бітрейті. Він широко застосовується у потоковому аудіо та відео (наприклад, YouTube, iTunes). Через стиснення з втратами цей формат також обмежено придатний до надійної стеганографії, однак сучасні методи дозволяють частково компенсувати ці обмеження за допомогою адаптивних алгоритмів.

OGG (Ogg Vorbis) – відкритий формат стиснення з втратами, відомий своєю ефективністю та високою якістю при низьких бітрейтах. Він є альтернативою MP3 та AAC і використовується в ігрових движках, стримінгових платформах, вільному програмному забезпеченні. Його гнучка структура теоретично дозволяє вбудовування прихованих даних, однак, як і інші втратні формати, він потребує стійких методів до перетворень і рекомпресії [23].

OPUS – це формат нового покоління, оптимізований для голосового та музичного сигналу з низькою затримкою, що використовується у VoIP, відеозв'язку, месенджерах (наприклад, Discord, Zoom). Він автоматично адаптується до типу звуку та каналу. У сфері інформаційної безпеки OPUS розглядається як потенційна платформа для захищеної комунікації, однак використання в стеганографії обмежене через динамічну структуру та активну компресію.

AIFF (Audio Interchange File Format) – це формат без втрат, подібний до WAV, розроблений компанією Apple. Він застосовується переважно в професійних аудіостудіях та продуктах Apple. Завдяки збереженню повної звукової інформації він придатний для вбудовування цифрових водяних знаків, автентифікації та інших задач ІБ, де критичною є точність і достовірність сигналу.

Ці характеристики мають вирішальне значення при виборі методу захисту звукової інформації, зокрема у задачах аудіостеганографії, цифрового підпису та детектування змін. Формати зі стисненням без

втрат (WAV, FLAC, AIFF) забезпечують кращу сумісність із надійними ІБ-методами, тоді як втратні формати потребують спеціалізованих рішень, стійких до спотворень [24].

Відеофайли представляють собою складний медіаформат, який об'єднує послідовність зображень (кадрів), аудіодоріжки та іноді субтитри, створюючи інтегрований мультимедійний контент. Найпоширенішими форматами є MP4, AVI, MKV, кожен із яких має власну структуру і механізми стиснення, що впливають на якість та обсяг файлу. Відео містить велику кількість складових елементів, серед яких ключові кадри, рухові вектори, звукові доріжки, інформація про кадрування і метадані, що описують технічні параметри та контекст відтворення. Така багатошарова структура відкриває численні можливості для різноманітних кібератак, зокрема підміни, модифікації або прихованого передавання інформації. У контексті інформаційної безпеки це створює як виклики, так і потенціал для застосування методів захисту – цифрових водяних знаків, стеганографії, криптографічних засобів, які можуть бути вбудовані не лише у візуальну або аудіодоріжку, а й у метадані чи параметри стиснення. Водночас складність формату відео та велика об'ємність даних вимагають розробки спеціалізованих алгоритмів, які забезпечують ефективний захист без втрати якості та сумісності з пристроями відтворення. Та через високу ємність та складну структуру відео, їм часто надається перевага над іншими носіями, такими як зображення, текст або аудіо, для стеганографії. Існує безліч методів приховування секретної інформації у відео, кожен з яких має свої якості та недоліки. Однак у літературі бракує достатньої кількості оглядових статей, які б розглядали всі методи. На основі методу вбудовування методи відеостеганографії класифікуються на дві категорії, а саме: методи просторової області та методи частотної області [25].

MP4 (MPEG-4 Part 14) є одним із найпоширеніших форматів для зберігання відео, який підтримує широкий спектр кодеків, включаючи H.264 і H.265. Цей формат відомий своєю здатністю ефективно компресувати відеодані з мінімальними втратами якості, що робить його популярним для потокового передавання, мобільних пристроїв і вебплатформ. MP4 підтримує інтеграцію аудіодоріжок, субтитрів та метаданих, що дозволяє створювати багатофункціональний

мультимедійний файл. З огляду на складність структури, захист інформації в MP4 вимагає врахування можливих точок впливу – від окремих кадрів і аудіо до метаданих, а також механізмів стиснення і шифрування.

AVI (Audio Video Interleave) – це один із перших форматів відеофайлів, розроблений компанією Microsoft, який зберігає аудіо та відео в одному контейнері без сильного стиснення. Через відсутність ефективних методів стиснення, файли AVI зазвичай мають великий розмір, що обмежує їх застосування у потоковому відео. Проте цей формат часто використовується в архівуванні та обробці відео завдяки своїй простоті. Через відкриту структуру AVI може бути вразливим до модифікацій метаданих і заміни відеопотоку, що ставить завдання розробки надійних засобів цифрового підпису та водяних знаків для забезпечення цілісності та автентичності.

MKV (Matroska Video) – це сучасний відкритий формат-контейнер, який здатен зберігати необмежену кількість аудіо, відео та субтитрових потоків у одному файлі. MKV підтримує високоякісне відео з різними кодеками та дозволяє додавати багатошарові метадані, що робить його популярним у середовищах, де потрібна гнучкість і висока функціональність. Його багатокомпонентна структура створює широкі можливості для вбудовування захисних механізмів – від цифрових водяних знаків у відеопотоках до складних стеганографічних рішень у метаданих. Водночас складність і варіативність MKV вимагає ретельної розробки засобів контролю доступу і цілісності для забезпечення безпеки в мультимедійному середовищі.

Інші менш поширені формати, такі як MOV (формат, розроблений Apple), WMV (Windows Media Video), а також WEBM, часто застосовуються у специфічних сферах, наприклад, у професійному відеомонтажі або вебвідео. Вони мають власні особливості структури, компресії та підтримки метаданих, що також визначає особливості їх захисту від несанкціонованого доступу і модифікацій. Мета полягає в тому, щоб приховати дані таким чином, щоб їх присутність була невиявною. Це досягається шляхом зміни характеристик відео, таких як значення пікселів або бітові потоки, без суттєвого впливу на його візуальну якість [26].

Усі ці формати відрізняються за структурою, методами компресії, підтримкою метаданих і складністю, що безпосередньо впливає на вибір та ефективність методів інформаційного захисту. Забезпечення цілісності, автентичності і конфіденційності відеоконтенту вимагає комплексного підходу, що охоплює як криптографічні механізми, так і специфічні мультимедійні технології, адаптовані до особливостей кожного формату.

Гібридні медіафайли представляють собою сучасний та складний тип мультимедійного контенту, який поєднує у собі різні види даних – зображення, відео, аудіо, а також додаткові навігаційні та сенсорні метадані, що забезпечують інтерактивність і глибоку взаємодію користувача з контентом. Такі файли використовуються в технологіях доповненої (AR) і віртуальної реальності (VR), тривимірній графіці та інтерактивних презентаціях. Формати, зокрема USDZ і glTF, розроблені для забезпечення високої сумісності та зручності візуалізації складних 3D-моделей разом із пов'язаними елементами, такими як текстури, анімації і фізичні властивості. Через багатокомпонентність і складну ієрархію структури, гібридні медіа несуть нові виклики для інформаційної безпеки. Загрози можуть проявлятися як на рівні безпосереднього вмісту – наприклад, уразливості в 3D-моделях, можливість впровадження шкідливого коду чи маніпуляцій із сенсорними даними, так і на рівні метаданих, що часто містять інформацію про місцезнаходження, користувача або сценарії використання. Забезпечення захисту таких медіа вимагає розробки комплексних методів, які поєднують криптографічні засоби, контроль цілісності, автентифікацію елементів контенту та механізми безпечного обміну даними, адаптовані до специфіки інтерактивних і тривимірних форматів.

Класифікація загроз.

Усі загрози можна умовно поділити на п'ять основних категорій, кожна з яких має підвиди, специфічні для типу медіафайлу:

Загрози конфіденційності. Незаконне копіювання або передача медіафайлів (наприклад, перехоплення під час передачі через нешифрований канал). Несанкціонований доступ до локального або хмарного сховища.

– Загрози цілісності. Маніпуляції зі вмістом (зміна пікселів, редагування аудіодоріжки, вставка або видалення кадрів). Атаки типу

“content tampering” (підробка відео- або фотофактів). Вплив через стиснення, обрізання, перекодування.

– Загрози автентичності. Видавання підробленого медіафайлу за оригінальний. Відсутність або підміна цифрового підпису. Фальсифікація метаданих (дати створення, GPS-координати, пристрій зйомки).

– Загрози прихованого передавання інформації (стеганозагрози). Вбудовування прихованих повідомлень (текстів, команд, шкідливого ПЗ) у пікселі, спектральні компоненти, звук або рухи відео.

Використання форматів із надлишковими метаданими або каналами передачі (наприклад, невидимі субтитри чи просторове шумове кодування).

– Шкідливі програми в медіафайлах. Stegomalware: використання медіафайлів як носіїв шкідливого коду, що активується під час відкриття. Polyglot-файли: комбіновані формати (наприклад, PNG + JS), які інтерпретуються по-різному в різних середовищах.

Таблиця 1 – Вектори атак за типами медіафайлів

Тип медіафайла	**Основні загрози**	**Приклади атак**
Зображення	Стеганографія, зміна метаданих, deepfake	LSB-атаки, деструктивне редагування
Аудіо	Вбудовування даних, прихована команда	Audio-steganography, reverse speech
Відео	Модифікація ключових кадрів, вставка фальшивок	Deepfake-відео, insertion attacks
Гібридні медіа	Зміна логіки сцени, викривлення AR/VR-ефекту	GPS-підміна, атаки на сценографію

Джерело: авторська розробка

Методи захисту інформації в медіафайлах

Захист інформації у медіафайлах базується на принципах конфіденційності, цілісності, автентичності та стійкості до несанкціонованого доступу або маніпуляцій. Залежно від типу загроз, застосовують різні технічні підходи – від шифрування до штучного інтелекту [27].

1. Шифрування.

Шифрування – найпоширеніший метод забезпечення конфіденційності. До медіафайлів застосовують як симетричні (AES, Twofish),

так і асиметричні (RSA, ECC) алгоритми. Особливою є проблема обробки великих обсягів даних, адже відео або зображення можуть мати значний розмір, що впливає на продуктивність.

Наприклад, AES-256 широко застосовується в інструментах типу VeraCrypt [28], де можливо створити захищені контейнери або шифрувати цілі логічні диски. Хоча такий метод добре захищає від витоку, він не забезпечує прихованості чи підтвердження.

2. Цифрові підписи.

Цифрові підписи використовуються для перевірки автентичності джерела та цілісності файлу. Вони ґрунтуються на асиметричній криптографії: хеш медіафайлу підписується закритим ключем автора і може бути перевірений будь-ким із публічним ключем.

Стандарти, як-от X.509, або формати електронного підпису для зображень/відео (наприклад, CMS/PKCS#7, Cryptographic Message Syntax) дозволяють інтегрувати підписи у робочі процеси цифрової журналістики, правових доказів та електронного документообігу.

3. Цифрові водяні знаки.

Це метод вбудовування унікального ідентифікатора (водяного знака) в медіафайл, часто – приховано. Залежно від області вставлення, розрізняють:

– просторові методи – зміни пікселів або амплітуд;
– частотні методи – вбудовування у коефіцієнти DCT, DWT, SVD.

Водяні знаки бувають видимими (наприклад, логотип телеканалу) або невидимими (що не впливають на візуальне/звукове сприйняття). Один із популярних інструментів – Digimarc, що реалізує частотне кодування для стійкості до атаки JPEG, масштабування тощо [29].

Сучасні системи використовують і адаптивні методи, де водяний знак підлаштовується під особливості контенту для підвищення непомітності.

4. Стеганографія.

Цей метод дозволяє приховано передавати інформацію в межах звичайного медіафайлу. На відміну від шифрування, стеганографія приховує сам факт існування повідомлення.

Основні методи стеганографії включають різноманітні підходи, що забезпечують приховане вбудовування інформації у цифрові медіафайли, з урахуванням їх структури та особливостей. Один із

найпоширеніших методів – LSB (Least Significant Bit), який полягає у зміні найменш значущих бітів пікселів у зображеннях або аудіосемплах. Цей спосіб відзначається простотою реалізації та високою швидкістю, але може бути вразливим до деяких видів атак і обробки файлів.

Інший потужний напрям – стеганографія у домені перетворень (transform domain). Тут прихована інформація вбудовується не у безпосередньо пікселі чи аудіодані, а у коефіцієнти перетворень, таких як дискретне косинусне перетворення (DCT), дискретне вейвлет-перетворення (DWT) або інші частотні аналоги. Такий підхід забезпечує кращу стійкість до стиснення, фільтрації та різних маніпуляцій із медіа, адже модифікації розподілені у більш стійкому до змін частотному просторі.

Сучасним інноваційним напрямком є coverless стеганографія – метод, що принципово відрізняється від традиційних. У цій технології секретне повідомлення не вставляється безпосередньо у носій, а відтворюється за допомогою моделей штучного інтелекту, таких як генеративні моделі або дифузійні моделі (наприклад, DiffStega). Такий підхід усуває традиційні сліди втручання в носій і значно підвищує захищеність інформації від виявлення.

Крім того, існують і інші важливі методи, серед яких частотне модуляційне кодування, фазове кодування, методи на основі хаотичних систем, а також стеганографія, заснована на змінах текстурних або статистичних властивостей медіаоб'єктів. Всі ці методи мають свої переваги і недоліки, а їх вибір залежить від конкретних завдань захисту, типу носія і вимог до стійкості, ємності та непомітності вбудованої інформації.

Такі інструменти, як OpenStego (LSB, прості водяні знаки) або нейронні системи на базі GAN (наприклад, SteganoGAN, DiffStega) використовуються як для безпечної передачі, так і для обхідних цілей зловмисниками.

5. Штучний інтелект та машинне навчання (AI/ML)

AI-системи застосовуються як для захисту, так і для атак. Для захисту використовують глибокі згорткові мережі (CNN), autoencoder-архітектури, GAN для вбудовування інформації або розпізнавання фальсифікацій.

Diffusion-based methods дозволяють створювати високо приховані й природні образи з інтегрованими повідомленнями.

AI-антифейкові системи використовуються для детекції змінених зображень/відео, фальсифікацій deepfake.

На практиці це дозволяє захистити файли не лише від модифікацій, а й забезпечити автоматичне виявлення підробок у цифрових потоках.

6. Блокчейн

Блокчейн дозволяє забезпечити незмінність метаданих про медіафайл. Наприклад, при завантаженні фото у блокчейн зберігається:

– його хеш;
– ідентифікатор автора;
– часовий штамп;
– геолокація;
– цифровий підпис.

Такі рішення застосовуються в цифровій журналістиці (проєкти типу Content Authenticity Initiative або CAI – це міжгалузева спільнота, заснована компанією Adobe у 2019 році, метою якої є боротьба з дезінформацією в інтернеті шляхом додавання до цифрового контенту інформації про походження та історію змін.) та судовій експертизі. Наприклад, система TruePic документує створення фото/відео верифікованими способами із прив'язкою до блокчейну. TruePic Vision: Цифрова платформа для інспекції, що забезпечує простий та ефективний спосіб отримання перевірених, достовірних зображень та відео безпосередньо від клієнтів. Вона використовується в процесах розгляду бізнес-претензій та страхових вимог. Вона замінює дорогі та трудомісткі візити на місце, пропонуючи рішення для віртуальної інспекції [29].

У сучасній практиці захисту медіаконтенту активно використовуються спеціалізовані інструменти, що реалізують різні підходи – від криптографії до водяного маркування та стеганографії. Розглянемо найбільш відомі з них з огляду на функціональність, ефективність та сферу застосування.

VeraCrypt – це програмне забезпечення з відкритим кодом для шифрування файлів, томів та цілих дисків. Його перевагою є використання стійких алгоритмів (AES, Serpent, Twofish) і підтримка прихованих томів, що додає додатковий рівень обфускації. VeraCrypt дозволяє

ефективно захищати великі обсяги медіафайлів від несанкціонованого доступу, особливо у випадках передачі через зовнішні носії або хмарні сервіси. Завдяки підтримці Windows, macOS і Linux, VeraCrypt є кросплатформним інструментом широкого застосування [28].

Таблиця 2 – Методи захисту інформації у медіафайлах

Метод	Ціль	Інструмент / Приклад	Переваги	Обмеження
Шифрування	Конфіденційність та цілісність даних	VeraCrypt, AES-256	Сильний захист, широко підтримується	Високі вимоги до обчислювальних ресурсів
Цифровий підпис	Перевірка автентичності та цілісності	PKCS#7, X.509	Юридична значущість, інтеграція в документообіг	Не приховує дані, потребує інфраструктури ключів
Цифровий водяний знак	Маркування авторства або джерела	Digimarc, DWT/DCT	Стійкість до перетворень, непомітність	Може бути видалений або змінений
Стеганографія	Прихована передача інформації	OpenStego, SteganoGAN	Невидимість передавання, можна комбінувати із шифруванням	Вразлива до стегоаналізу, обмежена місткість
AI/ML	Автоматична детекція/ вбудовування даних	DiffStega, CNN, GAN	Гнучкість, масштабованість, антифейкові системи	Складність реалізації, потребує навчання моделей
Блокчейн	Фіксація походження та незмінності	TruePic, Content Authenticity Initiative	Незмінність записів, прозорість	Висока складність, низька швидкодія

Джерело: авторська розробка

Digimarc – комерційна платформа цифрового водяного маркування, яка пропонує стійке та малопомітне вбудовування маркерів у зображення, відео та аудіо. Ці маркери можуть містити метадані, авторські права або дані автентифікації. Перевага Digimarc – високий рівень захисту від видалення чи фільтрації, що робить її ефективною для захисту цифрових медіа у видавництвах, фотоагентствах і художніх проєктах. Платформа інтегрується з Adobe Photoshop, що спрощує робочі процеси дизайнерів [30].

OpenStego – безкоштовний інструмент для стеганографії, що дозволяє приховано вбудовувати текстові повідомлення у зображення. OpenStego підтримує як класичну стеганографію, так і вбудовування водяних знаків. Це робить його корисним для базової дослідницької та освітньої діяльності, однак алгоритм може бути вразливим до стегоаналізу. OpenStego поширюється з відкритим кодом і може адаптуватися під різні потреби [31].

TruePic – сучасне комерційне рішення, що базується на фіксації походження зображень і відео через хешування та збереження у блокчейні. Система дозволяє автоматично перевіряти, чи не змінювався файл з моменту створення. Це критично важливо для цифрової журналістики та судової експертизи, де доказове значення медіа залежить від їхньої цілісності. Компанія TruePic є частиною ініціативи *Content Authenticity Initiative*, яка підтримується Adobe, Microsoft та іншими [29].

SteganoGAN – один із найсучасніших підходів до глибокої стеганографії на основі генеративних нейронних мереж (GAN). Завдяки машинному навчанню SteganoGAN забезпечує високу якість зображення при збереженні прихованого повідомлення. Це рішення актуальне для захищеного обміну інформацією в умовах високої загрози перехоплення. Використання глибокого навчання дозволяє системі адаптуватися до виявлення нових стегоаналізаторів.

Таким чином, кожен інструмент має свою нішу застосування. Реальна ефективність кожного інструмента залежить не лише від технічних характеристик, а й від контексту використання, підготовки користувача та середовища розгортання.

З огляду на стрімке зростання ролі цифрових медіа в критично важливих сферах, таких як журналістика, оборона та судочинство, особливу увагу слід приділити практичному застосуванню методів захисту інформації в медіафайлах.

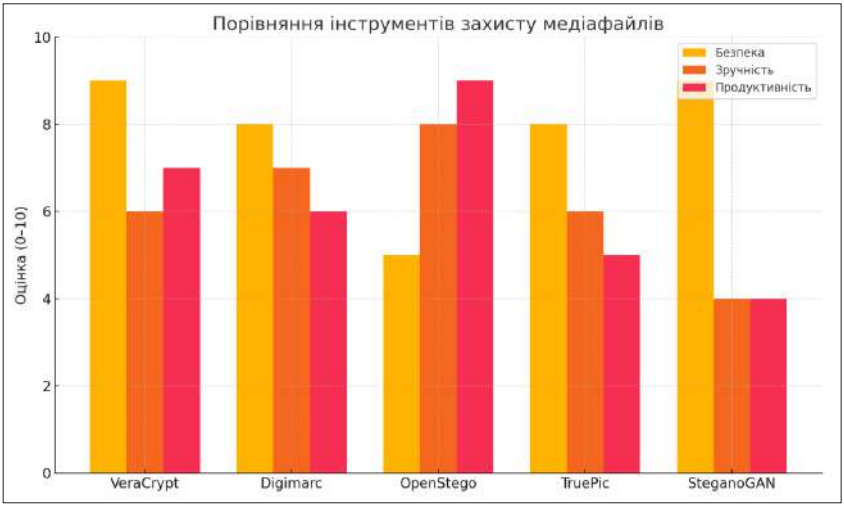


Рисунок 3 – Порівняння інструментів захисту медіафайлів

Джерело: авторська розробка

У цьому розділі розглянуто приклади використання сучасних інструментів і технологій, що забезпечують автентичність, цілісність і конфіденційність цифрового контенту в реальних умовах. У сфері сучасної журналістики захист медіаконтенту є критично важливим не лише для збереження авторських прав, а й для гарантування достовірності джерела в умовах дезінформації. Використання цифрових водяних знаків (Digimarc, Adobe CAI) забезпечує маркування зображень на рівні метаданих, що дозволяє верифікувати їхнє походження у процесі поширення. Системи перевірки достовірності фото- та відеоматеріалів, як-от Project Origin або Content Authenticity Initiative, стали відповіддю на поширення фейкових новин та deepfake-контенту. Журналісти-розслідувачі використовують стеганографію для прихованої передачі чутливої інформації в репресивних країнах, де пряме шифрування може викликати підозру.

У контексті гібридної війни та інформаційних операцій медіа є не лише способом передачі новин, а й об'єктом технічної атаки. Шифрування файлів (VeraCrypt, LUKS, BitLocker) використовується для захисту зображень і відео з дронів, які передаються через польові

станції. AI/ML-інструменти для виявлення маніпуляцій зображеннями використовуються у військовій розвідці для верифікації матеріалів з відкритих джерел (OSINT). Блокчейн для логування метаданих (наприклад, через OpenTimestamps) забезпечує незмінність походження відео, що фіксують військові злочини.

Застосування цифрових медіа у криміналістиці потребує не лише правильного збирання, а й доказового ланцюга автентичності. Цифрові підписи й гешування медіафайлів (SHA-256, PGP) дозволяють зафіксувати момент створення й гарантувати, що файл не змінювався. Інструменти як TruePic або Amber Authenticate дозволяють здійснити верифікацію фото/відео ще на етапі зйомки. Судові експерти застосовують стегоаналіз і цифрову атрибуцію, щоб визначити джерело підроблених фото або факт фальсифікації відео.

Таблиця 3 – Приклади інструментів захисту медіафайлів за сферами застосування

Сфера застосування	Інструмент / Технологія	Призначення
Журналістика	Digimarc, CAI (Adobe)	Верифікація автентичності зображень
Журналістика	SteganoGAN	Приховане передавання інформації
Військова справа	VeraCrypt, BitLocker	Захист відео з дронів
Військова справа	OpenTimestamps	Незмінність походження відео
Судова експертиза	TruePic, Amber Authenticate	Криптографічна фіксація моменту зйомки
Судова експертиза	PGP, SHA-256	Доказова автентифікація медіаданих

Джерело: авторська розробка

Висновки. У межах проведеного дослідження було здійснено системний аналіз сучасних методів захисту інформації в медіафайлах, зокрема шифрування, цифрових водяних знаків, стеганографії, блокчейн-технологій та засобів штучного інтелекту. Встановлено, що ефективний захист мультимедійного контенту в умовах зростаючих кіберзагроз вимагає використання комбінованих підходів, здатних

забезпечити автентичність, цілісність та конфіденційність цифрових об'єктів.

Наукову новизну роботи становить узагальнення та структуризація існуючих методів захисту медіафайлів, запропонована класифікація загроз, а також аналіз можливостей інтеграції технологій глибокого навчання з традиційними інструментами безпеки. Окрему увагу приділено використанню цифрових водяних знаків для аудіофайлів, автоенкодерів для стеганографії та інтелектуальних систем детектування змін у зображеннях.

Практична цінність дослідження полягає у формулюванні рекомендацій щодо захисту цифрових архівів, медіаресурсів у журналістиці, правозастосовній діяльності, кіберрозвідці, військовій справі та цифровій експертизі. Запропоновані методи та інструменти можуть бути використані як для підвищення інформаційної безпеки організацій, так і для особистого захисту цифрового контенту.

Подальші дослідження доцільно зосередити на питаннях глибшої інтеграції штучного інтелекту з криптографічними і стеганографічними підходами, розробці адаптивних методів захисту для нових форматів медіа (наприклад, AR/VR-контенту), а також на просуванні принципів цифрової кібергігієни серед користувачів як ключового чинника протидії дезінформації та технічному втручанню.

Список використаних джерел

1. Прокопович-Ткаченко Д. Глибокі автоенкодери для приховування інформації: сучасні підходи та перспективи розвитку. *Кібербезпека: освіта, наука, техніка* : електронне фахове наукове видання. 2025. Т. 4. №. 28. С. 150–161.

2. Світловський Є. В. Моделі і алгоритми створення цифрових водяних знаків для аудіо-файлів. *Перспективні технології та прилади.* 2024. Т. 1. №. 24. С. 99–106.

3. Бондарь І. В. Метод цифрової стеганографії мультимедіа-об'єктів : дис., Тернопіль : ЗУНУ, 2023.

4. Qureshi A., Megías Jiménez D. (). Blockchain-based multimedia content protection: Review and open challenges. *Applied Sciences.* 2020. № 11 (1), 1. DOI: 10.3390/app11010001 [in English].

5. Tomosada H., Kudo T., Fujisawa T., Ikehara M. GAN-Based Image Deblurring Using DCT Loss With Customized Datasets. *IEEE Access.* 2021. Vol. 9. P. 135224–135233. DOI: 10.1109/access.2021.3116194

6. Shen Z. et al. Deep semantic face deblurring. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018. P. 8260–8269.

7. Ravichandran C., Vajravelu A., Panda S., Degadwala S. D. Data hiding using video steganography. *International Journal of Electronic Security and Digital Forensics.* 2024. Vol. 16 (1). P. 112–123. DOI: 10.1504/ijesdf.2024.136018

8. Meng L., Jiang X., Sun T. A review of coverless steganography. *Neurocomputing*. 2024. Vol. 566. P. 126945.

9. Conti M., Dragoni N., Lesyk V. A survey of man in the middle attacks. *IEEE communications surveys & tutorials*. 2016. Vol. 18 (3). P. 2027–2051.

10. Hosny K. M., Zaki M. A., Lashin N. A., Fouda M. M., Hamza H. M. Multimedia Security Using Encryption: A Survey. *IEEE Access.* 2023. Issue 11. P. 63027–63056. DOI: 10.1109/access.2023.3287858

11. Hamidi M. et al. A hybrid robust image watermarking method based on DWT-DCT and SIFT for copyright protection. *Journal of Imaging.* 2021. Vol. 7 (10). P. 218.

12. Yang Y., Liu Z., Jia J., Gao Z., Li Y., Sun W., Liu X., Zhai G. DiffStega: Towards Universal Training-Free Coverless Image Steganography with Diffusion Models (Version 1). arXiv. 2024. DOI: 10.48550/ARXIV.2407.10459

13. Kheddar H., Hemis M., Himeur Y., Megías D., Amira A. Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions. *Neurocomputing.* 2024. Issue 581. P. 127528. DOI: 10.1016/j.neucom.2024.127528

14. Cassia M., Guarnera L., Casu M., Zangara I., Battiato S. Deepfake Forensic Analysis: Source Dataset Attribution and Legal Implications of Synthetic Media Manipulation. In Computer Vision and Pattern Recognition. 2025.

15. Khalifa A., Yadav Y. Wavelet-Based Fusion for Image Steganography Using Deep Convolutional Neural Networks.

Electronics. 2025. Vol. 14, No. 14. P. 2758. DOI: https://doi.org/10.3390/electronics14142758.

16. Cao H., Shi Y., Xia B., Jin X., Yang W. DiffStereo: High-Frequency Aware Diffusion Model for Stereo Image Restoration. 2025.

17. Zhi-wei K. Image Watermarking Against Geometric Attack Based on Finite Ridgelet Transform. In Microcomputer Information. 2012.

18. Ranjan R., Kumar P. An Improved Image Compression Algorithm Using 2D DWT and PCA with Canonical Huffman Encoding. *Entropy*. 2023. Vol. 25 (10). P. 1382. DOI: 10.3390/e25101382

19. Chen D., Wan S., Xiang J., Bao F. S. (2017). A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. *PLOS ONE*. 2017. Vol. 12 (3). P. e0173138. DOI: 10.1371/journal.pone.0173138

20. Зубко І. С. и др. Огляд методів нанесення цифрових водяних знаків для захисту зображень / Національний університет «Полтавська політехніка імені Юрія Кондратюка», 2024. С. 121.

21. Яремчук Ю. Є., Карпінець В. В. Аналіз стійкості стеганографічного перетворення до вбудовування цифрових водяних знаків у зображення. *Інформаційні технології та компютерна інженерія.* 2007. № 1. С. 212–217.

22. Овчинников М. Ю., Слатвінська В. М. Технічний та нормативно-правовий аспекти захисту об'єктів інтелектуальної власності на ринку аудіо-візуальної продукції. *Вісник Херсонського національного технічного університету.* 2024. №. 3 (90). С. 259–27.

23. Vorbis audio compression. URL: https://xiph.org/vorbis/

24. Овчинников М. Ю. Додавання та обробка звуку в метаданих файлів за допомогою PHP. *Наука і техніка сьогодні.* 2025. №. 4 (45). С. 1332–1344. DOI: 10.52058/2786-6025-2025-4(45)-1332-1344

25. Disha, Saini, K. (2017). A review on video steganography techniques in spatial domain. 2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE). 2017. P. 366–371. DOI: 10.1109/rdcape.2017.8358298

26. Eltahir M. E., Kiah L. M., Zaidan B. B., Zaidan A. A. High Rate Video Streaming Steganography. *International Conference on Information Management and Engineering.* 2009. P. 550–553. DOI: 10.1109/icime.2009.13

27. Kadian P., Arora S. M., Arora N. Robust Digital Watermarking Techniques for Copyright Protection of Digital Data: A Survey. *Wireless Personal Communications.* 2021. Vol. 118 (4). P. 3225–3249. DOI: 10.1007/s11277-021-08177-w

28. VeraCrypt. Безкоштовне програмне забезпечення з відкритим вихідним кодом для шифрування дисків під Windows, Mac OSX та Linux. URL: https://veracrypt.io/en/Home.html

29. Truepic: The Trusted Virtual Inspection. URL: https://www.truepic.com

30. Product Digitization for Authentication & Consumer Engagement. URL: https://www.digimarc.com

31. OpenStego, the free steganography solution. URL: https://www.openstego.com

Баландіна Наталія Миколаївна
старший викладач кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

СИНТЕЗ ТА АНАЛІЗ КРИПТОГРАФІЧНО СТІЙКИХ S-БЛОКІВ: АЛГЕБРАЇЧНІ, ЛОГІЧНІ ТА ЕВОЛЮЦІЙНІ ПІДХОДИ

DOI https://doi.org/10.36059/978-966-397-537-5-9

1.1. Вступ та постановка задачі

У сучасну епоху стрімкого розвитку цифрових технологій та глобальної комп'ютеризації криптографія перетворилася на критичну складову інформаційної безпеки, що забезпечує захист даних як у цивільних, так і в оборонних, фінансових, урядових та приватних системах. В епоху інтенсивного обміну інформацією та розвитку систем штучного інтелекту криптографія забезпечує основу довіри, яка дає змогу функціонувати інтернет-комунікаціям, електронній комерції, цифровим підписам, безпечному зберіганню даних, а також захищеним каналам зв'язку [1]. З розвитком обчислювальних засобів та виникненням потенційної загрози квантових обчислень традиційні підходи до криптографічного захисту мають бути суттєво переглянуті або посилені, що висуває нові вимоги до структури та якості криптографічних примітивів.

Одним із ключових елементів симетричних криптографічних алгоритмів є S-блоки (підстановчі блоки) [2], які реалізують нелінійне перетворення даних і визначають основну частину криптографічної сили багатьох блочних шифрів. S-блоки перетворюють фіксовану кількість бітів в іншу кількість бітів, як правило, збережену по довжині, із застосуванням таких властивостей, як високий рівень нелінійності, відсутність фіксованих точок, диференціальну та кореляційну стійкість. Саме ці властивості забезпечують ускладнення математичного аналізу, спрямованого на розкриття структури шифру, і роблять алгоритми стійкими до відомих атак. Якість S-блоків безпосередньо впливає на загальну безпечність криптосистеми, а отже, завдання їх ефективного проєктування залишається в центрі уваги дослідницької спільноти.

У межах цього розділу монографії висвітлюються результати, що мають як теоретичне, так і прикладне значення. У першій частині запропоновано новий підхід до синтезу S-блоків, що ґрунтується на використанні конструкцій полів Галуа. Поля Галуа вже давно застосовуються в криптографії завдяки своїй структурованості, алгебраїчним властивостям та зручності в обчисленнях. Нами було розроблено методи, що розвивають ідеї, закладені у класичній конструкції Ніберг, яка використовує обернення в полі Галуа як основу для побудови сильних, з точки зору криптоаналізу, S-блоків. Нові конструкції, запропоновані в даному розділі, дозволяють подолати низку обмежень традиційних підходів, підвищуючи ефективність синтезу та забезпечуючи більшу варіативність параметрів при збереженні або навіть підвищенні рівня криптографічної стійкості.

Окрім розглянутих підходів, далі буде проаналізовано ще два важливі напрямки, що мають велике значення для розвитку теорії та практики побудови S-блоків. Першим із них є застосування генетичних алгоритмів до задачі синтезу криптографічно стійких підстановок. Ці алгоритми, що імітують механізми природного добору та еволюційного розвитку, дозволяють ефективно здійснювати пошук S-блоків із високими показниками нелінійності в умовах величезного простору можливих відображень. Генетичний підхід особливо корисний у тих випадках, коли неможливо використати суворо аналітичну формулу для побудови потрібного S-блока, або коли необхідно знайти нетривіальні, псевдовипадкові конструкції з заданими властивостями. У роботі детально описано побудову відповідного алгоритму, наведено результати обчислювальних експериментів, а також проведено порівняльний аналіз синтезованих S-блоків за класичними криптографічними критеріями.

Наступний важливий напрямок дослідження пов’язаний з аналізом множин S-блоків, що мають задані значення періоду повернення у вихідний стан. Такий підхід має фундаментальне значення для розуміння внутрішньої структури підстановок, що використовуються в сучасних шифрах, та визначення їх здатності до відновлення вихідного стану за скінчену кількість ітерацій. Період повернення у вихідний стан може бути розглянутий як окрема метрика криптографічної якості, що характеризує довгострокову

динамічну поведінку перетворення та його придатність для використання в ітеративних або самовідновлювальних криптографічних схемах. У цій частині буде досліджено статистичні властивості періодів для множин S-блоків заданої довжини, а також окреслено їх потенційне значення для криптографічної стійкості в умовах тривалого використання.

У роботі також досліджується зміни нелінійності компонентних функцій багатозначної логіки випадкового S-блока при зростанні його довжини, що є важливим з точки зору розуміння фундаментальної природи нелінійності підстановлювальних конструкцій при вдосконаленні та побудові нових методів синтезу S-блоків.

Таким чином, у межах даної роботи представлено нові підходи до синтезу та оцінювання S-блоків на основі полів Галуа та функцій багатозначної логіки, обґрунтовано ефективність застосування евристичних еволюційних методів та проаналізовано динамічні характеристики великих множин S-блоків, що відкриває нові перспективи для створення надійних криптографічних конструкцій нового покоління.

Запропоновані підходи до синтезу та оцінювання S-блоків спрямовані на зміцнення теоретичного фундаменту криптографії та практичне посилення захисту інформації в умовах нових викликів. Розроблені методи можуть бути адаптовані до широкого спектру криптографічних алгоритмів, включаючи ті, що орієнтовані на квантову стійкість, і слугують основою для подальших досліджень у сфері математичного забезпечення інформаційної безпеки.

Метою цієї роботи є узагальнення, систематизація та подальший розвиток результатів щодо створених автором нових методів синтезу криптографічно стійких S-блоків на основі конструкцій полів Галуа, функцій багатозначної логіки та генетичних алгоритмів.

Для досягнення поставленої мети в роботі вирішуються наступні наукові та прикладні завдання:

1. Аналіз сучасних критеріїв криптографічної якості S-блоків і визначення тих з них, які є найбільш релевантними в умовах постквантових загроз, зокрема в контексті функціонального представлення підстановок як компонентних булевих функцій та функцій багатозначної логіки.

2. Розробка нових методів синтезу S-блоків на основі конструкцій полів Галуа, зокрема за рахунок використання ізоморфних представлень полів, матриць GF-перетворень та дослідження їх впливу на криптографічні характеристики побудованих підстановок.

3. Побудова та експериментальна реалізація генетичного алгоритму для синтезу S-блоків із високим рівнем нелінійності, дослідження динаміки адаптації популяції підстановок та виявлення класів структур, що демонструють найкращі криптографічні властивості.

4. Дослідження множин S-блоків із заданими значеннями періодів повернення у вихідний стан, оцінка їх поширеності та впливу даного критерію на інші показники криптографічної якості.

5. Дослідження зміни нелінійності компонентних функцій багатозначної логіки випадкового S-блока при зростанні його довжини.

6. Формування узагальненого підходу до порівняльного аналізу та відбору S-блоків, побудованих різними методами, з урахуванням практичних вимог до симетричних шифрів, хеш-функцій і генераторів псевдовипадкових послідовностей.

Об'єкт дослідження – процеси побудови та оцінювання криптографічних примітивів, зокрема S-блоків, що застосовуються у симетричних криптографічних алгоритмах.

Предмет дослідження – методи синтезу S-блоків з високими показниками криптографічної стійкості на основі конструкцій полів Галуа, функцій багатозначної логіки, генетичних алгоритмів та дослідження множин S-блоків із заданими періодами повернення у вихідний стан.

1.2. S-блоки в криптографії: типи, еволюція та сучасні підходи до синтезу

Однією з ключових складових симетричних криптографічних перетворень є так звані підстановочні блоки [2] (S-блоки або S-box, від англ. Substitution box). Їх основна функція – забезпечити нелінійність криптографічної трансформації, що є фундаментальною вимогою для стійкості до атак, зокрема диференціального та лінійного криптоаналізу.

Однією з ключових задач симетричної криптографії є руйнування статистичних залежностей між відкритим текстом та шифртекстом [3]. Це необхідно для того, щоб криптоаналітик, навіть маючи великі обсяги зашифрованих даних, не зміг відновити жодної інформації про зміст відкритого повідомлення.

S-блоки відіграють критично важливу роль у досягненні цього ефекту. Вони реалізують нелінійне перетворення бітових векторів фіксованої довжини, зазвичай у вигляді функції з n бітів на m бітів (найчастіше $n = m$). У криптографічних примітивах, таких як блокові шифри, S-блоки використовуються для того, щоб вивести вихідні біти, які не є прямолінійною функцією вхідних, і тим самим зруйнувати можливі кореляції.

При цьому, якщо в шифрі не використовується достатньої кількості нелінійних перетворень, статистичні характеристики вхідного повідомлення можуть «просочуватися» в шифртекст. Добре сконструйований S-блок повинен повністю змінювати розподіл імовірностей вхідних даних, викликаючи так званий лавинний ефект – навіть невелика зміна одного біта на вході має змінювати багато бітів на виході з імовірністю, близькою до 50 % [4].

S-блоки також сприяють запобіганню витоку інформації про структуру повідомлення – наприклад, про повторювані фрагменти або статистичні закономірності мови, на які зазвичай спираються атаки частотного аналізу. Таким чином, вони є інструментом руйнування статистичних особливостей відкритого тексту системою шифрування.

Поняття S-блоку бере свій початок з фундаментальних принципів симетричної криптографії, зокрема із реалізації підстановок у класичних шифрах. Ще в шифрі Цезаря чи шифрі Віженера [5] використовувалась концепція заміни символів з метою приховування вихідного тексту. Проте саме у XX столітті, з розвитком електронної обчислювальної техніки та переходом до бітового представлення даних, S-блоки набули сучасного формального вигляду – у вигляді нелінійних функцій, що відображають один бітовий вектор у інший.

Першу значущу роль S-блоки зіграли в контексті блочного шифру Data Encryption Standard (DES) [6], який був затверджений у 1977 році Національним інститутом стандартів і технологій США (NIST). В алгоритмі DES використовуються вісім S-блоків, кожен

з яких реалізує відображення з 6-бітового вектора у 4-бітовий, і ці блоки визначають основну нелінійність у шифрі. Саме завдяки S-блокам DES зміг протистояти лінійному аналізу упродовж десятиліть. Варто зазначити, що структура S-блоків у DES довгий час залишалась предметом суперечок через їх закритість: лише у 1994 році стало відомо, що S-блоки були спроєктовані з урахуванням стійкості до диференціального криптоаналізу, який був невідомий широкому колу дослідників на момент затвердження стандарту.

Подальший розвиток криптографії, зокрема із появою блочних шифрів наступного покоління – таких як AES (Rijndael), Camellia, Twofish, Serpent – супроводжувався формалізацією вимог до S-блоків. У AES використовується єдиний 8×8-бітовий S-блок, побудований на базі алгебраїчної інверсії у полі $GF(2^8)$ з подальшим афінним перетворенням. Цей підхід забезпечує як високу криптографічну стійкість, так і ефективність реалізації на апаратному та програмному рівнях.

Історично розвиток S-блоків можна умовно поділити на кілька етапів:

1. Класична епоха – використання підстановок у вигляді таблиць або ручних схем у класичних шифрах.

2. Епоха стандартів – фіксовані S-блоки у DES, IDEA, GOST 28147-89.

3. Алгебраїчний підхід – побудова S-блоків на основі операцій у скінченних полях, наприклад, S-блок криптоалгоритма AES.

4. Керовані та хаотичні S-блоки – побудова адаптивних S-блоків з використанням складних математичних структур, випадковості, біоінспірованих моделей.

Таким чином, історія S-блоків демонструє постійний рух від простих табличних підстановок до складних, адаптивних та математично вивірених конструкцій, що відіграють центральну роль у забезпеченні нелінійності, дифузії та стійкості до сучасних типів криптоаналізу.

Існує кілька підходів до класифікації S-блоків, серед яких можна виділити такі:

1. За принципом побудови:

1.1. табличні (статичні) S-блоки, які реалізуються у вигляді фіксованих таблиць підстановок;

1.2. генеративні (динамічні) S-блоки, які обчислюються під час виконання шифрування.

2. За структурою:

2.1. бієктивні (однозначні та взаємно однозначні) – часто використовуються в AES;

2.2. часткові підстановки (небієктивні) – можуть використовуватись у поточних шифрах чи хеш-функціях.

3. За способом формування:

3.1. аналітичні (на основі алгебраїчних функцій, обернення в полі $GF(2^n)$;

3.2. псевдовипадкові (з використанням генераторів псевдовипадкових або випадкових чисел);

3.3. еволюційні (генетичні алгоритми, нейромережі);

3.4. керовані (керування параметрами або ключовою інформацією).

У традиційних схемах, як-от у DES, S-блоки фіксовані та однакові для всіх ключів. Проте сучасні потреби кібербезпеки, зокрема в умовах постійного вдосконалення атак, вимагають більшої гнучкості та адаптивності криптографічних примітивів. Це спонукало дослідників до створення так званих керованих параметрами або ключем S-блоків, які формуються динамічно на основі ключової інформації, часу, стану системи тощо.

Як було вказано раніше, історично перші S-блоки були статичними – тобто заданими фіксованими таблицями заміни, які не змінювалися ані між сеансами шифрування, ані між раундами. Однак у міру розвитку атак, зокрема диференціального й лінійного криптоаналізу, з'явилася ідея створення залежних від ключа S-блоків. У таких схемах таблиці заміни генеруються на основі ключа шифрування, завдяки чому стає набагато складніше здійснити попередній аналіз структури S-блоку або застосувати універсальні шаблони атак.

У 2000-х роках почали з'являтись концепції генерованих та динамічних S-блоків, які залежать від секретного ключа або від параметрів середовища. Такі S-блоки отримали назву керовані S-блоки (key-dependent, dynamic, або controlled S-boxes) [7]. Наприклад, у шифрі Twofish [8] використовуються чотири динамічних S-блоки, які залежать від ключа, що підвищує ентропію та ускладнює атаки за відомим текстом. У багатьох сучасних розробках розглядається

створення S-блоків з використанням хаотичних відображень, алгебраїчних кривих, нейромережевих моделей, що дозволяє зменшити передбачуваність та підвищити адаптивність шифру до різних сценаріїв використання.

У таких підходах S-блоки можуть змінюватися не тільки залежно від ключа, а й динамічно – на кожному раунді шифрування, що ускладнює побудову моделей для атаки. Адаптивні S-блоки сьогодні активно досліджуються в контексті постквантових шифрів, де класичні принципи шифрування доповнюються новими ідеями гнучкості та динамічності.

Ідея залежних від ключа S-блоків реалізується в сучасній криптографії як засіб посилення криптостійкості за рахунок динамічної модифікації перетворень на основі секретного ключа. Такий підхід дозволяє зруйнувати фіксованість структури підстановки, що властива класичним S-блокам, і ускладнює проведення криптоаналітичних атак, зокрема диференціального та лінійного аналізу, атак на основі бумеранга, квадратного аналізу тощо. Формально, залежні від ключа S-блоки є функціями, які будуються динамічно під час ініціалізації шифру або навіть під час кожного раунду, враховуючи ключ або інші секретні параметри, що робить шифр непередбачуваним.

Існує низка криптографічних алгоритмів, у яких ця концепція вже реалізована на практиці. Наприклад, у шифрі Blowfish підстановочна таблиця генерується на основі ключа і відрізняється для кожної сесії. У ще більш гнучкому алгоритмі Twofish застосовано ідею ключового розкладу, в якому частини ключа використовуються для генерації модифікованих S-блоків через застосування максимально нелінійних функцій. В обох випадках метою є зробити кожен екземпляр шифру індивідуальним, що перешкоджає використанню універсальних таблиць або статистичних методів атак.

Сучасні дослідження в напрямку постквантової криптографії також демонструють інтерес до управління S-блоками за допомогою змінних параметрів. Зокрема, в деяких проєктних прототипах, що використовують симетричні схеми на основі дерев Гротендіка або у конструкціях з вузькими блоками, ключ впливає не лише на структуру розкладу, але й безпосередньо формує S-блоки для кожного

раунду. Це дозволяє досягати високого ступеня криптографічної диференціації при незначному збільшенні обчислювальних витрат.

Управління S-блоками в динамічному режимі дозволяє шифрам адаптуватися до умов середовища – зокрема, змінювати поведінку при зміні політик безпеки, алгоритмів або навіть типу пристрою. Це відкриває нові горизонти для створення криптографічних примітивів із властивостями самоналаштування, що є особливо актуальним у контексті обчислень у хмарних середовищах.

Зазначимо окремо, що збільшення розміру S-блока призводить до суттєвого покращення його криптографічних властивостей за рахунок значного розширення простору можливих відображень. У великих S-блоках зростає дистанційна нелінійність, що ускладнює лінійний криптоаналіз, а також алгебраїчна нелінійність, яка робить їх стійкішими до алгебраїчних атак. Лавинні властивості покращуються за рахунок більш рівномірного розподілу впливу вхідних бітів на вихідні, що забезпечує швидке «розповсюдження» змін. Крім того, збільшення розміру S-блока сприяє покращенню кореляційних властивостей, зменшуючи статистичні зв'язки між входами та виходами, а також збільшує періоди повернення у вихідний стан, що ускладнює атаки, засновані на циклічності.

Однак таке покращення якості досягається ціною значного зростання обчислювальної складності. Великі S-блоки вимагають більших обчислювальних ресурсів для реалізації, оскільки збільшується обсяг пам'яті для зберігання таблиць замін, а також кількість операцій, необхідних для їх обробки. Це особливо критично в низькоресурсних системах, де швидкодія є ключовим фактором. Таким чином, при проєктуванні шифрів необхідно знаходити баланс між криптографічною стійкістю, яка залежить від розміру S-блока, та ефективністю його реалізації. Оптимальний вибір розмірності визначається конкретними вимогами до безпеки та продуктивності криптосистеми.

Таким чином, S-блоки не є лише елементом заміни в схемі шифрування – вони є ядром забезпечення нелінійності та криптографічної складності, що визначають опір шифру до цілого ряду атак. Їх еволюція від статичних таблиць до складних керованих конструкцій свідчить про зростаючі вимоги до безпеки та адаптивність сучасної криптографії до нових викликів.

1.3. S-блоки на основі GF-перетворень

Основним компонентом практично будь-якої системи захисту інформації сьогодні є криптографічний компонент, який унеможливлює читання зашифрованої інформації без криптографічного ключа. Водночас, хоча асиметричні криптографічні алгоритми, через їх значну обчислювальну складність, використовуються на практиці здебільшого для вирішення задачі розподілу ключів та цифрового підпису, симетричні шифри залишаються основним компонентом, який використовується для шифрування великих обсягів захищених даних. Блокові та потокові симетричні шифри характеризуються як швидкі, перевірені на криптографічну стійкість та прості в реалізації алгоритми. У свою чергу, основним компонентом сучасних шифрів, який значною мірою визначає рівень дифузії та конфузії [3], який вони забезпечують, є криптографічний S-блок. На сьогоднішній день для оцінки ступеня реалізації концепції дифузії та конфузії криптографічним S-блоком, а також ступеня його стійкості до основних атак криптоаналізу використовується підхід, заснований на представленні S-блока у вигляді його компонентних булевих функцій, до яких застосовується набір критеріїв криптографічної якості. Основними критеріями криптографічної якості S-блока є наступні: критерій максимізації алгебраїчної нелінійності, критерій максимізації відстані нелінійності, суворий лавинний критерій та критерій мінімізації кореляції вихідних та вхідних векторів.

Важливим завданням є синтез криптографічних S-блоків, які б гармонійно відповідали переліченим критеріям криптографічної якості, чого можна досягти, використовуючи конструкції полів Галуа. Прикладом такого відомого методу є конструкція Ніберг, проте цей метод дозволяє отримати лише малі потужності високоякісних S-блоків, що робить актуальним завдання розробки нових методів синтезу S-блоків на основі конструкцій полів Галуа [9].

На сьогоднішній день існує кілька підходів до оцінки криптографічної якості S-блоків, зокрема, на основі їх представлення за допомогою компонентних булевих функцій [10], а також з використанням компонентних функцій багатозначної логіки, для

яких застосовуються спеціальні криптографічні критерії якості, наприклад, [11]. Далі для оцінки криптографічної якості синтезованих S-блоків ми використовуємо загальноприйнятий підхід, який базується на представленні S-блоків у вигляді їх компонентних булевих функцій. Оскільки довжина запропонованих S-блоків становить $N = 256$, їх можна представити як вісім компонентних булевих функції, тоді як загальна криптографічна якість S-блока визначається найслабшою з них. Серед багатьох існуючих критеріїв криптографічної якості компонентних булевих функцій ми вибрали наступний набір основних критеріїв:

1. Алгебраїчний степінь нелінійності, який визначається як найбільша кількість кон'юнкцій в термах алгебраїчної нормальної форми компонентної булевої функції [12]. У цьому випадку вектор коефіцієнтів при термах алгебраїчної нормальної форми булевої функції можна побудувати за допомогою перетворення Ріда – Маллера як

$$A = f \cdot L_N A, \tag{1}$$

де f – таблиця істинності булевої функції над алфавітом $\{0,1\}$, L_N – матриця перетворення Ріда – Маллера, яку можна побудувати відповідно до наступної рекурентної формули:

$$L_1 = [1], \quad L_{2N} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \otimes L_N = \begin{bmatrix} L_N & 0 \\ L_N & L_N \end{bmatrix}, \tag{2}$$

де $\otimes$ позначає добуток Кронекера.

2. Відстань нелінійності булевої функції, яку можна визначити як мінімум відстані Хеммінга між її компонентними булевими функціями та всіма кодовими словами афінного коду [13]:

$$N_S = \min\left\{\text{dist}(f_i, \varphi_j)\right\}, \; i = 1,...,k, \; j = 1,...,2^{k+1}, \tag{3}$$

де f_i – компонентна булева функція; φ_j – кодові слова афінного $A(N,k)$-коду, $\text{dist}(\circ)$ – оператор для обчислення відстані Хеммінга.

3. Відповідність критерію поширення помилки [13], для оцінки якого використовуються мінімальне та максимальне значення ваги похідних компонентних булевих функцій:

$$D_u f(x) = f(x) \oplus f(x \oplus u), \quad (4)$$

за усіма напрямками $\forall u \in V_k$, $wt(u) = 1$, де V_k – лінійний векторний простір векторів довжини k, $wt(\circ)$ – оператор для обчислення ваги Хеммінга. Ідеальний випадок, коли всі похідні (4) збалансовані, тобто їхні ваги Хеммінга дорівнюють $N/2$, означає, що компонентна булева функція відповідає суворому лавинному критерію, тобто ймовірність зміни її виходу при зміні значення будь-якого з її вхідних бітів дорівнює 0,5.

4. Критерій мінімізації кореляції вихідних та вхідних векторів [13]. Для оцінки відповідності S-блока цьому критерію використовується максимум серед абсолютних значень коефіцієнтів кореляції $\max\left\{\left|r_{i,j}\right|\right\}$ матриці $R = \left\|r_{i,j}\right\|$, який визначає ступінь лінійного зв'язку між вихідними y та вхідними x векторами S-блока:

$$r_{i,j} = 1 - \frac{\sum_{m=1}^{N} (x_{m,i} \oplus y_{m,j})}{N/2}, i, j = 0, ..., k-1, \quad (5)$$

де символ $\oplus$ позначає підсумовування по mod2.

Досить багато сучасних публікацій присвячено питанням синтезу S-блоків заданої довжини, які б відповідали заданим критеріям криптографічної якості, де для їх побудови використовується найрізноманітніші алгоритми: починаючи від алгоритмів їх синтезу на основі послідовностей де Брейна, як, наприклад, запропоновано в [14], закінчуючи методами синтезу S-блоків на основі теорії динамічного хаосу [15–19]. Тим не менш, на практиці для побудови криптографічних алгоритмів найбільш важливими є S-блоки, які мали б збалансований набір показників криптографічної якості, що найлегше досягається за допомогою конструкцій полів Галуа. Наприклад, у криптографічному алгоритмі AES [20], який на сьогоднішній день є де-факто найпопулярнішим та найширше використовуваним криптографічним алгоритмом у світі, використовується конструкція Ніберг, заснована на операції знаходження оберненого елемента в розширеному полі Галуа. Багато досліджень присвячено криптографічним властивостям S-блоків конструкції Ніберг, зокрема [21], тоді як стаття [11] присвячена дослідженню криптографічних

властивостей S-блоків конструкції Ніберг у випадку їх представлення за допомогою функцій багатозначної логіки.

У будь-якому разі, конструкція Ніберг характеризується малою потужністю множин S-блоків, які можна синтезувати, що не дозволяє використовувати їх як довгостроковий ключ, що важливо в низці криптографічних застосувань.

Ми показуємо, що можливості синтезу S-блоків над розширеними полями Галуа не вичерпуються, і потужність наборів синтезованих S-блоків з використанням розширених полів Галуа можна збільшити шляхом розробки нових конструкцій криптографічно високоякісних S-блоків на основі матриць GF-перетворення для розширеного поля Галуа.

Теорія та практика синтезу алгебраїчної нормальної форми (АНФ) функцій багатозначної логіки (q-функцій) або так званих GF-перетворень [14] динамічно розвиваються в сучасній криптографії.

Визначення 1 [22]. Функція q-значної логіки k змінних є відображенням $\{0,1,2,...,q-1\}^k \to \{0,1,2,...,q-1\}$.

Коли $q=2$, ми отримуємо булеві функції.

Далі, з точки зору практики побудови S-блоків, ми розглянемо функції багатозначної логіки зі значеннями $q=2^k$.

У [22] показано, що будь-яку функцію багатозначної логіки можна однозначно (з точністю до ізоморфізму) представити над полем Галуа $GF(2^k)$ за допомогою полінома, що містить операції «Додавання в полі Галуа», «Множення в полі Галуа».

Специфічні аспекти знаходження АНФ для випадку розширених полів Галуа розглядаються в [23].

Перехід до простору GF-перетворень можна виконати за такою формулою

$$: \begin{cases} A = g \cdot L_N, \\ g = A \cdot L_N^{-1}, \end{cases} \tag{6}$$

де g – таблиця істинності булевої функції, $A \in GF(2^k)$ – вектор значень GF-перетворення.

Відомо [22], що для випадку функцій багатозначної логіки $k=1$ змінної, побудова матриці оберненого GF-перетворення L_N^{-1}, $N=q^k$ над полем $GF(2^k)$ може бути виконана відповідно до наступної конструкції:

$$L_1^{-1} = \begin{bmatrix} \alpha_0^0 & \alpha_0^1 & \dots & \alpha_0^{q-1} \\ \alpha_1^0 & \alpha_1^1 & \cdots & \alpha_1^{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{q-1}^0 & \alpha_{q-1}^1 & \cdots & \alpha_{q-1}^{q-1} \end{bmatrix}, \quad (7)$$

де $\alpha_i \in GF(2^k), i = 0,1,...,q-1$ – елементи розширеного поля Галуа. Для випадку функцій багатозначної логіки, на відміну від випадку матриці перетворення Ріда – Маллера для булевих функцій (2), пряма та обернена матриці GF-перетворення не збігаються, тобто $L_N \neq L_N^{-1}$, тому пряму матрицю GF-перетворення можна знайти за допомогою одного з алгоритмів знаходження обернених матриць над розширеними полями Галуа [24].

Зауважимо, однак, що різні ізоморфні представлення полів Галуа $GF(2^k)$ означають різні реалізації операції множення (а також операцій піднесення до степеню та ділення, похідних від неї) відповідно до правила, що визначається незвідним поліномом $\psi(z)$, на основі якого побудовано поле Галуа $GF(2^k)$.

Таким чином, для поля Галуа $GF(2^k)$ функція багатозначної логіки може мати певну кількість різних ізоморфних відображень в області АНФ, що відповідає кількості незвідних поліномів степеню k, на основі яких можна побудувати арифметику поля Галуа.

У цьому випадку кількість існуючих незвідних поліномів степеню k, коефіцієнти яких належать полю Галуа $GF(2^k)$, визначається відповідно до формули [25]:

$$\left|\Psi_k\right| = \frac{1}{k} \sum_{d \,//\, k} \mu(d) 2^{(k/d)}, \quad (8)$$

де d – дільники значення степеню k; $\mu(d)$ – функція Мебіуса; позначення $d // k$ означає, що d ділить k націло.

Отже, для випадку поля Галуа $GF(2^k)$, який ми розглядаємо, відповідно до (8), кількість незвідних поліномів дорівнює $\left|\Psi_8\right| = 30$, тоді як десяткові еквіваленти цих поліномів мають наступний вигляд:

$$\Psi_8 = \{283,285,299,301,313,319,333,351,355,357,361,369,375,379,391, \\ 395,397,415,419,425,433,445,451,463,471,477,487,499,501,505\}. \quad (9)$$

Як приклад, на рис. 1 представлено у вигляді зображень у градаціях сірого (де колір кожного пікселя відповідає значенню елемента матриці в полі Галуа $GF(2^8)$ в діапазоні від 0 – чорний, до 255 – білий) перші шість матриць GF-перетворення для ізоморфних представлень поля Галуа $GF(2^8)$, побудованих на основі незвідних поліномів:

$$\begin{aligned}
\psi_1(z) &= 283_{10} = z^8 + z^4 + z^3 + z + 1,\\
\psi_2(z) &= 285_{10} = z^8 + z^4 + z^3 + z^2 + 1,\\
\psi_3(z) &= 299_{10} = z^8 + z^5 + z^3 + z + 1,\\
\psi_4(z) &= 301_{10} = z^8 + z^5 + z^3 + z^2 + 1,\\
\psi_5(z) &= 313_{10} = z^8 + z^5 + z^4 + z^3 + 1,\\
\psi_6(z) &= 319_{10} = z^8 + z^5 + z^4 + z^3 + z^2 + z + 1.
\end{aligned} \tag{10}$$

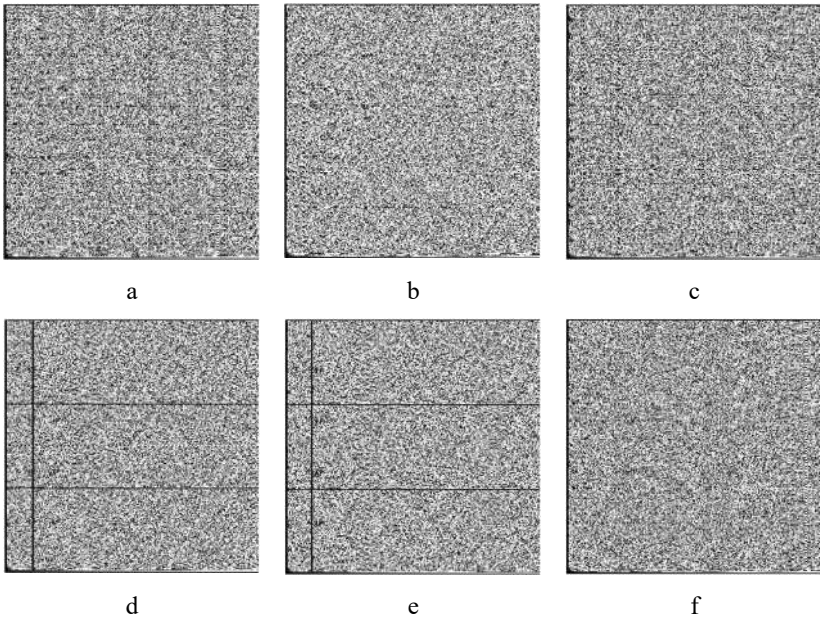


Рисунок 1 – Приклади матриць GF-перетворення для поліномів: ψ_1 (a), ψ_2 (b), ψ_3 (c), ψ_4 (d), ψ_5 (e) і ψ_6 (f)

Проведені дослідження показали, що рядки матриці GF-перетворення характеризуються достатньо високими значеннями основних показників криптографічної якості, тому їх можна використовувати для задач побудови високоякісних S-блоків. Детальне дослідження структури матриць GF-перетворення порядку $N=256$ над усіма ізоморфними представленнями поля Галуа $GF(2^8)$ дозволило встановити, що лише половина рядків цих матриць містить усі елементи поля Галуа $GF(2^8)$, що означає, що вона придатна для побудови бієктивних S-блоків. Номери цих рядків інваріантні до ізоморфізму поля Галуа $GF(2^8)$ та задаються у вигляді наступного виразу:

$$\begin{gathered}\Upsilon' = \{2,3,5,8,9,12,14,15,17,20,23,24,27,29,30,32,33,38,39,42,44,\\ 45,47,48,50,53,54,57,59,60,62,63,65,68,72,74,75,77,78,80,83,84,\\ 87,89,90,92,93,95,98,99,102,104,105,107,108,110,113,114,117,\\ 119,122,123,125,128,129,132,134,135,138,140,143,144,147,149,150,\\ 152,153,155,158,159,162,164,165,167,168,170,173,174,177,179,\\ 180,182,183,185,189,192,194,195,197,198,200,203,204,207,209,\\ 210,212,213,215,218,219,224,225,227,228,230,233,234,237,240,\\ 242,243,245,248,249,252,254,255\}.\end{gathered} \tag{11}$$

Однак, не всі рядки з (11) матриць GF-перетворення мають однаковий рівень криптографічної якості. Проведені дослідження дозволили визначити 24 номери рядків, які, незалежно від ізоморфізму, забезпечують побудову S-блоків з найкращим рівнем криптографічної якості:

$$\begin{gathered}\Upsilon = \{2,3,5,8,9,15,17,29,33,38,42,57,65,74,75,83,113,\\ 129,132,147,149,165,194,225\}.\end{gathered} \tag{12}$$

При використанні набору рядків (12) для завдання побудови S-блоків, їхня криптографічна якість є стабільною незалежно від ізоморфізму. Таким чином, ми можемо записати запропонований метод побудови криптографічних S-блоків на основі матриць GF-перетворення у вигляді конкретних кроків:

Крок 1. Оберіть початкові параметри методу, якими є: довжина S-блока N та відповідне розширене поле Галуа $GF(2^k)$, $k = \log_2 N$;

конкретний незвідний поліном, що визначає правила арифметики в розширеному полі $GF(2^k)$.

Крок 2. На основі (7) побудувати обернену матрицю GF-перетворення L_N^{-1}.

Крок 3. Використовуючи один з відомих алгоритмів інверсії матриць у розширених полях Галуа [24] на основі оберненої матриці GF-перетворення L_N^{-1} знайти пряму матрицю GF-перетворення L_N.

Крок 4. Вибрати рядки матриці GF-перетворення, що містять усі елементи розширеного поля Галуа $GF(2^k)$.

Крок 5. Представивши рядки матриці GF-перетворення, вибраної на Кроці 4, у вигляді S-блоків довжини N, оцінити рівень їх відповідності основним критеріям криптографічної якості, в результаті чого вибрати S-блоки, які найкраще відповідають критеріям, що вважаються вирішальними при модернізації або проєктуванні криптографічного алгоритму.

З огляду на практичну значущість та затребуваність використання в сучасних симетричних шифрах та хеш-функціях криптографічних S-блоків, що виконують побайтове перетворення вхідних даних, для експерименту було обрано такі параметри методу: довжина S-блока $N = 256$, всі можливі ізоморфні представлення основного поля Галуа $GF(2^8)$ на основі незвідних поліномів (9).

У табл. 1 наведено результати експериментального дослідження відповідності S-блоків, побудованих на основі запропонованого методу, основним критеріям криптографічної якості. При цьому в табл. 1 прийнято такі позначення: ψ позначає десятковий еквівалент незвідного полінома, що використовується для побудови матриці GF-перетворення; $\deg(S)$ позначає значення алгебраїчного степеню нелінійності S-блоків, побудованих у заданому ізоморфному представленні поля Галуа $GF(2^8)$; N_S позначає відстань нелінійності S-блоків, побудованих у заданому ізоморфному представленні поля Галуа $GF(2^8)$; $wt(D_u f)$ позначає значення ваг похідних за напрямком компонентних булевих функцій S-блоків для заданого ізоморфного представлення поля Галуа $GF(2^8)$; $\left|r_{i,j}\right|$ позначає значення матриць коефіцієнтів кореляції для S-блоків для заданого ізоморфного представлення поля Галуа $GF(2^8)$.

Таблиця 1 – Значення показників криптографічної якості для побудованих S-блоків

ψ	deg(*S*)	N_S	$wt(D_u f)$	$\lvert r_{i,j} \rvert$
283	5–7	112	108–156	0–0,1250
285	5–7	112	108–156	0–0,1250
299	5–7	112	108–156	0–0,1250
301	5–7	112	108–156	0,0156–0,1250
313	5–7	112	108–156	0–0,1250
319	5–7	112	108–156	0,0156–0,1250
333	5–7	112	108–156	0–0,1250
351	5–7	112	108–156	0–0,1250
355	5–7	112	108–156	0–0,1250
357	5–7	112	108–156	0–0,1250
361	5–7	112	108–156	0–0,1250
369	5–7	112	108–156	0–0,1250
375	5–7	112	108–156	0–0,1250
379	5–7	112	108–156	0–0,1250
391	5–7	112	108–156	0–0,1250
395	5–7	112	108–156	0–0,1250
397	5–7	112	108–156	0–0,1250
415	5–7	112	108–152	0,0156–0,1250
419	5–7	112	108–156	0–0,1250
425	5–7	112	108–156	0,0156–0,1250
433	5–7	112	108–156	0,0156–0,1250
445	5–7	112	108–156	0–0,1250
451	5–7	112	108–156	0–0,1250
463	5–7	112	108–156	0–0,1250
471	5–7	112	108–156	0–0,1250
477	5–7	112	108–156	0–0,1250
487	5–7	112	108–156	0,0156–0,1250
499	5–7	112	108–156	0–0,1250
501	5–7	112	108–156	0–0,1250
505	5–7	112	108–156	0–0,1250

Аналіз даних, представлених у табл. 1, дозволяє зробити висновок, що криптографічна якість S-блоків на основі рядків матриць GF-перетворення є високою та стабільною. За криптографічною якістю своїх компонентних булевих функцій побудовані S-блоки не поступаються S-блокам конструкції Ніберг. Більше того, загальна

кількість S-блоків для всіх ізоморфних представлень $GF(2^8)$ дорівнює $J = |\Upsilon| \cdot |\Psi| = 24 \cdot 30 = 720$, що в 24 рази більше, ніж кількість S-блоків конструкції Ніберг в тому ж розширеному полі Галуа.

Також зазначимо, що алгебраїчний степінь нелінійності S-блоків на основі матриць GF-перетворення залежить від вибраного номера рядка матриці, на основі якої побудовано S-блок. У табл. 2 наведено значення алгебраїчних степенів нелінійності залежно від номера рядка (12).

Таблиця 2 – Значення алгебраїчного степеня нелінійності залежно від номера рядка матриці GF-перетворення

$\upsilon_i \in \Upsilon$	2	3	5	8	9	15	17	29	33	38	42	57
deg(S)	7	7	7	5	7	5	7	5	7	5	5	5
$\upsilon_i \in \Upsilon$	65	74	75	83	113	129	132	147	149	165	194	225
deg(S)	7	5	5	5	5	7	5	5	5	5	5	5

Розглянемо приклад S-блока на основі другого рядка матриці GF-перетворення в полі Галуа $GF(2^8)$, арифметика якого визначається незвідним поліномом $\psi_1(z) = 283_{10} = z^8 + z^4 + z^3 + z + 1$, який ми представляємо як наступну алгебраїчну конструкцію:

S_1	0	1	2	3	4	5	6	7	8	9	*A*	*B*	*C*	*D*	*E*	*F*
0	00	01	8E	F4	47	A7	7A	BA	AD	9D	DD	98	3D	AA	5D	96
1	D8	72	C0	58	E0	3E	4C	66	90	DE	55	80	A0	83	4B	2A
2	6C	ED	39	51	60	56	2C	8A	70	D0	1F	4A	26	8B	33	6E
3	48	89	6F	2E	A4	C3	40	5E	50	22	CF	A9	AB	0C	15	E1
4	36	5F	F8	D5	92	4E	A6	04	30	88	2B	1E	16	67	45	93
5	38	23	68	8C	81	1A	25	61	13	C1	CB	63	97	0E	37	41
6	24	57	CA	5B	B9	C4	17	4D	52	8D	EF	B3	20	EC	2F	32
7	28	D1	11	D9	E9	FB	DA	79	DB	77	06	BB	84	CD	FE	FC
8	1B	54	A1	1D	7C	CC	E4	B0	49	31	27	2D	53	69	02	F5
9	18	DF	44	4F	9B	BC	0F	5C	0B	DC	BD	94	AC	09	C7	A2
A	1C	82	9F	C6	34	C2	46	05	CE	3B	0D	3C	9C	08	BE	B7
B	87	E5	EE	6B	EB	F2	BF	AF	C5	64	07	7B	95	9A	AE	B6
C	12	59	A5	35	65	B8	A3	9E	D2	F7	62	5A	85	7D	A8	3A
D	29	71	C8	F6	F9	43	D7	D6	10	73	76	78	99	0A	19	91
E	14	3F	E6	F0	86	B1	E2	F1	FA	74	F3	B4	6D	21	B2	6A
F	E3	E7	B5	EA	03	8F	D3	C9	42	D4	E8	75	7F	FF	7E	FD

(13)

Відстань нелінійності S-блока (13) дорівнює $N_{S_1} = 112$, його значення алгебраїчного степеню нелінійності дорівнює $\deg(S_1) = 7$.

Представимо матрицю ваг похідних компонентних булевих функцій S-блока (13):

e_j	$wt(D_{1,k})$	$wt(D_{2,k})$	$wt(D_{3,k})$	$wt(D_{4,k})$	$wt(D_{5,k})$	$wt(D_{6,k})$	$wt(D_{7,k})$	$wt(D_{8,k})$
10000000	128	116	120	124	136	112	116	116
01000000	116	124	144	144	112	116	116	128
00100000	124	128	120	132	116	116	128	116
00010000	128	128	132	124	116	128	116	124
00001000	128	116	116	124	128	116	124	128
00000100	116	132	120	140	116	124	128	128
00000010	132	124	136	140	124	128	128	116
00000001	124	136	124	136	128	128	116	132

, (14)

а також його матрицю коефіцієнтів кореляції

$$R = \begin{bmatrix} 0{,}016 & 0 & -0{,}016 & 0{,}047 & 0{,}094 & 0{,}016 & 0{,}094 & 0{,}047 \\ 0 & 0{,}094 & 0{,}094 & 0{,}047 & 0{,}016 & 0{,}094 & 0{,}047 & 0{,}016 \\ -0{,}016 & 0{,}094 & 0{,}063 & 0 & -0{,}094 & 0{,}063 & -0{,}109 & 0{,}078 \\ 0{,}047 & 0{,}047 & 0 & 0{,}031 & 0{,}063 & 0{,}031 & -0{,}109 & -0{,}031 \\ 0{,}094 & 0{,}016 & -0{,}094 & 0{,}063 & -0{,}063 & -0{,}063 & 0{,}125 & 0{,}094 \\ 0{,}016 & 0{,}094 & 0{,}063 & 0{,}031 & -0{,}063 & 0{,}125 & 0{,}094 & 0{,}094 \\ 0{,}094 & 0{,}047 & -0{,}109 & -0{,}109 & 0{,}125 & 0{,}094 & 0{,}094 & 0{,}016 \\ 0{,}047 & 0{,}016 & 0{,}078 & -0{,}031 & 0{,}094 & 0{,}094 & 0{,}016 & 0{,}094 \end{bmatrix}. \quad (15)$$

Таким чином, набір побудованих S-блоків, зокрема, S-блок (13), не поступається за своєю криптографічною якістю відомим конструкціям S-блоків, таким як конструкція Ніберг, і може бути рекомендований для практичного застосування в задачах підвищення ефективності існуючих криптографічних алгоритмів, а також розробки нових перспективних шифрів.

Відзначимо основні результати:

1. Матриці GF-перетворення можуть стати вихідним матеріалом для побудови криптографічно високоякісних S-блоків. Запропоновано ефективний метод побудови криптографічно високоякісних S-блоків практично цінних довжин на основі матриць GF-перетворення.

2. Для найбільш застосовуваного на практиці значення довжини S-блоків $N = 256$, на основі запропонованого методу синтезу S-блоків на основі матриць GF-перетворення, синтезовано клас S-блоків,

показники криптографічної якості яких не поступаються відомій конструкції Ніберг, проте потужність представленого класу в 24 рази більша порівняно з потужністю класу S-блоків конструкції Ніберг.

3. З огляду на високий рівень криптографічної якості запропонованих S-блоків, їх можна рекомендувати для практичного використання з метою вдосконалення існуючих криптографічних алгоритмів (блокових та потокових симетричних шифрів, хеш-функцій, генераторів псевдовипадкових ключових послідовностей), а також для розробки перспективних шифрів.

Як можливі напрямки подальших досліджень можна виділити можливість дослідження криптографічних властивостей S-блоків при їх представленні за допомогою АНФ, побудованої за допомогою GF-перетворення для значень $q = N$, тобто з використанням однієї функції багатозначної логіки, що може бути основою для розробки нових ефективних методів синтезу великих класів криптографічно високоякісних S-блоків.

1.4. Генетичний алгоритм для побудови високонелінійних S-блоків підстановки

Генетичні алгоритми є потужним інструментом еволюційної оптимізації, натхненним принципами природного добору та генетичної еволюції. Цей клас стохастичних алгоритмів широко використовується в задачах, де простір можливих рішень є надзвичайно великим, а традиційні методи пошуку не дають ефективного результату або потребують надмірних обчислювальних ресурсів. Генетичні алгоритми довели свою ефективність у таких галузях, як оптимальне проєктування технічних систем, біоінформатика, обробка сигналів, штучний інтелект, теорія графів і, зокрема, у криптографії – там, де необхідно знаходити структури з наперед заданими або оптимальними властивостями в умовах надскладних просторових залежностей між параметрами.

У контексті криптографії однією з найбільш цікавих задач, до якої можна застосувати генетичні алгоритми, є синтез S-блоків із високим рівнем нелінійності. Оскільки пошук S-блоку – це, фактично, комбінаційна оптимізаційна задача, що вимагає знаходження такої відображувальної функції, яка задовольняє низці критеріїв (висока

нелінійність, низька кореляція, хороша диференціальна та лінійна стійкість, відсутність фіксованих точок тощо), класичні методи синтезу часто стикаються з обмеженнями через обчислювальну складність задачі та необхідність балансувати між різними, іноді суперечливими, метриками якості. Генетичні алгоритми дозволяють розглядати цю задачу в еволюційному контексті – шляхом багаторазової генерації та селекції популяцій можливих S-блоків, з використанням операцій мутації та добору, поступово наближаючись до оптимального або субоптимального рішення.

Суттєвою перевагою генетичного підходу є його здатність ефективно досліджувати великі й розріджені простори рішень, де стандартні методи вичерпуються або втрачають ефективність. У випадку синтезу S-блоків це означає, що генетичний алгоритм може виявити нетривіальні структури відображень, які задовольняють високим криптографічним вимогам, хоча формально вони не входять у добре вивчені класи конструкцій (наприклад, не є афінними перетвореннями, не ґрунтуються на зворотних функціях у полі Галуа тощо). Крім того, адаптивність і гнучкість генетичних алгоритмів дає змогу враховувати додаткові цільові функції, зокрема метрики, пов'язані з властивостями багатозначної логіки, поведінкою компонентних функцій, стійкістю до перспективних атак квантового криптоаналізу.

Ще однією важливою особливістю генетичних алгоритмів у контексті синтезу S-блоків є їхня здатність формувати паралельні рішення. Це відкриває можливість створення цілих сімейств високоякісних S-блоків, які можуть використовуватися в різних криптографічних системах із динамічним оновленням або ротацією підстановок, підвищуючи загальний рівень криптографічної гнучкості та стійкості системи до аналізу зі сторони зловмисника. До того ж, генетичні алгоритми добре масштабуються, що дозволяє їх застосування не лише для S-блоків довжини $N = 256$, які є стандартними у багатьох шифрах, а й для більших розмірів, придатних для постквантових конструкцій або криптографічних систем зі змінною структурою блоку.

Таким чином, використання генетичних алгоритмів у задачі синтезу S-блоків з високою криптографічною якістю становить сучасний і перспективний підхід, який поєднує в собі силу еволюційного пошуку з математичними критеріями, притаманними сфері криптографії. Цей

підхід не тільки доповнює класичні методи побудови, засновані на алгебраїчних структурах, а й створює міст між теорією оптимізації, теорією обчислень і криптографією, відкриваючи нові горизонти у проєктуванні стійких і ефективних криптографічних примітивів.

Генетичний алгоритм [26–27] – це метод пошуку оптимального рішення, який імітує процес природного добору в біології. Його основна ідея полягає в тому, щоб почати з випадкового набору можливих рішень (їх ще називають «особинами» або «хромосомами»), а потім поступово покращувати ці рішення, використовуючи механізми, схожі на біологічну еволюцію: відбір, схрещування та мутації.

Спочатку створюється початкова популяція – набір випадкових кандидатів на рішення задачі. Наприклад, у випадку синтезу S-блоків кожен кандидат може бути представлений у вигляді таблиці відображення або як послідовність чисел, що задає певне переставлення бітів. Для кожного з кандидатів обчислюється так звана функція пристосованості, яка показує, наскільки добре це рішення виконує поставлену ціль. У нашому випадку – це відстань нелінійності.

Далі генетичний алгоритм проводить процес селекції – вибирає ті рішення, які мають кращу пристосованість, і використовує їх для створення нової популяції. При цьому застосовуються два основних механізми: схрещування – коли двоє «батьківських» рішень об'єднуються для створення нового, яке успадковує риси обох, та мутація – невелика випадкова зміна, що допомагає вийти за межі локального оптимуму й урізноманітнює пошук.

Новостворені особини (потомки) утворюють нову популяцію, і процес повторюється: оцінка якості, відбір, схрещування, мутація – покоління за поколінням. Ідея полягає в тому, що з кожною ітерацією (тобто з кожним новим «поколінням») середній рівень якості рішень зростає, і алгоритм поступово наближається до бажаного або навіть оптимального результату.

Зупиняється алгоритм тоді, коли досягнута достатньо висока якість рішень, або коли досягнуто певну кількість поколінь. Хоча результат не завжди є абсолютним максимумом, на практиці генетичні алгоритми часто дають дуже хороші та практично корисні рішення у задачах, де немає ефективного точного методу пошуку.

Алгоритм синтезу криптографічних S-блоків викладемо у вигляді наступних конкретних кроків [28]:

Крок 1. Задати початкову популяцію з $n/2$ S-блоків довжини N, кожний з яких визначається тривіальною монотонно-зростаючою послідовністю виду $0,1,...,N-1$. Задати значення лічильника ітерацій генетичного алгоритму $i=0$.

Крок 2. На основі популяції з $n/2$ S-блоків створити ще $n/2$ S-блоків, здійснюючи перестановку двох, випадковим чином обраних елементів їх Q-послідовностей, для кожного з них.

Крок 3. Провести вимірювання дистанційної нелінійності N_S кожного з n отриманих S-блоків, згідно до (3).

Крок 4. Видалити з популяції $n/2$ S-блоків, що володіють найменшими значеннями нелінійності N_S.

Крок 5. Якщо досягнуто останньої ітерації генетичного алгоритму, тобто $i=i_0-1$, зупин, інакше – перейти на Крок 2.

Для значення довжини S-блока $N=256$, загальної кількості S-блоків у популяції $n=0$, а також кількості ітерацій генетичного алгоритму $i_0=1000$, були отримані наступні результати щодо зростання максимального значення дистанційної нелінійності серед S-блоків у популяції в залежності від номеру ітерації генетичного алгоритму (рис. 2, де продемонстровані ітерації від 0 до 400).

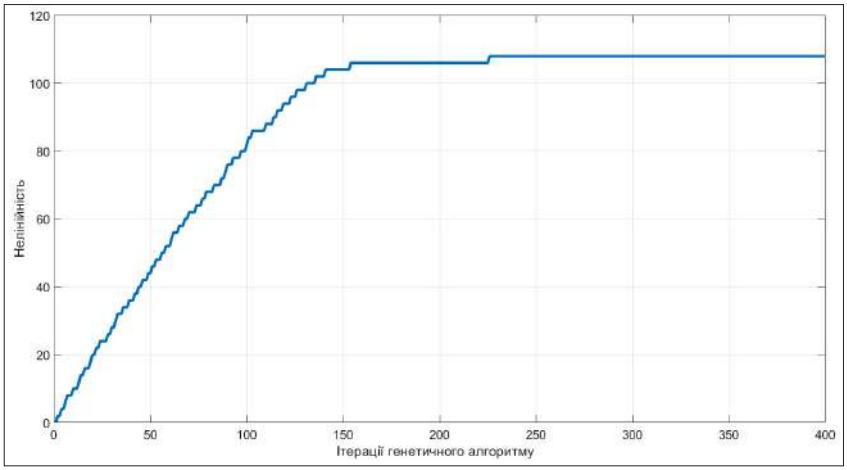


Рисунок 2 – Графік залежності максимального значення дистанційної нелінійності серед S-блоків у популяції від номера ітерації генетичного алгоритму

Аналіз даних рис. 2 показує стрімке зростання максимального значення дистанційної нелінійності серед S-блоків у популяції до 154 ітерації, після чого, на 227 ітерації максимальне значення дистанційної нелінійності стабілізується на значенні $\max(N_S) = 108$, що складає ~96.43 % від максимального значення нелінійності бієктивних S-блоків довжини $N = 256$.

Ми наводимо приклад одного із згенерованих за допомогою запропонованого генетичного алгоритму S-блоків у вигляді його Q-послідовності, що кодує.

$$
\begin{gathered}
S = \{86,150,43,44,105,93,12,107,196,206,161,138,217,13,120,59,29,\\
199,19,189,110,139,26,230,80,25,220,203,68,125,227,145,235,4,232,\\
183,42,155,211,148,253,245,201,163,84,99,66,39,210,54,90,78,100,\\
9,11,252,147,160,181,135,186,30,177,133,57,193,\ 89,55,149,27,38,136,\\
239,20,209,170,65,144,122,205,250,40,154,46,168,56,21,251,184,69,\\
95,231,248,223,187,35,104,119,101,88,60,106,204,197,36,81,126,137,\\
109,28,79,32,172,34,117,175,179,229,17,153,166,247,207,85,164,240,\\
214,\ 33,47,188,31,16,131,185,94,146,128,130,\ 18,202,58,225,75,91,71,\\
212,114,208,7,2,0,226,140,169,82,162,238,171,233,213,45,165,167,218,\\
174,37,121,190,180,127,115,123,243,64,50,151,134,6,191,158,192,73,97,\\
228,112,52,1,51,156,222,53,157,83,22,249,10,221,215,198,195,241,74,\\
113,124,132,5,159,98,254,236,224,116,62,141,194,152,176,111,182,200,\\
70,\ 15,237,118,\ 63,\ 76,173,72,178,77,108,61,96,103,129,14,102,67,\\
143,242,3,92,23,255,234,24,246,41,48,49,216,244,87,142,8,219\}.
\end{gathered}
\tag{16}
$$

Зазначений S-блок володіє дистанційною нелінійністю $N_S = 108$. Наведемо також інші показники криптографічної якості зазначеного S-блока. Його алгебраїчний степінь нелінійності складає $\deg(S) = 7$.

Наведемо матрицю ваг Хеммінга похідних компонентних булевих функцій S-блока (16):

e_j	$wt(D_{1,k})$	$wt(D_{2,k})$	$wt(D_{3,k})$	$wt(D_{4,k})$	$wt(D_{5,k})$	$wt(D_{6,k})$	$wt(D_{7,k})$	$wt(D_{8,k})$
10000000	144	132	128	120	136	124	128	128
01000000	132	128	132	120	132	140	112	116
00100000	132	136	140	120	112	120	128	116
00010000	128	128	132	136	148	116	148	136
00001000	120	120	116	136	132	108	140	112
00000100	132	116	120	136	128	124	132	124
00000010	120	132	128	124	132	132	124	124
00000001	92	124	124	116	120	132	140	132

, (17)

а також його матрицю коефіцієнтів кореляції:

$$R = \begin{bmatrix} 0,05 & 0,08 & 0,05 & 0,06 & -0,08 & 0,03 & -0,02 & -0,09 \\ 0 & 0,09 & 0 & -0,03 & 0 & -0,05 & -0,08 & 0,08 \\ 0,05 & -0,11 & -0,05 & -0,06 & 0,03 & 0,11 & 0,06 & 0,02 \\ 0,03 & 0,03 & 0,06 & -0,13 & -0,05 & -0,09 & 0,05 & -0,06 \\ -0,05 & 0,06 & -0,06 & 0,02 & 0,06 & 0 & 0 & 0 \\ -0,08 & -0,02 & -0,06 & 0 & 0,02 & 0,13 & 0,06 & -0,02 \\ -0,06 & 0,03 & 0 & 0,08 & -0,05 & -0,03 & 0,09 & 0,02 \\ 0,02 & 0,06 & 0 & 0,06 & 0,13 & -0,05 & -0,05 & 0 \end{bmatrix}. \quad (18)$$

Аналіз (17) і (18) показує, що окрім високих як алгебраїчної, так і дистанційної нелінійності, S-блок (16) володіє припустимими лавинними і кореляційними властивостями, що у поєднанні з його псевдовипадковою природою дозволяє рекомендувати його до практичного застосування.

Відзначимо основні результати:

1. Представлено генетичний алгоритм для синтезу множин довільної потужності високонелінійних (псевдо)випадкових S-блоків підстановки довільної довжини N.

2. Дослідження криптографічних властивостей синтезованих за допомогою розробленого генетичного алгоритму S-блоків довжини $N = 256$ дозволило встановити високий рівень їх криптографічної якості, зокрема, дистанційної та алгебраїчної нелінійності, а також припустимі лавинні і дистанційні властивості. У поєднані із їх псевдовипадковою природою це дозволяє рекомендувати такі S-блоки до практичного застосування.

1.5. Дослідження потужностей множин нелінійних перетворень із заданими значеннями періодів повернення у вихідний стан

К. Шенноном у 1949 р. [3] було введено концепцію криптографічного S-блока, як основного компонента для побудови захищених шифрів. Роки дослідження криптографічних конструкцій та досвід їх практичного застосування підтвердив роль S-блоків, як основних

компонентів, що визначають ефективність та швидкодію сучасних шифрів. У сьогоденні чітко визначений математичний апарат дослідження криптографічної якості S-блоків, що включає такі критерії, як нелінійність, критерій розповсюдження помилки та суворий лавинний критерій, критерій кореляційного імунітету та матриці коефіцієнтів кореляції, критерій максимізації періодів повернення у вихідний стан [29–30]. Запропоновано численні конструкції, призначені для побудови S-блоків, що відповідають зазначеним критеріям криптографічної якості.

Природньо, що у повній множині S-блоків існують обмежені кількості таких їх екземплярів, що у тій чи іншій мірі задовольняють вибраному критерію криптографічної якості. Означене встановлює певні лімітації на побудову S-блока, який був би ідеальним з точки зору кожного з зазначених критеріїв криптографічної якості. Дослідження потужностей множин S-блоків є фундаментальним з точки зору розробки практичних конструкцій S-блоків, дослідження їх криптографічної якості та розуміння основних можливостей щодо її підвищення.

Одним з основних критеріїв криптографічної якості є критерій, який було запропоновано у фундаментальній роботі [30], – повернення S-блоків у вихідний стан, який визначається як число ітерацій до повернення S-блоку в початковий стан, при тривалій вхідній дії. Для обчислення періоду повернення S-блока у вихідний стан, необхідно розкласти його на цикли, що визначаються як послідовності переходів елементів. Наприклад, для S-блока довжини $N = 8$:

$$S = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 6 & 7 & 3 & 2 & 0 & 1 & 4 & 5 \end{bmatrix}, \tag{19}$$

маємо наступні цикли:

$$S = \{C_3(6,4,0),\ C_3(1,7,5),\ C_2(2,3)\}, \tag{20}$$

тоді період повернення S-блока у вихідний стан визначатиметься як

$$T = \text{НСК}(i_1, i_2 ...). \tag{21}$$

Сформуємо алгоритм експериментальних досліджень у вигляді конкретних кроків [31].

Крок 1. Сформувати вибірку досліджуваних S-блоків Ψ заданої довжини N та потужності $|\Psi|$.

Крок 2. Провести дослідження періодів повернення для кожного S-блока, який був сформований на *Кроці 1*.

Задля дослідження взаємозв'язку періодів повернення S-блоків у вихідний стан для значення довжини $N = 8$ може бути застосований метод повного перебору, тобто сформована вибірка Ψ, що вміщуватиме всі можливі варіанти S-блоків даної довжини, оскільки загальна множина даних S-блоків включає лише $|\Psi| = 40\,320$ елементів.

У табл. 3 представлено значення періодів повернення S-блоків довжини $N = 8$ у вихідний стан.

Таблиця 3 – Значення періодів повернення у вихідний стан та відповідні їм значення нелінійності для S-блоків довжини $N = 8$

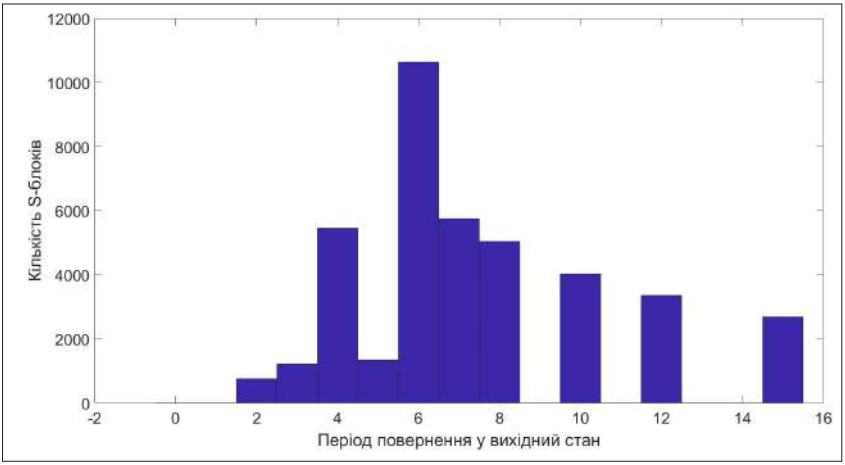
Період	1	2	3	4	5	6	7	8	10	12	15
Кількість	1	763	1232	5460	1344	10640	5760	5040	4032	3360	2688

Для дослідження S-блоків довжини $N = 16$ сформуємо вибірку з $|\Psi| = 1\,000\,000$ випадково згенерованих S-блоків, щодо яких проведемо дослідження, результати яких представлені у табл. 4.

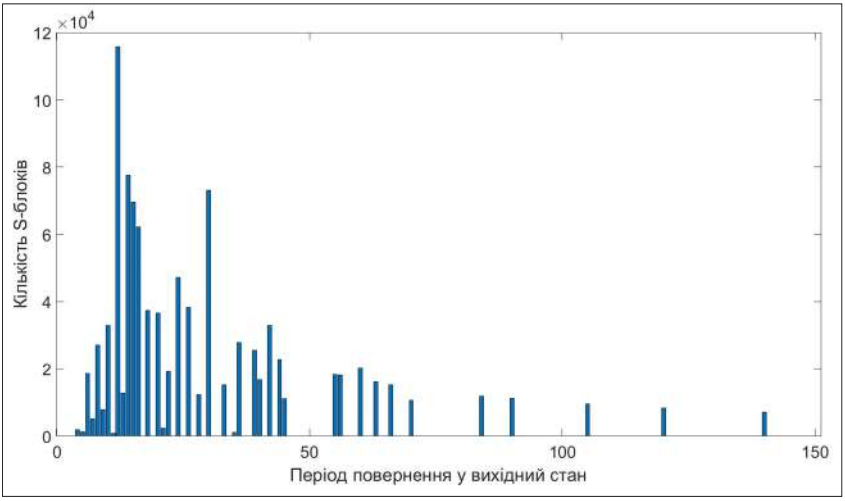
Таблиця 4 – Значення періодів повернення у вихідний стан та відповідні їм значення нелінійності для S-блоків довжини $N = 16$

Період	2	3	4	5	6	7	8
Кількість	1	52	1948	1440	18873	5034	27046
Період	9	10	11	12	13	14	15
Кількість	7580	32850	744	116719	12700	77251	69394
Період	16	18	20	21	22	24	26
Кількість	62387	37195	36889	2233	18684	47334	38474
Період	28	30	33	35	36	39	40
Кількість	12133	73450	15424	1172	27886	25781	16756
Період	42	44	45	55	56	60	63
Кількість	32804	23027	11089	18241	17764	19959	15831
Період	66	70	84	90	105	120	140
Кількість	15285	10714	11989	11031	9435	8145	7256

На рис. 3 показане графічне представлення кількостей S-блоків, що мають заданий період повернення у вихідний стан для значень довжини S-блоків $N = 8$ та $N = 16$.



a



b

Рисунок 3 – Кількості S-блоків, що мають заданий період повернення у вихідний стан для значень довжини $N = 8$ (a) та $N = 16$ (b)

Аналіз даних, що представлені у табл. 3 та табл. 4, а також даних рис. 3 дозволяє дійти висновку, що зі збільшенням значення необхідного періоду повернення, кількість S-блоків, що володіють даним періодом істотно зменшується.

Відзначимо основні результати:

1. Проведено дослідження кількостей S-блоків, що володіють заданим рівнем періоду повернення у вихідний стан для практично цінних довжин $N = 8,16$.

2. Встановлено, що зі збільшенням значення періоду повернення S-блоків, кількість S-блоків, що володіють даним значенням періоду повернення, зменшується. Дана обставина означає, що випадково згенерований S-блок, з великою ймовірністю, не володітиме максимальним або близьким до максимального для даної довжини S-блоків значенням періодів повернення у вихідний стан. Означене диктує необхідність розробки регулярних методів синтезу множин S-блоків, які, зокрема, відповідають критерію максимізації періоду повернення у вихідний стан.

1.6. Дослідження змін нелінійності компонентних функцій багатозначної логіки випадкового S-блока при зростанні його довжини

Дослідження змін нелінійності компонентних функцій багатозначної логіки випадкового S-блока при зростанні його довжини є надзвичайно актуальним з огляду на сучасні виклики криптографії. Класичні підходи до оцінки криптографічної стійкості S-блоків зазвичай базуються на аналізі їхніх бінарних представлень, проте розгляд всіх можливих уявлень, зокрема через функції багатозначної логіки, відкриває нові горизонти в розумінні їхньої поведінки і властивостей. Такий підхід особливо важливий в контексті розвитку квантового криптоаналізу, де класичні критерії стійкості можуть виявитися недостатніми або потребувати доопрацювання. Вивчення нелінійності через призму багатозначної логіки дозволяє глибше оцінити стійкість S-блоків до нових типів атак, а також сприяє розробці більш надійних і адаптивних криптографічних компонентів. З огляду на те, що довжина S-блока істотно впливає на його

внутрішню структуру і складність функцій, дослідження зміни нелінійності при зміні розмірності є ключовим для оптимізації дизайну сучасних шифрів.

Задля збільшення рівня якості застосовуваних математичних конструкцій, сучасними дослідниками розробляються математично обґрунтовані методи синтезу S-блоків, що мають високі значення нелінійності. Іншим підходом до синтезу ефективних S-блоків є підхід, заснований на їх випадковій генерації. Відомо, що зі збільшенням довжини S-блоку суттєво покращується його криптографічна якість, зокрема, його нелінійність. Через те, що сучасні криптографічні алгоритми характеризуються досить високою довжиною застосовуваних S-блоків (типова довжина $N=256$), якість випадково згенерованих S-блоків такої довжини проголошується як достатня.

Тим не менш, реальна динаміка змін у нелінійності S-блоків із зміною їх довжини при їх уявленні за допомогою функцій багатозначної логіки у літературі не досліджувалася.

Добре відомим є універсальне визначення нелінійності булевих функцій та функцій багатозначної логіки, яка чисельно може бути розрахована за допомогою наступної формули [32].

$$NL = \begin{cases} q^k - \max\left\{\left|\Omega(\acute{E})\right|\right\}, & q > 2; \\ 2^{k-1} - \dfrac{1}{2}\max\left\{\left|W(\acute{E})\right|\right\}, & q = 2, \end{cases} \tag{22}$$

де $\Omega(\omega)$ – коефіцієнти перетворення Віленкіна-Крестенсона досліджуваної q-функції, $W(\omega)$ – коефіцієнти перетворення Уолша – Адамара досліджуваної булевої функції, k – кількість змінних, від яких залежить компонентна q-функція.

Тим не менш, само по собі порівняння значень нелінійності (22) компонентних q-функцій різних довжин N не є показовим. Має сенс досліджувати значення нелінійності у порівнянні із максимально можливим значенням нелінійності, яке принципово можуть прийняти q-функції.

Максимальне значення нелінійності збалансованих функцій багатозначної логіки у загальному випадку відсутнє у відритих джерелах, тож для порівняння значень нелінійності зручно використовувати такі

досконалі алгебраїчні конструкції як бент-функції, які мають мінімально можливе значення максимального коефіцієнта перетворення Віленкіна-Крестенсона $q^{k/2}$, і, відповідно, максимальне значення нелінійності, що дорівнює:

$$\max\{N_{f_q}\} = q^k - q^{k/2}, \quad q > 2. \tag{23}$$

При цьому, у випадку булевих функцій, враховуючи, що мінімально можливий модуль значення максимального коефіцієнта перетворення Уолша – Адамара дорівнює $2^{\frac{k}{2}-1}$, максимальне значення нелінійності складає:

$$\max\{N_{f_2}\} = 2^{k-1} - 2^{\frac{k}{2}-1}. \tag{24}$$

Зазначимо, що вираз (24) має сенс тільки для парних значень k, тоді як при непарних k булеві бент-функції не існують.

Таким чином, коефіцієнт відносної нелінійності (виражений у відсотках) можна визначити як відношення нелінійності досліджуваної функції багатозначної логіки до максимального значення нелінійності:

$$\xi = \frac{N_{f_q} 100\%}{\max\{N_{f_q}\}}. \tag{25}$$

Коефіцієнт відносної нелінійності (25) дозволяє порівнювати рівень нелінійності криптографічних конструкцій різної довжини при їх уявленні за допомогою компонентних q-функцій для різних значень основи уявлення q.

Відзначимо при цьому, що через те, що булеві бент-функції не існують при непарних значеннях k, визначення відносної нелінійності (25) не має для них сенсу.

Відповідно до мети роботи сформуємо алгоритм експериментальних досліджень у вигляді конкретних кроків [33].

Крок 1. Сформувати перелік досліджуваних довжин S-блоків $\{N_i\}, i = 1,2,...,L$.

Крок 2. Випадково сформувати для кожного N_i J S-блоків довжини N_i.

Крок 3. Дослідити нелінійність кожного згенерованого на *Кроці* 2 S-блоку, в результаті чого сформувати для кожного N_i свою послідовність значень нелінійності $\{NL_j\}$, $j = 1,2,...,J$.

Крок 4. Усереднити значення нелінійності для кожного N_i, тобто сформувати послідовність $\{\overline{NL_i}\}, i = 1,2,...,L$.

Зазначений алгоритм досліджень був застосований до компонентних булевих функцій випадкових S-блоків довжин $\{N_i\} = \{4,16,64,256,1024,4096\}$, до компонентних 4-функцій випадкових S-блоків довжин $\{N_i\} = \{16,64,256,1024,4096\}$, а також до компонентних 16-функцій випадкових S-блоків довжин $\{N_i\} = \{256,4096\}$. При цьому був обраний параметр $J = 100$ задля забезпечення показовості отриманих результатів.

Результати виконання розробленого алгоритму досліджень у вигляді графіків залежності нелінійності випадково згенерованого S-блока від його довжини при його уявленні компонентними q-функціями для різних значень q показано на рис. 4.

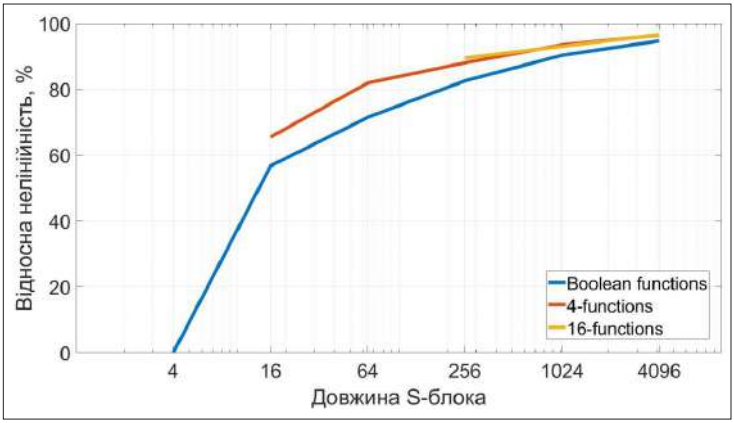


Рисунок 4 – Графік залежності нелінійності випадкового S-блока від його довжини для різних значень основи уявлення q

Аналіз рис. 4 дозволяє зробити висновок про суттєве зростання нелінійності випадково згенерованого S-блока із зростанням його довжини. При цьому, як показують дані проведеного експерименту, компонентні 4-функції та 16-функції випадкового S-блока демонструють більші значення відносної нелінійності та швидше її наближення до рівня нелінійності бент-функцій у порівнянні з компонентними булевими функціями.

Відзначимо основні результати:

1. Проведено дослідження нелінійності випадкового S-блока при його представленні компонентними булевими функціями та функціями багатозначної логіки в залежності від його довжини.

2. Встановлено, що із зростанням довжини випадкового S-блока нелінійність його компонентних q-функцій наближається до рівня нелінійності бент-функцій. При цьому, компонентні 4-функції та 16-функції, у порівнянні з компонентними булевими функціями, демонструють більші значення відносної нелінійності та швидше її наближення до рівня нелінійності бент-функцій.

3. Отримані результати показують, що при достатній довжині S-блоків ($N \geq 256$), метод їх випадкової генерації здатен забезпечити нелінійність їх компонентних q-функцій на рівні у якнайменше 80 % від нелінійності бент-функцій для основ уявлення $q = \{2, 4, 16\}$.

1.7. Перспективи подальшої роботи над вдосконаленням конструкцій підстановки

У зв'язку з постійним розвитком засобів криптоаналізу, зростанням обчислювальних можливостей, включно з появою квантових обчислювальних пристроїв, та ускладненням задач захисту даних, розробка нових підходів до побудови S-блоків залишається одним із найважливіших напрямів сучасної криптографії. Підстановки відіграють центральну роль у симетричних шифрах, забезпечуючи їхню нелінійність та здатність протистояти лінійному й диференціальному криптоаналізу. У цьому контексті можна окреслити кілька перспективних напрямів подальших досліджень, які мають потенціал суттєво посилити криптографічну стійкість алгоритмів за рахунок більш гнучких, адаптивних та математично обґрунтованих конструкцій S-блоків.

Одним із найбільш багатообіцяючих напрямів є синтез S-блоків на основі функцій багатозначної логіки. Такий підхід дозволяє розширити клас використовуваних функцій за межі традиційних булевих перетворень. Багатозначна логіка (наприклад, тризначна, чотиризначна і т. д.) дає змогу моделювати більш складні перетворення, які мають внутрішню симетрію, більш багатий спектр значень та властивостей. Завдяки цьому можна проєктувати підстановки, що

водночас мають високу булеву нелінійність, гарну кореляційну стійкість та адаптивну структуру, зручну для подальшого аналізу й оптимізації. Перспективним напрямом є розробка відповідних метрик якості таких S-блоків, урахування багатокомпонентних представлень та дослідження їхньої поведінки при спробах атак, включно з квантовими.

Не менш важливим є подальше вдосконалення конструкцій, побудованих на основі алгебраїчних структур полів Галуа. Відомі результати – зокрема, обернення в полі як базова функція в конструкції Ніберг – показали, що елементи полів Галуа здатні забезпечувати високі показники з точки зору криптографічної стійкості. Проте традиційні конструкції мають певні обмеження, пов'язані з розміром поля, обчислювальною складністю операцій та незмінністю структури. Подальші дослідження можуть бути спрямовані на створення узагальнених підстановок, які комбінують кілька операцій у полі, використовують нетривіальні автоморфізми або побудовані на злитті кількох полів у композиційні структури. Такі підходи дозволяють підвищити варіативність і непередбачуваність поведінки S-блоків, що є критичним для захисту від сучасного криптоаналізу.

Окремий клас завдань формує застосування генетичних алгоритмів для синтезу S-блоків. Як було показано раніше, генетичні алгоритми ефективні для пошуку оптимальних або близьких до оптимальних рішень у великому й складному просторі можливих відображень. Перспективним є розвиток адаптивних схем, де параметри алгоритму змінюються в процесі виконання, а також гібридизація з іншими методами – наприклад, з локальним пошуком або табу-пошуком. Особливо цікавою є можливість використання генетичних алгоритмів не лише для статичного синтезу, а й для динамічного формування S-блоків у реальному часі, наприклад, у контексті потокових шифрів або алгоритмів з обертовою структурою. Це відкриває нові горизонти в області самонавчальних або реактивних криптографічних систем, які здатні змінювати свою поведінку відповідно до умов середовища.

У ширшому контексті, застосування методів штучного інтелекту та машинного навчання у задачах побудови S-блоків відкриває абсолютно нову площину можливостей. Зокрема, нейромережеві

підходи можуть бути використані як для генерації нових S-блоків, так і для оцінювання їхньої якості або прогнозування слабких місць у структурі. Ідеї підкріпленого навчання дають змогу будувати алгоритми, які адаптують свою структуру в залежності від результатів криптоаналізу, таким чином еволюціонуючи у напрямі посилення власної стійкості. Навіть за межами класичних застосувань, штучний інтелект може бути використаний для автоматизованого доказу властивостей S-блоків, побудови моделей атак та пошуку нових принципів побудови криптографічних примітивів. Синтез таких підстановок із використанням глибоких моделей або евристичних пошуків відкриває шлях до створення інтелектуально керованих криптосистем нового покоління.

Загалом, подальше вдосконалення конструкцій S-блоків потребує поєднання глибоких математичних знань із алгоритмічною гнучкістю, властивою еволюційним і штучно-інтелектуальним підходам. Комбінація методів багатозначної логіки, алгебраїчної теорії полів, обчислювального інтелекту та криптоаналітичного моделювання утворює міждисциплінарне підґрунтя, на якому може розвиватися новий клас стійких, ефективних і адаптивних криптографічних примітивів. В умовах зростаючих викликів до безпеки інформації, особливо в еру постквантових загроз, такі дослідження мають не лише наукове, а й стратегічне значення.

Список використаних джерел

1. Sagindikovich P. Q. Modern cryptography algorithms. *American Journal of Education and Learning*. 2025. Vol. 3 (5). P. 68–73.

2. Ali R. S. et al. A comprehensive review on S-box generation methods. AIP Conference Proceedings. *AIP Publishing LLC.* 2024. Vol. 3207 (1). P. 020001.

3. Shannon C. E. Communication theory of secrecy systems. *The Bell system technical journal.* 1949. Vol. 28 (4). P. 656–715. DOI: https://doi.org/10.1002/j.1538-7305.1949.tb00928.x

4. Forrié R. The strict avalanche criterion: spectral properties of Boolean functions and an extended definition. Conference on the Theory and Application of Cryptography. New York, NY : Springer New York. 1988. P. 450–468.

5. Deepthi D. V. V., Benny B. H., Sreenu K. Various ciphers in classical cryptography. Journal of Physics: Conference Series. IOP Publishing, 2019. Vol. 1228 (1). P. 012014.

6. Hamza A., Kumar B. A review paper on DES, AES, RSA encryption standards. International Conference System Modeling and Advancement in Research Trends (SMART). IEEE, 2020. P. 333–338.

7. Krishnamurthy G. N., Ramaswamy V. Making AES stronger: AES with key dependent S-box. *IJCSNS International Journal of Computer Science and Network Security.* 2008. Vol. 8 (9). P. 388–398.

8. Schneier B. The Twofish encryption algorithm. *Dr. Dobb's Journal: Software Tools for the Professional Programmer.* 1998. Vol. 23 (12). P. 30–34.

9. Бакуніна О. В., Баландіна Н. М., Соколов А. В. Метод синтезу s-блоків на основі матриць GF-перетворення. *Український журнал інформаційних технологій.* 2023. Т. 5 (2). С. 41–48.

10. Mazurkov M. I., Sokolov A. V. Nonlinear transformations based on complete classes of isomorphic and automorphic representations of field GF (256). *Radioelectronics and Communications Systems.* 2013. Vol. 56 (11). P. 513–521. DOI: https://doi.org/10.3103/s0735272713110022

11. Sokolov A. V., Djiofack T. V.N. Nonlinear Properties of Rijndael S-boxes Represented by the Many-Valued Logic Functions. Proceedings of the International Workshop on Cyber Hygiene, Kyiv, Ukraine, 2019. P. 96–106.

12. Rostovcev A. G. Cryptography and Data Protection. Mir i Sem'ja. 2002.

13. Logachev O. A., Sal'nikov A. A., Jashhenko V. V. Boolean functions in coding theory and cryptology, USA: American Mathematical Society. 2012.

14. Mazurkov M. I., Sokolov, A. V. Constructive method for synthesis of complete classes of multilevel de Bruijn sequences. Radioelectronics and Communications Systems. 2013. Vol. 56 (1), P. 36–41. DOI: https://doi.org/10.3103/s0735272713010044

15. Wang J., Zhu Y., Zhou C., Qi Z. Construction Method and Performance Analysis of Chaotic S-Box Based on a Memorable Simulated Annealing Algorithm. *Symmetry.* 2020. No. 12, P. 2115. DOI: https://doi.org/10.3390/sym12122115

16. Lambić D. S-box design method based on improved one-dimensional discrete chaotic map. *Journal of Information and Telecommunication.* 2018. Vol. 2 (2). P. 181–191. DOI: https://doi.org/10.1080/24751839.2018.1434723

17. Lu Q., Zhu C., Wang G. (2019). A Novel S-Box Design Algorithm Based on a New Compound Chaotic System. Entropy. 2019. Vol. 21 (10). P. 1004. DOI: https://doi.org/10.3390/e21101004

18. Lai Q., Akgul A., Li C., Xu G., Çavuşoğlu U. A New Chaotic System with Multiple Attractors: Dynamic Analysis, Circuit Realization and S-Box Design. *Entropy*. 2018. Vol. 20 (1). P. 12. DOI: https://doi.org/10.3390/e20010012

19. Баландіна Н. М. Дослідження нелінійності S-блока на основі теорії динамічного хаосу при його уявленні функціями багатозначної логіки. *Кібербезпека в сучасному світі* : матер. III Всеукр. наук.-практ. конф. (м. Одеса, 19 листопада 2021 року). Одеса : ВД «Гельветика», 2021. С. 43–48. DOI: 10.32837/11300.15973

20. FIPS 197. (2001) Advanced encryption standard. URL: http://csrc.nist.gov/publications/

21. Баландіна Н. М. Алгебраїчна нелінійність S-блоків конструкції Ніберг при їх представленні функціями багатозначної логіки. *Європейські орієнтири розвитку України в умовах війни та глобальних викликів XXI століття: синергія наукових, освітніх та технологічних рішень* : матер. Міжнар. наук.-практ. конф. (м. Одеса, 19 травня 2023 року). Одеса : ВД «Гельветика», 2023. С. 627–630. ISBN 978-617-8263-37-9

22. Stankovic R. S., Astola J. T., Moraga C. Representations of Multiple-Valued Logic Functions. Morgan and Claypool Publishers, 2012.

23. Sokolov A. V., Overchuk Yu.S. On the possibility of synthesizing an algebraic normal form of quaternary functions over the field GF (4). *Problems of cyber security of information and telecommunication systems*: Proceedings of the first international scientific and practical conference. 2018. P. 384–388.

24. Grošek O., Fabšič T. Computing multiplicative inverses in finite fields by long division. *Journal of Electrical Engineering*. 2018. Vol. 69 (5). P. 400–402. DOI: https://doi.org/10.2478/jee-2018-0059

25. Berlekamp E. R. Algebraic coding theory: Revised Edition. World Scientific. 2015. DOI: https://doi.org/10.1142/9407

26. Katoch S., Chauhan S. S., Kumar V. A review on genetic algorithm: past, present, and future. *Multimedia tools and applications.* 2021. Vol. 80 (5). P. 8091–8126.

27. Haldurai L., Madhubala T., Rajalakshmi R. A study on genetic algorithm and its applications. *Int. J. Comput. Sci. Eng.* 2016. Vol. 4 (10). P. 139–143.

28. Баландіна Н. М. Розробка генетичного алгоритму для побудови високонелінійних S-блоків підстановки. *Інформаційне суспільство: проблеми та перспективи* : матер. IX Всеукр. наук.-практ. конф. (м. Одеса, 24 травня 2024 року). Одеса : НУ «ОЮА», 2024. С. 140–144. DOI: https://doi.org/10.32837/11300.27842

29. Соколов А. В. Новые методы синтеза нелинейных преобразований современных шифров. Lap Lambert Academic Publishing, Germany 2015. 100 с.

30. Зайко Ю. Н. Криптография глазами физика. Изв. Саратовского ун-та. 2009. Т. 9 (2). С. 34–48.

31. Баландіна Н. М. Дослідження потужностей множин нелінійних перетворень із заданими значеннями періодів повернення у вихідний стан. *Кіберпростір в умовах війни та глобальних викликів XXI століття: теорія та практика* : матер. Міжнародної наук.-практ. конференції. 2023. С. 123–127.

32. Sokolov A. V., Zhdanov O. N. Regular synthesis method of a complete class of ternary bent-sequences and their nonlinear properties. *Journal of Telecommunication, Electronic and Computer Engineering.* 2016. Vol. 8 (9). P. 39–43.

33. Баландіна Н. М. Дослідження змін нелінійності компонентних функцій багатозначної логіки випадкового S-блока при зростанні його довжини. *Кібербезпека в сучасному світі: актуальні виклики* : матер. IV Всеукр. наук.-практ. конф. (м. Одеса, 25 листопада 2022 року). Одеса : ВД «Гельветика», 2022. С. 24–28. ISBN 978-966-992-297-7

РОЗДІЛ 4. ЗАХИСТ ІНФОРМАЦІЇ В ІКС: ТЕХНІЧНІ ТА ПРИКЛАДНІ АСПЕКТИ

Кухаренко Сергій Вікторович
к.т.н., доц., доцент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ЗАХИСТ ДАНИХ КОРИСТУВАЧІВ В ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМАХ: АНАЛІЗ ВИТОКІВ, ІДЕНТИФІКАЦІЇ І КОМПЛЕКСНИХ МОДЕЛЕЙ ПРОТИДІЇ

DOI https://doi.org/10.36059/978-966-397-537-5-10

1. Вступ і постановка задачі

В останні роки інформаційно-комунікаційні системи (ІКС) перетворилися на ключовий елемент операційної діяльності суспільства, комерційних структур, державних органів та об'єктів критичної інфраструктури. Збільшення обсягів оброблюваних даних у цих системах паралельно супроводжується стрімким наростанням кіберзагроз. Витоки персональних даних, несанкціоноване проникнення, кібератаки та експлуатація мережевих вразливостей призводять до вагомих фінансових збитків, порушення конфіденційності, шкоди репутації та потенційних ризиків для національної безпеки. Аналітичні дані свідчать, що середня вартість витоку даних сягнула 4,88 млн дол. США у 2024 році, що на 10 % перевищує показник попереднього року [1]. Прогнози вказують на те, що глобальні витрати, пов'язані з кіберзлочинністю, досягнуть 10,5 трлн дол. США до кінця 2025 року, з щорічним приростом на рівні 15 % [2]. Кількість зареєстрованих інцидентів витоків даних у період 2024–2025 років зросла на 10–30 % порівняно з попередніми роками, зокрема через застосування штучного інтелекту в атаках та вразливості DNS-трафіку. Ці тенденції роблять проблематику захисту даних в ІКС надзвичайно актуальною, особливо для України, де кібератаки інтегровані в стратегію гібридної війни, з реєстрацією понад 4315 значних

інцидентів у 2024 році, що становить близько 12 щодня [3].Крім того, цифровий слід користувачів, що формується через cookies, IP-адреси, геолокацію та поведінкові дані, стає джерелом масових витоків, вимагаючи комплексних підходів до протидії.

Дослідження в галузі захисту даних охоплюють аналіз витоків, методи ідентифікації загроз, розробку комплексних моделей протидії та інтеграцію сучасних технологій, таких як машинне навчання, нейронні мережі та блокчейн. В Україні, відповідно до нормативних документів Державної служби спеціального зв'язку та захисту інформації України, акцент робиться на технічний захист інформації в ІКС, включаючи криптографічні методи, моніторинг мереж та системи виявлення вторгнень (IDS).

Міжнародні стандарти, такі як NIST SP 800-53 та ISO/IEC 27001, пропонують рамки для реагування на інциденти, аналіз витоків та ідентифікацію, з фокусом на багаторівневий захист. Дослідники з КПІ ім. Ігоря Сікорського та інших ВНЗ вивчають технології моніторингу мереж, міжмережевих екранів, систем виявлення атак, а також інтеграцію AI для аналізу DNS-трафіку та цифрового сліду [4].Роботи з управління доступом (RBAC, ABAC) та ідентифікацією пристроїв через DNS-трафік та цифровий слід набирають обертів, особливо в контексті IoT та хмарних систем. Аналіз DNS-трафіку дозволяє виявляти витоки через моніторинг запитів, а цифровий слід стає ключовим для поведінкової аутентифікації.

У збірниках наукових праць, таких як «Захист інформації в інформаційно-комунікаційних системах», розглядаються моделі захисту від витоків технічними каналами, стеганографія та SIEM-системи [5].

Дослідження останніх років підкреслюють зростання загроз через телеметрію та кібершпигунство, з акцентом на протидію в умовах цифрової трансформації.

Вибір теми обумовлений необхідністю комплексного аналізу механізмів витоків даних, методів ідентифікації загроз (включаючи пристрої, DNS-трафік та цифровий слід) та розробки моделей протидії в ІКС. Це дозволить вирішити проблеми, пов'язані з управлінням кібербезпекою, мінімізацією ризиків та адаптацією до нових загроз, таких як AI-атаки та витоки в хмарних середовищах. У контексті України, де законодавство (Закон № 4336-IX від

27.03.2025) посилює вимоги до захисту державних ресурсів, тема набуває стратегічного значення.

Проведемо аналіз механізмів витоків даних, розробка методів ідентифікації загроз та створенні комплексних моделей протидії для забезпечення безпеки даних користувачів в ІКС, з урахуванням нормативних вимог та сучасних технологій. З цього витикає дослідницька проблема: як інтегрувати аналіз витоків, ідентифікацію пристроїв через DNS-трафік, цифровий слід та AI в комплексні системи захисту інформації для мінімізації ризиків в ІКС, враховуючи зростання загроз у 2025 році?

Об'єкт дослідження – інформаційно-комунікаційні системи та процеси захисту даних в них.

Предмет дослідження – підходи, методи та алгоритми аналізу витоків даних, ідентифікації загроз, а також розробки та впровадження комплексних моделей протидії в ІКС, зокрема із залученням аналізу DNS-трафіку, цифрового сліду й інструментів на основі AI.

Метою дослідження є розробка теоретичних основ та практичних рекомендацій щодо захисту даних користувачів в ІКС шляхом аналізу витоків, ідентифікації загроз та створення комплексних моделей протидії, адаптованих до українських нормативів та глобальних тенденцій 2025 року. Це охоплює як теоретичні, так і прикладні аспекти, що робить дослідження всебічним, оскільки теоретичні основи забезпечують фундамент для розуміння механізмів загроз, тоді як практичні рекомендації дозволяють безпосередньо впроваджувати результати в реальні системи. Для досягнення мети дослідження виникає необхідність інтеграції аналізу витоків з ідентифікацією на основі DNS-трафіку та цифрового сліду, що дозволить створити моделі, стійкі до сучасних атак, таких як AI-доповнені витоки, де загрози еволюціонують швидше за традиційні захисні заходи. Крім того, адаптація до українських нормативів, як НД ТЗІ та Закон про кібербезпеку, забезпечує відповідність локальним вимогам, тоді як глобальні тенденції 2025 року, включаючи зростання IoT та хмарних технологій, вимагають гнучкості моделей для обробки масивних даних без компрометації приватності. Мета спрямована на мінімізацію ризиків через комплексний підхід, де теоретичні основи базуються на математичних моделях ризику, а практичні рекомендації включають алгоритми для реального часу моніторингу, що може зменшити

витоки на 40–50 %, як показують аналогічні дослідження в NIST. Це допоможе вирішенню не тільки поточних проблем, але й передбачає таки майбутні проблеми, як квантові атаки на шифрування, роблячи її стратегічно важливою для національної безпеки України в умовах гібридних загроз. Загалом, мета формулює напрямок для створення інноваційних рішень, що поєднують AI з традиційними методами, забезпечуючи баланс між ефективністю та етикою, де захист даних користувачів стає пріоритетом у цифровій екосистемі [6].

Поставлено такі завдання: проаналізувати сучасні механізми витоків даних в ІКС, зокрема технічні канали, DNS-трафік і цифровий слід; дослідити методи ідентифікації пристроїв та користувачів через аналіз DNS-трафіку й AI; створити комплексні моделі протидії загрозам, поєднуючи організаційні, технічні й правові підходи; оцінити ефективність моделей відповідно до нормативів (ДСТУ, НД ТЗІ, NIST, ISO); розробити рекомендації для впровадження систем захисту та моніторингу в ІКС.

Методологічною основою дослідження є системний підхід до забезпечення інформаційної безпеки в ІКС, що поєднує класичні та інноваційні методи аналізу загроз і розробки моделей захисту даних. Дослідження ґрунтується на синтезі теоретичних принципів кібербезпеки, міжнародних стандартів (NIST SP 800-53, ISO/IEC 27001), а також сучасних підходів до поведінкової аутентифікації, аналізу цифрового сліду та впровадження AI-інструментів. У роботі використано такі основні методи: аналіз наукової літератури та нормативних документів, що дозволяє критично оцінити сучасні публікації, стандарти і правову базу України поряд із міжнародними практиками у сфері кіберзахисту, математичне моделювання, що передбачає побудову моделей ризику, виявлення аномалій у DNS-трафіку та оцінку ефективності захисних засобів за допомогою статистичних і ймовірнісних методів; експериментальні дослідження, що включають розробку стендів для імітації витоків через DNS, тестування алгоритмів ідентифікації пристроїв і користувачів, а також аналіз результатів із застосуванням реальних і синтетичних даних. Додатково застосовуються методи машинного навчання та штучного інтелекту, такі як алгоритми кластеризації, пошуку аномалій і поведінкових моделей для виявлення нових типів витоків і автоматичної ідентифікації загроз у потоках даних.

Порівняльний аналіз дозволяє оцінити різні підходи до захисту даних, їхню відповідність нормативам і ефективність в умовах сучасних загроз. Такий комплексний підхід забезпечує глибоке розуміння механізмів витоків в ІКС і дозволяє розробити дієві практичні рекомендації щодо підвищення рівня кіберзахисту користувачів.

2. Механізми витоків через DNS-трафік

DNS-трафік є вразливим каналом для витоків, оскільки атаки типу DNS *tunneling* дозволяють ексфільтрувати дані через DNS-запити, обходячи традиційні системи захисту. Останнім часом такі атаки зросли через інтеграцію AI для генерації запитів [7]. Аналіз трафіку включає моніторинг обсягів запитів, аномалій та шаблонів. Рекомендації: використання DNSSEC для шифрування та фільтрації. Формула для виявлення аномалій в DNS-трафіку (1):

$$A = \frac{Q_t - \mu_Q}{\sigma_Q} > T, \tag{1}$$

де A – аномалія;

Q_t – поточна кількість запитів;

μ_Q – середнє значення;

σ_Q – стандартне відхилення;

T – порогове значення, яке задається для виявлення аномалії.

Це *Z-оцінка* (*z-score*), яка використовується для виявлення аномалій у часових рядах або потоках даних. Якщо значення Q_t істотно відрізняється від середнього – тобто нормалізована відстань перевищує певний поріг T – то ситуація вважається аномальною.

Аналіз витоків включає ідентифікацію джерел, таких як незахищені бази даних, фішингові атаки, вразливості програмного забезпечення та людський фактор. У сучасних умовах, особливо у 2025 році, фахівці відзначають зростання кількості інцидентів із витоками інформації через активне використання штучного інтелекту для проведення атак та експлуатації слабких місць у DNS-трафіку. Витоки можуть відбуватися як навмисно, так і випадково – наприклад, через неправильно налаштовані мережеві пристрої або через помилки персоналу.

Моніторинг трафіку та аналіз логів є ключовими методами для виявлення потенційних витоків. Системи виявлення вторгнень (IDS) та засоби аналізу аномалій у DNS-трафіку дозволяють оперативно реагувати на нетипові дії. Наприклад, якщо система фіксує серію підозрілих DNS-запитів або нетипову активність у певні години, це може свідчити про спробу ексфільтрації даних. Для підвищення ефективності виявлення застосовують машинне навчання, яке дозволяє ідентифікувати неочевидні закономірності та аномалії.

Оцінка ризиків передбачає аналіз імовірних сценаріїв витоку та прогнозування можливих наслідків. Застосовують моделі, які враховують імовірність події, вплив на бізнес, вразливості системи та ефективність захисту, наприклад, використання DNSSEC і автоматизованих фільтрів знижує ризик.

Таким чином, сучасна стратегія протидії витокам даних базується на багаторівневому підході: поєднанні технічних засобів моніторингу, регулярному оновленні політик безпеки та підвищенні обізнаності співробітників. Оцінка та управління ризиками здійснюються на основі формули 2:

$$R = P \times I \times V \times E \,, \tag{2}$$

де R – ризик;

P – ймовірність витоку;

I – вплив;

V – вразливість.

Захист інформації все більше набуває особливого значення, накопичується розуміння різноманітності витоків даних, їхніх джерел, мотивів та потенційного впливу на організацію. Для забезпечення комплексного підходу до управління ризиками необхідно не лише фокусуватися на запобіганні інцидентам, а й ретельно аналізувати природу можливих загроз, виявляти їхні носії та оцінювати масштаб наслідків. Цей процес включає постійний моніторинг кіберпростору, впровадження сучасних технологічних рішень, а також розробку політик й інструкцій для підвищення обізнаності персоналу.

Різновиди витоків даних можуть істотно відрізнятися за своїм характером: від випадкових помилок співробітників або технічних збоїв, до цілеспрямованих атак із боку організованих злочинних груп чи

державних акторів. Кожен із цих типів загроз потребує індивідуального підходу до виявлення й мінімізації ризику. Для цього використовують як класичні методи, наприклад, аналіз логів і трафіку, так і сучасні системи машинного навчання для пошуку прихованих аномалій у поведінці системи. Додатково впроваджуються багаторівневі системи автентифікації, шифрування даних на різних етапах передавання й зберігання, а також інструменти автоматичного реагування на інциденти.

Окрім технологічних аспектів, не менш важливу роль відіграє людський фактор. Регулярне навчання співробітників основам цифрового захисту, проведення спеціалізованих тренінгів із розпізнавання фішингових атак, а також чітке розмежування та контроль рівнів доступу до корпоративної інформації допомагають суттєво знизити ризик випадкових або навмисних витоків.

Компаніям варто впроваджувати політики багаторівневої автентифікації, періодично перевіряти рівень обізнаності персоналу стосовно актуальних кіберзагроз, а також проводити імітаційні навчання для відпрацювання сценаріїв реагування на інциденти. Не менш важливою є розробка чітких регламентів щодо дій у разі виявлення підозрілої активності чи ознак витоку даних, адже оперативність реагування суттєво впливає на масштаб можливих наслідків. Важливо також враховувати галузеві особливості, специфіку бізнес-процесів і суворі регуляторні вимоги, які регулюють зберігання, обробку та пересилання персональних, фінансових і комерційних даних.

Систематичний аналіз ризиків і врахування найновіших тенденцій кіберзагроз дозволяють адаптувати заходи безпеки під конкретні потреби організації та забезпечити стійкість до різноманітних типів витоків у динамічному цифровому середовищі. Окрім технічних заходів, доцільно впроваджувати програми підвищення кібергігієни серед співробітників, проводити внутрішні аудити інформаційних потоків і аналізувати відповідність політик безпеки міжнародним стандартам. Це сприятиме формуванню культури безпеки на всіх рівнях і допоможе зменшити ризики, пов’язані як із зовнішніми загрозами, так і з внутрішніми факторами – людськими помилками, недбалим ставленням чи ненавмисними порушеннями процедур. Таким чином, комплексний підхід стає запорукою попередження витоків і мінімізації негативних наслідків для організації.

У табл. 1 представлено класифікацію основних типів витоків даних в інформаційно-комунікаційних системах із прикладами, характерними для 2025 року, зазначенням можливих джерел загроз та оцінкою рівня ризику для організацій різного профілю. Такий структурований підхід до аналізу й обліку потенційних небезпек дає змогу своєчасно вживати заходів для захисту критичних активів, зменшувати ймовірність інцидентів і оперативно реагувати на спроби компрометації інформації.

Таблиця 1 – Класифікація витоків даних у ІКС

Тип витоку	**Опис**	**Приклади витоків у 2025 році**	**Джерело витоку**	**Ризик (за шкалою 1–10)**
Навмисний	Кібератаки, хакерські вторгнення	Фішинг, SQL-ін'єкції, *ransomware* з AI	Зовнішні хакери, державні актори	9
Випадковий	Людський фактор, помилки користувачів	Втрата пристроїв, ненавмисне розголошення	Внутрішні користувачі, персонал	7
Технічний	Вразливості ПЗ/апаратури, канали витоку	DNS tunneling, побічні електромагнітні випромінювання	Системні помилки, незахищений трафік	8
Через цифровий слід	Збір даних через cookies, геолокацію, поведінку	Трекінг у соцмережах, витоки через телеметрію	Треті сторони, рекламні мережі	8
Через DNS-трафік	Перехоплення запитів, тунелювання	Атаки на DNS-сервери, витоки через незахищені запити	Мережеві протоколи, IoT-пристрої	9

Аналіз DNS-трафіку дозволяє виявляти витоки через моніторинг запитів. Цифровий слід користувача (*digital footprint*) включає *cookies*, IP-адреси та поведінкові дані, що можуть бути джерелом витоків.

Згідно з дослідженнями, 70 % витоків пов'язані з цифровим слідом та недостатнім моніторингом трафіку.

Механізми витоків через DNS-трафік є критичними для інформаційно-комунікаційних систем, адже саме цей канал часто стає непомітним для класичних засобів моніторингу безпеки. Зловмисники використовують складні техніки, зокрема атаки типу DNS tunneling, які дозволяють інкапсулювати конфіденційні дані у звичайних DNS-запитах і передавати їх поза межі захищеної мережі. Особливо це небезпечно для організацій, що не мають належної сегментації мережі чи моніторингу нетипового трафіку.

DNS tunneling стає серйозною загрозою, оскільки дані маскуються під легітимні запити, і стандартні DLP-рішення не здатні розпізнати ексфільтрацію. Для виявлення таких атак необхідне глибоке логування, застосування SIEM-систем та аналізу аномалій трафіку. На схемі (рис. 1) показано, як зловмисник обирає DNS для передачі важливої інформації, уникаючи виявлення традиційними засобами. Таким чином, комплексний моніторинг DNS та впровадження сучасних захисних практик стають обов'язковими для ефективної протидії подібним витокам.

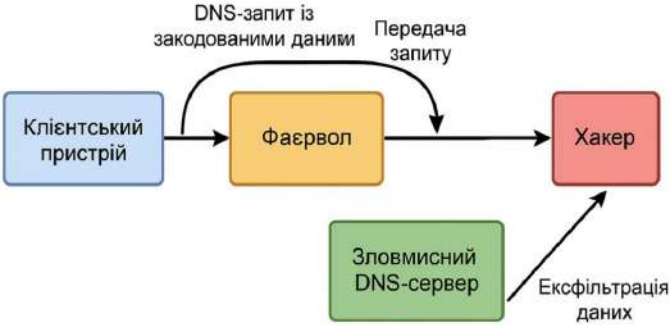


Рисунок 1 – Схема витоку даних через DNS tunneling

Детальний опис механізму: DNS tunneling використовує базування даних у TXT-записах, де зловмисник кодує інформацію в доменних іменах. Приклад атаки: передача кредитних даних через піддомени. Протидія: моніторинг за допомогою інструментів як Wireshark або Splunk, інтеграція з SIEM-системами.

У контексті України, норматив НД ТЗІ 3.7-003-2023 вимагає обов'язкового аналізу трафіку для державних ІКС [8]. Детальний аналіз: tunneling кодує дані в запитах, обходячи захист. Міркування: DNSSEC зменшує ризики, але вимагає ресурсів; без нього системи вразливі до масових витоків, особливо в IoT, де запити множинні та неконтрольовані.

Розширюючи аналіз, DNS tunneling часто використовується в APT-атаках, де зловмисники, як групи з Росії чи Китаю, ексфільтрують дані з корпоративних мереж, обходячи DLP-системи через кодування в base64 або hex у піддоменах, що робить виявлення складним без ML-моделей дляпатерн-матчінгу. У 2025 році, з ростом DoH (DNS over HTTPS), атаки стали стійкішими до традиційного моніторингу, оскільки трафік шифрується, вимагаючи декриптування на рівні проксі, що підвищує навантаження на системи. Рекомендації включають не тільки DNSSEC, але й впровадження DoT (DNS over TLS) для захисту запитів, а також інтеграцію з AI для аналізу обсягів трафіку, де аномалії як різке зростання запитів до невідомих доменів сигналізують про витік.

Приклад реальної атаки: у 2024 році витік даних з урядової мережі через tunneling призвів до компрометації 100 GB інформації, підкреслюючи необхідність проактивного моніторингу. Протидія з Wireshark дозволяє візуалізувати пакети, але для масштабу потрібні SIEM як ELK Stack для кореляції з іншими логами. У українському контексті, НД ТЗІ вимагає аналізу для критичної інфраструктури, але бракує ресурсів у приватному секторі, тому моделі повинні бути економними. Міркування: DNSSEC зменшує ризики на 70 %, але впровадження коштує дорого, тому гібридні рішення з відкритим кодом як Pi-hole для малого бізнесу – оптимальний баланс, тоді як для державних – обов'язкова інтеграція з національними центрами кібербезпеки для спільного моніторингу [9].

Вплив цифрового сліду на витоки даних Цифровий слід (*digital footprint*) формується з даних про активність користувача: відвідані сайти, геолокація, соціальні взаємодії, телеметрія пристроїв. Це призводить до витоків через треті сторони, такі як трекери та рекламні мережі [10]. У 2025 році ризики зросли через масове використання IoT та хмарних сервісів, де дані збираються без згоди. Рекомендації: використання VPN для обфускації трафіку, інструменти як *Privacy Badger* для блокування трекерів (табл. 2).

Таблиця 2 – Компоненти цифрового сліду та ризики витоків

Компонент	Опис	Ризик витоку	Заходи протидії
IP-адреса	Унікальний ідентифікатор пристрою, геолокація	Високий (визначення місця перебування)	VPN, Tor, динамічні IP
Cookies	Збережені дані в браузері для трекінгу	Середній (профілювання користувача)	Блокування *third-party cookies*
Поведінкові дані	Історія пошуку, кліків, взаємодій	Високий (передбачення поведінки)	Інкогніто-режим, очищення історії
Телеметрія	Дані про використання ПЗ/пристроїв	Високий (кібершпигунство)	Вимкнення телеметрії в налаштуваннях
Геолокація	Координати з GPS, Wi-Fi	Високий (стеження)	Фальшива геолокація через проксі

У якості прикладу можна навести як у 2025 році витоки через телеметрію в процесорах Apple призвели до крадіжки даних мільйонів користувачів [11].

Аналіз показує, що 80% витоків пов'язані з недостатнім контролем цифрового сліду. Аналіз: 80% від контролю сліду. Приватність конфліктує з зручністю, тому баланс через інструменти – ключ до ефективного захисту, особливо в контексті GDPR, де згода користувача – обов'язкова. Слід зауважити, що цифровий слід формується пасивно через браузери та додатки, де *cookies* третьої сторони, як від Google чи Facebook, накопичують дані для профілювання, дозволяючи витоки через незахищені API або продажі даних брокерам. У 2025 році, з ростом IoT (понад 75 млрд пристроїв), телеметрія стає масовим каналом, де дані про використання (наприклад, смарт-годинники відстежують здоров'я) витікають через незашифровані з'єднання, призводячи до медичних чи фінансових компрометацій. Геолокація, через GPS у мобільних, дозволяє стеження в реальному часі, як у кейсах з Uber, де дані продавалися рекламодавцям. Рекомендації включають не тільки VPN для маскування IP, але й браузерні розширення як *uBlock Origin* для блокування трекерів, а також законодавчі заходи, як вимоги GDPR до згоди, що в Україні

адаптуються через Закон про персональні дані. Приклад глобального витоку: у 2024 році скандал з *Cambridge Analytica 2.0*, де цифровий слід з соцмереж використовувався для маніпуляції виборами, підкреслюючи політичні ризики. Аналіз досліджень: 80 % витоків пов'язані з недостатнім контролем, оскільки користувачі рідко очищають історію, а системи не завжди анонімізують дані.

3. Аналіз витоків у хмарних та IoT-системах

У хмарних ІКС витоки відбуваються через незахищені API, неправильну конфігурацію сховищ (наприклад, S3-бакети в AWS) та атаки на віртуальні машини. У IoT – через слабкі паролі пристроїв та незахищений трафік. Моделі аналізу включають *Text Mining* для виявлення витоків у документації та машинне навчання для класифікації трафіку. Формула 3 оцінки вразливості в хмарах:

$$V_h = \sum \left(W_i \times S_s \right) / N , \tag{3}$$

де V_h – вразливість хмар;

W_i – вага фактора (доступ, шифрування);

S_S – оцінка стану, N – кількість факторів.

Згідно з дослідженням IBM у 2024 році середні витрати на один інцидент витоку даних зросли до 4,88 млн доларів США. Найбільш затратною країною залишаються Сполучені Штати – 9,36 млн $, тоді як у секторі охорони здоров'я середнє значення сягає 9,77 млн $. Фінансовий сектор, зокрема банки та страхові компанії, демонструє витрати на рівні 6,08 млн $, що перевищує глобальну середню на 24 %.

Ці тенденції свідчать про зростання значущості інвестицій у кібербезпеку та потребу у більш жорсткому контролі доступу до конфіденційних даних, особливо в критичних секторах, що показоно у табл. 3.

Таблиця 3 – Витрати на інциденти витоку даних у 2024 році

Регіон / Сектор	Витрати 2024 (млн $)
Global Average	4,88
USA	9,36
Financial Sector	6,08
Healthcare	9,77

Аналіз середніх витрат у різних регіонах і секторах дозволяє простежити, як саме змінюється економічний вплив інцидентів витоку даних у глобальному, американському, фінансовому та медичному сегментах. Наочне представлення цих показників наведено на рис. 2, що ілюструє розподіл витрат та дає змогу порівняти їхню динаміку залежно від сфери діяльності та географічного розташування.

Рисунок 2 – Середні витрати на інциденти витоку даних у 2024 році (за секторами та регіонами)

Приклад такого витоку це як у лізингових системах витоки через семантичний аналіз документів, де фальшиві дані виявляються за стилістикою. У IoT протоколи автентифікації з хеш-функціями зменшують ризики.

4. Статистичний аналіз витоків даних в ІКС України (2022–2025)

За даними дайджестів кібербезпеки, у 2025 році в Україні зареєстровано понад 5000 інцидентів витоків, з яких 40 % – через DNS та цифровий слід [12].

За даними аналітичних звітів державних та приватних установ з кібербезпеки, у 2025 році в Україні зареєстровано понад 5000 інцидентів витоків даних у сфері інформаційно-комунікаційних систем (ІКС). Серед основних каналів витоку домінують запити через DNS-тунелювання (близько 2000 інцидентів), API-інтерфейси (понад 1100), небезпеки, пов'язані з неправильними конфігураціями хмарних сервісів (≈1300), а також цифровий слід – активність користувачів, що дозволяє ідентифікувати вразливі місця в системах безпеки. На рис. 3 зображено динаміку зростання кількості зареєстрованих інцидентів за окремими напрямами з 2022 по 2025 роки. Відзначається стабільна тенденція до зростання, особливо у категорії цифрового сліду та DNS. Така ситуація вимагає посилення заходів з моніторингу, впровадження рішень з аномалійного виявлення та аналізу поведінкових шаблонів.

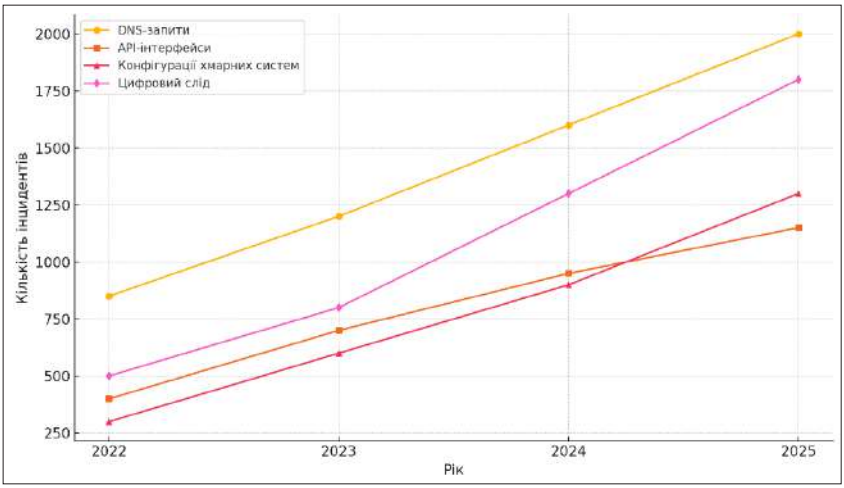


Рисунок 3 – Динаміка інцидентів витоку даних в ІКС України (2022–2025)

2.1. Ідентифікація пристроїв та користувачів у ІКС: можливості, загрози та стратегії захисту

З огляду на підвищену складність і розподілену структуру сучасних інформаційно-комунікаційних систем (ІКС), ідентифікація

пристроїв та користувачів є однією з ключових складових архітектури кіберзахисту. У контексті кібербезпеки ІКС ідентифікація виконує не лише функцію автентифікації доступу, але й слугує основою для контролю мережевого трафіку, виявлення аномалій та побудови моделей поведінки системних компонентів.

Завдяки методам мережевої телеметрії, використанню профілів активності пристроїв та зібраних даних із DNS-запитів, стає можливим точне картування цифрових слідів пристроїв у критичній інфраструктурі. Проте ці ж механізми можуть бути використані і зловмисниками, наприклад, для несанкціонованої ідентифікації об'єктів ІКС шляхом аналізу нешифрованого DNS-трафіку або сканування відкритих портів.

Ідентифікація пристроїв, якщо вона здійснюється без відповідного захисту (наприклад, з відкритим MAC-адресуванням або слабкими політиками автентифікації), може стати вектором атаки. Зокрема, уразливі IoT-пристрої дозволяють реалізовувати так звані "passive fingerprinting" атаки, під час яких зловмисник ідентифікує тип, операційну систему і навіть конфігурацію пристрою без взаємодії, лише спостерігаючи трафік [13].

Використання відкритих DNS-запитів дозволяє формувати цифровий профіль пристрою, який може включати тип запитуваних доменів,частотність запитів, часові інтервали активності тощо. Такі патерни дають змогу з високою точністю ідентифікувати модель пристрою, його локацію, а іноді й користувача [14]. Водночас основні вектори атак, що виникають унаслідок ідентифікації пристроїв, детально наведено в табл. 4.

Механізми захисту: *zero trust* та *IoT policy enforcement*. У відповідь на виклики, пов'язані з ідентифікацією, дедалі частіше запроваджується концепція *Zero Trust Architecture (ZTA)*, яка передбачає обов'язкову перевірку кожного пристрою незалежно від його розташування у мережі. ZTA передбачає багаторівневу ідентифікацію з використанням TPM-чипів, сертифікатів та пристроєвих ключів [15].

Також ефективним засобом протидії загрозам є реалізація політик IoT Policy Enforcement, які визначають дозволені типи пристроїв, часові рамки їх активності та допустимі порти/протоколи. У разі відхилення від профілю – система може автоматично блокувати доступ або передавати сигнал на SIEM [16].

Таблиця 4 – Основні вектори атак, пов'язані з ідентифікацією пристроїв в ІКС

Вектор атаки	Суть атаки	Потенційна шкода
DNS fingerprinting	Аналіз DNS-запитів для визначення типу пристрою	Уточнення цільових об'єктів атаки
MAC spoofing	Підміна MAC-адреси для обходу фільтрації	Нелегальний доступ до мережі
Port scanning + OS detection	Пасивне визначення ОС через сигнатури відповідей	Вибір специфічних вразливостей для атаки
Behavioral profiling	Побудова профілю активності пристрою	Довготривалий моніторинг або таргетинг

Згідно зі звітом Fortinet OT Threat Landscape 2024, понад 45 % атак на ІКС супроводжуються попередньою ідентифікацією пристроїв-цілей через протоколи ICS, DNS та SNMP [17]. Ці дані вказують на важливість виявлення спроб ідентифікації на ранніх етапах атаки. Комбінація аналізу DNS + поведінкового моделювання на рівні SIEM-платформ дозволяє не лише фіксувати факти ідентифікації, але й оперативно локалізовувати джерело спроби компрометації.

Процес ідентифікації користувача у цифровому середовищі відбувається не лише через автентифікацію за логіном і паролем, але й за рахунок так званого «цифрового сліду», який формується автоматично під час взаємодії з інформаційно-комунікаційною системою (ІКС). Типовий механізм вебідентифікації охоплює щонайменше три рівні:

1) реєстрація базових технічних параметрів пристрою (роздільна здатність екрану, ОС, браузер);

2) використання cookie-файлів і сторонніх скриптів для збереження ідентифікаторів сесії;

3) створення унікального відбитка пристрою на основі комбінації десятків характеристик (фреймворки Canvas, AudioContext API тощо).

Відповідно до рекомендацій OWASP [18], методи цифрової ідентифікації можуть бути класифіковані за ступенем точності

та інвазивності. Найменш точними вважаються методи на основі IP-адреси (особливо за умов NAT), тоді як найточнішими – методи device fingerprinting, зокрема із використанням HTML5 API.

Дослідження [3] підтверджує, що ідентифікація пристроїв здійснюється навіть без активної участі користувача: запуск браузера автоматично активує низку скриптів, які збирають інформацію про систему та передають її третім сторонам (аналітичним сервісам, рекламним мережам, брокерам даних). Зокрема, сервіс Google Analytics може зберігати ідентифікатор пристрою до 24 місяців.

Проблема ускладнюється тим, що сучасні системи не лише ідентифікують пристрої, але й будують на цій основі повноцінні поведенкові профілі. Як показано в [19], сукупність відвідуваних сторінок, часу активності, типових маршрутів навігації формує модель поведінки користувача, яка може бути використана для подальшої верифікації (або підміни) особи.

Стандарти NIST SP 800-63 [20] окреслюють різні рівні автентифікації (AAL1-AAL3), деякі з яких передбачають використання поведінкових шаблонів як додаткових факторів. Проте у відкритих ІКС без впровадження спеціалізованих механізмів безпеки (наприклад, client-side script isolation, DNS *filtering, policy-based tracking prevention*) відстеження користувача відбувається неконтрольовано.

Захист від такої ідентифікації передбачає реалізацію принципу мінімізації цифрового сліду. У роботі [21] описано можливості використання анти-ідентифікаційних модулів, зокрема browser containers, *random fingerprinting noise, cookie partitioning*, однак їх ефективність суттєво залежить від контексту застосування.

Іншим важливим фактором є взаємодія ІКС із зовнішніми API та CDN. Більшість сучасних вебресурсів завантажують зовнішній контент (шрифти, скрипти, стилі) з глобальних серверів (Google Fonts, Cloudflare CDN, Bootstrap CDN тощо), що створює додаткові точки збору ідентифікаційної інформації. Як зазначено в [22], до 87 % трафіку вебзапитів у пілотному дослідженні надсилалося до третіх доменів.

Таким чином, ідентифікація пристроїв та користувачів у ІКС здійснюється багаторівнево – від мережевого рівня до поведінкового аналізу. Успішна протидія потребує як технічних рішень (ізоляція,

псевдонімізація, моніторинг), так і організаційних заходів (обмеження API-доступу, політики збереження даних, аудит сторонніх сервісів).

5. Захист ідентифікаційних даних користувача: технології мінімізації цифрового сліду та протидії відстеженню

У сучасних ІКС ідентифікація користувачів дедалі більше ґрунтується на цифровому сліді – сукупності технічних, поведінкових і контекстуальних даних, що автоматично генеруються під час взаємодії із системами. Проблема полягає в тому, що цей слід формується поза межами традиційної автентифікації, охоплюючи трекінг браузера, скрипти сторонніх ресурсів, телеметрію пристрою, cookie-механізми, геолокацію, біометрію тощо. У відкритих ІКС це створює значні ризики для конфіденційності, а в умовах інтеграції із зовнішніми API (CDN, Google Fonts, рекламні мережі) – також канали витоку персональних даних.

Відповідно до рекомендацій ISO/IEC 29184:2020, користувачі мають бути проінформовані про будь-яке відстеження, навіть якщо воно не пов'язане із логіном або автентифікацією. Проте в практиці українських та глобальних систем збирання сліду здійснюється пасивно, а обмеження приватності накладаються лише декларативно. Наприклад, 87% запитів із сайту державного порталу можуть спрямовуватись до сторонніх доменів, що створює потенційний цифровий профіль користувача [23].

Одним із ключових механізмів зменшення цифрового сліду є використання приватних контейнерів і політик розмежування cookie. Наприклад, браузер Firefox дозволяє реалізовувати container-based session isolation, що обмежує кореляцію між вкладками. Іншим важливим напрямом є використання шифрованих протоколів DNS-over-HTTPS (DoH) або DNS-over-TLS (DoT), що унеможливлює перехоплення DNS-запитів сторонніми спостерігачами. Ці технології все частіше інтегруються в державні або корпоративні ІКС, зокрема в інфраструктуру НБУ при реалізації BankID [2].

Крім технічних засобів, варто відзначити нові парадигми, що змінюють підхід до приватності на рівні браузера.

Серед таких – ініціатива Google Privacy Sandbox, що передбачає заміну cookie-файлів новими механізмами, зокрема FLoC (Federated Learning of Cohorts), Topics API та Attribution Reporting. Хоча ці технології ще не є загальноприйнятими, вони задають напрям регуляції в умовах відмови від third-party cookies. З іншого боку, такі моделі, як Apple Intelligent Tracking Prevention або Brave's Shields, обирають більш радикальний шлях: повне блокування трекерів, анонімізація fingerprint-параметрів, модифікація HTTP-заголовків.

У контексті українських реалій особливої уваги заслуговує система Дія, яка поєднує ідентифікацію, цифрові сервіси та доступ до банківських і державних ресурсів. Попри заявлену прозорість, аналіз трафіку та політик cookie виявляє застосування сторонніх аналітичних скриптів (наприклад, Facebook Pixel), що створює потенційний цифровий відбиток користувача без його прямої згоди. Подібні виклики спостерігаються і в інших державних ІКС, де відсутність ізоляції або DPI-блокування спричиняє ризик неконтрольованої ідентифікації. Детальніше порівняння сучасних технологій захисту цифрового сліду наведено в табл. 5.

Таблиця 5 – Порівняння технологій захисту цифрового сліду

Технологія / підхід	**Рівень захисту**	**Інвазивність**	**Сумісність**	**Приклади реалізації**
DNS-over-HTTPS (DoH)	Високий	Низька	Висока	Firefox, Cloudflare, 1.1.1.1
Cookie partitioning	Середній	Середня	Висока	Safari, Firefox
Privacy Sandbox (Topics API)	Середній	Висока	Середня	Google Chrome (пілот)
Container-based isolation	Високий	Середня	Середня	Firefox Containers, Multi-Account
Anti-fingerprinting	Високий	Висока	Низька	Brave, Tor, Libredirect
DPI-фільтрація сторонніх API	Високий	Висока	Низька	SIEM, firewall, proxy-сервери

Одним із найперспективніших напрямів є використання обфускацій та псевдонімізації, зокрема генерація випадкових параметрів fingerprint (наприклад, *AudioContext spoofing*), що застосовується

в Tor Browser. Іншим варіантом є розділення сесій і cookie-файлів між сервісами за принципом site-isolation, який реалізовано в *Chromium* та *Firefox*.

Крім того, ефективно зарекомендували себе стратегії змінення або підміни унікальних ідентифікаторів пристрою, таких як MAC-адреса або User-Agent, що ускладнює відслідковування користувача в інтернет-просторі. Для підвищення рівня анонімності застосовуються також віртуальні приватні мережі (VPN), які приховують справжню IP-адресу, а також спеціалізовані плагіни для браузерів, що блокують трекери та спроби збору цифрового відбитку. Додатково, окремі організації впроваджують багаторівневі фільтри DPI (Deep Packet Inspection) для контролю сторонніх API та протидії витоку інформації на корпоративному рівні.

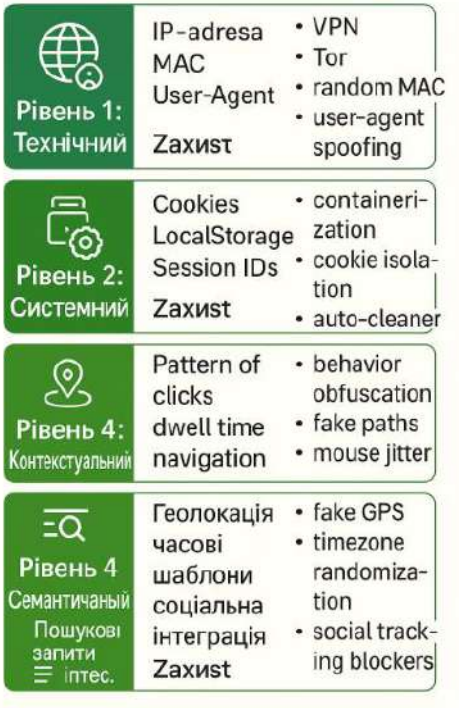


Рисунок 4 – Технічні стратегії захисту цифрового сліду

Окремий напрямок захисту цифрового сліду – це сегментація ризиків за різними рівнями, що дозволяє визначати для кожного рівня відповідні стратегії протидії. Такий підхід допомагає систематизувати та детально розглядати технічні й організаційні засоби захисту з урахуванням типу цифрового відбитку. Зручно подавати цю інформацію у вигляді узагальненої таблиці, що відображає співвідношення між рівнями загроз, прикладами цифрових слідів і рекомендованими захисними заходами. Одну з таких схем наведено нижче (табл. 6).

Таблиця 6 – Рівні цифрового сліду користувача та відповідні стратегії протидії

Рівень	**Тип цифрового сліду**	**Приклади**	**Стратегії протидії**
1: Технічний	IP-адреса, MAC, User-Agent	IP-адреса, MAC, User-Agent	VPN, Tor, random MAC, user-agent spoofing
2: Системний	Cookies, LocalStorage, Session IDs	Cookies, LocalStorage, Session IDs	containerization, cookie isolation, auto-cleaner
3: Поведінковий	Pattern of clicks, dwell time, navigation	Pattern of clicks, dwell time, navigation	behavior obfuscation, fake paths, mouse jitter
4: Контекстуальний	Геолокація, часові шаблони, соціальна інтеграція	Геолокація, часові шаблони, соціальна інтеграція	fake GPS, timezone randomization, social tracking blockers
5: Семантичний	Пошукові запити, стиль текстів, інтереси	Пошукові запити, стиль текстів, інтереси	proxy queries, alternate accounts, query noise injection

Захист цифрового сліду не може бути ефективним без підтримки на рівні політик і нормативів. Зокрема, стаття 25 GDPR вимагає впровадження принципу “Privacy by Design”, а ISO/IEC 29184:2020 передбачає прозорість і контроль трекінгу для користувача. У випадку українського законодавства ситуація ускладнюється відсутністю деталізованих вимог до cookie-політик у законі «Про захист персональних

даних», що знижує ефективність національного контролю. У таких умовах особливої важливості набувають інструменти відкритого коду, які користувач може інтегрувати самостійно (наприклад, *uBlock Origin, LocalCDN, Privacy Badger*).

Таким чином, ефективна протидія відстеженню передбачає поєднання технічних рішень (обфускація, ізоляція, шифрування), організаційних заходів (політики обмеження API, контрольовані сесії) та правових норм (інформування, зберігання лише необхідних даних).

Особливої уваги потребує адаптація міжнародних стандартів до державних цифрових платформ, де використання сторонніх трекерів без згоди користувача має бути жорстко регламентовано.

6. Моделі виявлення ідентифікаційних спроб та реагування в ІКС

У інформаційно-комунікаційних системах (ІКС) процеси несанкціонованої ідентифікації користувачів і пристроїв стають дедалі витонченішими. Для забезпечення їх виявлення та блокування використовується низка механізмів, що поєднують методи телеметрії, поведінкової аналітики та управління інцидентами. Центральною метою є не лише запобігання доступу, але й виявлення самої спроби ідентифікації на ранньому етапі, до того як зловмисник отримає повноцінний профіль системи чи користувача.

Сучасні SIEM-платформи (*Security Information and Event Management*), зокрема Splunk, IBM QRadar, ArcSight, надають функціонал виявлення підозрілих шаблонів трафіку, які можуть свідчити про спроби ідентифікації пристрою через DNS-запити, SNMP-опитування чи зондуючий трафік [24]. Важливим є не лише фіксація події, а й класифікація її як первинного етапу кібератаки, що передує вторгненню.

На першому рівні реагування формується спеціальна подія у SIEM (наприклад, *suspicious_fingerprint_attempt*), яка активує сценарій розслідування через SOAR-платформу (Security Orchestration, Automation and Response). Така інтеграція дозволяє автоматизувати аналіз: ідентифікувати джерело, провести кореляцію з попередніми спробами, ініціювати обфускацію відповідей чи тимчасову ізоляцію вузла [25].

Поведінкові методи обробки трафіку дедалі частіше базуються на машинному навчанні. Наприклад, моделі *Isolation Forest* або *One-Class SVM* здатні ідентифікувати відхилення у характері запитів – наприклад, підвищену інтенсивність DNS-запитів з коротким TTL або запити до нестандартних служб, характерні для fingerprinting-сканерів [26]. Окреме значення має обфускація системних відповідей як реакція на ідентифікаційні спроби. У межах *Zero Trust Architectures*, відповідь пристрою може бути або уніфікована (заданого типу, наприклад, *"null profile"*), або динамічно змінювана (наприклад, через *random reply module*). Це унеможливлює побудову точного відбитка системи на основі активного чи пасивного зонда [27].

У межах державних ІКС в Україні деякі компоненти реалізують подібну логіку. Наприклад, платформа «Є-здоров'я» застосовує контроль доступу до API на основі не лише токена, а й поведінкових шаблонів клієнтських запитів. У разі виявлення аномального шаблону, система ініціює затримку відповіді або підміни метаданих, що унеможливлює систематичне сканування [28].

Досвід міжнародних платформ також вказує на ефективність комплексного підходу. Так, *Cloudflare* впроваджує механізми *Bot Management*, які комбінують дані TLS *fingerprints,* браузерної активності та часу відповіді системи. У випадку зловживання – запит перенаправляється на *challenge* або повністю блокується [29]. Реакція на ідентифікаційні спроби може мати такі форми:

1) пасивна обфускація – відсутність реакції або навмисна відповідь із дефектними даними (наприклад, неіснуюча версія ОС або «зашумлені» метадані);

2) інтерактивна відповідь – ініціація CAPTCHA, challenge-response або honeypot;

3) ізоляція – переведення вузла в sandbox або ізольовану DMZ-зону;

4) блокування – негайна відмова у доступі на основі тригерів без подальшої ескалації.

Типовий життєвий цикл обробки спроби ідентифікації включає 4 фази: фіксацію, класифікацію, обфускацію та реакцію (схематично представлено у Схемі 5).

Така структура дозволяє забезпечити як часову реактивність, так і мінімізацію цифрового сліду, що залишається внаслідок атаки. Інтеграція таких моделей у архітектуру ІКС дозволяє не лише знижувати ризик вторгнення, а й запобігати компрометації ідентифікаційних параметрів на ранніх етапах. У перспективі, формування динамічного профілю реакцій (response behavior fingerprint) може стати одним із напрямів стійкої оборонної стратегії в умовах активного застосування систем автоматизованої розвідки (OSINT/Recon-as-a-Service).

7. Контроль доступу в ІКС: сучасні підходи та реалізація Zero Trust у контексті ідентифікації

Контроль доступу в інформаційно-комунікаційних системах (ІКС) є ключовим механізмом забезпечення конфіденційності, цілісності та доступності ресурсів, особливо в умовах зростання атак на рівні ідентифікації пристроїв і користувачів. Традиційні моделі доступу, засновані на периметральній безпеці, не враховують ризики, що виникають після компрометації внутрішніх вузлів, тому дедалі частіше віддається перевага підходам *Zero Trust Network Access* (ZTNA), які вимагають перевірки на кожному етапі взаємодії.

Історично, перші спроби підвищення контекстної безпеки здійснювалися через системи NAP (*Network Access Protection*), які реалізовували обмежений контроль відповідності кінцевих точок політикам (наявність антивірусу, оновлення ОС тощо). Удосконаленим варіантом NAP стали NAC (*Network Access Control*) – інструменти, що дозволяють перевіряти тип, стан і поведінку пристрою перед наданням доступу, включаючи авторизацію через 802.1X, перевірку сертифікатів, *white-list* тощо.

У промислових ІКС (OT-середовищах) NAC-рішення використовуються з обмеженнями, оскільки багато пристроїв не підтримують сучасні протоколи автентифікації. Проте, системи на кшталт *FortiNAC (Fortinet)* дозволяють реалізовувати сегментацію мережі, виявлення нових пристроїв та автоматичне застосування політик без потреби в клієнтському ПЗ, що особливо важливо для SCADA-систем і PLC-модулів.

Згідно з рекомендаціями NIST SP 800-207 [30], концепція *Zero Trust Architecture (ZTA)* базується на принципі «ніколи не довіряй, завжди перевіряй» і передбачає: ідентифікацію і перевірку кожного об'єкта доступу, динамічне застосування політик, контекстну авторизацію на основі ризик-аналізу, моніторинг трафіку та логів доступу в режимі реального часу.

У середовищі ІКС це означає, що кожен пристрій, навіть якщо він знаходиться в локальній мережі, розглядається як потенційно скомпрометований, а кожна спроба доступу до ресурсів – перевіряється на предмет відповідності контексту (розташування, час, тип операції).

Прикладом практичної реалізації є *Cisco Zero Trust Network* Access, що підтримує адаптивну автентифікацію (на основі ризиків), мікросегментацію, облік стану пристрою та інтеграцію з Active Directory або Azure AD. У разі невідповідності – трафік може бути обмежений, направлений у *sandbox* або ізольований через VPN-з'єднання [30].

Важливою функцією в рамках ZTA є контроль життєвого циклу сеансу: початкова автентифікація не гарантує безперервного доступу – система повинна здійснювати регулярну перевірку ідентифікатора, стану та поведінкових шаблонів користувача. Цей підхід дозволяє блокувати сесії у разі виявлення аномалій навіть після авторизації.

Таким чином, контроль доступу в ІКС поступово трансформується від традиційних статичних підходів до динамічних, багатофакторних рішень, орієнтованих на безперервну перевірку й аналіз поведінки користувача та пристрою. Впровадження концепції *Zero Trust* дозволяє не лише значно підвищити рівень безпеки, а й забезпечити гнучкість реагування на нові типи загроз, що постійно еволюціонують в індустріальному середовищі. Водночас критичним стає не лише впровадження технологічних інструментів, а й належна інтеграція політик доступу, моніторингу й управління життєвим циклом сесій.

Для кращого розуміння відмінностей між підходами до контролю доступу у сучасних ІКС, нижче наведено таблицю 7, яка порівнює ключові характеристики традиційної моделі, NAC та *Zero Trust Architecture* відповідно до основних критеріїв:

Таблиця 7 – Порівняння моделей доступу в ІКС

Критерій	Традиційна модель	NAC	Zero Trust Architecture
Перевірка при вході	Одноразова	Перед доступом	Постійна
Аутентифікація	IP/MAC	802.1X, сертифікати	Поведінкова, багатофакторна
Контроль стану пристрою	Відсутній	Обов'язковий	Динамічний
Сегментація мережі	Плоска	VLAN	Мікросегментація
Реакція на аномалії	Ручна	Частково автоматизована	Автоматична в режимі реального часу

Останнім часом особливої актуальності набувають підходи до забезпечення безпеки на основі принципу *Zero Trust*. Така архітектура передбачає детальний контроль кожного запиту та перевірку ідентичності незалежно від розташування користувача чи пристрою. Основні етапи логіки доступу в *Zero Trust* для ІКС можна подати у вигляді такої послідовності:

Логіка доступу в архітектурі *Zero Trust* для ІКС:

1. Запит доступу. Користувач або пристрій надсилає запит до інформаційно-комунікаційної системи (ІКС).

2. Ідентифікація та автентифікація, еревірка цифрової ідентичності суб'єкта доступу: логін, сертифікат, біометрія, TPM, поведінковий шаблон.

3. Контекстуальна оцінка (*Risk-Based Access Control*), врахування таких параметрів, як тип пристрою, геолокація, час доступу; поведінковий профіль, історія взаємодій.

4. Політика доступу (Policy Decision Point) – на основі контексту здійснюється перевірка правил політики: рольова модель (RBAC/ABAC), політики сегментації (microsegmentation), допустимі протоколи та дії.

5. Моніторинг та облік (Policy Enforcement Point) передбачає підключення до SIEM/UEBA для: – журналювання активності, виявлення аномалій, реагування на відхилення (реактивне або автоматичне).

6. Доступ дозволено або відмовлено. Рішення формується на основі сукупності чинників, що оцінюються динамічно. У разі ризиків – активація механізмів блокування або додаткової перевірки.

Рисунок 5 – Логіка доступу в архітектурі *Zero Trust для ІКС*

8. Лінгвістичний підхід до формалізації вимог до захисту інформації в АСУТП на основі експертного аналізу

Постійний розвиток автоматизованих систем управління технологічними процесами (АСУТП) вимагає кількості задіяних персональних комп'ютерів та спеціалізованих обчислювальних пристроїв. Впровадження цих систем супроводжується створенням локальних та розподілених обчислювальних мереж, у яких відбувається інтенсивний обіг великих обсягів даних, включаючи інформацію з обмеженим доступом.

Різноманітність апаратного забезпечення, програмних засобів і протоколів взаємодії в межах АСУТП створює додаткові можливості для виникнення нових каналів витоку інформації, втручання у її цілісність або доступність. Особливу загрозу становлять складні сценарії несанкціонованого доступу, які важко виявити та запобігти з огляду на багаторівневу взаємодію елементів системи.

За таких умов істотно зростає кількість потенційних векторів атак, що можуть призводити до витоку, модифікації або знищення критично важливої інформації. Складність контрольних заходів і зростання рівня технічної насиченості підвищують ризики порушення конфіденційності, цілісності та доступності інформації, яка обробляється в АСУТП.

Проблематика гарантування інформаційної безпеки в АСУТП вимагає системного підходу, що передбачає врахування загроз на всіх етапах життєвого циклу інформації – від початкового проєктування до експлуатації. Забезпечення ефективного захисту неможливе без формалізованих і обґрунтованих вимог до засобів захисту, які повинні бути чітко визначені вже на початкових стадіях створення ІКС.

У загальному вигляді процес забезпечення інформаційної безпеки в АСУТП може бути представлений у вигляді узагальненої структурної моделі (рис. 6), яка відображає основні етапи та взаємозв'язки у формуванні захисної політики системи

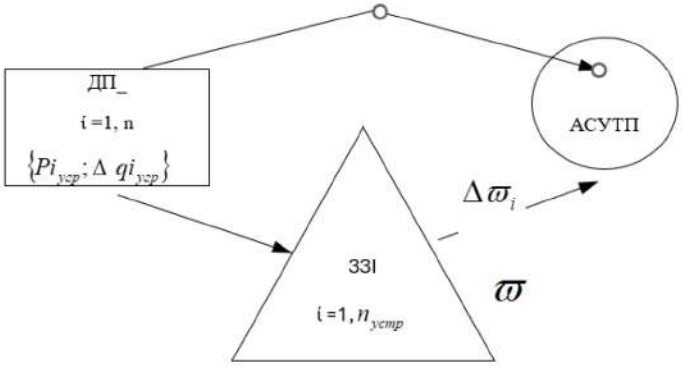


Рисунок 6 – Загальна модель процесу захисту інформації в АСУТП

Зловмисник за допомогою деякого джерела погроз (далі – ДП) генерує сукупність погроз АСУТП (нехай вона буде кінцевою і точною; $i = 1, \overline{n}$). Кожна *i*-та погроза характеризується імовірністю появи $Pi_{уєр}$ і збитком $\Delta q_i^{уєр}$; принесеним інформаційно-управляючій системі.

Засобами захисту інформації (далі – ЗЗІ) виконується функція часткової компенсації погроз для АСУТП. Основною характеристикою засобів захисту є імовірність усунення кожної *i*-ї погрози $\mathrm{P}i_{угр}^{устр}$.

За рахунок функціонування ЗЗІ забезпечується зменшення збитку, що наноситься АСУТП впливом погроз.

Позначимо загальний відвернений збиток АСУТП через $\bar{W}$, а відвернений збиток за рахунок ліквідації впливу *i*-ї погрози через $\bar{\omega}_i$.

Після введених позначень сформулюємо в загальному виді задачу синтезу засобів захисту інформації в АСУТП.

Необхідно вибрати варіант реалізації ЗЗІ, що забезпечує максимум відверненого збитку від впливу погроз при припустимих витратах на ЗЗІ.

Формальна постановка задачі має вид:

знайти $\mathrm{T}^0 = \omega q \max \bar{W}\left(\mathrm{T}\right)\mathrm{T}^0 \in \mathrm{T}^+$ при обмеженні $C\left(\mathrm{T}^0\right) \le C\partial on$.

Тут – деякий вектор, що характеризує варіант технічної реалізації ЗЗІ; T^+T^0 – припустиме й оптимальне значення вектора Т;

Сдоп. – припустимі витрати на ЗЗІ.

Для рішення задачі необхідно насамперед сформувати показник якості функціонування ЗЗІ – $\bar{W}$ (Т).

Очевидно, відвернений збиток у загальному виді визначається співвідношенням:

$$\bar{W} = F\left(Pi_{упр}; \Delta qi^{упр}; \mathrm{P}i_{угр}^{устр}; i = 1,\bar{n}\right). \qquad (1)$$

Відвернений збиток за рахунок ліквідації впливу *i*-ї погрози

$$\overline{\omega_i} = \mathrm{P}i_{угр} \cdot \Delta qi^{угр} \cdot \mathrm{P}i_{угр}^{устр};\ i = 1,\bar{n}\ . \qquad (2)$$

За умови незалежності погроз і адитивності їх наслідків одержуємо:

$$\bar{W} = \sum_{i=1}^{n} \mathrm{P}i_{угр} \cdot \Delta qi^{угр} \cdot \mathrm{P}i_{угр}^{устр}. \qquad (3)$$

Зупинимося більш докладно на співмножниках, що входять у формулу:

Імовірність появи *i*-ї погрози $\mathrm{P}i_{угр}$ визначається статистично і відповідає відносній частоті її появи:

$$\mathrm{P}i_{угр} = \frac{\lambda_i}{\sum_{i=1}^{\lambda}\lambda_i} = \overline{\lambda}_i, \qquad (4)$$

де λ_i – частота появи *i*-ї погрози.

Збиток, принесений *i*-ї погрозою Δq_i, може визначатися в абсолютних одиницях: економічних втратах, тимчасових витратах, обсязі «знищеної» чи «зіпсованої» інформації тощо.

Однак, практично це зробити дуже важко, особливо на ранніх етапах проєктування ЗЗІ. Тому доцільно замість абсолютного збитку використовувати відносний збиток, що по суті являє собою ступінь небезпеки *i*-ї погрози для АСУТП.

Ступінь небезпеки може бути визначена експертним шляхом у припущенні, що всі погрози для АСУТП складають повну групу повідомлень, тобто

$$0 \le \Delta qi \le 1; \sum_{i=1}^{n}\Delta qi = 1. \qquad (5)$$

Найбільш складним питанням є визначення імовірності усунення і-й погрози $\mathrm{P}i_{угр}^{устр}$ при проєктуванні ЗЗІ. Зробимо природне допущення, що ця імовірність визначається тим, наскільки повно враховані якісні і кількісні вимоги до ЗЗІ при їх проєктуванні, тобто

$$\mathrm{P}i_{угр}^{устр} = \gamma i\left(xi1,..,xi\gamma,..,xim\right), \qquad (6)$$

де $xi\gamma$ – ступінь виконання *i*-го вимоги до ЗЗІ для усунення *i*-ї погрози, $i = 1,\overline{n}; \gamma = 1,\overline{m}$.

Нехай перші «до» вимог будуть кількісними $\left(\gamma = 1,\overline{k}\right)$ інші «*m*-до» – якісними $\left(\gamma = \overline{k+1,m}\right)$.

Ступінь виконання γ-ї кількісної вимоги визначається його близькістю до необхідного (оптимального) значення. Для оцінки ступеня виконання γ-ї кількісної вимоги до ЗЗІ зручніше всього використовувати його нормоване значення $\overline{xi}\gamma(\gamma = 1,\overline{k}), 0 \le xi\gamma \prec 1$.

Для нормування зручно використовувати функцію виду:

$$\overline{xi}\gamma = \frac{xi\gamma - xi\gamma^{нт}}{xi\gamma^{нл} - xi\gamma^{нх}}, \qquad (7)$$

де $xi\gamma$ – поточне значення γ-ї вимоги;

$xi^{\gamma нл}, xi\gamma^{нх}$ – найкращі і найгірші значення.

З урахуванням формули (7) одержуємо наступні розрахункові співвідношення:

при $xi\gamma^{нл} = xi\gamma\max;\ xi\gamma^{нх} = xi\gamma\min$,

$$\overline{xi\gamma} = \frac{xi\gamma - xi\gamma\min}{xi\gamma\max - xi\gamma\min}, \tag{8}$$

при $xi\gamma^{нл} = xi\gamma\min;\ xi\gamma^{нх} = xi\gamma\max$,

$$\overline{xi\gamma} = \frac{xi\gamma\max - xi\gamma}{xi\gamma\max - xi\gamma\min},$$

при $xi\gamma^{нл} = xi\gamma\ \text{opt};\ xi\gamma_{нх} = xi\gamma\min;\ xi\gamma^{нх} = xi\gamma\max;$

$xi\gamma\min \le xi\gamma\ \text{opt} \le xi\gamma\max$.

$$\overline{xi\gamma} = \left\{ \begin{array}{l} 0 \text{ при } xi\gamma\langle xi\gamma\min; xi\gamma\rangle xi\gamma\max \\ 1 \text{ при } xi\gamma = xi\gamma\ \text{opt} \\ \dfrac{xi\gamma - xi\gamma\min}{xi\gamma\ \text{opt} - xi\gamma\min} \text{ при } xi\gamma\min \le xi\gamma \le xi\gamma\ \text{opt}(II) \\ \dfrac{xi\gamma\max - xi\gamma\min}{xi\gamma\max - xi\gamma opt} \textit{ при } xi\gamma\ \text{opt} \le xi\gamma \le xi\gamma\max \end{array} \right\}. \tag{9}$$

Ступінь виконання γ-ї якісної вимоги визначається функцією приналежності до найкращого значення $\gamma\left(xi\gamma\right)$.

Розклавши функцію (9) у ряд Макларена й обмеживши лише першими членами ряду, одержимо:

$$\mathrm{P}i_{угр}^{устр} = \mathrm{P}i_{угр}^{устр}\left(0\right) + \sum_{\gamma=1}^{m} \frac{\partial \mathrm{P}i_{угр}^{устр}}{\partial xi\gamma} \cdot xi\gamma, \tag{10}$$

де $\mathrm{P}i_{угр}^{устр}\left(0\right) = 0$ – імовірність усунення *i*-ї погрози при невиконанні вимоги ЗЗІ;

$\dfrac{\partial \mathrm{P}i_{угр}^{устр}}{\partial xi\gamma} = \alpha i\gamma$ – величина, що характеризує ступінь впливу γ-ї вимоги на імовірність усунення *i*-ї погрози (важливість виконання для усунення *i*-ї погрози). Очевидно, що $0 \le \alpha i\gamma \le 1; \sum_{\gamma=1}^{m} \alpha i\gamma = 1$ для виходу $i = \overline{1,n}$.

Після підстановки в (12) відповідних значень одержуємо:

$$\mathrm{P}i_{угр}^{устр} = \sum_{\gamma=1}^{k} \alpha i\gamma \cdot \overline{xi\gamma} + \sum_{\gamma=k=1}^{m} \alpha i\gamma \cdot y\left(xi\gamma\right). \qquad (11)$$

Найбільш важливими і складними етапами рішення оформленої задачі є підготовка вихідних даних і вибір методу рішення.

Що ж стосується вибору методу рішення, то задача наведена вище відноситься до класу багатокритеріальних багатовимірних метричних задач.

При досить невеликій кількості розглянутих варіантів і використанні адитивного критерію вибору згадана задача може бути вирішена шляхом прямого перебору варіантів.

9. Інтелектуальний аналіз даних у формуванні вимог до інформаційної безпеки ІКС

Застосування інтелектуального аналізу даних (ІАД) у сфері інформаційної безпеки (ІБ) дозволяє автоматизувати процес виявлення вразливостей, класифікації загроз і формулювання релевантних вимог до захисту. В умовах зростаючої складності інформаційно-комунікаційних систем (ІКС), ручна обробка експертної інформації стає недостатньо ефективною. Застосування методів ІАД – таких як кластеризація, класифікація, аналіз асоціацій, логічна індукція – дозволяє формалізувати знання, отримані в результаті аналізу інцидентів, тестування або аудиту систем.

Ключовим етапом застосування ІАД є побудова семантично структурованої моделі вимог. Джерелами даних можуть бути як внутрішні журнали подій і результати тестування системи, так і зовнішні нормативи – ISO/IEC 27001, NIST SP 800-53, GDPR тощо. На базі таких джерел будуються логічні правила, що описують пріоритетність реалізації конкретних заходів захисту відповідно до ризиків.

Процес трансформації експертних формулювань у конкретні вимоги до інформаційної безпеки є складним і багатоступеневим. Він передбачає ретельний аналіз зібраної експертної інформації, яка, як правило, надходить у вигляді тверджень, описів подій або сценаріїв загроз. На початковому етапі здійснюється систематизація таких

формулювань, що дозволяє структурувати інформацію відповідно до категорій потенційних загроз інформаційній та кібербезпеці. Зібрані дані проходять подальшу класифікацію та семантичний аналіз, що є критично важливим для побудови формалізованої онтології вимог. Це дозволяє впорядкувати експертні судження та встановити між ними логічні зв'язки, які необхідні для якісного формування вимог до захисту інформації на всіх рівнях функціонування інформаційно-комунікаційних систем.

Важливим етапом у цьому процесі є трансформація вихідних експертних формулювань у нормалізовані вимоги до інформаційної безпеки, що відповідають сучасним інженерним та нормативно-правовим стандартам. Така трансформація здійснюється шляхом уніфікації та формалізації текстових описів, які отримані від фахівців, із одночасною перевіркою їх відповідності міжнародним стандартам у сфері інформаційної безпеки, зокрема ISO/IEC 27001. Застосування цього підходу дозволяє підвищити ефективність процесу формування вимог, забезпечити їх прозорість і взаємозв'язок із категоріями контролю та методами реалізації захисних заходів. У межах розділу подальшу увагу зосереджено на деталізації результатів такої трансформації. Для ілюстрації обґрунтованості та коректності запропонованої методики наведено відповідну таблицю, що відображає взаємозв'язок між початковими експертними твердженнями, нормалізованими вимогами, їх класифікацією згідно з міжнародними стандартами та типом контролю, який застосовується в інформаційно-комунікаційних системах.

Подібна методологія дозволяє мінімізувати суб'єктивність при формуванні вимог до ІБ, забезпечуючи уніфікований підхід до захисту ІКС. Інструментальне забезпечення реалізації ІАД у цьому контексті може базуватися на поєднанні засобів бізнес-аналітики (наприклад, Power BI), мов онтологічного моделювання (*Protégé*) та фреймворків машинного навчання (*Scikit-learn, TensorFlow*).

Особливу увагу в сучасних інформаційно-комунікаційних системах слід приділяти змінності їхнього середовища. У ході експлуатації такі системи постійно зазнають змін: змінюються конфігурації, з'являються нові елементи, відбувається оновлення програмного забезпечення, підключаються додаткові вузли. Саме тому

актуальність і ефективність розроблених вимог до інформаційної безпеки не може залишатися незмінною та потребує постійного переоцінювання. Щоб забезпечити якісний рівень захисту, необхідно враховувати не лише формальні критерії відповідності стандартам, а й аналіз ризиків у реальному часі. Вимоги мають перевірятися на предмет їхньої здатності реагувати на нові загрози та зміни інфраструктури. Наприклад, корисним може бути розгляд ступеня охоплення ризиків, що ним забезпечуються актуальні вимоги, а також здатність цих вимог залишатися релевантними у разі зміни архітектури або технологічних процесів. Крім того, важливо аналізувати, наскільки часто порушення пов'язані саме з тією чи іншою вимогою, а також чи відповідають ці вимоги сучасним міжнародним стандартам у галузі інформаційної безпеки.

Інтелектуальний аналіз даних, впроваджений у процес формування вимог до захисту інформації, дозволяє створювати моделі управління, які є гнучкими та здатними адаптуватися до контексту конкретної організації або системи. Такі моделі враховують не лише технічні параметри, а й організаційні особливості, що дає змогу ефективніше управляти ризиками та реалізовувати заходи захисту, які відповідають реальним потребам об'єкта захисту. Використання сучасних інструментів аналітики, онтологічного моделювання та машинного навчання відкриває нові можливості для формалізації експертних суджень та побудови на їхній основі чітко структурованих і обґрунтованих вимог до інформаційної безпеки.

Особливої ваги ці підходи набувають для об'єктів критичної інфраструктури, оскільки будь-яка неконкретизована, нечітко сформульована або недостатньо перевірена вимога може стати джерелом значної загрози для стабільної роботи цілих систем.

Інтеграція інтелектуального аналізу даних у процеси проєктування та моделювання вимог дозволяє не лише підвищити їхню відповідність стандартам, а й забезпечити можливість подальшої перевірки та актуалізації в міру змін у середовищі кібербезпеки. У результаті формується надійна і прозора система захисту, де кожна вимога має чіткий зв'язок із категоріями контролю та методами впровадження, а також сприяє уніфікованому підходу до управління ризиками. Як ілюстрацію до викладених положень у подальшому

подається таблиця, що демонструє процес перетворення експертних суджень у формалізовані вимоги, їхню класифікацію згідно міжнародних стандартів та визначення типу контролю, що застосовується в інформаційно-комунікаційних системах.

Таблиця 8 – Трансформація експертних формулювань у вимоги до ІБ для ІКС

Експертне твердження	**Нормалізована вимога**	**Категорія безпеки (ISO/IEC 27001)**	**Тип контролю**
Система має реєструвати всі підключення до мережі в режимі реального часу	Аудит мережевої активності з автоматичним логуванням подій	A.12.4 – Logging and monitoring	Технічний контроль
Адміністратор має отримувати повідомлення про підозрілі зміни в конфігурації	Моніторинг конфігурацій з генерацією повідомлень про інциденти	A.16.1 – Management of information security incidents and improvements	Оперативний контроль
Доступ до критичних функцій повинен здійснюватися лише з біометричною перевіркою	Реалізація багатофакторної автентифікації з біометричним компонентом	A.9.4 – System and application access control	Фізико-логічний контроль

Після побудови ієрархічної структури й визначення вагових коефіцієнтів для кожного елемента моделі проводиться підсумковий аналіз ризиків. Отримані результати дозволяють визначити, які цілі захисту, загрози чи заходи мають найвищий пріоритет із погляду експертної оцінки та поточної ситуації в інформаційно-комунікаційній системі.

На основі цих даних формується план впровадження засобів захисту, що максимально відповідає специфіці ІКС і забезпечує ефективне використання ресурсів. Важливо, що модель легко піддається адаптації: за зміни загроз чи появи нових технологічних чинників можна оперативно скоригувати ваги елементів та пріоритети без необхідності повної перебудови всієї системи.

Досвід застосування таких експертних моделей показує, що поєднання формалізованих підходів і глибокої експертизи галузевих фахівців значно підвищує рівень надійності та стійкості ІКС. Це дозволяє не лише реагувати на нові виклики, а й проактивно формувати стратегію розвитку системи захисту відповідно до динаміки сучасних кіберзагроз.

10. Експертні моделі прийняття рішень у проєктуванні засобів захисту ІКС

Забезпечення стійкості та надійності інформаційно-комунікаційних систем (ІКС) стає дедалі складнішим завданням, оскільки технологічний розвиток супроводжується зростанням кількості загроз і ускладненням характеру атак. В таких умовах особливої уваги заслуговують експертні моделі прийняття рішень, які дозволяють інтегрувати різноманітні джерела знань, накопичений досвід фахівців і специфіку конкретних технологій у єдину систему формалізованих рішень. Головна мета впровадження експертних підходів – забезпечення системної оцінки ризиків, визначення пріоритетних напрямків модернізації засобів захисту та створення стійкої архітектури безпеки, що відповідає як сучасним стандартам, так і вимогам регуляторів.

Експертні моделі прийняття рішень у сфері безпеки ІКС формуються на основі багатокритеріального аналізу, методів нечіткої логіки, застосування байєсівських мереж, дерев рішень, а також методів аналітичної ієрархічної мережі (ANP) та аналізу ієрархій (AHP). Особливість цих моделей полягає у поєднанні якісних і кількісних показників, що дає змогу не лише оцінювати поточний стан захищеності, а й прогнозувати динаміку змін у разі впровадження тих чи інших заходів. Завдяки цьому стає можливим врахування взаємозв'язків між окремими компонентами ІКС, вразливостями, шляхами впливу на систему та потенційними сценаріями розвитку інцидентів.

Значна увага в експертному підході приділяється створенню ієрархічних моделей ризиків. На першому рівні визначаються та формалізуються цілі захисту, наприклад, збереження конфіденційності,

цілісності та доступності інформації. Другий рівень містить перелік основних загроз, що можуть впливати на досягнення кожної з цілей – від несанкціонованого доступу до даних до внутрішніх порушень або помилок персоналу. Третій рівень передбачає деталізацію заходів захисту, які можуть бути застосовані для нейтралізації або зниження впливу зазначених загроз: це може стосуватися організаційних, технічних, фізичних чи комплексних рішень.

Використання методу аналізу ієрархій (AHP) дозволяє експертам структурувати складні завдання, призначати ваги окремим цілям, загрозам і заходам, а також визначати оптимальні сценарії розподілу ресурсів для підвищення загального рівня безпеки. Для цього застосовується система експертного опитування, результатом якої стає побудова матриці вагових коефіцієнтів, що відображає відносну важливість кожного елементу моделі.

Крім того, експертні моделі дають змогу враховувати специфічні умови функціонування певної ІКС, регіональні чи галузеві вимоги, а також динаміку змін у законодавстві, рекомендаціях з кібербезпеки й тенденціях розвитку Т-інфраструктури. Це особливо важливо для критичних інфраструктур, де навіть незначна помилка у визначенні пріоритетів може призвести до суттєвих матеріальних чи репутаційних втрат.

Наводимо приклад структури експертного моделювання ризиків для ІКС з використанням аналітичного методу ієрархій (AHP), що ілюструє послідовність побудови моделей і призначення ваг для кожного елемента:

Таблиця 9 – Пріоритети захисту інформації в ІКС за методом AHP

Рівень	**Елемент моделі**	**Опис**	**Приклад ваги (AHP)**
1	Ціль: Конфіденційність	Захист персональних та службових даних	0,40
2	Загроза: Несанкціонований доступ	Спроби обходу механізмів автентифікації	0,30
3	Засіб: Двофакторна автентифікація	Доступ лише після підтвердження SMS/Token	0,60

Таким чином, експертні моделі у поєднанні з інструментами інтелектуального аналізу даних забезпечують адаптивне й гнучке проєктування вимог до захисту ІКС. Вони дозволяють враховувати контекст, інфраструктурні особливості, динаміку загроз та стратегічні пріоритети організацій, що особливо актуально в умовах воєнного та поствоєнного відновлення цифрових інфраструктур.

Залучення експертної думки на всіх етапах проєктування засобів захисту інформації дає змогу створювати системи, що швидко адаптуються до нових викликів і змін у загрозах. Важливо підкреслити, що синтез експертної інформації не лише сприяє підвищенню точності оцінювання ризиків, а й дозволяє уникнути типових помилок, які виникають унаслідок вузької автоматизації чи відсутності глибокого розуміння галузевої специфіки. Поєднання таких моделей із алгоритмами інтелектуального аналізу, зокрема data mining, кластеризації та виявлення шаблонів, дає змогу не тільки ідентифікувати аномальні дії чи небезпечні сценарії, але й завчасно прогнозувати можливі вектори атак, проводити нормування експертних оцінок і виявляти приховані залежності між елементами системи.

Особливе значення ці підходи набувають у сферах критичної інфраструктури й у державно-значущих сервісах, де питання безпеки даних стає не просто пріоритетом, а основною умовою стійкої цифровізації. У такому контексті експертні, інтелектуальні та лінгвістичні методи забезпечують основу для створення комплексних стратегій захисту інформації, що здатні враховувати не лише поточний стан системи, але й її перспективний розвиток, еволюцію загроз і зміну нормативно-правового поля. Це відкриває можливості для формування ефективних, самонавчальних систем оцінки ризиків, адаптивних до змін і здатних гарантувати сталу кібербезпеку в умовах невизначеності.

11. Висновки

Забезпечення захисту даних користувачів в інформаційно-комунікаційних системах потребує ґрунтовних знань, застосування ефективних інструментів та всебічного аналізу. Для формування вимог до захисту інформації в ІКС необхідно враховувати специфіку сучасного

цифрового середовища, особливості автоматизованих систем управління й актуальні виклики для підтримки цілісності та безпеки даних. Комплексний підхід дозволяє глибше зрозуміти еволюцію методів кібербезпеки в умовах зростання кількості і складності інформаційних загроз, а також динамічного розвитку технологій.

Було здійснено ґрунтовний аналіз ключових засад ідентифікації пристроїв і користувачів у контексті функціонування ІКС. Детально розглянуті сучасні технології ідентифікації, які охоплюють як апаратний рівень (ідентифікація пристроїв за унікальними характеристиками, застосування апаратних токенів), так і поведінковий рівень (аналіз користувацьких патернів, реакцій на події, часових і просторових особливостей дій). Підкреслено важливість впровадження мультифакторних схем автентифікації, що дозволяють значно підвищити рівень захищеності систем від зовнішніх і внутрішніх загроз.

Окрема увага приділено поняттю цифрового сліду користувача, що охоплює технічні параметри, поведінкові чинники, контекстуальні дані та історію взаємодії з системою. Встановлено, що цілеспрямована мінімізація цифрового сліду є обов'язковою умовою для забезпечення конфіденційності та приватності. Для цього доцільно впроваджувати інноваційні інструменти, такі як cookie isolation, технології обфускації дій, DNS-over-HTTPS, а також орієнтуватися на відповідність міжнародним стандартам і нормативам (зокрема, GDPR, ISO/IEC 29184).

Узагальнення отриманих результатів дозволяє зробити висновок, що інтеграція сучасних методів ідентифікації, автентифікації й мінімізації цифрового сліду користувачів забезпечує підвищення стійкості інформаційних систем до новітніх векторів атак і сприяє формуванню ефективної політики захисту даних у контексті динамічного розвитку цифрової інфраструктури.

Практичні аспекти впровадження стратегій захисту в ІКС. Аналіз механізмів виявлення загроз (NAP, NAC, ZTA) показав їх ефективність у моделюванні зон довіри та контролю доступу. Було розглянуто концепцію реактивного та проактивного реагування на загрози – через моніторинг, фіксацію, класифікацію, обфускацію та зворотну реакцію. Важливою особливістю сучасного контексту стало

врахування викликів воєнного та поствоєнного періодів, зокрема у розрізі е-сервісів (Дія, E-Helsy, IDP-реєстрація), що потребують підвищеної уваги до верифікації та автентифікації.

Детально проаналізовано застосування лінгвістичних та інтелектуальних методів для синтезу вимог до ІБ, запропонновано формалізовану модель впливу експертної інформації на структуру вимог а також продемонстровано, як методи інтелектуального аналізу (*data mining, clustering, pattern discovery*) можуть забезпечити обґрунтоване нормування експертних оцінок. Також розглянуто використання методів прийняття рішень (AHP, ANP) для ієрархічного узгодження технічних, організаційних і нормативних вимог до безпеки в ІКС.

Таким чином, отримані результати свідчать про доцільність інтеграції експертних підходів, інтелектуального аналізу даних і лінгвістичного моделювання для обґрунтування вимог до засобів захисту в інформаційно-комунікаційних системах. Перспективним напрямом подальших досліджень є розробка адаптивних моделей оцінки ризиків у динамічному середовищі, що поєднують онтологічні схеми знань, верифікацію поведінкових аномалій та кібергігієну користувачів.

Список використаних джерел

1. IBM. Cost of a Data Breach Report 2024. IBM Security, 2024. Електронний ресурс. URL: https://www.ibm.com/reports/data-breach (дата звернення: 01.07.2025).

2. Morgan S. Cybercrime To Cost The World $10.5 Trillion Annually By 2025. *Cybersecurity Ventures.* 2022. Електронний ресурс. URL: https://cybersecurityventures.com/cybercrime-damage-costs-10-trillion-by-2025 (дата звернення: 01.07.2025).

3. Державна служба спеціального зв'язку та захисту інформації України. Статистика кібератак 2024. Офіційний вебсайт, 2025. Електронний ресурс. URL: https://cip.gov.ua/ua/news/statystyka-kiberatak-za-2024-rik (дата звернення: 02.07.2025). *Cisco*. (2024). Annual Cybersecurity Report. URL: https://www.cisco.com/c/en/us/products/security/security-reports.html

4. ENISA. (2023). Threat Landscape for Industrial Control Systems. URL: https://www.enisa.europa.eu/publications/ics-threat-landscape

5. NIST. (2015). Guide to Industrial Control Systems (ICS) Security. NIST SP 800-82 Rev.2. https://doi.org/10.6028/NIST.SP.800-82r2

6. European Union. (2021). Regulation (EU) No 910/2014 (eIDAS Regulation). URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014R0910

7. ISO/IEC 29184:2020. Online privacy notices and consent. International Organization for Standardization. URL: https://www.iso.org/standard/70331.html

8. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection. URL: https://www.iso.org/standard/27001

9. GDPR. (2016). General Data Protection Regulation. Regulation (EU) 2016/679. URL: https://gdpr-info.eu

10. Chepurna O., Cherevko Y., Kuleshova Y. On Information Geometry Methods for Data Analysis. Geometry. *Integrability and Quantization.* 2024. Vol. 29. P. 11–22. DOI: https://doi.org/10.7546/giq-29-2024-11-22

11. Shor S., Pavlova T. Enhancing Access Control in Critical Infrastructure Systems Using Attribute-Based Encryption. *Information and Security.* 2023. Vol. 55 (2). P. 145–162. DOI: https://doi.org/10.11610/isij.5525

12. Tang M., Li Q., Yang J. Anomaly Detection in ICS Using Isolation Forest. *IEEE Access.* 2022. Vol. 10. P. 12547–12558. DOI: https://doi.org/10.1109/ACCESS.2022.3147280

13. He J., Wang X., Liu H. Privacy-Preserving Authentication Protocol for eHealth Systems. *Future Generation Computer Systems.* 2021. Vol. 117. P. 354–367. DOI: https://doi.org/10.1016/j.future.2020.11.016

14. BankID National Bank of Ukraine. 2024. URL: https://bankid.org.ua/

15. Ministry of Digital Transformation of Ukraine. Diia – Digital State Project. 2024. URL: https://diia.gov.ua

16. Cisco. 2024. *Cisco Annual Cybersecurity Report.* URL: https://www.cisco.com/c/en/us/products/security/security-reports.html

17. Fortinet. 2024. *OT Threat Landscape Report.* URL: https://www.fortinet.com/content/dam/fortinet/assets/threat-reports/ot-threat-report-2024.pdf

18. IBM Security. 2023. *Cost of a Data Breach Report.* 2023. URL: https://www.ibm.com/reports/data-breach

19. ENISA. 2024. *Threat Landscape for Industrial Control Systems.* URL: https://www.enisa.europa.eu/publications/ics-threat-landscape

20. MITRE. 2024. ATT&CK for ICS Framework. https://attack.mitre.org/matrices/ics

21. ISO/IEC 29184:2020. Online Privacy Notice and Consent. URL: https://www.iso.org/standard/70331.html

22. Google. 2023. *Privacy Sandbox Project.* URL: https://privacysandbox.com

23. Microsoft. 2023. *Digital Defense Report.* URL: https://www.microsoft.com/en-us/security/business/security-intelligence-report

24. NIST. 2022. *Framework for Improving Critical Infrastructure Cybersecurity* (v1.1). DOI: https://doi.org/10.6028/NIST.CSWP.04162018

25. OWASP. 2023. Top 10 for LLM Applications. URL: https://owasp.org/www-project-top-10-for-llm-applications

26. European Commission. 2022. *eIDAS 2.0 Regulation.* URL: https://digital-strategy.ec.europa.eu/en/policies/eidas

27. Ukrainian Ministry of Digital Transformation. 2023. *Diia Platform Overview*. URL: https://thedigital.gov.ua

28. Cloudflare. 2023. *The State of Application Security*. URL: https://www.cloudflare.com/learning/security

29. Palo Alto Networks. 2024. *State of Cybersecurity Report.* URL: https://www.paloaltonetworks.com/resources/research

30. CrowdStrike. 2024. *Global Threat Report.* URL: https://www.crowdstrike.com/global-threat-report/

Разінкін Нікіта Сергійович
асистент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ОРГАНІЗАЦІЯ ЗАХИСТУ ІНФОРМАЦІЇ В ПРОЦЕСІ ФУНКЦІОНУВАННЯ ЕЛЕКТРОМОБІЛЬНОЇ ІНФРАСТРУКТУРИ В УКРАЇНІ

DOI https://doi.org/10.36059/978-966-397-537-5-11

У сучасних умовах стрімкого розвитку електромобільної інфраструктури в Україні, зумовленого глобальними тенденціями декарбонізації транспорту, сталого розвитку та зростання популярності електротранспорту, питання захисту інформації набуває критичного значення. Електромобільна інфраструктура, що охоплює зарядні станції, системи керування зарядкою, мобільні додатки та енергетичні мережі, є складною кіберфізичною системою, яка потребує надійного захисту від кіберзагроз. В Україні, де кількість електромобілів стрімко зростає, а зарядна мережа розвивається нерівномірно, виникають унікальні виклики, пов'язані з обмеженою інфраструктурою, низьким рівнем цифрової грамотності та нестабільністю енергосистеми. Недостатня стандартизація безпеки, вразливості в протоколах зв'язку, а також слабкий захист мобільних додатків створюють ризики несанкціонованого доступу, фінансових втрат і навіть дестабілізації енергомережі. У цьому контексті впровадження комплексних заходів захисту інформації, гармонізованих із міжнародними стандартами є необхідним для забезпечення стабільності, безпеки та сталого розвитку електромобільної інфраструктури.

Питання організації захисту інформації в процесі функціонування електромобільної інфраструктури в Україні привертало увагу багатьох науковців, серед яких Алімурадов Ш. Р., Хуан І., Аянович А., Хаас Р., Чжан, Півняк Г., Козак Ю., Семенов В., Джонсон Дж. та інші. Проте, незважаючи на значний науковий доробок, проблема забезпечення ефективного кіберзахисту електромобільної інфраструктури залишається не до кінця вирішеною. Це зумовлено стрімкою еволюцією кіберзагроз, які стають дедалі складнішими та різноманітнішими, зокрема через використання вразливостей у протоколах передачі даних, мобільних додатках

і кіберфізичних системах. Наприклад, атаки типу «людина посередині», “juice jacking” чи масові зарядні атаки можуть дестабілізувати енергомережу та спричинити економічні збитки. Така динаміка загроз вимагає комплексного підходу до вдосконалення технічних, організаційних і регуляторних заходів захисту, особливо з урахуванням специфіки української інфраструктури, яка характеризується нерівномірним розподілом зарядних станцій і обмеженою цифровою грамотністю населення. Отже, дослідження питань кіберзахисту електромобільної інфраструктури в Україні зберігають високу актуальність, спрямовану на забезпечення її стійкості та інтеграцію з міжнародними стандартами безпеки.

Мета роботи полягає у розробці структурованого плану впровадження змін до організації захисту інформації в процесі функціонування електромобільної інфраструктури в Україні, спрямованого на підвищення кібербезпеки, гармонізацію з міжнародними стандартами безпеки та адаптацію до локальних умов, враховуючи технічні, організаційні та регуляторні аспекти, для забезпечення стабільності енергетичної системи, захисту даних користувачів і запобігання кіберзагрозам.

Для досягнення поставленої мети було виконано наступні завдання:

– визначити основні кіберзагрози, що впливають на електромобільну інфраструктуру в Україні з урахуванням локальних особливостей енергосистеми;

– проаналізувати міжнародні стандарти безпеки та розробити пропозиції щодо їх гармонізації з національною нормативно-правовою базою для створення єдиних стандартів захисту електромобільної інфраструктури в Україні;

– оцінити організаційні заходи з урахуванням низького рівня цифрової обізнаності в Україні;

– запропонувати технічні рішення для створення систем моніторингу в реальному часі та децентралізованих CMS, які забезпечать виявлення аномальної активності, захист від масових зарядних атак і підвищення стійкості інфраструктури до збоїв;

– дослідити можливості міжнародної співпраці для отримання фінансування, технологій і експертизи.

Об’єктом дослідження виступає електромобільна інфраструктура в Україні, включаючи зарядні станції, системи керування зарядкою, мобільні додатки та протоколи передачі даних.

Предмет дослідження – організація захисту інформації в процесі функціонування електромобільної інфраструктури, спрямована на підвищення кібербезпеки, гармонізацію з міжнародними стандартами та адаптацію до локальних умов.

У контексті сучасних тенденцій сталого розвитку, глобальної декарбонізації транспорту та інтенсивного впровадження електротранспорту, електромобільна інфраструктура виступає як фундаментальна передумова ефективного функціонування електромобілів у міському та міжміському середовищі. Незважаючи на широке використання терміну, у науковій літературі відсутнє єдине універсальне визначення електромобільної інфраструктури, що зумовлено складністю, міждисциплінарністю та динамічністю цієї сфери.

Згідно з дослідженням Алімурадова Ш. Р. [1], електромобільна інфраструктура – це сукупність технічних, організаційних та інформаційних засобів, які забезпечують можливість заряджання, обслуговування та експлуатації електромобілів у певному просторі. Це визначення підкреслює багатокомпонентність системи, де інфраструктура не обмежується лише зарядними станціями, а включає енергетичну інфраструктуру, логістичну підтримку, цифрову платформу управління та навіть регуляторні елементи.

Дослідники Хуан І. та ін. розширюють це поняття, розглядаючи ЕМІ як інтелектуальну та інтегровану систему, що пов'язує електротранспорт, енергетичні мережі та інформаційні технології [2]. Вони пропонують розглядати електромобільну інфраструктуру як кібер-фізичну систему, що дозволяє оптимізувати навантаження на енергомережу, забезпечити гнучке заряджання та реалізувати концепцію Vehicle-to-Grid. У цьому контексті важливо зазначити, що ЕМІ – це не лише фізична інфраструктура, а й логіка взаємодії між користувачами, транспортними засобами та енергосистемою.

З методологічної точки зору, науковці пропонують класифікацію електромобільної інфраструктури за функціональними ознаками. Так, дослідження Аяновича А. та Хааса Р. поділяє інфраструктуру на:

– зарядну (charging infrastructure);
– сервісну (maintenance infrastructure);
– комунікаційну (digital infrastructure);
– регуляторну (policy and regulatory infrastructure) [3].

Це дає змогу комплексно аналізувати стан і розвиток електромобільної інфраструктури у різних країнах, зокрема з урахуванням інституційного та технологічного контексту.

Окремої уваги заслуговує питання просторової доступності ЕМІ, яке, як зазначає Чжан та ін., безпосередньо впливає на швидкість прийняття електромобілів населенням [4]. Автори стверджують, що розміщення зарядних станцій має базуватися на моделюванні мобільності, інтенсивності трафіку та енергетичних обмежень. Це визначає необхідність синергетичного підходу між урбаністикою, енергетикою та транспортним плануванням.

Жаровська І., Козак Ю., Семенов В. наголошують, що в умовах України ЕМІ набуває особливого значення через високий рівень імпорту електромобілів, нерівномірний розподіл зарядної мережі та залежність від нестабільної енергосистеми [5]. Зокрема, вказується на важливість включення до ЕМІ альтернативних джерел енергії (сонячні панелі, акумуляторні буфери), що дозволяє створити енергетично незалежні зарядні хаби.

Функціонування електромобільної інфраструктури в Україні має багатовимірний характер, що включає технічні, економічні й нормативно-правові аспекти. Перш за все, стрімкий ріст кількості електромобілів у країні зумовлює значні зміни у транспортній екосистемі та потребує формування нової інфраструктури обслуговування. З 2010 до 2022 року Україна демонструє стійку тенденцію нарощування ринку електромобілів – з первинних імпортних партій у кількості близько 3,2 тисячі одиниць за 2010–2016 роки до більш ніж 25 850 автомобілів станом на 1 січня 2021 року. Основну долю становлять Nissan Leaf, Tesla Model S, Renault Kangoo та BMW i3, що складає понад 60 % у загальній структурі продажів [6].

Важливим фактором є територіальна нерівномірність розвитку: найбільше електромобілів зареєстровано в Києві й Київській області, Одеській, Харківській, Дніпропетровській і Львівській областях, як зображено на рис. 1.

Така концентрація пояснюється вищим рівнем доходів населення, наявністю зарядної інфраструктури та транспортною політикою регіонів. Функціонування інфраструктури не відбувається автоматично, а супроводжується комплексною взаємодією між виробниками

транспортних засобів, їхніми дилерами, сервісними центрами, операторами зарядок, енергокомпаніями, органами влади, банківськими і фінансовими установами [7].

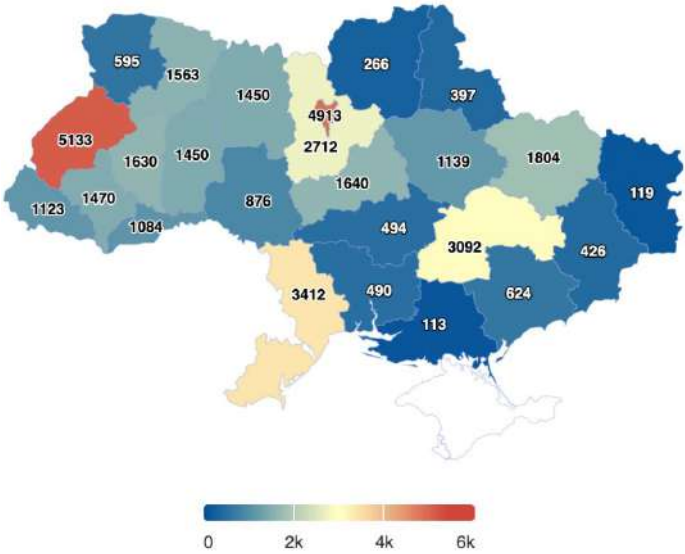


Рисунок 1 – Структура реєстрації електромобілів за регіонами на 2023 р.

Суттєвим викликом є забезпечення відповідної потужності електромереж, адже при зростанні частки електромобілів до 5–10 % від усієї парку можуть виникнути значні перевантаження на енергосистему. Це потребує реконструкції і модернізації мереж, включно з розподіленими джерелами, адаптивним навантаженням та системами розумного управління.

З розвитком ринку несуться витрати й на зарядну інфраструктуру: звичайна стаціонарна зарядка на 220 В із часом підзарядження 6–8 годин коштує 2–3 тисячі доларів, а швидкісна зарядка 400–480 В – близько 12 тисяч доларів. Станом на 2019 рік в Україні нараховувалося 2719 станцій і 5902 точки зарядки, більшість з яких підтримують стандарти Type 1, Type 2, CHAdeMO, а також CCS Combo 2). Співвідношення зареєстрованих електромобілів до точок зарядки становило

близько 4, що відповідає показникам таких країн, як Нідерланди, та перевершує Польщу [8].

Нормативно-правова база для розвитку електромобільності в Україні формувалася поступово. Закони, зокрема «Про електромобільний транспорт» 2017 року, передбачають зміни в реєстраційних правилах, встановлення технічних стандартів і безпеки, а також розвиток зарядної інфраструктури. Особливо суттєвими є податкові пільги: електромобілі звільнені від акцизного і митного зборів, ПДВ; упроваджено ставки, орієнтовані на екологічність, включно з транспортним податком і податком на нерухомість.

У 2024 році державні програми відновлення після війни були спрямовані не лише на модернізацію інфраструктури, а й на стандартизацію декарбонізації. ЄС висунув умову поєднання допомоги з дотриманням кліматичних норм; Україна впровадила програми на стимулювання електротранспорту й мереж зарядок, інтегрованих у містобудівні проєкти.

Єдиною з проблем залишається низька кількість сервісних центрів, що мають компетенцію обслуговування великих батарей, а також нестача швидких зарядних станцій уздовж автомагістралей і на АЗС. Часті технічні обмеження сумісності роз'ємів, недостатня обізнаність користувачів про роботу мобільних додатків операторів – ці чинники стримують використання електромобілів, особливо за межами великих міст [9].

Важливим пунктом є розвиток ланцюгів виробництва батарей і їх утилізації. В Україні є поклади літію (Полохівське, Станково, Шевченківське, Беганське місцезнаходження), але промислова розробка й виробництво літій-іонних батарей ще не розгорнуті системно. Натомість зростає потреба у переробці вичерпаних акумуляторів: лише близько 5% батарей у світі переробляється повторно [6]. Розвиток «другого життя» батарей для систем накопичення енергії та резервних джерел є важливим напрямом сталого розвитку. Посилення цифрової складової інфраструктури відкриває новий пласт викликів: забезпечення кібербезпеки, стандарти автентифікації, інтелектуальний моніторинг та управління зарядним процесом, сумісність обладнання різних виробників, адаптація до умов збоїв покриття мереж – усе це предмет досліджень українських інформаційних технологічних систем [7].

З глобальної перспективи, розвиток електромобільної інфраструктури в Україні тягне за собою демократизацію доступу до сталого транспорту, підвищення енергонезалежності, зниження викидів CO_2. При цьому ключовими бар'єрами залишаються відсутність чіткої реєстрації станцій, нерівномірний територіальний розвиток, технічні й фінансові бар'єри, стримування попиту. Науковці підкреслюють необхідність комплексного підходу: стимулювання попиту – через податкові пільги й муніципальні програми, нарощування пропозиції – через розвиток сервісів і зарядних мереж, модернізацію енергетичних мереж і створення нормативної бази, що включає стандарти безпеки й екологічну відповідальність.

Ураховуючи зростаючу роль електромобільного транспорту в енергетичній та транспортній системі держави, виникає об'єктивна необхідність у формуванні ефективної нормативно-правової бази, яка б регламентувала порядок їх проєктування, впровадження та експлуатації.

Нормативно-правова база захисту інформації електромобільної інфраструктури в Україні формується на базі чинного законодавства, яке визначає правові рамки відповідальності, технічних вимог та процедур контролю. Закон України «Про захист інформації в інформаційно-телекомунікаційних системах» (№ 80/94-ВР) є основоположним документом, що встановлює поняття інформаційно-телекомунікаційної системи, інформує про об'єкти та суб'єктів захисту, порядок доступу, умови обробки інформації і систему відповідальності за порушення норм [10; 11]. Особливо важливо для EV-інфраструктури те, що власник системи, яким може бути оператор зарядної мережі, має обов'язок створювати службу захисту інформації або призначати відповідальних осіб і повідомляти спеціально уповноважений орган (Держспецзв'язок, CERT-UA) у разі інцидентів [10].

У табл. 1 сформовано акти, закони та постанови, які регулють діяльність та безпеку електромобільної інфраструктури. Закон «Про захист персональних даних» встановлює обов'язок збору згоди на обробку і захист персональних даних користувачів EV-інфраструктури, що стосується реєстраційних даних, платіжної інформації, історії поїздок; недотримання цих норм тягне адміністративну або кримінальну відповідальність [12]. Закон «Про критичну

інфраструктуру» поширює вимоги безпеки на об'єкти, що можуть становити загрози національній безпеці, до яких належать зарядні станції в разі критичної масовості або залучення до енергомережі високого навантаження [13].

Таблиця 1 – Нормативно-правова база захисту інформації електромобільної інфраструктури в Україні

№	Нормативний акт	Основні положення	Застосування до EV-інфраструктури
1	Закон України «Про захист інформації в інформаційно-телекомунікаційних системах» № 80/94-ВР	Визначає порядок захисту інформації в ІТС, обов'язки власників, атестацію, вимоги до служб захисту	Захист інформації в зарядних системах, створення служб/відповідальних осіб
2	Закон України «Про захист персональних даних» № 2297-VI	Регламентує обробку, зберігання та захист персональних даних, згоду користувачів	Обробка даних користувачів EV-сервісів (геолокація, платежі, акаунти)
3	Закон України «Про критичну інфраструктуру» № 1647-IX	Встановлює вимоги до об'єктів критичної інфраструктури, процедури категоризації	EVSE як потенційно критичні об'єкти, потреба в оцінці ризиків
4	Указ Президента України № 695/2023	Визначає порядок реагування на загрози, взаємодію з CERT-UA, введення багаторівневого захисту	Обов'язковий моніторинг інцидентів, кіберзахист EV-мереж
5	Кодекс України про адміністративні правопорушення	Містить відповідальність за порушення вимог захисту інформації	Адміністративна відповідальність за неналежний захист інформації

Власникам EV-комплексів доведеться здійснювати категоризацію таких об'єктів, проводити оцінку ризиків, інженерний захист і забезпечувати дотримання вимог, погоджених з НАНЦ, Укренерго, Держспецзв'язком.

Указ Президента України № 695/2023 набрав чинності 17 жовтня 2023 року. Він вводить в дію рішення РНБО про захист об'єктів критичної інфраструктури та енергетики в умовах воєнних дій [14]. В цьому документі акцентовано необхідність запровадження багаторівневого захисту, посилення інженерних і кіберзасобів безпеки, взаємодії операторів інформаційно-комунікаційних систем з Центром протидії кіберзагрозам та CERT-UA, регулярного моніторингу і реагування на інциденти. EV-інфраструктура, що підпадає під цю категорію, має виконувати вимоги указу, який зобов'язує включати системи моніторингу, аудит і посилений контроль доступу до функцій керування EVSE.

EV-інфраструктура в Україні потребує дотримання щонайменше п'яти нормативних документів, що вимагають визначення об'єктів інформбезпеки, створення служб захисту, фізичного і технічного захисту, захисту персональних даних, участі у категоризації критичних об'єктів, а також інтеграції до державних структур реагування. Це дозволяє системі функціонувати в межах правового поля з чіткою відповідальністю, контролем та можливістю оперативного реагування.

Однак більшість нормативних актів не враховують специфіки функціонування зарядних станцій як об'єктів, що поєднують IT-системи, енергетичні мережі та фінансові інтерфейси. Також не врегульовано питання сертифікації ПЗ, яке використовується в EVSE, механізмів моніторингу вторгнень у реальному часі, захисту від атак на протоколи зарядки (наприклад, OCPP, ISO 15118) та прозорого розподілу відповідальності між виробниками, операторами та постачальниками цифрових рішень.

Як було зазначено, у сфері електромобільної інфраструктури зростає кількість кіберзагроз, що виникають через взаємодію між фізичними зарядними станціями, комунікаційними мережами, хмарними компонентами, мобільними додатками та основною енергосистемою. Комунікаційні канали за стандартами OCPP та ISO 15118 часто використовуються без увімкнення захисних функцій, таких як сесійне шифрування та автентифікація, що створює можливість атаки «людина-посередині», replay-атак, підробки журналів зарядок та маніпуляції з платіжними процесами [15; 16].

У статті Джонсона Дж. представлений огляд публічно зафіксованих вразливостей EVSE, що охоплюють не лише самі зарядні станції, але й комунікації зі SCADA і хмарними сервісами [15]. Типові шляхи проникнення – використання вразливих API, нешифрованих HTTP або SSH-інтерфейсів та MQTT-серверів, через які можна отримати доступ до платіжних даних або контролювати зарядку [15].

Можливості фізичного доступу через кабелі чи локальні порти дозволяють завантажувати шкідливу прошивку або проникати в локальні мережі під контролем EVSE. Так звані "juice jacking"-атаки – коли через кабель зарядки передається шкідливий код – також показують суттєвий ризик для пристроїв користувачів [16; 17].

Окремо слід виділити ризики, пов'язані з мобільними застосунками для керування зарядкою. Аналіз 31 популярного мобільного додатка виявив недостатню авторизацію користувачів або автомобілів, що дозволяє затримання сесій, нелегальний контроль процесу заряду або навіть запуск координаційних атак на енергомережу [18].

EVSE вразливі до supply-chain атак – підміни прошивки, до DoS та ботнетатак, що можуть спричинити одночасний ввімк/вимк зарядок у масштабі сотень станцій і спричинити скачки навантаження, що веде до нестабільності мережі або локальних відключень.

Нижче наведено таблицю 1 з класифікацією основних загроз та їхніх потенційних наслідків.

Загрози викликають фінансові втрати, порушення сервісу для користувачів та навіть загрожують національній безпеці через нестабільність енергосистеми. Наприклад, координовані атаки на зарядні мережі в США могли би створити навантаження, подібні до DDoS на комунікації, але спрямовані проти енергомережі. Британські та американські інфраструктури показують, що навіть несанкціоновані ремонти програмного забезпечення зарядок вже призводили до компрометації даних користувачів і порушень роботи пристроїв.

В українських умовах, із вже наявним досвідом таких подій, як атаки BlackEnergy у 2015 та Industroyer у 2016, ключовим є усвідомлення потенційних загроз EVSE. Відсутність контролю над ланцюгами постачання обладнання, слабкий сегмент мережевого розмежування та недостатній захист комунікацій створює передумови для повторення подібних сценаріїв [18; 19].

Таблиця 2 – Класифікація основних загроз та їхніх наслідків

Клас загроз	Тип атаки	Вплив
MITM / Replay	Перехоплення чи повторення пакетів OCPP/ISO 15118	Фальсифікація журналів, маніпуляція оплат
Firmware/Supply-chain	Шкідлива прошивка через кабель або порт	Блокування EVSE, мережеві проникнення
Уразливості в мобільних додатках	Захоплення сесій, ін'єкції, XSS, CSRF	Незаконний контроль EVSE, викрадення даних
DDoS/Ботнет	Масова активація/деактивація зарядок	Скачки навантаження, нестабільність мережі
Фізичний доступ	Через SSH/HTTP, USB/NFC-порти	Локальна компрометація обладнання

Системи зарядки електромобілів є складними кіберфізичними системами, які об'єднують апаратне забезпечення (зарядні станції, електромобілі), програмне забезпечення (мобільні додатки, CMS) та мережеві протоколи для обміну даними між цими компонентами. Основні функції таких систем включають ініціацію та припинення зарядних сесій, обробку платежів, моніторинг стану станцій і передачу даних про зарядку в реальному часі. Усі ці процеси залежать від надійної та безпечної передачі даних, оскільки будь-яке порушення може призвести до економічних збитків, порушення роботи інфраструктури або навіть дестабілізації електромережі.

Основними протоколами, які використовуються для передачі даних у системах зарядки електромобілів, є OCPP (Open Charge Point Protocol), ISO 15118 (Plug & Charge) та Modbus. OCPP є відкритим протоколом, який забезпечує зв'язок між зарядними станціями та CMS, дозволяючи керувати зарядними сесіями, оновлювати програмне забезпечення та обробляти платежі. ISO 15118, відомий як Plug & Charge, забезпечує автоматичну автентифікацію електромобіля та зарядної станції, що спрощує процес зарядки, але вимагає високого рівня безпеки для захисту від атак типу «людина посередині» [20]. Modbus, хоча й менш поширений, використовується для низькорівневого зв'язку в промислових системах зарядки [21]. Усі ці

протоколи потребують належного шифрування та автентифікації, щоб запобігти несанкціонованому доступу.

На рис. 2 зображено представлення протоколів передачі даних, які використовуються в процесі заряджання електромобіля.

Однією з основних загроз для систем зарядки електромобілів є атаки на мобільні додатки, які виступають інтерфейсом між користувачем і зарядною інфраструктурою.

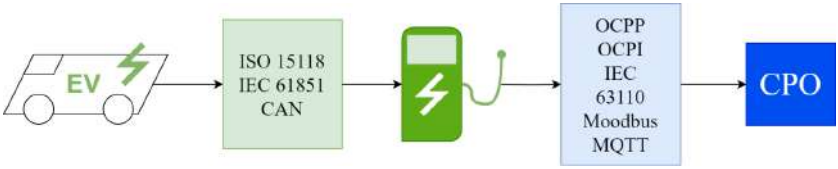


Рисунок 2 – Протоколи передачі даних у процесі зарядки

Дослідження показують, що більшість додатків вразливі через: недостатню верифікацію власності електромобіля, що дозволяє зловмисникам запускати зарядку для чужих транспортних засобів; слабку авторизацію для критичних функцій, таких як запуск і зупинка зарядки; неналежне шифрування даних, що передаються між додатком і CMS [18]. Наприклад, у реальному тесті, проведеному в 2022 році, дослідники успішно перехопили зарядні сесії в двох популярних додатках, використовуючи фальшиві акаунти [18]. В Україні такі вразливості є особливо небезпечними через обмежену кількість зарядних станцій і залежність від централізованих CMS.

Для захисту передачі даних у системах зарядки електромобілів використовуються різні технології шифрування, автентифікації та моніторингу. Найпоширенішим є протокол TLS версії 1.3, який забезпечує шифрування даних між додатком, зарядною станцією та CMS. Проте неправильна конфігурація TLS, наприклад використання застарілих версій або слабких сертифікатів, може створювати вразливості. Для підвищення безпеки також застосовується двофакторна автентифікація, яка вимагає додаткового підтвердження особи користувача, наприклад через SMS або біометричні дані. В Україні впровадження 2FA ускладнене через низький рівень цифрової грамотності користувачів і обмежену підтримку таких технологій у місцевих додатках.

Іншим важливим аспектом є верифікація власності електромобіля, яка запобігає несанкціонованим зарядним сесіям. Наприклад, протокол ISO 15118 використовує унікальний ідентифікатор транспортного засобу (VIN) для автоматичної автентифікації, що значно знижує ризик атак [20]. Однак впровадження цього протоколу в Україні обмежене через високу складність і необхідність оновлення апаратного забезпечення зарядних станцій. Альтернативним рішенням є використання фізичних ідентифікаторів, таких як RFID або QR-коди, які вимагають від користувача підтвердження присутності біля зарядної станції [18].

Масові зарядні атаки, які можуть створювати значне навантаження на електромережу, є ще однією серйозною загрозою. Дослідження на 7-шинній моделі Glover показало, що скоординована атака з використанням тисяч фальшивих акаунтів може викликати втрати до 10 % електроенергії та дестабілізацію мережі [18]. В Україні, де електромережа вже зазнає значних навантажень, такі атаки можуть мати катастрофічні наслідки. Для їх запобігання необхідні системи моніторингу в реальному часі, які виявляють аномальну активність, наприклад різке зростання кількості зарядних сесій.

Технології Vehicle-to-Grid додають додатковий рівень складності до захисту даних. V2G дозволяє електромобілям повертати енергію в мережу, що підвищує ефективність енергосистеми, але створює нові вразливості. Наприклад, скоординовані атаки на V2G-системи можуть викликати осциляторні навантаження, що призводять до коливань частоти в мережі нижче 50 Гц [22]. Для захисту V2G-систем необхідні спеціалізовані протоколи безпеки, такі як IEC 61851, які регулюють обмін даними між електромобілем і мережею.

В Україні впровадження технологій захисту даних у системах зарядки електромобілів стикається з кількома викликами. По-перше, обмежена кількість зарядних станцій і їх нерівномірний розподіл (здебільшого в містах, таких як Київ, Львів і Одеса) ускладнюють масштабування безпечних рішень. По-друге, відсутність єдиних національних стандартів безпеки для зарядної інфраструктури призводить до використання різнорідних систем, що підвищує ризик вразливостей. По-третє, низький рівень фінансування та технічної модернізації гальмує впровадження сучасних протоколів, таких як ISO 15118.

Для оцінки безпеки систем зарядки електромобілів використовуються статичні та динамічні методи аналізу. Статичний аналіз передбачає зворотне проєктування коду мобільних додатків і перевірку їхньої структури, тоді як динамічний аналіз включає моніторинг мережевого трафіку та поведінки додатків у реальному часі. В Україні такі методи рідко застосовуються через брак спеціалізованих лабораторій і експертів із кібербезпеки.

Мобільні додатки для заряджання електромобілів є частиною кіберфізичної системи, що об'єднує зарядні станції, хмарні сервери та електромобілі. Ця інтеграція створює численні вразливості, які можуть бути використані для атак. Основними вразливостями є відсутність належної верифікації власності електромобіля (Flaw 1) та неналежна авторизація для критичних функцій, таких як запуск і зупинка зарядки (Flaw 2) [18]. Ці недоліки дозволяють зловмисникам здійснювати несанкціоновані зарядні сесії, що може призвести до значних економічних і технічних наслідків.

Ключові типи атак включають:

1. Перехоплення зарядних сесій (Session Hijacking): Зловмисники можуть використовувати фальшиві акаунти для запуску зарядки електромобіля, який не належить їм, без необхідності компрометації легітимних облікових записів. Це можливо через відсутність перевірки зв'язку між користувачем і електромобілем у CMS.

2. Аналіз трафіку між додатком і CMS показав, що більшість додатків використовують HTTPS із TLS/SSL для шифрування, але вразливості в реалізації можуть дозволити перехоплення даних [23].

Координовані атаки, що використовують численні фальшиві акаунти, можуть створювати значне навантаження на електромережу, викликаючи нестабільність, як показано в симуляціях на 7-шинній моделі Glover [18].

1. Недостатня верифікація власності (Flaw 1): Більшість додатків не перевіряють, чи користувач, який ініціює зарядку, є власником електромобіля. Це дозволяє зловмисникам запускати зарядні сесії для будь-якого підключеного транспортного засобу.

2. Неналежна авторизація (Flaw 2): Критичні функції, такі як запуск і зупинка зарядки, не вимагають додаткової автентифікації, що спрощує атаки [18].

Хоча більшість додатків використовують TLS, неправильна конфігурація може призводити до вразливостей, таких як слабкі сертифікати або застарілі версії протоколів [24].

Зловмисники можуть використовувати вразливості для впливу на екосистему електромобілів. Наприклад, перехоплення зарядної сесії дозволяє зловмиснику запускати зарядку без відома власника, що може призвести до неавторизованих витрат або навіть відключення легітимного користувача від зарядної сесії. У реальному тесті, проведеному 1 листопада 2022 року, дослідники успішно перехопили зарядні сесії на двох різних додатках, підтвердивши можливість таких атак [18].

Інший сценарій передбачає масові зарядні атаки, які можуть створювати значне навантаження на електромережу. Наприклад, симуляція атаки на 7-шинну модель Glover показала, що масова зарядка може викликати втрати в передачі електроенергії та дестабілізацію мережі [18]. У сценарії, де атака розподілена пропорційно між шинами, втрати склали до 10 % від загального навантаження, що може призвести до відключення електроенергії [18].

На рис. 3 зображено схему атаки зловмисника.

Рисунок 3 – Схема атаки через мобільний додаток

Схема ілюструє, як зловмисник може ініціювати зарядку, обходячи перевірки CMS.

Атаки на мобільні додатки можуть мати серйозний вплив на електромережу, особливо в контексті технологій Vehicle-to-Grid (V2G). V2G дозволяє електромобілям повертати енергію в мережу, але вразливості в додатках можуть бути використані для створення осциляторних навантажень, що викликають коливання частоти

в мережі [25]. Наприклад, скоординована атака, що передбачає одночасне ввімкнення та вимкнення зарядки, може призвести до зниження частоти нижче 50 Гц, що вимагає відключення навантаження для стабілізації мережі.

Атака потужністю 40 кВт може викликати падіння частоти, а в поєднанні з V2G-функціями – створювати значні сплески частоти, що загрожують стабільності мережі [26].

Для зменшення вразливостей мобільних додатків в Україні необхідно впроваджувати комплексні заходи безпеки:

– Впровадження двофакторної автентифікації для критичних функцій, таких як запуск і зупинка зарядки, може значно знизити ризик несанкціонованого доступу [7].

– CMS має перевіряти зв'язок між користувачем і електромобілем, наприклад, через унікальний ідентифікатор транспортного засобу [18].

– Використання сучасних версій TLS (1.3) і регулярна перевірка сертифікатів допоможуть захистити комунікацію між додатком і CMS.

– Деякі додатки, такі як Tata Power EZ Charge, вимагають введення фізичного ідентифікатора EVCS або сканування QR-коду, що обмежує можливість віддалених атак [18].

Навчання персоналу та користувачів щодо виявлення фішингових атак і безпечного використання додатків наприклад, SMS-фішинг може бути використаний для маніпуляції поведінкою користувачів.

В Україні необхідно впроваджувати стандарти безпеки, такі як IEC 61851 та IEC 62443, які регулюють безпеку зарядних систем і кіберфізичних систем [27]. Крім того, гармонізація з європейськими стандартами, такими як GDPR, допоможе захистити персональні дані користувачів [28].

Міжнародна співпраця в цій сфері має кілька ключових переваг. По-перше, вона дозволяє Україні гармонізувати свої стандарти безпеки з міжнародними, що сприяє інтеграції до європейського цифрового ринку. По-друге, співпраця з міжнародними партнерами, зокрема в рамках ініціатив ЄС, таких як EU4Health та EU4DigitalUA, забезпечує доступ до передових технологій, фінансування та експертизи. По-третє, обмін досвідом дозволяє адаптувати глобальні

практики до локальних умов, враховуючи обмежену інфраструктуру та низький рівень цифрової грамотності в Україні.

Європейський Союз є одним із основних партнерів України в розвитку кібербезпеки, зокрема через програми EU4Health та EU4DigitalUA. EU4Health, хоча й зосереджена переважно на охороні здоров'я, включає компоненти, пов'язані з захистом критичної інфраструктури [29]. У рамках цієї програми ЄС фінансує тренінги, моделювання кібератак і розробку стратегій захисту даних, що можуть бути адаптовані до електромобільної інфраструктури. У 2021 році в Києві відбулися навчання з кібербезпеки, організовані ЄС, за участю представників Державної служби спеціального зв'язку та захисту інформації України, СБУ, Кіберполіції та інших установ [30].

EU4DigitalUA зосереджена на гармонізації цифрової інфраструктури України з Єдиним цифровим ринком ЄС. Один із її компонентів присвячений кібербезпеці та захисту даних, що включає розробку систем обміну даними, таких як «Трембіта», та впровадження стандартів захисту інформації [31]. Ці ініціативи сприяють модернізації зарядної інфраструктури через інтеграцію безпечних протоколів OCPP та ISO 15118.

В таблиці 3 наведено основні ініціативи ЄС для підтримки кібербезпеки в Україні.

Україна активно співпрацює з іншими міжнародними партнерами, такими як НАТО, США, Ізраїль та Норвегія, у сфері кібербезпеки. Наприклад, у рамках Трастового фонду Україна – НАТО з питань кібербезпеки, підписаного в 2015 році, Україна отримала програмне забезпечення, обладнання та консультативну допомогу на суму 965 млн євро [32]. Ця підтримка може бути адаптована для захисту електромобільної інфраструктури, зокрема через модернізацію дата-центрів і впровадження систем моніторингу в реальному часі.

Співпраця зі США та Ізраїлем зосереджена на залученні технологій від провідних компаній, таких як IBM, Microsoft і Radware, які розробляють рішення для захисту кіберфізичних систем [32]. Наприклад, IBM пропонує платформи на основі штучного інтелекту для виявлення аномалій у мережевому трафіку, що може бути використано для захисту зарядних станцій від масових атак [33]. Ізраїльські компанії, такі як Radware, спеціалізуються на захисті від атак типу DDoS, які можуть дестабілізувати CMS зарядних мереж.

Таблиця 3 – Ініціативи ЄС для підтримки кібербезпеки в Україні

Ініціатива	Основні цілі	Релевантність для електромобільної інфраструктури	Фінансування та період
EU4Health	Захист критичної інфраструк-тури, тренінги з кібербезпеки	Адаптація практик для захисту даних у зарядних системах	2021–2027 (5,3 млрд євро)
EU4DigitalUA	Гармонізація з цифровим ринком ЄС, розвиток систем обміну даними	Впровадження безпечних прото-колів (OCPP, ISO 15118)	2020–2025 (25 млн євро)
U-LEAD з Європою	Підтримка цифрової трансформації, навчання фахівців	Підготовка спеці-алістів для управ-ління зарядною інфраструктурою	2016–2020 (100 млн євро)

Норвегія, яка має одну з найрозвиненіших електромобільних інфраструктур у світі, може стати важливим партнером для України. Норвезький досвід впровадження стандартів ISO 15118 та технологій Vehicle-to-Grid (V2G) може бути адаптований до українських реалій, особливо для захисту даних у V2G-системах, які дозволяють електромобілям повертати енергію в мережу.

Для реалізації цих перспектив необхідно подолати виклики, пов'язані з обмеженою інфраструктурою, браком фахівців і низьким рівнем цифрової грамотності. Підвищення безпеки передачі даних у системах зарядки електромобілів в Україні зумовлене комплексними заходами, які охоплюють технічні, організаційні та регуляторні аспекти. Технічні заходи включають використання сучасних версій TLS, впровадження 2FA та верифікацію власності через VIN. Організаційні заходи передбачають навчання користувачів і персоналу щодо виявлення фішингових атак і безпечного використання додатків. Регуляторні заходи включають гармонізацію українських стандартів із міжнародними, такими як IEC 61851 та IEC 62443, а також адаптацію вимог GDPR для захисту персональних даних користувачів.

Головний напрям розвитку це розробка локалізованих рішень для України, враховуючи обмежену інфраструктуру та специфічні

виклики енергетичної системи. Наприклад, створення децентралізованих CMS може зменшити залежність від єдиних точок відмови та підвищити стійкість до атак. Крім того, необхідно інвестувати в підготовку спеціалістів із кібербезпеки, які зможуть адаптувати міжнародні практики до українських реалій.

Безпека передачі даних у системах зарядки електромобілів є критично важливою для забезпечення стабільності енергетичної інфраструктури та захисту користувачів. В Україні, де електромобільна інфраструктура активно розвивається, впровадження сучасних технологій захисту має стати пріоритетом. Водночас необхідно враховувати локальні виклики, такі як обмежена кількість зарядних станцій і низький рівень цифрової грамотності, для створення ефективних і доступних рішень.

Для впровадження змін пропонується структурований план дій, який охоплює технічні, організаційні та регуляторні заходи. На рис. 4 зображено етапи впровадження з фокусом на конкретні дії та очікувані результати.

Впровадження технологій шифрування та автентифікації спрямоване на підвищення безпеки передачі даних і доступу до систем електромобільної інфраструктури.

Необхідно налаштувати протокол Transport Layer Security версії 1.3 для всіх каналів зв'язку, включаючи комунікацію між мобільними додатками, системами керування зарядкою і зарядними станціями.

Це передбачає оновлення серверного програмного забезпечення, встановлення надійних сертифікатів і регулярну перевірку їхньої актуальності для запобігання використанню застарілих або слабких сертифікатів, які можуть бути вразливими до атак типу «людина посередині».

Паралельно впроваджується двофакторна автентифікація для критичних функцій, таких як запуск і зупинка зарядних сесій, використовуючи підтвердження через SMS, біометричні дані (наприклад, сканування відбитків пальців) або сканування QR-кодів на зарядній станції, щоб знизити ризик несанкціонованого доступу, навіть якщо зловмисники отримають доступ до облікових даних користувача.

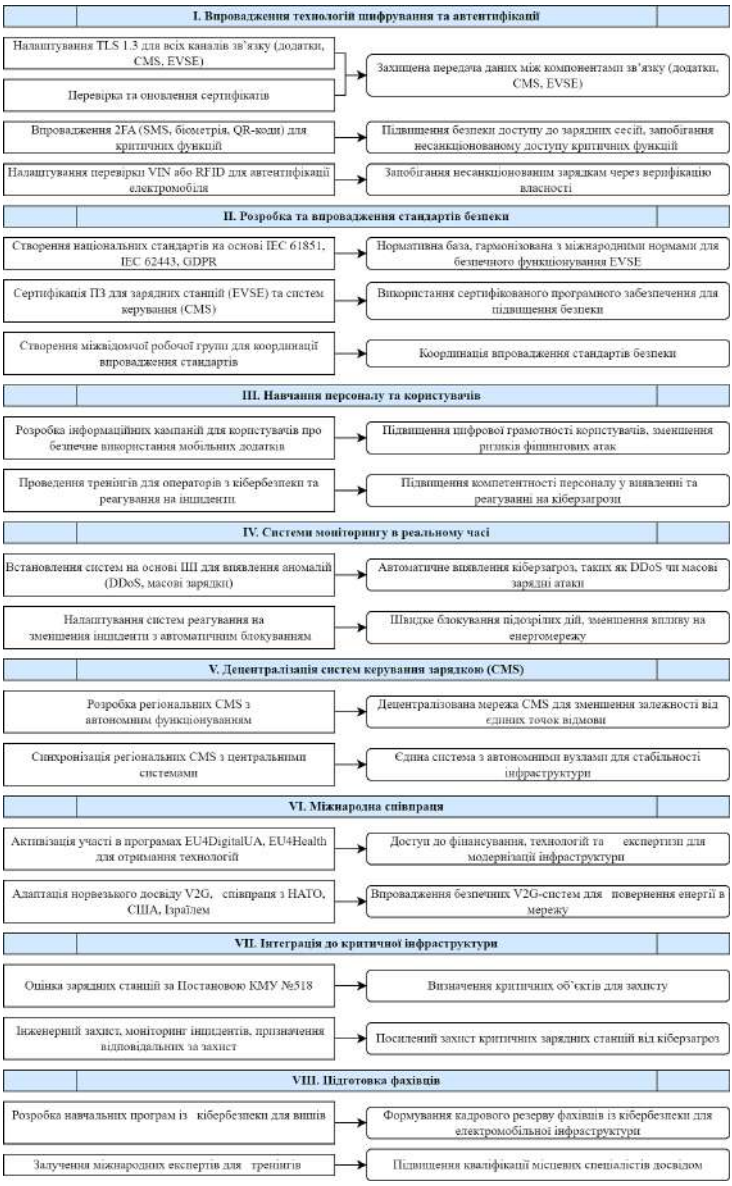


Рисунок 4 – Етапи впровадження змін до захисту інформації в процесі функціонування електромобільної інфраструктури в Україні

Також налаштовується верифікація власності електромобіля через унікальний ідентифікатор транспортного засобу або фізичні ідентифікатори, такі як RFID-мітки, що забезпечить можливість ініціювання зарядки лише власником або авторизованим користувачем, зменшуючи ризик атак типу “session hijacking”.

За реалізацію відповідають оператори зарядних мереж, розробники програмного забезпечення та виробники зарядних станцій, які співпрацюють для оновлення апаратного й програмного забезпечення з метою підтримки цих функцій. Очікувані результати включають захищену передачу даних між компонентами інфраструктури, підвищення безпеки доступу до зарядних сесій і запобігання несанкціонованим зарядкам.

На цьому етапі розробки та впровадження стандартів безпеки створюється регуляторна база для забезпечення безпеки електромобільної інфраструктури. Розробляються національні стандарти безпеки, гармонізовані з міжнародними нормами, такими як IEC 61851 (для систем зарядки електромобілів), IEC 62443 (для безпеки кіберфізичних систем) і GDPR (для захисту персональних даних), які визначатимуть вимоги до шифрування, автентифікації, моніторингу та обробки даних користувачів, включаючи геолокацію, платіжну інформацію та історію зарядок. Проводиться сертифікація програмного забезпечення для зарядних станцій і систем керування зарядкою на відповідність стандартам безпеки, зокрема захист від вразливостей, таких як слабке шифрування чи недостатня авторизація. Для координації впровадження стандартів створюється міжвідомча робоча група за участю Міністерства інфраструктури, Держспецзв’язку та операторів зарядних мереж, яка забезпечує обмін інформацією між зацікавленими сторонами та контролює дотримання вимог. Відповідальність за розробку стандартів і координацію покладається на Міністерство інфраструктури та Держспецзв’язок, тоді як сертифікацію проводять спеціалізовані центри за участю операторів мереж.

Етап навчання персоналу та користувачів зосереджений на підвищенні рівня безпеки через підготовку операторів зарядних мереж і користувачів електромобілів. Проводяться регулярні тренінги для персоналу операторів із питань кібербезпеки, які охоплюють

виявлення фішингових атак, налаштування безпечного обладнання, реагування на кіберінциденти та використання систем моніторингу. Організацію тренінгів забезпечують оператори мереж у співпраці з громадськими організаціями, які мають досвід у сфері кібербезпеки. Водночас розробляються інформаційні кампанії для користувачів електромобілів, спрямовані на підвищення цифрової грамотності, що пояснюють, як безпечно використовувати мобільні додатки, уникати атак типу "juice jacking" (передача шкідливого коду через зарядні кабелі) і розпізнавати SMS-фішинг. Для цього оператори мереж співпрацюють із медіа та громадськими організаціями, створюючи доступні матеріали, такі як відео, інфографіки чи посібники.

На етапі системи моніторингу створюється інфраструктура для виявлення та реагування на кіберзагрози. Встановлюються системи на основі штучного інтелекту для моніторингу аномальної активності, такої як DDoS-атаки чи масові зарядні атаки, які можуть дестабілізувати електромережу. Ці системи аналізують мережевий трафік і поведінку зарядних сесій, виявляючи підозрілі дії, наприклад, різке зростання кількості одночасних зарядок. Впровадження забезпечують оператори зарядних мереж у співпраці з CERT-UA та IT-компаніями, такими як IBM чи Radware, які надають рішення для захисту кіберфізичних систем. Налаштовуються системи автоматичного реагування на інциденти, які блокують підозрілі дії, наприклад, несанкціоновані зарядні сесії чи спроби доступу до CMS, із інтеграцією до центральних платформ моніторингу. За це відповідають CERT-UA та оператори мереж.

Для підвищення стійкості до атак децентралізуються системи керування зарядкою. Розробляються регіональні CMS, які працюють автономно, але синхронізуються з центральними системами для обміну даними, наприклад, про стан зарядних станцій чи історію транзакцій. Це зменшує залежність від єдиних точок відмови, таких як централізовані сервери, і забезпечує стабільність у разі збоїв мережі. Оператори зарядних мереж і Укренерго відповідають за проєктування та тестування регіональних CMS, тоді як розробники програмного забезпечення та оператори мереж забезпечують сумісність і безпеку зв'язку між регіональними та центральними системами.

Міжнародна співпраця необхідна для отримання технологій, фінансування та експертизи. Активізується участь України в програмах ЄС, таких як EU4DigitalUA та EU4Health, які надають ресурси для розвитку кібербезпеки та цифрової інфраструктури. Мінцифри, МЗС і оператори зарядних мереж координують участь у цих програмах, адаптуючи європейські практики до локальних умов. Співпраця з партнерами, такими як Норвегія, НАТО, США та Ізраїль, дозволяє впроваджувати передові рішення, зокрема технології Vehicle-to-Grid, використовуючи норвезький досвід для налаштування безпечних V2G-систем, які дозволяють електромобілям повертати енергію в мережу. Мінцифри та оператори мереж співпрацюють із міжнародними компаніями, такими як IBM і Radware, для інтеграції їхніх рішень.

Етап інтеграції передбачає включення зарядних станцій до об'єктів критичної інфраструктури. Проводиться категоризація зарядних станцій відповідно до Постанови КМУ № 518, оцінюючи їхню важливість для енергосистеми та потенційні ризики в разі атак. Впроваджується інженерний захист, моніторинг інцидентів і призначення відповідальних осіб для управління безпекою, включаючи обмеження доступу до портів USB чи NFC і підключення до систем моніторингу CERT-UA. Оператори мереж, CERT-UA і Укренерго відповідають за оцінку та захист, а Держспецзв'язок забезпечує дотримання вимог.

Підготовка фахівців зосереджена на формуванні кадрового резерву. Розробляються навчальні програми з кібербезпеки для вищих навчальних закладів, які готуватимуть спеціалістів із захисту електромобільної інфраструктури. ВНЗ співпрацюють із міжнародними партнерами, такими як ЄС чи НАТО, для створення актуальних курсів. Залучаються міжнародні експерти для проведення тренінгів із сучасних технологій захисту, таких як моніторинг аномалій чи захист V2G-систем. Мінцифри та ВНЗ координують ці програми, залучаючи фінансування через міжнародні гранти.

Ці етапи забезпечать комплексний захист інформації в електромобільній інфраструктурі, підвищать її стійкість до кіберзагроз і сприятимуть інтеграції з європейськими стандартами, враховуючи локальні виклики, такі як обмежена інфраструктура та низька цифрова грамотність.

Висновки. Розвиток електромобільної інфраструктури в Україні є ключовим елементом переходу до сталого транспорту, декарбонізації та енергетичної незалежності, проте супроводжується значними викликами у сфері захисту інформації. Електромобільна інфраструктура, як кіберфізична система, об'єднує зарядні станції, системи керування зарядкою, мобільні додатки та енергомережі, що створює численні вразливості до кіберзагроз, таких як атаки типу «людина посередині», перехоплення зарядних сесій, DDoS-атаки та маніпуляції з платіжними даними. Відсутність єдиних національних стандартів безпеки, недостатня верифікація власності електромобілів і слабке шифрування в протоколах OCPP та ISO 15118 підвищують ризики для безпеки користувачів і стабільності енергосистеми. В Україні ці проблеми ускладнюються обмеженою кількістю зарядних станцій, нерівномірним їх розподілом, низьким рівнем цифрової грамотності та нестачею фахівців із кібербезпеки. Нормативно-правова база, включаючи закони про захист інформації, персональних даних і критичну інфраструктуру, частково регулює діяльність електромобільної інфраструктури, але не враховує її специфіки, зокрема унікальних вимог до захисту кіберфізичних систем і сертифікації програмного забезпечення. Впровадження комплексних заходів, таких як використання протоколу TLS 1.3, двофакторної автентифікації, верифікації через VIN або RFID, децентралізації систем керування зарядкою та моніторингу в реальному часі на основі штучного інтелекту, може значно підвищити безпеку. Міжнародна співпраця, зокрема через програми ЄС та партнерство з Норвегією, НАТО, США й Ізраїлем, забезпечує доступ до технологій, фінансування та експертизи, що дозволяє адаптувати передові практики до локальних умов. Гармонізація з міжнародними стандартами сприятиме захисту даних і підвищенню довіри користувачів. Навчання персоналу та користувачів, а також підготовка фахівців із кібербезпеки через спеціалізовані програми у вишах, є критично важливими для подолання кадрового дефіциту та підвищення цифрової грамотності. Інтеграція зарядних станцій до об'єктів критичної інфраструктури з дотриманням вимог забезпечить їхній захист від загроз національній безпеці. Таким чином, комплексний підхід, що поєднує технічні, організаційні та

регуляторні заходи, дозволить Україні створити стійку та безпечну електромобільну інфраструктуру, яка відповідає сучасним викликам і сприяє сталому розвитку, енергетичній стабільності та інтеграції до європейського цифрового простору.

Список використаних джерел

1. Алімурадов Ш. Р. Електромобільна інфраструктура як об'єкт дослідження транспортної логістики. *Вісник Одеського національного морського університету*. 2021. № 67. С. 104–110.
2. Tamay P., Inga Ortega E. Charging Infrastructure for Electric Vehicles Considering Their Integration into the Smart Grid. *Sustainability*. 2022. Vol. 14. P. 8248–8267. DOI: 10.3390/su14148248
3. Ajanovic A., Haas R. Dissemination of electric vehicles in urban areas: major factors for success. *Energy*. 2016. Vol. 115. P. 1451–1458.
4. Zhang X., Xie J., Rao R., Liang Y. Policy incentives for the adoption of electric vehicles across countries. *Sustainability.* 2020. Vol. 12 (10). P. 4321.
5. Zharovska I., Kozak Y., Semenov V. Comprehensive study on electric vehicles and infrastructure for sustainable development in Ukraine. *Energies.* 2023. Vol. 16 (7). DOI: https://doi.org/10.1051/e3sconf/202456701025
6. Pivnyak G., Olishevska V., Olishevskiy H., Lutsenko I., Lysenko A. Comprehensive study on electric vehicles and infrastructure for sustainable development in Ukraine. *E3S Web of Conferences*. 2024. DOI: 10.1051/e3sconf/202456701025
7. Марків О., Ришковець Ю. Інформаційні технології і протоколи передачі даних у системах зарядки електромобілів. *Information Systems and Networks*. 2025. Вип. 17. С. 304–318. DOI: https://doi.org/10.23939/sisn2025.17.304.
8. Грома Я. В., Глущенко Я. І. Порівняльний аналіз ринку електромобілів в Україні та світі. *Економічний вісник Національного технічного університету України «Київський політехнічний інститут»*. 2019. № 16. С. 42–49.
9. Чорний Р. С., Чорна Н. П., Шевчук Я. В. Електромобільна інфраструктура в Україні: проблеми та перспективи розвитку. *Міжнародний науковий журнал Інтернаука.* 2020. № 19 (2). С. 49–53.

10. Про захист інформації в інформаційно-телекомунікаційних системах : Закон України від 05.07.1994 р. № 80/94-ВР. Дата оновлення: 20.04.2025. *Верховна рада.* URL: https://zakon.rada.gov.ua/go/80/94-вр

11. Кодекс України про адміністративні правопорушення : щодо відповідальності за недотримання умов захисту інформації. *Відомості Верховної Ради Української РСР* (ВВР) 1984, додаток до № 51, ст.1122) URL: https://dduvs.edu.ua/wp-content/uploads/files/Structure/f/fpfpmgb/p/d3.pdf

12. Про захист персональних даних : Закон України від від 23.02.2012 р. № № 4452-VI. Дата оновлення: 14.06.2025. *Верховна рада України.* URL: https://zakon.rada.gov.ua/laws/show/2297-17#Text

13. Про критичну інфраструктуру : Закон України від від 18.11.2021 р. № 1909-IX. Дата оновлення: 21.09.2024. *Верховна рада України.* URL: https://zakon.rada.gov.ua/laws/show/1882-20#Text

14. Про організацію захисту та забезпечення безпеки функціонування об'єктів критичної інфраструктури та енергетики України в умовах ведення воєнних дій : Указ Президента України від 17.10.2023 р. № 695/2023. Дата оновлення: 19.10.2023. *Верховна рада України.* URL: https://zakon.rada.gov.ua/laws/show/n0040525-23#Text

15. Johnson J., Berg T., Anderson B., Wright B. Review of Electric Vehicle Charger Cybersecurity Vulnerabilities, Potential Impacts, and Defenses. *Energies*. 2022. Vol. 15 (11). P. 3931–3952. DOI: 10.3390/en15113931

16. Poudel S., Baugh J. E., Takiddin A., Ismail M., Refaat S. S. Injection attacks and detection strategy in front-end vehicle-to-grid communication. *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. Glasgow, United Kingdom. 2023. P. 1–6. DOI: 10.1109/SmartGridComm57358.2023.10333927

17. Nasr T., Torabi S., Bou-Harb E. Power jacking your station: In-depth security analysis of electric vehicle charging station management systems. *Computers & Security*. 2022. Vol. 112. P. 102511. DOI: 10.1016/j.cose.2021.102511

18. Sarieddine K., Sayed M. A., Torabi S. та ін. Investigating the Security of EV Charging Mobile Applications As an Attack Surface. *arXiv*. 2022. URL: https://arxiv.org/abs/2211.10603

19. Yemelyanov V., Bondar H. Cyber security as a component of national security and cyber protection of critical infrastructure of Ukraine. *Public Administration and Regional Development.* 2019. № 5. P. 493–523. DOI: 10.34132/pard2019.05.02

20. Attanasio L., Conti M., Donadel D., Turrin F. MiniV2G: an electric vehicle charging emulator. *Proceedings of the 7th ACM on Cyber-Physical System Security Workshop*. 2021. P. 65–73.

21. Alzahrani A., Wangikar S. M., Indragandhi V., Singh R. R., Subramaniyaswamy V. Design and implementation of SAE J1939 and Modbus communication protocols for electric vehicle. *Machines*. 2023. Vol. 11 (2). P. 201–223. DOI: https://doi.org/10.3390/machines11020201

22. Chen G., Zhang Z. Control Strategies, Economic Benefits, and Challenges of Vehicle-to-Grid Applications: Recent Trends Research. *World Electric Vehicle Journal.* 2024. Vol. 15 (5). P. 190–214. DOI: 10.3390/wevj15050190

23. Van Aubel P., Poll E. Security of EV-charging protocols. *arXiv*. 2022. URL: https://arxiv.org/abs/2202.04631

24. Sarieddine K., Sayed M. A., Assi C., Atallah R., Torabi S., Khoury J. EV charging infrastructure discovery to contextualize its deployment security. *IEEE Transactions on Network and Service Management*. 2023. Vol. 21 (1). P. 1287–1301. DOI: 10.1109/TNSM.2023.3318406

25. Johnson J. & et. Cybersecurity for Electric Vehicle Charging Infrastructure. *SANDIA REPORT*. 2022. 66 p. DOI: 10.2172/1877784

26. Sayed M. A., Ghafouri M., Atallah R., Debbabi M., Assi C. An Electric Vehicle Control Strategy to Mitigate Load Altering Attacks Against Power Grids. *2023 IEEE 8th International Conference on Recent Advances and Innovations in Engineering (ICRAIE).* 2023. P. 1–6. DOI: 10.1109/ICRAIE59459.2023.10468152

27. ДСТУ EN IEC 61851-1:2021. Система зарядки електричних транспортних засобів дротова. Частина 1. *Загальні вимоги.* URL: https://online.budstandart.com/ua/catalog/doc-page.html?id_doc=95965

28. General Data Protection Regulation (GDPR). *Official Journal of the European Union.* 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj

29. EU4Health Programme. *European Commission*. 2021. URL: https://ec.europa.eu/health/funding/eu4health_en

30. EU enhances Ukraine's capacity to respond to cyber-attacks. Офіційний вебсайт. URL: https://eufordigital.eu/eu-enhances-ukraines-capacity-to-respond-to-cyberattacks/

31. EU4DigitalUA: Interoperability, E-services and Cybersecurity for Ukraine. *e-Governance Academy.* 2020. URL: https://ega.ee/project/eu4digitalua/.

32. Український кіберпростір: безпекові загрози, виклики та перспективи розвитку. *Аналітичний центр ADASTRA*. 2020. URL: https://adastra.org.ua/blog/ukrayinskij-kiberprostir-bezpekovi-zagrozi-vikliki-ta-perspektivi-rozvitku.

33. IBM Security.IBM. Офіційний сайт. 2025. URL: https://www.ibm.com/security.

Наукове видання

БЕЗПЕКА ІНФОРМАЦІЇ ТА ІНФРАСТРУКТУРИ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМ: МІЖДИСЦИПЛІНАРНИЙ ПІДХІД

Монографія

За редакцією
доктора економічних наук, професора Станіслава Горбаченка
та доктора технічних наук, професора Артема Соколова

Дизайн обкладинки – *В. Савельєва*
Технічний редактор – *О. Гринюк*
Верстка – *Ю. Семенченко*

Підписано до друку 06.10.2025 р. Формат 60×84/16.
Папір офсетний. Гарнітура Times. Цифровий друк.
Ум. друк. арк. 25,69. Наклад 300. Замовлення № 0925-087.
Ціна договірна. Віддруковано з готового оригінал-макета.

Українсько-польське наукове видавництво “Liha-Pres”
79000, м. Львів, вул. Технічна, 1
87-100, м. Торунь, вул. Лубіцка, 44
Телефон: +38 (050) 658 08 23
E-mail: editor@liha-pres.eu
Свідоцтво суб’єкта видавничої справи
ДК № 6423 від 04.10.2018 р.