НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
Навчально-науковий фізико-технічний інститут
Кафедра математичного моделювання та аналізу даних

«На правах рукопису»
УДК 004.94:519.6/519.7

«До захисту допущено»
В.о. завідувача кафедри
_____ Іван ТЕРЕЩЕНКО
«__» _____ 2026 р.

Дипломна робота

на здобуття ступеня бакалавра
за освітньо-професійною програмою
«Математичні методи моделювання, розпізнавання образів та комп'ютерного зору»

зі спеціальності: 113 Прикладна математика
на тему: **«Математичні моделі та методи оцінки ризиків кіберзагроз в умовах невизначеності даних»**

Виконав:
студент IV курсу, групи ФІ-21
Мелоян Мирослав Вагаршакович _____

Керівник:
доктор філософії, доцент
Яйлимова Ганна Олексіївна _____

Рецензент:
канд. техн. наук, зав. кафедри ММЗІ
Яковлєв Сергій Володимирович _____

Засвідчую, що у цій дипломній роботі немає запозичень з праць інших авторів без відповідних посилань.
Студент _____

Київ — 2026

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
Навчально-науковий фізико-технічний інститут
Кафедра математичного моделювання та аналізу даних

Рівень вищої освіти — перший (бакалаврський)
Спеціальність — 113 Прикладна математика,
ОПП «Методи математичного моделювання, розпізнавання образів та комп'ютерного зору»

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
_____ Іван ТЕРЕЩЕНКО
«__» _____ 2026 р.

ЗАВДАННЯ
на дипломну роботу

Студент: Мелоян Мирослав Вагаршакович

1. Тема роботи: «*Математичні моделі та методи оцінки ризиків кіберзагроз в умовах невизначеності даних*»,

керівник: доктор філософії, доцент Яйлимова Ганна Олексіївна,

затверджені наказом по університету №__ від «__» _____ 2026 р.

2. Термін подання студентом роботи: «__» _____ 2026 р.

3. Вихідні дані до роботи: *стандарти та методології оцінювання ризиків інформаційної безпеки; математичні методи моделювання невизначеності; нечіткі системи виведення; байєсівські мережі; теорія доказів Демпстера–Шейфера; набір даних CIC-IDS2017; потокові ознаки мережевого трафіку; програмні засоби Python та Jupyter Notebook.*

4. Зміст роботи: *аналіз сучасних підходів до оцінювання кіберризиків в умовах невизначеності даних; формулювання математичної постановки задачі та визначення факторів ризику; побудова гібридної моделі на основі нечіткої системи виведення, байєсівської мережі та теорії доказів Демпстера–Шейфера; програмна реалізація моделі та експериментальна перевірка на даних CIC-IDS2017;*

аналіз результатів, конфлікту між джерелами свідчень, впливу неповноти факторної інформації та обмежень запропонованого підходу.

5. Перелік ілюстративного матеріалу: *презентація доповіді.*

6. Дата видачі завдання: 10 вересня 2025 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання	Примітка
1	Узгодження теми дипломної роботи з науковим керівником	01–15 вересня 2025 р.	Виконано
2	Огляд джерел за тематикою дослідження та аналіз підходів до оцінювання кіберризиків	Вересень–жовтень 2025 р.	Виконано
3	Формулювання мети, завдань, об'єкта, предмета та методів дослідження	Жовтень–листопад 2025 р.	Виконано
4	Розроблення математичної постановки задачі та формалізація гібридної моделі	Листопад–грудень 2025 р.	Виконано
5	Побудова FIS-, BN- та DST-компонентів моделі оцінювання ризику	Грудень 2025 р. – січень 2026 р.	Виконано
6	Підготовка експериментальних даних та параметризація факторного відображення	Січень–лютий 2026 р.	Виконано
7	Програмна реалізація моделі та проведення експериментального дослідження	Лютий–квітень 2026 р.	Виконано
8	Аналіз результатів, обмежень моделі та формулювання висновків	Квітень–травень 2026 р.	Виконано
9	Оформлення роботи, підготовка доповіді та презентаційних матеріалів	Травень–червень 2026 р.	Виконано

Студент _____ Мелоян М.В.

Керівник _____ Яйлимова Г.О.

РЕФЕРАТ

Кваліфікаційна робота містить: 128 стор., 6 рисунків, 13 таблиць, 38 джерел.

Метою роботи є побудова та дослідження гібридної математичної моделі нормованого оцінювання ризику кіберзагроз за мережевими ознаками в умовах неповноти, нечіткості та суперечливості даних. Об'єктом дослідження є процес оцінювання ризиків кіберзагроз у мережевих інформаційних системах за умов невизначеності даних. Предметом дослідження є математичні моделі та методи інтеграції нечітких, імовірнісних і доказових оцінок ризику.

У роботі проаналізовано сучасні підходи до оцінювання ризиків інформаційної безпеки, сформульовано математичну постановку задачі та побудовано гібридну модель, що поєднує нечітку систему виведення, байєсівську мережу та теорію доказів Демпстера–Шейфера. Запропоновано відображення мережевих ознак у нормовані фактори ризику, сформовано базові функції призначення маси для FIS- та BN-компонентів і виконано їх агрегацію з урахуванням надійності джерел та конфлікту між свідченнями. Програмну реалізацію виконано мовою Python на основі набору даних CIC-IDS2017. Проведено порівняння моделей, аналіз фінального індексу ризику, конфлікту, чутливості до параметрів та впливу неповноти факторної інформації. Показано, що запропонована модель дозволяє отримувати нормовану оцінку ризику та явно враховувати залишкову невизначеність і суперечливість різнорідних джерел свідчень.

КІБЕРРИЗИК, НЕВИЗНАЧЕНІСТЬ ДАНИХ, НЕЧІТКА СИСТЕМА ВИВЕДЕННЯ, БАЙЄСІВСЬКА МЕРЕЖА, ТЕОРІЯ ДОКАЗІВ ДЕМПСТЕРА–ШЕЙФЕРА, БАЗОВА ФУНКЦІЯ ПРИЗНАЧЕННЯ МАСИ, АГРЕГАЦІЯ СВІДЧЕНЬ, CIC-IDS2017

ABSTRACT

The qualification work contains: 128 pages, 6 figures, 13 tables, 38 references.

The aim of the work is to construct and study a hybrid mathematical model for normalized cyber threat risk assessment based on network features under conditions of incomplete, fuzzy and contradictory data. The object of the research is the process of cyber threat risk assessment in network information systems under conditions of data uncertainty. The subject of the research is mathematical models and methods for integrating fuzzy, probabilistic and evidential risk assessments.

The work analyzes modern approaches to information security risk assessment, formulates the mathematical problem statement and constructs a hybrid model that combines a fuzzy inference system, a Bayesian network and Dempster–Shafer evidence theory. A mapping from network features to normalized risk factors is proposed, basic probability assignments for the FIS and BN components are formed, and their aggregation is performed with account of source reliability and conflict between evidence. The software implementation was carried out in Python using the CIC-IDS2017 dataset. The study includes model comparison, analysis of the final risk index, conflict, sensitivity to parameters and the influence of incomplete factor information. It is shown that the proposed model makes it possible to obtain a normalized risk assessment and explicitly account for residual uncertainty and contradiction between heterogeneous sources of evidence.

CYBER RISK, DATA UNCERTAINTY, FUZZY INFERENCE SYSTEM, BAYESIAN NETWORK, DEMPSTER–SHAFER EVIDENCE THEORY, BASIC PROBABILITY ASSIGNMENT, EVIDENCE AGGREGATION, CIC-IDS2017

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CIC-IDS2017 — набір даних мережевого трафіку, використаний для експериментального дослідження

FIS — нечітка система виведення

BN — байєсівська мережа

DST — теорія доказів Демпстера–Шейфера

BPA — базова функція призначення маси

CPT — таблиця умовних ймовірностей

BetP — пігністичне перетворення

Θ — простір гіпотез рівня ризику

z — вектор ознак мережевого спостереження

$\phi(z)$ — відображення мережевих ознак у фактори ризику

T — фактор активності загрози

V — фактор вразливості або експозиції

A — фактор критичності активу

I — фактор потенційного впливу інциденту

R_{FIS} — оцінка ризику, отримана нечіткою системою виведення

R_{BN} — числова оцінка ризику, отримана з байєсівської мережі

R_{final} — фінальний нормований індекс ризику

K — коефіцієнт конфлікту між джерелами свідчень

$\rho_{\text{FIS}}, \rho_{\text{BN}}$ — коефіцієнти надійності FIS- та BN-компонентів

ВСТУП

Актуальність дослідження. Сучасні інформаційні системи функціонують в умовах зростання складності кіберзагроз, великих обсягів мережевих даних і швидкої зміни сценаріїв атак. За таких умов оцінювання ризиків інформаційної безпеки має спиратися не лише на організаційні процедури ризик-менеджменту, а й на формальні математичні моделі, здатні працювати з неповними, нечіткими та суперечливими даними.

Міжнародні стандарти й методології, зокрема ISO/IEC 27005:2022 [1], NIST SP 800-30 Rev. 1 [2], NIST Cybersecurity Framework 2.0 [3] та Open FAIR [4, 5], задають процесну основу для ідентифікації, аналізу та обробки ризиків. Водночас вони не визначають єдиного математичного механізму, який би дозволяв узгоджено поєднувати експертні оцінки, статистичні залежності та конфліктні свідчення від різних джерел.

Однією з ключових проблем є семантичний розрив між низькорівневими ознаками мережевого трафіку та високорівневими факторами ризику. Дані моніторингу містять технічні характеристики мережевих потоків, тоді як ризик-орієнтовані моделі оперують поняттями загрози, вразливості, критичності активу та потенційного впливу. Без формального відображення між цими рівнями складно забезпечити узгодженість між експериментальними даними та математичною моделлю ризику.

Для моделювання такої невизначеності доцільно поєднувати кілька математичних підходів. Нечіткі системи виведення дають змогу формалізувати експертні правила та лінгвістичні категорії [6, 7, 8], а байєсівські мережі забезпечують імовірнісне моделювання залежностей між факторами ризику [9, 10, 11]. Оскільки ці підходи мають різну математичну природу, їх результати не можна коректно об'єднувати

простим усередненням. Для формальної агрегації різнорідних і потенційно конфліктних свідчень у роботі використовується апарат теорії доказів Демпстера–Шейфера [12, 13, 14, 15].

Отже, актуальність дослідження зумовлена потребою в гібридній математичній моделі оцінювання кіберризику, орієнтованій на узгоджене використання нечіткого виведення, байєсівського моделювання та теорії доказів за умов неповноти, нечіткості та конфліктності даних.

Метою дослідження є побудова та дослідження гібридної математичної моделі нормованого оцінювання ризику кіберзагроз за мережевими ознаками, у якій результати нечіткої системи виведення та байєсівської мережі перетворюються у базові функції призначення маси й агрегуються засобами теорії доказів Демпстера–Шейфера з урахуванням конфлікту та надійності джерел.

Для досягнення поставленої мети й обґрунтування запропонованого підходу необхідно розв'язати наступні **завдання**:

1) проаналізувати сучасні стандарти, методології та математичні підходи до оцінювання ризиків інформаційної безпеки;

2) формалізувати простір гіпотез рівня ризику та множину високорівневих факторів, що характеризують кіберризик;

3) визначити процедуру відображення низькорівневих ознак мережевого трафіку у нормовані фактори ризику;

4) побудувати нечітку систему виведення та байєсівську мережу для отримання окремих оцінок ризику;

5) розробити процедуру перетворення виходів FIS і BN у базові функції призначення маси на спільному просторі гіпотез;

6) синтезувати алгоритм агрегації свідчень із контролем конфлікту між джерелами;

7) виконати програмну реалізацію запропонованого підходу та експериментально дослідити поведінку моделі на мережевих даних, зокрема порівняти результати FIS-, BN- та DST-компонентів, проаналізувати конфлікт між джерелами свідчень і вплив неповноти

факторної інформації на фінальну оцінку ризику.

Об'єктом дослідження є процес оцінювання ризиків кіберзагроз у мережевих інформаційних системах за умов невизначеності, неповноти та суперечливості даних.

Предметом дослідження є математичні моделі та методи інтеграції нечітких, імовірнісних і доказових оцінок ризику, зокрема процедури побудови факторів ризику, формування функцій мас довіри та агрегації свідчень.

У роботі використано такі **методи дослідження**: математичне моделювання, теорію нечітких множин і нечітке виведення [6, 7, 8], байєсівські мережі та ймовірнісні графові моделі [9, 10, 16], теорію доказів Демпстера–Шейфера [12, 13, 14], а також методи попередньої обробки, нормалізації та експериментального аналізу даних [17, 18, 19]. Програмна реалізація виконується мовою Python із використанням бібліотек для обробки даних, числових обчислень, попередньої обробки, оцінювання метрик та візуалізації; FIS-, BN- та DST-компоненти реалізовано програмно відповідно до математичних формул, наведених у роботі.

Наукова новизна роботи полягає в такому:

– запропоновано структуру гібридної математичної моделі нормованого оцінювання кіберризику за мережевими ознаками, у якій нечітка система виведення та байєсівська мережа розглядаються як різнорідні джерела свідчень про рівень ризику, а теорія доказів Демпстера–Шейфера використовується для їх формального узгодження;

– формалізовано відображення мережевих ознак у систему високорівневих факторів ризику, що охоплюють активність загрози, вразливість або експозицію, критичність активу та потенційний вплив інциденту;

– розроблено процедуру перетворення виходів нечіткої системи виведення та байєсівської мережі у базові функції призначення маси на спільному просторі ризикових гіпотез із урахуванням коефіцієнтів надійності джерел;

– запропоновано процедуру агрегації різнорідних свідчень із контролем конфлікту між джерелами та подальшим переходом до фінального нормованого індексу ризику.

Практичне значення роботи полягає в тому, що запропонована модель може бути використана як основа для обчислювального компонента систем підтримки прийняття рішень у сфері інформаційної безпеки. Розроблений підхід дає змогу узгоджено оцінювати кіберризик на основі мережевих ознак, експертних правил та ймовірнісних залежностей між факторами ризику. Окрему практичну цінність має можливість явного врахування залишкової невизначеності та конфлікту між джерелами свідчень, що є важливим для побудови аналітичних модулів SOC- та SIEM-систем.

Структура роботи. Дипломна робота складається зі вступу, трьох розділів, висновків, списку використаних джерел і додатків. У першому розділі розглянуто стандарти, методології та математичні підходи до оцінювання ризиків інформаційної безпеки. У другому розділі подано формальну постановку задачі та математичну модель гібридного оцінювання кіберризику. У третьому розділі наведено програмну реалізацію, опис експериментальних даних і результати дослідження поведінки моделі.

1 ОГЛЯД ЛІТЕРАТУРИ ТА ПОСТАНОВКА ДОСЛІДЖЕННЯ

У цьому розділі розглянуто сучасний стан проблеми оцінювання ризиків кіберзагроз, проаналізовано нормативні, методологічні та математичні підходи до управління ризиками інформаційної безпеки, а також обґрунтовано потребу в гібридній моделі, придатній для роботи з неповними, нечіткими та суперечливими даними.

Оцінювання кіберризику є міждисциплінарною задачею: з одного боку, воно спирається на стандарти й організаційні процедури ризик-менеджменту, а з іншого – потребує формальних методів аналізу даних, моделювання невизначеності та прийняття рішень. Реальні дані кібербезпеки часто є неповними, неоднорідними та отриманими з різних джерел, тому самих лише процесних підходів недостатньо для побудови стійкої обчислювальної оцінки ризику.

З огляду на це в розділі послідовно розглядаються стандарти та методології ризик-менеджменту, нечіткі системи виведення, байєсівські мережі та теорія доказів Демпстера–Шейфера. Нечіткі системи дають змогу формалізувати лінгвістичну нечіткість експертних правил, байєсівські мережі – описати ймовірнісні залежності між факторами ризику, а апарат теорії доказів – агрегувати частково узгоджені або суперечливі свідчення. Саме поєднання цих підходів є основою гібридної моделі, що розглядається в роботі.

1.1 Сучасні підходи та стандарти управління ризиками інформаційної безпеки

Управління ризиками інформаційної безпеки є систематичним процесом виявлення, аналізу, оцінювання та обробки ризиків, пов’язаних

із можливим порушенням конфіденційності, цілісності та доступності інформаційних активів. У загальному випадку кіберризик виникає внаслідок взаємодії кількох складових: наявності цінного активу, існування загрози, здатної завдати йому шкоди, наявності вразливості або умов для реалізації загрози, а також потенційних наслідків інциденту. Такі наслідки можуть мати технічний, фінансовий, операційний, репутаційний або правовий характер.

У нормативних документах ризик зазвичай трактується не лише як числовий показник, а як елемент процесу прийняття рішень. Міжнародні стандарти описують повний цикл роботи з ризиком: встановлення контексту, ідентифікацію активів, загроз і вразливостей, аналіз можливих наслідків, визначення прийнятності ризику, вибір заходів обробки та моніторинг залишкового ризику. Такий підхід є необхідним з управлінської точки зору, однак він не задає єдиного математичного механізму для обчислення ризику за умов неповноти, нечіткості та суперечливості даних.

Загальну методологічну основу управління ризиками задає стандарт ISO 31000:2018 [20], у якому ризик розглядається як вплив невизначеності на цілі організації. Для задач кібербезпеки це означає, що оцінка ризику не повинна розглядатися ізольовано від контексту активу та пов'язаних із ним бізнес-процесів: один і той самий технічний інцидент може мати різну значущість залежно від критичності системи, ролі сервісу та очікуваних наслідків.

Спеціалізованою основою для управління ризиками інформаційної безпеки є ISO/IEC 27005:2022 [1], узгоджений із вимогами системи управління інформаційною безпекою ISO/IEC 27001:2022 [21]. ISO/IEC 27005 визначає процес менеджменту ризиків інформаційної безпеки, зокрема встановлення контексту, оцінювання ризиків, їх обробку, прийняття залишкового ризику, комунікацію, моніторинг і перегляд. Важливо, що стандарт не нав'язує одного способу числового розрахунку ризику, а допускає якісні, кількісні та змішані методи залежно від мети

оцінювання, доступності даних і зрілості організації.

Методологія NIST SP 800-30 Rev. 1 [2] деталізує проведення оцінки ризиків для інформаційних систем і організацій. Вона явно розрізняє джерела загроз, події загроз, вразливості, умови реалізації загроз, імовірність небажаної події та масштаб потенційного впливу. З математичної точки зору це створює основу для подання ризику як функції кількох факторів:

$$R = f(T,V,A,I), \tag{1.1}$$

де T характеризує інтенсивність або релевантність загрози, V – рівень вразливості або експозиції, A – цінність чи критичність активу, а I – можливий вплив інциденту. Формула (1.1) не визначає конкретного алгоритму обчислення, але фіксує багатофакторну природу кіберризику.

Сучасні документи NIST розширюють це розуміння в напрямі інтеграції кіберризику з управлінням ризиками підприємства. Зокрема, NIST Cybersecurity Framework 2.0 [3] розглядає кібербезпеку як складову загального управління ризиками організації, а NISTIR 8286 Rev. 1 [22] описує зв'язок між кіберризиком та enterprise risk management. Для цієї роботи такі джерела важливі тим, що підкреслюють необхідність не лише технічно коректної, а й інтерпретованої оцінки, придатної для підтримки прийняття рішень.

Поряд зі стандартами використовуються прикладні методології ризик-аналізу. OCTAVE Allegro [23] орієнтована на ідентифікацію критичних інформаційних активів, аналіз загроз і формування сценаріїв ризику з урахуванням бізнес-контексту. Метод CRAMM, представлений в інвентарях ENISA [24], застосовувався для структурованого аналізу активів, загроз, вразливостей і вибору контрзаходів, а рекомендації ENISA щодо методів оцінювання ризиків [25] систематизують різні підходи до організації такого процесу. Сучасні аналітичні звіти ENISA щодо ландшафту загроз [26] додатково підкреслюють динамічність

середовища кіберзагроз, що ускладнює побудову сталих детермінованих моделей ризику.

Окреме місце займає методологія Open FAIR [4, 5], яка пропонує кількісну таксономію аналізу інформаційного ризику. На відміну від багатьох матричних підходів, FAIR розділяє частоту подій втрат і величину втрат, наближаючи оцінювання ризику до ймовірнісного моделювання. Водночас практичне застосування такого підходу потребує достатньої кількості статистичних даних або узгоджених експертних оцінок параметрів, що в кібербезпеці часто є складною умовою.

Для оцінювання технічної серйозності вразливостей широко використовується CVSS v4.0 [27]. Ця система стандартизує опис характеристик вразливості та формує числовий показник її критичності. Проте CVSS не є повною моделлю організаційного ризику: високий бал вразливості не завжди означає високий ризик, якщо актив не є критичним або відсутній реалістичний сценарій експлуатації. Навпаки, вразливість із помірним технічним балом може створювати значний ризик для критичного активу. Отже, оцінювання кіберризику потребує поєднання технічних характеристик, контексту активу, імовірності реалізації загрози та потенційного впливу.

У найпростішому випадку ризик часто подають як добуток імовірності інциденту та величини його наслідків:

$$R = P \cdot L,$$

де P – імовірність реалізації небажаної події, а L – очікувана величина втрат або впливу. Така формула є інтуїтивною, однак для складних інформаційних систем вона надто спрощена. Імовірність інциденту не завжди можна надійно оцінити за частотною інтерпретацією, оскільки релевантні події можуть бути рідкісними, а середовище загроз швидко змінюється. Крім того, наслідки інциденту не завжди зводяться до одного скалярного показника, а фактори ризику часто є взаємозалежними.

У практичних методиках також поширені матриці ризику, у яких імовірність та вплив поділяються на дискретні рівні, наприклад «низький», «середній» і «високий». Формально такий підхід можна подати як відображення

$$g : \mathcal{P} \times \mathcal{L} \to \mathcal{R},$$

де $\mathcal{P}$ – множина категорій імовірності, $\mathcal{L}$ – множина категорій впливу, а $\mathcal{R}$ – множина категорій ризику. Недоліком такого відображення є втрата інформації: різні за змістом комбінації факторів можуть переходити в одну й ту саму категорію ризику, хоча управлінські рішення щодо них мають бути різними. Крім того, межі між категоріями часто встановлюються експертно, що посилює суб'єктивність оцінювання.

З позиції прикладної математики традиційні підходи мають кілька обмежень. По-перше, якісні та напівкількісні шкали зручні для комунікації ризику, але не завжди придатні для побудови обчислювальних моделей, чутливих до змін у вхідних даних. По-друге, прості матричні схеми недостатньо враховують залежності між факторами: спостережувані ознаки атаки, рівень вразливості, тип активу та доступні контрзаходи можуть змінювати оцінку ризику спільно, а не ізольовано. По-третє, дані кібербезпеки містять невизначеність різної природи: випадковість подій, нечіткість експертних понять, неповноту та конфліктність свідчень. Модель, побудована лише на класичних імовірностях або лише на експертних правилах, не завжди здатна відобразити всю структуру такої невизначеності.

Таким чином, сучасні стандарти та методології управління ризиками інформаційної безпеки є необхідною концептуальною і термінологічною основою для постановки задачі, але не розв'язують повністю проблему формального математичного оцінювання ризику за умов неповноти, нечіткості та суперечливості даних. У межах цієї роботи вони задають загальний контекст, тоді як безпосереднє обчислення ризику потребує спеціалізованих математичних моделей. Тому подальший

аналіз зосереджується на інтелектуальних методах моделювання невизначеності, насамперед на нечітких системах виведення та байєсівських мережах.

1.2 Методи моделювання невизначеності при оцінці ризиків

Задача оцінювання кіберризику є складною не лише через велику кількість технічних факторів, а й через неоднорідність невизначеності, яка виникає під час аналізу. У класичній імовірнісній моделі невизначеність зазвичай пов'язують із випадковістю: подія може відбутися або не відбутися, а її ймовірність оцінюється на основі даних, статистичних припущень або експертних суджень.

У задачах кібербезпеки поряд із випадковістю наявні також нечіткість понять, неповнота інформації та конфліктність свідчень. Наприклад, твердження «актив має високу критичність» не є випадковою подією у вузькому сенсі, а описує лінгвістичну категорію з розмитими межами. Натомість твердження «зафіксований мережевий потік належить до атаки» має ймовірнісний характер і може уточнюватися при надходженні нових ознак.

У межах цієї роботи доцільно розрізняти три типи невизначеності, важливі для побудови моделі ризику:

1) *стохастичну невизначеність*, пов'язану з випадковістю реалізації загроз, непередбачуваністю поведінки атакувальника та варіативністю мережевого трафіку;

2) *лінгвістичну або семантичну нечіткість*, що виникає через використання експертних категорій без точних числових меж;

3) *епістемічну невизначеність*, зумовлену неповнотою, недостатньою надійністю або суперечливістю доступних даних.

Кожен із цих типів невизначеності потребує відповідного математичного інструментарію. Далі розглянуто два класи моделей, які часто застосовуються в системах підтримки прийняття рішень: нечіткі

системи виведення та байєсівські мережі. Їх аналіз є необхідним для подальшого обґрунтування гібридної моделі оцінювання кіберризику.

1.2.1 Нечіткі системи виведення

Теорія нечітких множин, запропонована Л. Заде [6], є математичним апаратом для формалізації понять із розмитими межами. На відміну від класичної множини, для якої характеристична функція набуває лише значень 0 або 1, нечітка множина $\tilde{A}$ на універсумі X задається функцією належності

$$\mu_{\tilde{A}} : X \to [0, 1],$$

де $\mu_{\tilde{A}}(x)$ інтерпретується як ступінь належності елемента x до нечіткого поняття $\tilde{A}$.

Наприклад, якщо $X = [0, 1]$ є нормованою шкалою критичності активу, то нечіткі множини «низька критичність», «середня критичність» і «висока критичність» можуть перекриватися. Завдяки цьому модель не створює штучного розриву між близькими значеннями лише через жорстко заданий поріг.

Для задач оцінювання ризику особливий інтерес становлять нечіткі системи виведення, які перетворюють набір вхідних факторів у результуючу оцінку на основі бази правил. Класична система типу Мамдані [7, 8] містить множину вхідних лінгвістичних змінних, функції належності для їхніх термів, базу нечітких правил, механізм нечіткого виведення, процедуру агрегування результатів правил і метод дефазифікації.

У контексті кіберризику вхідними змінними можуть бути, зокрема, рівень загрози T, рівень вразливості V, критичність активу A та потенційний вплив I. Кожна з цих змінних може бути описана набором

лінгвістичних термів:

$$T,V,A,I \in \{\text{низький, середній, високий}\}.$$

Типове нечітке правило тоді має вигляд:

$$\text{IF } T \text{ is High AND } V \text{ is High AND } A \text{ is High THEN } R \text{ is High,}$$

де R позначає рівень ризику. Такі правила є зрозумілими для аналітиків інформаційної безпеки, оскільки безпосередньо відтворюють природну логіку аналізу інцидентів.

Перевага нечітких систем полягає в тому, що вони переводять неформальні експертні судження в обчислювальну процедуру. Якщо значення вхідних змінних нормовано до проміжку $[0, 1]$, то функції належності задають, наскільки кожне значення відповідає певному лінгвістичному рівню. Після застосування бази правил система формує нечіткий висновок щодо ризику, який за потреби перетворюється у скалярну оцінку, наприклад методом центра ваги. Тому нечіткі моделі є доцільними в задачах, де суттєву роль відіграють експертні правила та якісні категорії [28, 29].

Водночас нечітка логіка не усуває всіх проблем оцінювання ризику. Форма функцій належності та параметри термів часто задаються експертно; якщо вони не обґрунтовані статистично або предметно, результат моделі може бути надмірно суб'єктивним. Крім того, кількість правил швидко зростає зі збільшенням кількості вхідних змінних і термів. Якщо є n вхідних факторів, кожен із яких має k лінгвістичних термів, то повна база правил може містити k^n комбінацій. Для чотирьох факторів і трьох термів це вже $3^4 = 81$ правило, а для більш деталізованої моделі кількість правил стає ще більшою.

Нарешті, нечітка система сама по собі не задає механізму ймовірнісного оновлення оцінки ризику при надходженні нових свідчень. Отже, нечіткі системи є доцільним інструментом для формалізації

лінгвістичних категорій і експертних правил, але їх недостатньо для повного опису стохастичної природи кіберінцидентів і причинно-наслідкових залежностей між подіями.

1.2.2 Байєсівські мережі

Байєсівські мережі є класом імовірнісних графових моделей, призначених для компактного подання спільного розподілу випадкових величин через умовні залежності [9, 10]. Формально байєсівська мережа задається парою

$$\mathcal{B} = (G,\boldsymbol{\theta}),$$

де $G = (\mathcal{V},\mathcal{E})$ – орієнтований ациклічний граф, вершини $\mathcal{V}$ якого відповідають випадковим змінним $X_1,\ldots,X_n$, а ребра $\mathcal{E}$ відображають безпосередні залежності між ними; $\boldsymbol{\theta}$ – множина параметрів, тобто умовних розподілів $P(X_i \mid Pa(X_i))$, де $Pa(X_i)$ позначає множину батьківських вершин змінної X_i.

Ключова властивість байєсівської мережі полягає у факторизації спільного розподілу:

$$P(X_1,\ldots,X_n) = \prod_{i=1}^{n} P(X_i \mid Pa(X_i)).$$

Ця формула має важливе прикладне значення. Замість повного задання спільного розподілу, розмірність якого зростає експоненційно з кількістю змінних, достатньо задати локальні умовні розподіли для кожної вершини графа. Якщо структура залежностей вибрана коректно, байєсівська мережа зменшує кількість параметрів і водночас зберігає інтерпретованість моделі.

У задачах кібербезпеки вершинами байєсівської мережі можуть бути фактори загрози, наявність або відсутність вразливості, спостережувані ознаки мережевої активності, стан захисних механізмів,

тип атаки та рівень ризику. Наприклад, підвищена інтенсивність мережевого трафіку, нетипові порти або аномальні TCP-прапорці можуть змінювати апостеріорну ймовірність певного сценарію атаки.

Якщо R – цільова змінна, що описує рівень ризику, а e – набір спостережених свідчень, то основною задачею є обчислення апостеріорного розподілу

$$P(R \mid e).$$

Саме ця властивість робить байєсівські мережі придатними для динамічного уточнення оцінки ризику при надходженні нової інформації.

На відміну від нечітких систем, байєсівські мережі не описують ступінь належності об'єкта до розмитої категорії, а задають імовірнісний розподіл над можливими станами змінних. Якщо нечітка логіка відповідає на питання «наскільки даний стан можна вважати високоризиковим», то байєсівська мережа відповідає на питання «яка ймовірність певного рівня ризику за наявних свідчень». Це розрізнення є принциповим: значення функції належності не є ймовірністю і не повинно безпосередньо ототожнюватися з нею.

Байєсівські мережі активно застосовуються для аналізу ризиків, зокрема для моделювання кібербезпеки, залежностей між вразливостями та оцінювання сценаріїв атак [16, 30, 11, 31]. Основними перевагами є:

– формальне врахування причинно-наслідкових або статистичних залежностей між факторами ризику в межах єдиної ймовірнісної моделі;

– оновлення оцінки ризику при появі нових свідчень за допомогою байєсівського виведення та умовних імовірностей;

– обробка неповних даних через маргіналізацію неспостережуваних змінних у байєсівському виведенні без їх явного відновлення;

– компактне подання багатовимірного розподілу за умови коректно заданої структури графа та таблиць умовних імовірностей.

Водночас застосування байєсівських мереж у кібербезпеці має низку обмежень. Побудова структури графа часто потребує поєднання

експертних знань і статистичного навчання. Якщо структура задається експертно, виникає ризик суб'єктивності; якщо вона навчається з даних, потрібна достатньо велика й репрезентативна вибірка, що не завжди доступно для рідкісних або нових типів атак.

Подібна проблема виникає і для таблиць умовних ймовірностей. Їх можна оцінювати за даними, але для цього потрібні якісні мітки та достатня кількість спостережень для кожної комбінації станів батьківських змінних. Якщо таких даних недостатньо, доводиться використовувати експертні оцінки, регуляризацію або спрощення структури мережі.

Ще одне обмеження полягає в тому, що класична байєсівська мережа працює з чітко визначеними станами змінних. Наприклад, якщо змінна «критичність активу» має стани {низька,середня,висока}, то для використання в мережі потрібно або дискретизувати числовий показник, або задати ймовірнісний розподіл над цими станами. Такий перехід може бути штучним, якщо сама змінна має нечітку семантичну природу. Саме тому в задачах кіберризику природно виникає потреба поєднати байєсівське та нечітке моделювання.

1.2.3 Порівняння нечітких систем і байєсівських мереж

Нечіткі системи виведення та байєсівські мережі не слід розглядати як взаємозамінні інструменти. Вони моделюють різні аспекти невизначеності та мають різну інтерпретацію вихідних величин. Нечітка система добре підходить для формалізації експертних правил і роботи з лінгвістичними категоріями, але не забезпечує природного механізму ймовірнісного оновлення. Байєсівська мережа, навпаки, надає строгий апарат для роботи з умовними ймовірностями, але потребує чіткої структури змінних, достатньої кількості даних або надійно заданих умовних розподілів.

З погляду прикладної математики важливо не змішувати ці підходи

на рівні інтерпретації. Значення $\mu_{\tilde{A}}(x) = 0.8$ означає високий ступінь належності x до нечіткого поняття $\tilde{A}$, але не означає, що ймовірність події $\tilde{A}$ дорівнює 0.8. Аналогічно значення $P(R = \text{High} \mid e) = 0.8$ є апостеріорною ймовірністю високого ризику за наявних свідчень, але не описує розмитість самого поняття «високий ризик». Некоректне ототожнення цих величин може призвести до помилкових висновків і математично необґрунтованої агрегації результатів.

Доцільність інтеграції нечітких систем і байєсівських мереж полягає в тому, що вони можуть виконувати різні функції в межах єдиного контуру оцінювання. Нечітка логіка може використовуватися для перетворення сирих або експертних показників у лінгвістично інтерпретовані рівні факторів ризику, тоді як байєсівська мережа може моделювати залежності між цими факторами та обчислювати апостеріорну оцінку ризику.

Однак навіть таке поєднання не повністю розв'язує проблему конфлікту між джерелами оцінок. Якщо нечітка система й байєсівська мережа дають різні висновки щодо рівня ризику, потрібен додатковий механізм формального агрегування свідчень.

Таким механізмом у цій роботі розглядається теорія доказів Демпстера–Шейфера. Вона дозволяє розподіляти масу довіри не лише між окремими гіпотезами, а й між їхніми підмножинами, тобто явно моделювати неповноту знань. На цьому ґрунтується подальша ідея гібридної моделі: нечітка система виведення та байєсівська мережа формують окремі оцінки ризику, після чого ці оцінки перетворюються у функції мас довіри та агрегуються в межах теорії доказів. У такій постановці результати різних моделей розглядаються не як взаємовиключні відповіді, а як часткові свідчення про стан ризику, що можуть бути узгодженими або конфліктними.

1.3 Теорія доказів для агрегації конфліктних свідчень

Нечіткі системи виведення та байєсівські мережі формують оцінки ризику з різних математичних позицій. Нечітка система описує лінгвістичну нечіткість експертних правил і дає змогу працювати з поняттями на кшталт «низький», «середній» або «високий» ризик. Байєсівська мережа, навпаки, задає імовірнісну модель залежностей між факторами та дозволяє отримувати апостеріорний розподіл рівня ризику за наявних свідчень. У практичній задачі ці два джерела не завжди дають узгоджені висновки. Наприклад, нечітка система може оцінювати ситуацію як високоризикову через велику критичність активу та значний потенційний вплив, тоді як байєсівська мережа, спираючись на статистично неповні або неоднозначні ознаки атаки, може надавати більшу підтримку середньому рівню ризику.

У такій ситуації просте усереднення результатів є недостатньо обґрунтованим, оскільки воно не враховує різну природу вихідних величин, ступінь довіри до джерел та можливий конфлікт між ними. Для формального поєднання таких свідчень доцільно використовувати апарат, який дозволяє явно представити не лише підтримку конкретних гіпотез, а й залишкову невизначеність. Одним із таких підходів є теорія доказів Демпстера–Шейфера, або Dempster–Shafer theory (DST), основи якої були закладені в роботах А. Демпстера [12] та Г. Шейфера [13].

На відміну від класичної ймовірнісної моделі, у DST міра довіри може призначатися не лише окремим елементарним гіпотезам, а й їхнім підмножинам. Нехай

$$\Theta = \{H_1, H_2, \ldots, H_q\}$$

є простором розрізнення, тобто скінченною множиною взаємовиключних і вичерпних гіпотез. Для задачі оцінювання ризику природним вибором є

$$\Theta = \{H_{\text{low}}, H_{\text{medium}}, H_{\text{high}}\},$$

де H_{low}, H_{medium} та H_{high} відповідають низькому, середньому та високому рівням ризику. Основним об'єктом DST є базова функція призначення маси, або basic probability assignment (BPA),

$$m : 2^{\Theta} \to [0,1],$$

яка задовольняє умови

$$m(\varnothing) = 0, \qquad \sum_{A \subseteq \Theta} m(A) = 1.$$

Значення $m(A)$ інтерпретується як маса довіри, безпосередньо призначена твердженню про те, що істинна гіпотеза належить множині A. Якщо A є одноелементною множиною, наприклад $\{H_{\text{high}}\}$, то така маса безпосередньо підтримує конкретний рівень ризику. Якщо ж маса призначена всій множині Θ, вона відображає не підтримку окремого рівня, а залишкову невизначеність або брак інформації. Саме це розрізнення є важливим для задач кібербезпеки: система може не лише видати категоричну оцінку, а й показати, що наявних даних недостатньо для впевненого вибору між кількома рівнями ризику.

На основі BPA можуть визначатися функції довіри та правдоподібності, які задають нижню та верхню межі підтримки гіпотези або множини гіпотез. У прикладному сенсі це дозволяє розглядати оцінку ризику не як єдине точкове значення, а як результат, що супроводжується певною мірою невизначеності. Для систем підтримки прийняття рішень така властивість є суттєвою: ситуація впевненого високого ризику та ситуація значної невизначеності щодо рівня ризику можуть вимагати різних управлінських дій.

Центральним механізмом DST є комбінування свідчень від кількох джерел. Нехай m_1 та m_2 – базові функції призначення маси, отримані з двох джерел, наприклад з нечіткої системи виведення та байєсівської мережі. Класичне правило Демпстера поєднує узгоджену частину свідчень і нормалізує результат, відкидаючи конфліктну масу. При цьому

коефіцієнт конфлікту

$$K = \sum_{B \cap C = \varnothing} m_1(B) m_2(C)$$

характеризує сумарну масу взаємовиключних тверджень, підтриманих різними джерелами. Якщо K є малим, джерела переважно узгоджуються. Якщо K наближається до одиниці, значна частина свідчень є суперечливою, і застосування класичного правила Демпстера може призводити до нестійких або контрінтуїтивних результатів. Ця проблема відома, зокрема, у зв'язку з парадоксом Заде [32, 14].

Для задач кібербезпеки проблема конфлікту є особливо важливою, оскільки джерела свідчень можуть мати різну природу: експертні правила, статистичні моделі, сигнатурні або аномалійні детектори, журнали подій, мережеві ознаки чи зовнішні індикатори компрометації. Такі джерела можуть бути частково залежними, мати різну якість і по-різному реагувати на неповні або зашумлені дані. Тому механічне комбінування оцінок без аналізу конфлікту може створити ілюзію високої впевненості там, де насправді потрібно фіксувати суперечливість інформації.

Одним зі способів пом'якшення цієї проблеми є правило Ягера [14]. Його ідея полягає в тому, що конфліктна маса не відкидається через нормалізацію, а переноситься на повну множину гіпотез Θ. Тобто конфлікт інтерпретується як додаткова невизначеність. Такий підхід є консервативнішим: за високої суперечливості джерел модель не формує надмірно категоричний висновок, а збільшує частку невизначеності. Для оцінювання кіберризику це є природною властивістю, оскільки суперечливі сигнали від різних компонентів системи не повинні автоматично перетворюватися на впевнену оцінку одного рівня ризику.

У літературі також розглядаються інші правила комбінування та модифікації DST, зокрема підходи, засновані на зважуванні джерел, дисконтуванні ненадійних свідчень і transferable belief model [33, 15]. Для

цієї роботи особливо важливим є механізм дисконтування. Якщо джерело має коефіцієнт надійності $\alpha \in [0,1]$, то зменшення довіри до нього не повинно автоматично підтримувати протилежну гіпотезу. Натомість відповідна частина маси переноситься в область невизначеності. Це узгоджується з природною інтерпретацією: якщо модель або дані є менш надійними, система має збільшувати рівень незнання, а не робити необґрунтовано впевнений висновок.

У гібридній системі оцінювання кіберризику DST не замінює нечітку логіку чи байєсівські мережі, а виконує роль формального рівня агрегації їхніх результатів. Нечітка система може формувати скалярну оцінку ризику на основі експертної бази правил, тоді як байєсівська мережа формує апостеріорний розподіл над рівнями ризику. Після цього обидва результати мають бути приведені до спільного простору гіпотез Θ у вигляді базових функцій призначення маси.

Важливо, що виходи FIS і BN не можна безпосередньо ототожнювати з масами довіри без додаткового перетворення. Скалярна оцінка ризику $R \in [0,1]$, ступінь належності $\mu(R)$, апостеріорна ймовірність $P(H_i \mid e)$ і маса довіри $m(A)$ мають різну математичну природу. Тому коректність гібридної моделі залежить від явно заданої процедури переходу від виходів окремих моделей до BPA. У подальшому розділі цей перехід буде формалізовано окремо для нечіткої системи виведення та байєсівської мережі.

Отже, теорія доказів Демпстера–Шейфера є придатним апаратом для агрегації результатів різнорідних моделей оцінювання ризику, оскільки дозволяє явно враховувати залишкову невизначеність, вимірювати конфлікт між джерелами свідчень, використовувати коефіцієнти надійності та уникати надмірно категоричних висновків у ситуаціях суперечливих даних. Саме тому в цій роботі DST використовується як третій компонент гібридної моделі, що узгоджує результати нечіткої системи виведення та байєсівської мережі в межах спільного простору ризикових гіпотез.

1.4 Прогалини досліджень та постановка мети дослідження

Проведений аналіз стандартів, методологій та математичних підходів до оцінювання ризиків інформаційної безпеки показує, що наявні підходи розв'язують лише окремі частини загальної задачі. Стандарти та практичні рамки ризик-менеджменту, зокрема ISO/IEC 27005:2022 [1], NIST SP 800-30 Rev. 1 [2], NIST Cybersecurity Framework 2.0 [3] та Open FAIR [4, 5], задають зрілу процесну основу для ідентифікації, аналізу й обробки ризиків. Водночас вони не визначають єдиної математичної процедури, яка б дозволяла формально поєднувати експертні правила, ймовірнісні залежності, неповноту інформації та конфліктні свідчення.

Нечіткі моделі є ефективними для формалізації якісних понять і експертних правил [28, 29], однак залежать від вибору функцій належності, бази правил і процедури дефазифікації. Байєсівські мережі забезпечують строгий імовірнісний апарат для моделювання залежностей між факторами та оновлення оцінок за наявних свідчень [16, 30, 11], проте потребують визначення структури графа й оцінювання таблиць умовних імовірностей. Теорія доказів Демпстера–Шейфера дозволяє агрегувати свідчення та враховувати конфлікт між джерелами [13, 14, 15], але її застосування потребує коректного переходу від виходів окремих моделей до базових функцій мас довіри.

Окремі сучасні дослідження демонструють перспективність поєднання нечітких моделей, байєсівського виведення та теорії доказів у задачах ризик-аналізу [34, 35, 36]. Проте для оцінювання кіберризику за мережевими ознаками залишається недостатньо формалізованим цілісний обчислювальний контур, який охоплює всі ключові етапи у межах єдиної моделі: перехід від технічних ознак трафіку до факторів ризику, отримання окремих оцінок за допомогою FIS і BN, перетворення цих оцінок у BPA та подальше комбінування з урахуванням конфлікту й надійності джерел. Саме цей розрив між наявністю окремих

математичних інструментів і потребою в узгодженій інтеграційній процедурі визначає дослідницьку проблему цієї роботи.

На основі проведеного огляду виділено ключові прогалини, що визначають потребу в гібридній моделі оцінювання кіберризику.

1. Семантичний розрив між мережевими ознаками та факторами ризику. Дані мережевого моніторингу містять низькорівневі технічні характеристики: тривалість потоку, кількість пакетів, обсяг переданих байтів, часові параметри, TCP-ознаки та інші характеристики з'єднання. Натомість моделі ризик-менеджменту оперують високорівневими поняттями: загроза, вразливість, критичність активу та потенційний вплив. Тому необхідно явно задати відображення

$$\phi : D \to X, \qquad X \subseteq [0,1]^d,$$

яке пов'язує простір технічних ознак D із простором нормованих факторів ризику X. Без такого відображення практична частина роботи не має строгого зв'язку з теоретичною моделлю ризику.

2. Різна математична природа виходів FIS і BN. Нечітка система працює зі ступенями належності до лінгвістичних термів, тоді як байєсівська мережа формує апостеріорні ймовірності станів ризику. Ці величини не можна безпосередньо ототожнювати або механічно усереднювати. Отже, потрібен спільний формалізм, який дозволяє узгодити результати нечіткої та ймовірнісної компонент без змішування їхньої інтерпретації.

3. Необхідність формалізованого переходу до функцій мас довіри. Для використання теорії доказів результати FIS і BN мають бути подані у вигляді базових функцій призначення маси на спільному просторі гіпотез Θ. Цей перехід не є автоматичним: скалярна оцінка ризику, ступінь належності, апостеріорна ймовірність і маса довіри мають різний зміст. Тому потрібно явно визначити процедури, за якими вихід FIS перетворюється у маси через належність до термів ризику, а вихід

BN – через апостеріорний розподіл на множині ризикових гіпотез.

4. Проблема конфлікту між джерелами свідчень. У гібридній системі різні компоненти можуть підтримувати різні рівні ризику. Такий конфлікт може виникати через неповноту даних, обмеженість статистичної вибірки, різну чутливість моделей до факторів або неточність експертних правил. Класичне правило Демпстера є ефективним за помірного конфлікту, але за високих значень коефіцієнта K може давати нестійкі або контрінтуїтивні результати [32, 14]. Тому модель має містити механізм контролю конфлікту та можливість консервативного комбінування, за якого суперечливість свідчень інтерпретується як додаткова невизначеність.

5. Потреба врахування надійності джерел. Якість окремих джерел свідчень може бути різною. Експертні правила можуть краще описувати відомі сценарії, але слабше реагувати на нетипові дані; статистична модель може бути інформативною для добре представлених класів, але нестійкою для рідкісних комбінацій факторів. Тому гібридна модель не повинна фіксовано надавати перевагу одному джерелу. У межах DST це можна врахувати за допомогою коефіцієнтів надійності α_{FIS} та α_{BN}, які визначають, яка частина маси призначається конкретним гіпотезам, а яка переноситься в область залишкової невизначеності.

Зазначені прогалини визначають постановку подальшого дослідження. У роботі розглядається задача побудови гібридної математичної моделі оцінювання ризику кіберзагроз, у якій нечітка система виведення та байєсівська мережа виступають як два різнорідні джерела свідчень про рівень ризику, а теорія доказів Демпстера–Шейфера використовується для їх формального узгодження та агрегації.

Метою роботи є побудова та дослідження гібридної математичної моделі нормованого оцінювання ризику кіберзагроз за мережевими ознаками, у якій результати нечіткої системи виведення та байєсівської мережі перетворюються у базові функції призначення маси й агрегуються

засобами теорії доказів Демпстера–Шейфера з урахуванням конфлікту та надійності джерел.

Досягнення цієї мети передбачає послідовне розв'язання кількох пов'язаних підзадач: формалізацію простору ризикових гіпотез і факторів ризику, побудову відображення від мережевих ознак до нормованих факторів, отримання окремих оцінок за допомогою FIS і BN, перетворення цих оцінок у функції мас довіри та їх подальшу агрегацію з урахуванням конфлікту між джерелами.

Очікуваний прикладний результат роботи полягає не в заміні чинних стандартів ризик-менеджменту, а в доповненні їх формальним обчислювальним механізмом для оцінювання кіберризику за умов неповноти та конфліктності даних. Запропонована модель має забезпечити узгоджене використання експертних правил, імовірнісних залежностей і механізмів обробки конфліктних свідчень, зберігаючи інтерпретованість результатів для аналітика інформаційної безпеки.

Таким чином, наступний розділ присвячено формальній постановці задачі та побудові математичної моделі. У ньому буде визначено простір факторів ризику, описано нечітку систему виведення, структуру байєсівської мережі, процедуру формування BPA та алгоритм комбінування свідчень у межах теорії доказів.

Висновки до розділу 1

У першому розділі проаналізовано основні стандарти, методології та математичні підходи до оцінювання ризиків інформаційної безпеки. Показано, що сучасні рамки ризик-менеджменту, зокрема ISO/IEC 27005:2022 [1], NIST SP 800-30 Rev. 1 [2] та Open FAIR [4, 5], задають необхідну процесну й термінологічну основу, однак не визначають єдиного формального механізму для обчислювального оцінювання ризику за умов неповноти, нечіткості та конфліктності даних.

Встановлено, що нечіткі системи виведення, байєсівські мережі та

теорія доказів Демпстера–Шейфера описують різні, але взаємодоповнювальні аспекти невизначеності. Нечітка логіка є придатною для формалізації експертних правил і лінгвістичних категорій, байєсівські мережі забезпечують імовірнісне моделювання залежностей між факторами ризику, а DST дає змогу агрегувати різнорідні свідчення з урахуванням залишкової невизначеності та конфлікту між джерелами.

Таким чином, проведений огляд обґрунтовує доцільність побудови гібридної моделі, у якій нечітка система виведення та байєсівська мережа формують окремі оцінки ризику, а теорія доказів використовується для їх формального узгодження. Це визначає зміст наступного розділу, присвяченого математичній постановці задачі, формалізації факторів ризику, побудові FIS- і BN-компонентів та процедурі формування й агрегації функцій мас довіри.

2 МАТЕМАТИЧНА ФОРМАЛІЗАЦІЯ ГІБРИДНОЇ МОДЕЛІ ОЦІНЮВАННЯ РИЗИКІВ

У цьому розділі побудовано формальну математичну модель гібридного оцінювання ризику кіберзагроз. Основна увага приділяється послідовному переходу від мережевого спостереження до фінального нормованого індексу ризику: спочатку низькорівневі ознаки трафіку перетворюються у високорівневі фактори ризику, далі ці фактори використовуються нечіткою системою виведення та байєсівською мережею, після чого результати обох компонент агрегуються засобами теорії доказів Демпстера–Шейфера.

Розділ має не лише описовий, а й конструктивний характер: у ньому визначаються вхідні та вихідні величини моделі, простір ризикових гіпотез, відображення ознак у фактори ризику, математичні форми FIS- і BN-компонентів, правила побудови базових функцій призначення маси та процедура отримання фінального числового показника. Конкретні числові параметри цієї схеми, пов'язані з вибором ознак, порогів, ваг, функцій належності, таблиць умовних ймовірностей і коефіцієнтів надійності, задаються в наступному розділі під час опису програмної реалізації та експериментального сценарію.

2.1 Математична постановка задачі оцінювання ризику

У цьому підрозділі формалізується загальна задача, яку розв'язує запропонована гібридна модель. Надалі окремо розглядаються її основні етапи: побудова відображення від мережевих ознак до факторів ризику, нечітке виведення, байєсівське моделювання та агрегація свідчень у межах теорії доказів.

У стандартах і методологіях управління ризиками інформаційної

безпеки ризик пов'язується з невизначеністю щодо реалізації небажаної події та її можливих наслідків для інформаційних активів [1, 2, 5]. У цій роботі під оцінкою кіберризику розуміється не ймовірність атаки сама по собі й не очікувана грошова величина втрат, а нормований індекс ризику

$$R \in [0, 1],$$

який узагальнює інформацію про активність загрози, вразливість або експозицію цільового середовища, критичність активу та потенційний вплив інциденту. Значення, близькі до нуля, відповідають низькому рівню ризику, а значення, близькі до одиниці, – високому рівню ризику.

Безпосередньою одиницею обчислення в моделі є мережеве спостереження

$$z = (z_1, z_2, \ldots, z_p) \in \mathcal{D},$$

де $\mathcal{D}$ – простір вхідних мережевих даних, а $z_1, \ldots, z_p$ – технічні ознаки мережевого потоку або агрегованого фрагмента трафіку. Якщо в експериментальній частині використовується сценарна інтерпретація, то сценарій $s \in \mathcal{S}$ розглядається як контекст, що може відповідати типу атаки, цільовому активу або умовам реалізації загрози. У такому разі може бути введене відображення $g : \mathcal{D} \to \mathcal{S}$, однак у математичній постановці базовим аргументом оцінювання залишається саме спостереження z.

Щоб перейти від низькорівневих технічних ознак до змінних, придатних для ризик-орієнтованого моделювання, вводиться відображення

$$\phi : \mathcal{D} \to \mathcal{X}, \qquad \mathcal{X} \subseteq [0, 1]^4,$$

таке, що

$$x = \phi(z) = \big(T(z), V(z), A(z), I(z)\big).$$

Компоненти вектора x мають такий зміст:

– $T(z) \in [0, 1]$ – нормований показник активності або релевантності

загрози, що відображає інтенсивність, аномальність або характер мережевої поведінки, пов'язаної з можливим інцидентом;

– $V(z) \in [0,1]$ – нормований показник потенційної вразливості або експозиції цільового середовища. Якщо відсутні прямі дані про CVE, конфігурації, патчі чи результати сканування, ця величина трактується як проксі-показник, побудований на основі доступних технічних ознак;

– $A(z) \in [0,1]$ – критичність або цінність активу, якого стосується спостереження. На відміну від мережевих ознак, ця компонента може задаватися контекстно або експертно, оскільки датасети мережевого трафіку зазвичай не містять безпосередньої інформації про бізнес-цінність активів;

– $I(z) \in [0,1]$ – очікуваний вплив конкретного інциденту або класу інцидентів на конфіденційність, цілісність чи доступність інформаційного ресурсу.

Отже, A описує важливість самого активу, тоді як I характеризує можливі наслідки реалізації конкретної загрози щодо цього активу. Таке розмежування потрібне для того, щоб уникнути дублювання факторів у моделі.

Вектор $x = \phi(z)$ є спільним входом для двох різних за природою компонентів оцінювання. Нечітка система виведення формує числову оцінку

$$R_{\mathrm{FIS}}(z) = f_{\mathrm{FIS}}(x), \qquad R_{\mathrm{FIS}}(z) \in [0,1],$$

використовуючи лінгвістичні терми та експертні правила [6, 7, 28]. Байєсівська мережа, зі свого боку, задає апостеріорний розподіл над дискретними рівнями ризику за наявних свідчень:

$$P_{\mathrm{BN}}(H \mid e(z)), \qquad H \in \Theta,$$

де $e(z)$ – набір спостережень, отриманий на основі вектора факторів x, а

$$\Theta = \{H_{\mathrm{low}}, H_{\mathrm{med}}, H_{\mathrm{high}}\}$$

є спільним простором гіпотез про рівень ризику. Такий імовірнісний компонент дає змогу враховувати залежності між факторами ризику та оновлювати оцінку за наявності нових свідчень [9, 10, 16, 11].

Оскільки вихід нечіткої системи та вихід байєсівської мережі мають різну математичну природу, їх не слід поєднувати прямим усередненням. Для узгодження обидва результати надалі перетворюються у базові функції призначення маси на спільному просторі гіпотез Θ. Після цього виконується агрегація свідчень із контролем конфлікту:

$$m(z) = \mathcal{C}\big(m_{\mathrm{FIS}}(z), m_{\mathrm{BN}}(z)\big),$$

де m_{FIS} та m_{BN} – функції мас, отримані відповідно з нечіткої системи та байєсівської мережі, а $\mathcal{C}$ – правило комбінування свідчень у межах теорії доказів Демпстера–Шейфера [13, 15]. За значного конфлікту між джерелами доцільно використовувати правило Ягера або іншу консервативну схему, яка переносить конфліктну масу в область невизначеності [14].

Загальну структуру задачі можна подати як послідовність відображень

$$\begin{aligned} z &\xrightarrow{\phi} x = (T, V, A, I) \\ &\xrightarrow{f_{\mathrm{FIS}},\, f_{\mathrm{BN}}} \big(R_{\mathrm{FIS}},\, P_{\mathrm{BN}}(H \mid e)\big) \\ &\xrightarrow{\psi_{\mathrm{FIS}},\, \psi_{\mathrm{BN}}} \big(m_{\mathrm{FIS}}, m_{\mathrm{BN}}\big) \\ &\xrightarrow{\mathcal{C}} m \xrightarrow{\mathcal{U}} R. \end{aligned}$$

Тут ψ_{FIS} та ψ_{BN} позначають процедури переходу від результатів окремих моделей до функцій мас, а $\mathcal{U}$ – процедуру отримання фінального нормованого індексу ризику з агрегованої функції мас. Детальне означення цих етапів подано в наступних підрозділах.

Для наочності загальну структуру запропонованої гібридної моделі подано на рисунку 2.1. Схема відображає логіку математичного переходу від вхідного мережевого спостереження до фінального нормованого індексу ризику.

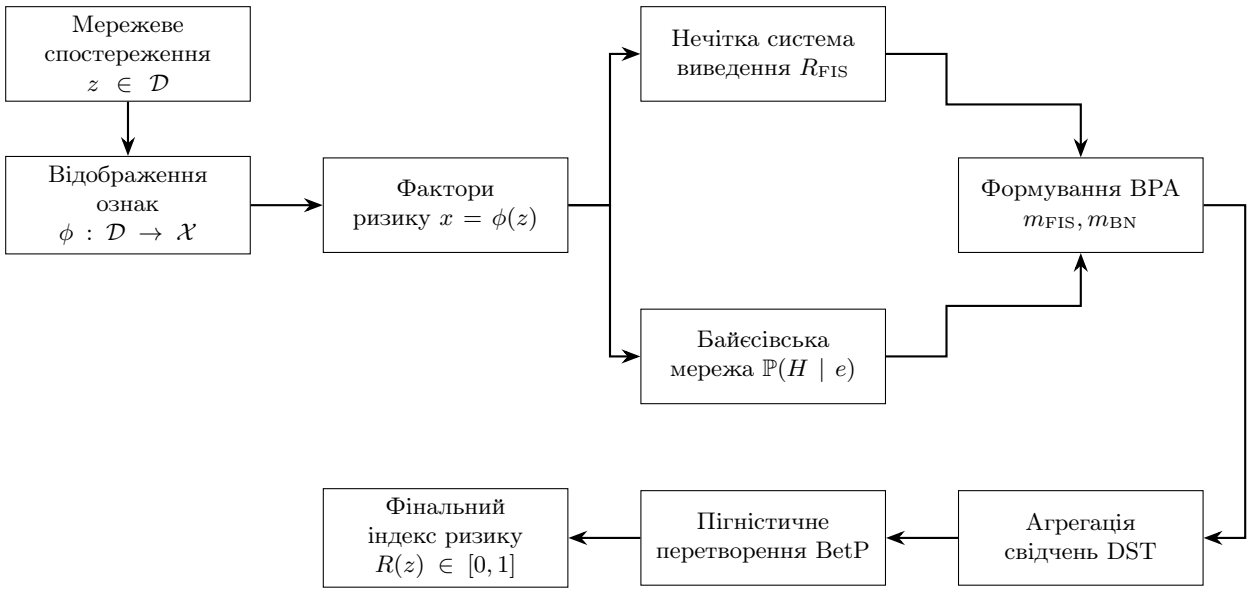

Рисунок 2.1 – Загальна схема гібридного оцінювання кіберризику

З урахуванням наведеної схеми математична задача цієї роботи полягає в побудові гібридного оператора

$$\mathcal{R} : \mathcal{D} \to [0, 1], \qquad \mathcal{R}(z) = R(z),$$

який для кожного мережевого спостереження формує узгоджену оцінку кіберризику на основі нормованих факторів ризику, нечіткого виведення, байєсівського ймовірнісного моделювання та доказової агрегації свідчень. Коректність і прикладна придатність такого оператора надалі оцінюються не як абстрактна мінімізація похибки в постановці задачі, а через експериментальне дослідження його поведінки: стійкість до шуму та неповноти даних, рівень конфлікту між джерелами, частку залишкової невизначеності та узгодженість фінальної оцінки з обраною експериментальною інтерпретацією ризику.

2.2 Формалізація вхідних даних та функція відображення ϕ

У попередньому підрозділі введено загальне відображення $\phi : \mathcal{D} \to \mathcal{X}$, яке переводить мережеве спостереження у вектор нормованих

факторів ризику. Далі уточнюється структура цього відображення та визначається, які типи даних можуть бути використані для побудови компонент

$$x = \phi(z) = (T(z),V(z),A(z),I(z)).$$

Цей етап потрібен насамперед для подолання семантичного розриву між низькорівневими технічними ознаками мережевого трафіку та високорівневими змінними, що використовуються в моделях оцінювання кіберризику.

Нехай

$$z = (z_1,z_2,\ldots,z_p) \in \mathcal{D}$$

є вектором ознак мережевого потоку або агрегованого фрагмента трафіку. У наборах даних на кшталт CIC-IDS2017 такі ознаки формуються на основі характеристик мережевих потоків: тривалості з'єднання, кількості пакетів у прямому та зворотному напрямах, обсягу переданих байтів, статистик довжин пакетів, швидкості передавання, міжпакетних часових інтервалів та окремих параметрів TCP-з'єднань [18, 19, 37]. Частина цих ознак безпосередньо отримується інструментами аналізу потоків, зокрема CICFlowMeter [38].

Перед формуванням факторів ризику вхідні ознаки потрібно привести до уніфікованого числового вигляду. Нехай $\mathcal{D}_{train} \subset \mathcal{D}$ позначає навчальну частину вибірки, на якій визначаються параметри попередньої обробки. Для числової ознаки z_j задаються параметри масштабування a_j та b_j, обчислені лише за $\mathcal{D}_{train}$. У найпростішому випадку це мінімальне та максимальне значення ознаки, а за наявності викидів – відповідні нижній і верхній перцентилі [17]. Нормоване значення ознаки визначається як

$$\widetilde{z}_j = \begin{cases} \min\left\{1, \max\left\{0, \dfrac{z_j - a_j}{b_j - a_j}\right\}\right\}, & b_j > a_j, \\ 0, & b_j = a_j. \end{cases}$$

Використання параметрів, обчислених лише на навчальній вибірці, дає змогу уникнути витоку інформації з тестових даних. Обрізання до інтервалу $[0, 1]$ дозволяє коректно обробляти значення, що виходять за межі діапазону, зафіксованого на навчальній вибірці.

Після нормалізації для кожного фактора ризику задається підмножина релевантних ознак. Позначимо через

$$\mathcal{I}_T, \quad \mathcal{I}_V, \quad \mathcal{I}_A, \quad \mathcal{I}_I$$

множини індексів ознак, що використовуються відповідно для побудови показників загрози, вразливості або експозиції, критичності активу та потенційного впливу. У загальному вигляді кожний фактор $Q \in \{T,V,A,I\}$ можна подати як зважену нормовану агрегацію:

$$Q(z) = \sum_{j \in \mathcal{I}_Q} w_{Qj}\, \zeta_{Qj}(z), \qquad w_{Qj} \geqslant 0, \qquad \sum_{j \in \mathcal{I}_Q} w_{Qj} = 1,$$

де $\zeta_{Qj}(z)$ є нормованим внеском j-ї ознаки у відповідний фактор. Якщо зростання ознаки підвищує значення фактора, то $\zeta_{Qj}(z) = \widetilde{z}_j$; якщо ж зростання ознаки знижує відповідний ризиковий показник, можна використовувати обернене перетворення $\zeta_{Qj}(z) = 1 - \widetilde{z}_j$. Ваги w_{Qj} можуть задаватися експертно або визначатися за результатами попереднього аналізу даних; у будь-якому разі їх значення мають бути зафіксовані перед проведенням експерименту.

У цій постановці важливо розрізняти фактори, які можуть бути наближено побудовані з мережевих ознак, і фактори, що потребують зовнішнього контексту. Показник $T(z)$ може спиратися на ознаки інтенсивності та аномальності трафіку. Показник $V(z)$ у разі відсутності даних про CVE, конфігурації, патчі або результати сканування не слід трактувати як фактичну вразливість системи; у такому випадку він є проксі-показником потенційної експозиції сервісу. Компонента $A(z)$ характеризує критичність активу і зазвичай не виводиться безпосередньо з датасету мережевого трафіку, а задається на основі контекстної

інформації про роль вузла або сегмента мережі. Компонента $I(z)$ описує очікуваний вплив конкретного інциденту і може визначатися через поєднання технічних ознак трафіку та контекстної оцінки наслідків для конфіденційності, цілісності або доступності.

Відповідність між групами ознак і факторами ризику не є єдино можливою специфікацією моделі, проте вона задає узгоджену основу для подальшої реалізації відображення ϕ. У межах роботи використовуються такі групи ознак.

Загроза (T). До цього фактора відносяться ознаки *Flow Duration*, *Total Fwd Packets*, *Total Backward Packets*, *Flow Bytes/s*, *Flow Packets/s* та статистики міжпакетних інтервалів. У моделі T інтерпретується як нормований показник активності, інтенсивності або аномальності мережевої поведінки, пов'язаної з можливою атакою.

Вразливість / експозиція (V). Для цього фактора можуть використовуватися *Destination Port*, *Protocol*, TCP flags, ознаки доступності сервісу або зовнішні дані про вразливості, якщо вони наявні. У моделі V розглядається як проксі-показник потенційної експозиції цільового сервісу; за наявності даних про CVE або конфігурації він може уточнюватися як показник фактичної вразливості.

Критичність активу (A). Цей фактор може задаватися на основі таких контекстних ознак, як мережевий сегмент, роль вузла, тип сервісу, належність до DMZ, бази даних, користувацького або службового сегмента. У моделі A інтерпретується як контекстна або експертно задана оцінка важливості активу для функціонування системи.

Вплив (I). Для оцінювання цього фактора можуть використовуватися *Bwd Packet Length Mean*, *Total Length of Fwd/Bwd Packets*, *Flow Bytes/s*, тип потенційного інциденту та контекстні оцінки наслідків. У моделі I є нормованим показником можливого впливу інциденту на конфіденційність, цілісність або доступність інформаційного ресурсу.

Таким чином, функція відображення

$$\phi(z) = \big(T(z), V(z), A(z), I(z)\big) \in [0,1]^4$$

формує неперервний нормований вектор факторів ризику, який надалі використовується як спільний вхід для нечіткої системи виведення та байєсівської мережі.

Окремо від ϕ вводиться процедура дискретизації, необхідна для тих компонентів моделі, які працюють із категоріальними станами. Для кожного фактора $Q \in \{T,V,A,I\}$ задаються пороги

$$0 \leqslant \tau_{Q,1} < \tau_{Q,2} \leqslant 1,$$

після чого визначається відображення

$$q_Q(Q) = \begin{cases} Low, & 0 \leqslant Q \leqslant \tau_{Q,1}, \\ Medium, & \tau_{Q,1} < Q \leqslant \tau_{Q,2}, \\ High, & \tau_{Q,2} < Q \leqslant 1. \end{cases}$$

Пороги можуть задаватися експертно або оцінюватися за навчальною вибіркою, наприклад за 33-м і 66-м перцентилями розподілу відповідного фактора [17]. Таке дискретне подання буде використане для задання станів байєсівської мережі, тоді як неперервні значення факторів залишаються придатними для фаззифікації у нечіткій системі.

Отже, у цьому підрозділі формалізовано перехід від мережевого спостереження z до вектора факторів ризику $x = \phi(z)$ та, за потреби, до його дискретного подання. Саме ця конструкція забезпечує узгодженість шкал між подальшими компонентами гібридної моделі й дозволяє пов'язати експериментальні мережеві дані з математичною постановкою задачі оцінювання кіберризику.

2.3 Математична модель нечіткої системи виведення

Нечітка система виведення використовується в роботі для формалізації експертної логіки оцінювання ризику за умов лінгвістичної невизначеності. На відміну від байєсівської мережі, яка розглядається в наступному підрозділі, нечітка модель не задає ймовірнісний розподіл над станами ризику, а перетворює нормований вектор факторів

$$x = (T,V,A,I) \in [0,1]^4$$

у скалярну оцінку ризику

$$R_{\text{FIS}} = f_{\text{FIS}}(x), \qquad R_{\text{FIS}} \in [0,1].$$

Ця оцінка відображає рівень ризику з позиції експертних правил і надалі використовується як один із виходів гібридної моделі. Побудова нечіткої системи ґрунтується на підході Мамдані [7], зручному для задач, у яких важливо зберегти інтерпретованість правил виду «якщо–то» [8]. Застосування нечітких систем у задачах оцінювання кіберризиків також пов'язане з їхньою здатністю працювати з якісними експертними категоріями та неповними даними [28, 29].

Нехай множина вхідних змінних позначається через

$$\mathcal{Q} = \{T,V,A,I\}.$$

Для кожної змінної $Q \in \mathcal{Q}$ задається множина лінгвістичних термів

$$\mathcal{L}_Q = \{Low,Medium,High\},$$

а для вихідної змінної ризику – множина термів

$$\mathcal{L}_R = \{Low,Medium,High\}.$$

Кожному терму відповідає нечітка множина на універсумі $[0,1]$ з функцією належності

$$\mu_{Q,l} : [0,1] \to [0,1], \qquad Q \in \mathcal{Q}, \quad l \in \mathcal{L}_Q.$$

Аналогічно для вихідної змінної ризику задаються функції належності

$$\mu_{R,l} : [0,1] \to [0,1], \qquad l \in \mathcal{L}_R.$$

Значення $\mu_{Q,l}(q)$ показує ступінь належності числового значення q до лінгвістичного терма l, але не є ймовірністю відповідного стану.

У цій моделі доцільно використовувати трикутні та трапецієподібні функції належності, оскільки вони мають просту параметризацію і добре інтерпретуються в прикладних задачах [6, 8]. Трикутна функція належності з параметрами $a \leqslant b \leqslant c$ може бути задана як

$$\mathrm{tri}(q; a,b,c) = \begin{cases} 0, & q \leqslant a, \\ \dfrac{q-a}{b-a}, & a < q \leqslant b, \\ \dfrac{c-q}{c-b}, & b < q < c, \\ 0, & q \geqslant c, \end{cases}$$

а трапецієподібна функція з параметрами $a \leqslant b \leqslant c \leqslant d$ – як

$$\mathrm{trap}(q; a,b,c,d) = \begin{cases} 0, & q \leqslant a, \\ \dfrac{q-a}{b-a}, & a < q \leqslant b, \\ 1, & b < q \leqslant c, \\ \dfrac{d-q}{d-c}, & c < q < d, \\ 0, & q \geqslant d. \end{cases}$$

Параметри цих функцій можуть задаватися експертно або визначатися за навчальною вибіркою, наприклад на основі перцентильних порогів, введених на етапі формалізації факторів ризику [17]. Функції належності сусідніх термів мають перекриватися: це забезпечує плавний перехід між

рівнями ризику та запобігає ситуації, коли для деякого вхідного значення не активується жодне правило.

Фазифікація вхідного вектора полягає в обчисленні ступенів належності кожної компоненти x до відповідних термів:

$$\alpha_{Q,l} = \mu_{Q,l}(Q), \qquad Q \in \mathcal{Q}, \quad l \in \mathcal{L}_Q.$$

Отримані значення далі використовуються для активації бази нечітких правил.

База правил задається скінченною множиною

$$\mathcal{B}_{\text{FIS}} = \{l_1, l_2, \ldots, l_M\}.$$

Кожне правило l_k має вигляд

$$l_k : \quad \text{IF } T \text{ is } A_T^{(k)} \wedge V \text{ is } A_V^{(k)} \wedge A \text{ is } A_A^{(k)} \wedge I \text{ is } A_I^{(k)} \text{ THEN } R \text{ is } B^{(k)},$$

де

$$A_T^{(k)} \in \mathcal{L}_T, \quad A_V^{(k)} \in \mathcal{L}_V, \quad A_A^{(k)} \in \mathcal{L}_A, \quad A_I^{(k)} \in \mathcal{L}_I, \quad B^{(k)} \in \mathcal{L}_R.$$

Такий запис дозволяє уникнути некоректного використання одного індексу для термів різних змінних. Якщо для кожного з чотирьох факторів використовується по три терми, повна база правил може містити $3^4 = 81$ правило.

На практиці база може бути скороченою, якщо частина комбінацій є малоймовірною, неінформативною або не має змістовної інтерпретації в контексті кіберризику. Водночас принцип формування правил має залишатися однаковим: зростання активності загрози, експозиції, критичності активу та потенційного впливу не повинно зменшувати результуючу оцінку ризику без явного предметного обґрунтування.

Типовими прикладами правил є:

$$l_1 : \text{IF } T \text{ is } High \wedge V \text{ is } High \wedge A \text{ is } High \wedge I \text{ is } High \text{ THEN } R \text{ is } High,$$

$$l_2 : \text{IF } T \text{ is } Low \wedge V \text{ is } Low \wedge A \text{ is } Low \wedge I \text{ is } Low \text{ THEN } R \text{ is } Low.$$

Ці правила не вичерпують базу знань, а лише ілюструють її загальну логіку.

Нехай у правилі l_k вхідним змінним (T,V,A,I) відповідають лінгвістичні терми ($A_T^{(k)}$,$A_V^{(k)}$,$A_A^{(k)}$,$A_I^{(k)}$). Тоді за Т-норми мінімуму ступінь активації правила дорівнює

$$w_k = \min\left\{\mu_{A_T^{(k)}}(T), \mu_{A_V^{(k)}}(V), \mu_{A_A^{(k)}}(A), \mu_{A_I^{(k)}}(I)\right\}, \qquad k = 1, \ldots, M.$$

Величина $w_k \in [0,1]$ визначає, наскільки поточний вхідний вектор відповідає передумові k-го правила.

У моделі Мамдані імплікація реалізується операцією відсікання вихідної функції належності:

$$\mu_R^{(k)}(r) = \min\left\{w_k, \mu_{R,B^{(k)}}(r)\right\}, \qquad r \in [0,1].$$

Після застосування всіх правил отримані нечіткі висновки агрегуються за допомогою S-норми максимуму:

$$\mu_R^{\text{agg}}(r) = \max_{1 \leqslant k \leqslant M} \mu_R^{(k)}(r), \qquad r \in [0,1].$$

Функція μ_R^{agg} задає результуючу нечітку множину для вихідної змінної ризику.

Для переходу від нечіткого висновку до чіткого числового значення використовується центроїдний метод дефазифікації:

$$R_{\text{FIS}} = \frac{\displaystyle\int_0^1 r\, \mu_R^{\text{agg}}(r)\, dr}{\displaystyle\int_0^1 \mu_R^{\text{agg}}(r)\, dr}. \tag{2.1}$$

За умови коректного перекриття функцій належності та достатнього покриття бази правил знаменник у (2.1) є додатним для допустимих вхідних векторів.

Отже, нечітка система виведення задає відображення

$$f_{\mathrm{FIS}} : [0,1]^4 \to [0,1], \qquad x \mapsto R_{\mathrm{FIS}}.$$

Отримане значення є інтерпретованою експертною оцінкою рівня ризику, побудованою на основі лінгвістичних термів і правил. У наступному підрозділі розглядається інша компонента моделі – байєсівська мережа, яка описує ймовірнісні залежності між факторами ризику. Подальше узгодження виходів цих двох компонентів виконується вже на рівні функцій мас у межах теорії доказів.

2.4 Математична модель на основі байєсівських мереж

Байєсівська мережа використовується в роботі як імовірнісний компонент гібридної моделі. Її призначення полягає в описі залежності між дискретизованими факторами ризику та рівнем ризику у вигляді апостеріорного розподілу над гіпотезами

$$\Theta = \{H_{\mathrm{low}}, H_{\mathrm{med}}, H_{\mathrm{high}}\}.$$

На відміну від нечіткої системи, розглянутої в попередньому підрозділі, байєсівська мережа не оперує ступенями належності до лінгвістичних термів, а задає умовні ймовірності станів ризику за наявних свідчень. Такий підхід є природним для задач, у яких потрібно враховувати статистичні залежності між факторами та оновлювати оцінку після надходження нової інформації [9, 10, 16].

У загальному вигляді байєсівська мережа задається парою

$$\mathcal{B} = (G, \boldsymbol{\theta}),$$

де

$$G = (\mathcal{V},\mathcal{E})$$

є орієнтованим ациклічним графом, вершини якого відповідають випадковим змінним, а ребра задають безпосередні залежності між ними. Позначення $\mathcal{V}$ використовується саме для множини вершин графа, щоб не змішувати його з фактором вразливості V. Множина параметрів $\boldsymbol{\theta}$ складається з локальних умовних розподілів

$$\mathbb{P}(X_i \mid \mathrm{Pa}_G(X_i)),$$

де $\mathrm{Pa}_G(X_i)$ – множина батьківських вершин змінної X_i у графі G.

Якщо $\mathcal{V} = \{X_1, \ldots, X_n\}$, то за локальною марковською властивістю спільний розподіл факторизується у вигляді

$$\mathbb{P}(X_1, \ldots, X_n) = \prod_{i=1}^{n} \mathbb{P}\big(X_i \mid \mathrm{Pa}_G(X_i)\big). \tag{2.2}$$

Завдяки такій факторизації байєсівські мережі є зручними для прикладних задач ризик-аналізу: замість повного задання багатовимірного розподілу достатньо визначити локальні залежності між змінними [10, 30].

У межах цієї роботи використовується дискретна байєсівська мережа, побудована на факторах ризику, отриманих у підрозділі 2.2. Неперервні значення

$$T(z), \quad V(z), \quad A(z), \quad I(z)$$

спочатку переводяться у категоріальні стани за допомогою дискретизації:

$$T_d = q_T(T(z)), \quad V_d = q_V(V(z)), \quad A_d = q_A(A(z)), \quad I_d = q_I(I(z)).$$

Для кожної з цих змінних використовується множина станів

$$\mathcal{S}_T = \mathcal{S}_V = \mathcal{S}_A = \mathcal{S}_I = \{\text{Low},\text{Medium},\text{High}\}.$$

Цільовою змінною мережі є дискретний рівень ризику

$$H \in \Theta.$$

Мінімальна структура байєсівської мережі, достатня для поточної постановки задачі, має вигляд

$$\mathcal{V} = \{T_d, V_d, A_d, I_d, H\}, \tag{2.3}$$

$$\mathcal{E} = \{(T_d, H), (V_d, H), (A_d, H), (I_d, H)\}. \tag{2.4}$$

Така структура означає, що рівень ризику безпосередньо залежить від станів загрози, вразливості або експозиції, критичності активу та потенційного впливу. Вона є спрощеною, але інтерпретованою: кожен із чотирьох факторів розглядається як безпосередній аргумент імовірнісної оцінки ризику. За наявності додаткових даних структуру можна розширити проміжними вузлами, наприклад для типу атаки або стану захисних механізмів. У межах цієї моделі, однак, використовується наведена базова структура, щоб уникнути необґрунтованого ускладнення.

Для графа (2.3)–(2.4) факторизація спільного розподілу має вигляд

$$\mathbb{P}(T_d, V_d, A_d, I_d, H) = \mathbb{P}(T_d)\mathbb{P}(V_d)\mathbb{P}(A_d)\mathbb{P}(I_d)\mathbb{P}(H \mid T_d, V_d, A_d, I_d).$$

Основним параметром, що визначає ризик у такій мережі, є таблиця умовних ймовірностей

$$\mathbb{P}(H \mid T_d, V_d, A_d, I_d), \tag{2.5}$$

яка задає розподіл рівня ризику для кожної комбінації станів вхідних факторів. Якщо кожен фактор має три стани, то така таблиця містить 3^4 комбінацій батьківських станів, для кожної з яких задається розподіл на множині Θ.

Параметри байєсівської мережі можуть задаватися експертно, оцінюватися за розміченими даними або формуватися комбіновано. У задачах кіберризику це особливо важливо, оскільки для рідкісних або

нових сценаріїв атак статистичної інформації може бути недостатньо [16, 11]. Якщо параметри оцінюються за вибіркою, для умовної таблиці (2.5) можна використати оцінку максимальної правдоподібності зі згладжуванням:

$$\widehat{\mathbb{P}}(H = h \mid \pi) = \frac{N(h,\pi) + \lambda}{N(\pi) + \lambda|\Theta|}, \qquad h \in \Theta,$$

де $\pi = (t,v,a,i)$ – конкретна комбінація станів змінних T_d,V_d,A_d,I_d, $N(h,\pi)$ – кількість спостережень з рівнем ризику h та комбінацією π, $N(\pi)$ – загальна кількість спостережень з цією комбінацією, а $\lambda > 0$ – параметр згладжування. Таке згладжування запобігає появі нульових імовірностей для комбінацій, які рідко трапляються у вибірці [17].

Для нового мережевого спостереження z формується набір свідчень

$$e(z) = \{T_d = t,\, V_d = v,\, A_d = a,\, I_d = i\},$$

де t,v,a,i є результатами дискретизації відповідних факторів. Основним результатом імовірнісного виведення є апостеріорний розподіл

$$\mathbb{P}(H = h \mid e(z)), \qquad h \in \Theta.$$

У випадку повністю спостережених батьківських змінних цей розподіл безпосередньо задається відповідним рядком таблиці умовних ймовірностей (2.5). Якщо ж частина свідчень відсутня, апостеріорний розподіл обчислюється шляхом маргіналізації неспостережуваних змінних згідно з факторизацією (2.2):

$$\mathbb{P}(H = h \mid e) = \frac{\displaystyle\sum_{\omega:\, H=h,\ \omega \models e} \prod_{X_i \in \mathcal{V}} \mathbb{P}\big(X_i \mid \mathrm{Pa}_G(X_i)\big)}{\displaystyle\sum_{\omega:\, \omega \models e} \prod_{X_i \in \mathcal{V}} \mathbb{P}\big(X_i \mid \mathrm{Pa}_G(X_i)\big)},$$

де ω пробігає допустимі конфігурації змінних мережі, а запис $\omega \models e$ означає, що конфігурація узгоджується з наявними свідченнями.

Для скалярної інтерпретації байєсівського результату вводиться числове кодування рівнів ризику:

$$u(H_{\text{low}}) = 0, \qquad u(H_{\text{med}}) = 0.5, \qquad u(H_{\text{high}}) = 1.$$

Тоді числова оцінка ризику, отримана з байєсівської мережі, визначається як математичне сподівання цього кодування за апостеріорним розподілом:

$$R_{\text{BN}}(z) = \sum_{h \in \Theta} u(h)\, \mathbb{P}(H = h \mid e(z)). \tag{2.6}$$

Без такого кодування математичне сподівання для категоріальних станів ризику не має коректного змісту.

Отже, байєсівська мережа задає відображення

$$f_{\text{BN}} : [0,1]^4 \to \Delta(\Theta), \qquad x \mapsto \mathbb{P}(H \mid e),$$

де $\Delta(\Theta)$ позначає множину ймовірнісних розподілів на просторі гіпотез Θ. Скалярна величина R_{BN} у (2.6) може використовуватися для інтерпретації та порівняння з виходом нечіткої системи. Для подальшої агрегації, однак, важливішим є сам апостеріорний розподіл $\mathbb{P}(H \mid e)$. Його перетворення у функцію маси та узгодження з результатом нечіткої системи розглядаються в наступному підрозділі.

2.5 Апарат теорії доказів для агрегації результатів моделей

Після побудови нечіткої та байєсівської компонент постає задача їх коректного узгодження. Вихід нечіткої системи $R_{\text{FIS}} \in [0,1]$ є скалярною експертною оцінкою, тоді як байєсівська мережа формує апостеріорний розподіл $\mathbb{P}(H \mid e)$ на просторі гіпотез ризику. Тому такі результати не слід безпосередньо усереднювати або ототожнювати. Для їх агрегації в роботі використовується апарат теорії доказів Демпстера–Шейфера, який дозволяє представляти не лише підтримку окремих гіпотез, а й

залишкову невизначеність [12, 13].

Простором розрізнення є множина

$$\Theta = \{H_{\text{low}}, H_{\text{med}}, H_{\text{high}}\},$$

де H_{low}, H_{med} та H_{high} відповідають низькому, середньому та високому рівням ризику. Базова функція призначення маси, або BPA, визначається як відображення

$$m : 2^{\Theta} \to [0, 1],$$

що задовольняє умови

$$m(\varnothing) = 0, \qquad \sum_{A \subseteq \Theta} m(A) = 1.$$

Значення $m(A)$ інтерпретується як маса довіри, безпосередньо призначена твердженню про те, що істинний рівень ризику належить множині A. Якщо $A = \Theta$, така маса відображає не підтримку конкретного рівня ризику, а залишкову невизначеність.

Для кожної гіпотези або множини гіпотез можна визначити функції довіри та правдоподібності:

$$Bel(A) = \sum_{B \subseteq A} m(B),$$

$$Pl(A) = \sum_{B \cap A \neq \varnothing} m(B) = 1 - Bel(\Theta \setminus A).$$

Вони задають нижню та верхню межі підтримки множини A, що є корисним для інтерпретації невизначеності фінальної оцінки ризику.

Перетворення виходу нечіткої системи у BPA виконується не за самим числовим значенням R_{FIS}, а через його належність до термів ризику. Нехай $\mu_{R,k}(r)$, $k \in \{\text{low,med,high}\}$, – функції належності вихідної змінної ризику, введені для FIS. Для скалярної оцінки R_{FIS} визначимо

нормовані коефіцієнти підтримки:

$$\gamma_k(R_{\text{FIS}}) = \frac{\mu_{R,k}(R_{\text{FIS}})}{\sum\limits_{j\in\{\text{low,med,high}\}} \mu_{R,j}(R_{\text{FIS}})}, \qquad k \in \{\text{low,med,high}\}. \tag{2.7}$$

За коректного покриття універсуму функціями належності знаменник у (2.7) є додатним. Тоді BPA, отримана з нечіткої системи, задається як

$$m_{\text{FIS}}(\{H_k\}) = \rho_{\text{FIS}}\, \gamma_k(R_{\text{FIS}}), \qquad k \in \{\text{low,med,high}\},$$

$$m_{\text{FIS}}(\Theta) = 1 - \rho_{\text{FIS}},$$

а для інших підмножин $A \subseteq \Theta$ покладається $m_{\text{FIS}}(A) = 0$. Тут $\rho_{\text{FIS}} \in [0, 1]$ є коефіцієнтом надійності нечіткої компоненти. Він визначає, яка частина маси призначається конкретним гіпотезам, а яка залишається на повній множині Θ як невизначеність. Такий підхід відповідає ідеї дисконтування свідчень у теорії доказів [15].

Для байєсівської мережі природним джерелом підтримки гіпотез є апостеріорний розподіл $\mathbb{P}(H \mid e)$. Тому BPA для BN задається формулами

$$m_{\text{BN}}(\{H_k\}) = \rho_{\text{BN}}\, \mathbb{P}(H_k \mid e), \qquad k \in \{\text{low,med,high}\},$$

$$m_{\text{BN}}(\Theta) = 1 - \rho_{\text{BN}},$$

де $\rho_{\text{BN}} \in [0, 1]$ – коефіцієнт надійності байєсівської компоненти. Якщо статистичні дані є неповними, шумними або недостатньо репрезентативними, значення ρ_{BN} може бути зменшене. У такому разі частина маси переноситься в область невизначеності, а не на протилежні гіпотези.

Важливо, що коефіцієнти ρ_{FIS} та ρ_{BN} не є рівнями ризику. Вони характеризують довіру до відповідних джерел свідчень. У цій роботі FIS і BN розглядаються як різнорідні, але не повністю незалежні джерела, оскільки обидві компоненти використовують фактори, побудовані з одного вектора $\phi(z)$. Формальне правило Демпстера передбачає

незалежність джерел, тому застосування дисконтування та подальший контроль конфлікту є необхідними елементами моделі [15].

Коефіцієнт конфлікту між двома джерелами визначається як сумарна маса несумісних тверджень:

$$K = \sum_{A \cap B = \varnothing} m_{\mathrm{FIS}}(A) m_{\mathrm{BN}}(B).$$

Якщо K є малим, свідчення моделей узгоджуються між собою. Якщо ж K наближається до одиниці, значна частина маси припадає на взаємовиключні гіпотези, і нормалізація в класичному правилі Демпстера може призвести до нестійких або контрінтуїтивних результатів. Саме з цим пов'язаний відомий парадокс Заде [32, 14].

Для помірного конфлікту застосовується класичне правило Демпстера:

$$m_D(C) = \frac{\sum\limits_{A \cap B = C} m_{\mathrm{FIS}}(A) m_{\mathrm{BN}}(B)}{1 - K}, \qquad C \neq \varnothing.$$

При цьому $m_D(\varnothing) = 0$. Для високого конфлікту використовується правило Ягера. Воно не відкидає конфліктну масу, а переносить її на повну множину гіпотез Θ, тобто інтерпретує конфлікт як додаткову невизначеність [14]. Для $C \neq \varnothing$ і $C \neq \Theta$ це правило має вигляд

$$m_Y(C) = \sum_{A \cap B = C} m_{\mathrm{FIS}}(A) m_{\mathrm{BN}}(B),$$

а для повної множини гіпотез:

$$m_Y(\Theta) = \sum_{A \cap B = \Theta} m_{\mathrm{FIS}}(A) m_{\mathrm{BN}}(B) + K.$$

У цій роботі використовується саме класичне правило Ягера, тому воно не називається модифікованим.

Для формалізації вибору правила комбінування вводиться поріг

конфлікту $\tau \in [0,1]$. Агрегована функція мас визначається як

$$m(C) = \begin{cases} m_D(C), & K \leqslant \tau, \\ m_Y(C), & K > \tau, \end{cases} \qquad C \subseteq \Theta.$$

Значення τ є параметром моделі та може задаватися експертно або аналізуватися в експериментальній частині. Така схема дозволяє використовувати інформативність правила Демпстера за узгоджених свідчень і водночас уникати надмірно категоричних висновків за високого конфлікту.

Оскільки агрегована функція мас m у загальному випадку може призначати частину маси не лише одноелементним гіпотезам, а й множині Θ, фінальний числовий ризик не можна коректно визначати як звичайне математичне сподівання без додаткового перетворення. Для переходу від BPA до ймовірнісного розподілу використовується pignistic transformation:

$$\mathrm{BetP}(H_k) = \sum_{\substack{A \subseteq \Theta \\ H_k \in A}} \frac{m(A)}{|A|}, \qquad H_k \in \Theta.$$

Після цього фінальний нормований індекс ризику обчислюється за числовим кодуванням рівнів ризику, введеним для байєсівської компоненти:

$$R(z) = \sum_{H_k \in \Theta} u(H_k)\,\mathrm{BetP}(H_k),$$

де

$$u(H_{\mathrm{low}}) = 0, \qquad u(H_{\mathrm{med}}) = 0.5, \qquad u(H_{\mathrm{high}}) = 1.$$

Отримане значення $R(z) \in [0,1]$ є фінальною оцінкою кіберризику для мережевого спостереження z.

Таким чином, у межах підрозділу побудовано формальний механізм переходу

$$R_{\mathrm{FIS}},\ \mathbb{P}(H \mid e) \longrightarrow m_{\mathrm{FIS}},\ m_{\mathrm{BN}} \longrightarrow m \longrightarrow R(z),$$

який узгоджує результати нечіткої та байєсівської компонент без

змішування їхньої математичної природи. Теорія доказів у цій схемі виконує роль рівня агрегації свідчень, а не замінює попередні моделі оцінювання ризику.

Висновки до розділу 2

У другому розділі побудовано формальну математичну модель гібридного оцінювання кіберризику. На відміну від загального огляду методів, розділ подає послідовну обчислювальну схему переходу від мережевого спостереження $z \in \mathcal{D}$ до фінального нормованого індексу ризику $R(z) \in [0,1]$.

Сформульовано математичну постановку задачі та уточнено, що в межах роботи оцінюється не ймовірність атаки сама по собі і не абсолютна величина фінансових втрат, а нормований індекс кіберризику. Введено простір вхідних мережевих спостережень $\mathcal{D}$, простір факторів ризику $\mathcal{X} \subseteq [0,1]^4$ та відображення

$$\phi : \mathcal{D} \to \mathcal{X}, \qquad \phi(z) = (T(z),V(z),A(z),I(z)),$$

яке пов'язує низькорівневі ознаки мережевого трафіку з високорівневими факторами ризику. Окремо розмежовано зміст факторів загрози, вразливості або експозиції, критичності активу та потенційного впливу інциденту.

Для моделювання лінгвістичної невизначеності побудовано нечітку систему виведення типу Мамдані, яка задає відображення

$$f_{\mathrm{FIS}} : [0,1]^4 \to [0,1]$$

та формує експертно інтерпретовану оцінку R_{FIS}. Для ймовірнісного моделювання залежностей між факторами ризику введено байєсівську мережу з дискретними станами факторів та цільовою змінною ризику $H \in \Theta$. Результатом байєсівської компоненти є апостеріорний розподіл

$\mathbb{P}(H \mid e)$, а за потреби також числова оцінка R_{BN}, отримана через задане кодування рівнів ризику.

Оскільки виходи FIS і BN мають різну математичну природу, у розділі обґрунтовано їх перехід до спільного формалізму теорії доказів Демпстера–Шейфера. Визначено процедури побудови базових функцій призначення маси m_{FIS} та m_{BN}, введено коефіцієнти надійності джерел, коефіцієнт конфлікту K і правило вибору між класичним правилом Демпстера та правилом Ягера. Для отримання фінального числового індексу ризику використано pignistic transformation, що забезпечує коректний перехід від агрегованої функції мас до значення $R(z) \in [0, 1]$.

Таким чином, у другому розділі сформовано завершену математичну схему:

$$z \xrightarrow{\phi} (T,V,A,I) \xrightarrow{\mathrm{FIS,\,BN}} R_{\mathrm{FIS}},\ \mathbb{P}(H \mid e) \xrightarrow{\mathrm{DST}} m \xrightarrow{\mathrm{BetP}} R(z).$$

Ця схема є теоретичною основою для подальшої програмної реалізації та експериментального дослідження.

Конкретні числові параметри моделі – ваги у відображенні ϕ, пороги дискретизації, параметри функцій належності, таблиці умовних ймовірностей, коефіцієнти надійності джерел і поріг конфлікту – задаються та обґрунтовуються в третьому розділі відповідно до використаних даних і експериментального сценарію. Це дозволяє відокремити загальну математичну структуру моделі від її прикладної параметризації та реалізації.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТА ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ ГІБРИДНОЇ МОДЕЛІ

У попередньому розділі було сформульовано математичну схему гібридного оцінювання кіберризику, що задає перехід від мережевого спостереження $z \in \mathcal{D}$ до нормованого індексу ризику $R(z) \in [0,1]$. Вона охоплює побудову факторного подання $\phi(z) = (T,V,A,I)$, отримання оцінок за допомогою нечіткої системи виведення та байєсівської мережі, формування базових функцій маси, DST-агрегацію свідчень і обчислення фінального показника через pignistic transformation.

У цьому розділі подано практичну конкретизацію моделі та результати її експериментальної перевірки на мережевих даних. Розглянуто постановку експериментального сценарію, підготовку даних, параметризацію ϕ, реалізацію FIS- та BN-компонентів, формування функцій маси, DST-агрегацію й аналіз результатів.

Практична реалізація не трактується як завершена промислова система управління кіберризиками. Її мета полягає в обчислювальній перевірці запропонованої схеми: аналізі поведінки компонентів на потокових ознаках мережевого трафіку, формуванні фінального індексу ризику та оцінюванні впливу конфлікту між джерелами свідчень і неповноти факторної інформації на результат.

3.1 Експериментальний сценарій та програмне середовище

Експериментальна частина роботи спрямована на реалізацію математичної моделі, сформульованої в розділі 2, та перевірку її прикладної придатності на потокових мережевих даних. Одиницею обчислення є окреме мережеве спостереження z, подане вектором числових ознак мережевого потоку.

Для кожного спостереження послідовно будуються фактори ризику, обчислюються результати нечіткої та ймовірнісної компонент, після чого ці результати узгоджуються в межах теорії доказів. У практичному розділі фінальний індекс позначається як $R_{\text{final}}(z)$; за змістом він відповідає нормованій оцінці $R(z)$, введеній у теоретичній частині.

Загальний експериментальний pipeline має вигляд

$$\begin{aligned} z &\longrightarrow \phi(z) = (T,V,A,I) \\ &\longrightarrow R_{\text{FIS}},\, P(H \mid e) \\ &\longrightarrow m_{\text{FIS}}, m_{\text{BN}} \longrightarrow m \longrightarrow \text{BetP} \longrightarrow R_{\text{final}}. \end{aligned}$$

У цій схемі $\phi(z)$ задає перехід від низькорівневих ознак мережевого потоку до факторів ризику. Величина R_{FIS} є скалярним виходом нечіткої системи виведення, а $P(H \mid e)$ позначає апостеріорний розподіл байєсівської мережі на просторі гіпотез ризику Θ. Функції m_{FIS} та m_{BN} є базовими функціями призначення маси, сформованими на основі відповідних компонент; m – агрегована функція маси після комбінування свідчень; BetP – результат pignistic transformation; $R_{\text{final}} \in [0,1]$ – фінальний нормований індекс ризику.

У межах експерименту дані виконують дві різні функції. Мережеві ознаки використовуються як вхід для побудови факторів T,V,A,I. Натомість мітки класів трафіку не подаються на вхід гібридної моделі, а застосовуються лише для формування експериментальної цільової змінної рівня ризику та подальшого оцінювання результатів. Отже, початкові мітки атак виконують роль зовнішньої розмітки для перевірки поведінки побудованого індексу ризику.

Програмну реалізацію виконано мовою Python в інтерактивному середовищі Jupyter Notebook. Для роботи з табличними даними використовувалися бібліотеки `pandas` та `numpy`; для попередньої обробки, поділу вибірки та обчислення метрик якості – засоби `scikit-learn`; для побудови графіків і візуального аналізу результатів – `matplotlib`. Нечітка система виведення, таблиці умовних ймовірностей байєсівської

мережі та процедури формування й комбінування базових функцій маси реалізовувалися програмно відповідно до формул, наведених у другому розділі.

Структура програмного проєкту була організована так, щоб відокремити вхідні дані, проміжні результати, параметри моделі та матеріали для аналізу. Початкові CSV-файли розміщувалися в директорії `data/raw`, очищені й підготовлені набори даних – у `data/processed`, збережені таблиці результатів – у `outputs/tables`, графіки – у `outputs/figures`, а службові об'єкти та параметри експерименту – у `outputs/models`.

Основні етапи обчислень виконувалися в окремих notebook-файлах, що відповідали логіці експерименту: первинний огляд даних, попередня обробка, побудова факторів ризику, реалізація FIS-, BN- та DST-компонентів, а також аналіз отриманих результатів. Така організація дозволила зберегти послідовність експерименту та відокремити підготовку даних від етапів моделювання й інтерпретації результатів.

У цьому розділі основна увага приділяється не тексту програмного коду, а відтворюваній обчислювальній схемі, параметрам експерименту, числовим результатам та їх інтерпретації. Фрагменти програмної реалізації або повна структура notebook-файлів можуть бути винесені до додатків. В основному тексті подаються лише ті елементи реалізації, які безпосередньо впливають на математичну постановку, експериментальну перевірку та змістовне тлумачення результатів.

3.2 Підготовка експериментальних даних

Для експериментальної перевірки запропонованої моделі використано набір даних CIC-IDS2017 у форматі `MachineLearningCSV`. Він містить потокові ознаки мережевого трафіку, сформовані на основі характеристик з'єднань: тривалості потоку, кількості пакетів, обсягів переданих даних, статистик довжин пакетів, часових інтервалів та

параметрів TCP-з'єднань [18, 19, 37, 38]. Кожне спостереження в такому поданні розглядається як вектор ознак

$$z_i = (z_{i1}, z_{i2}, \ldots, z_{ip}) \in \mathcal{D},$$

що відповідає окремому мережевому потоку.

У межах базового експериментального сценарію було використано п'ять CSV-файлів, наведених у табл. 3.1.

Таблиця 3.1 – CSV-файли, використані в базовому експериментальному сценарії

№	Назва файлу
1	`Monday-WorkingHours.pcap_ISCX.csv`
2	`Tuesday-WorkingHours.pcap_ISCX.csv`
3	`Wednesday-workingHours.pcap_ISCX.csv`
4	`Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv`
5	`Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv`

Початкова мітка класу `Label` не використовувалась як вхідна ознака гібридної моделі. Вона застосовувалась лише для формування експериментальної цільової змінної `risk_level`, необхідної для параметризації байєсівської компоненти та подальшого оцінювання результатів. Формально це задається відображенням

$$h : \mathcal{L} \to \Theta,$$

де $\mathcal{L}$ – множина початкових класів трафіку, а

$$\Theta = \{H_{\text{low}}, H_{\text{med}}, H_{\text{high}}\}$$

– простір гіпотез рівня ризику, введений у розділі 2.

У межах цього відображення клас `BENIGN` було віднесено до низького рівня ризику; класи `PortScan`, `FTP-Patator` та `SSH-Patator` – до

середнього; DoS/DDoS-атаки – до високого.

Змінна `risk_level` є експериментальною розміткою для перевірки поведінки моделі на доступному датасеті, а не повною бізнес-оцінкою кіберризику. Це зумовлено тим, що CIC-IDS2017 не містить інформації про цінність активів, фінансові втрати, організаційний контекст або фактичні наслідки інцидентів. Тому рівень ризику в цьому експерименті трактується як нормована сценарна інтерпретація класів мережевої активності.

Частину класів CIC-IDS2017 не було включено до базового сценарію. Клас `Heartbleed` містив лише 11 записів, що недостатньо для стабільного навчання й тестування. Класи `Bot`, `Infiltration` і групу `Web Attack` також залишено поза межами експерименту, оскільки вони потребують окремої інтерпретації в контексті факторів ризику. Їх включення могло б змішати перевірку базової математичної схеми з аналізом рідкісних або неоднорідних класів.

Розподіл використаних класів та їх зіставлення з експериментальними рівнями ризику наведено в табл. 3.2.

Таблиця 3.2 – Розподіл класів CIC-IDS2017 за рівнями ризику

Рівень ризику	**Клас**	**Кількість записів**
Low	`BENIGN`	30 000
Medium	`PortScan`	30 000
	`FTP-Patator`	7 938
	`SSH-Patator`	5 897
High	`DDoS`	30 000
	`DoS Hulk`	30 000
	`DoS GoldenEye`	10 293
	`DoS slowloris`	5 796
	`DoS Slowhttptest`	5 499

Така структура дозволяє відокремити нормальний трафік від атак середнього й високого рівнів ризику та використати ці групи як основу для

побудови цільової змінної.

Для зменшення дисбалансу між класами кількість записів одного класу було обмежено значенням 30 000. Менші класи використовувалися в повному доступному обсязі. Після фільтрації класів і застосування цього обмеження сформований експериментальний набір містив 155 423 записи.

Попередня обробка даних включала такі етапи:

1) об'єднання CSV-файлів;
2) фільтрацію класів;
3) вилучення службових полів;
4) обробку нескінченних і пропущених значень;
5) формування цільової змінної `risk_level`;
6) поділ даних на навчальну та тестову вибірки;
7) імпутацію пропусків і нормалізацію ознак.

Обробка нескінченних значень була необхідною, оскільки в потокових мережевих ознаках можуть виникати некоректні значення, пов'язані з діленням на нуль або граничними характеристиками окремих потоків. Перед нормалізацією такі значення приводилися до коректного числового подання.

Поділ на навчальну та тестову вибірки виконувався після формування експериментального набору даних. Навчальна вибірка містила 108 796 записів, а тестова – 46 627 записів. Параметри імпутації пропущених значень і нормалізації ознак визначалися лише на навчальній вибірці, після чого застосовувалися до тестової. Такий порядок запобігає витоку інформації з тестових даних у процес попередньої обробки та забезпечує коректнішу експериментальну перевірку моделі [17].

Підготовлений у такий спосіб набір даних використовується в наступному підрозділі для побудови факторного відображення

$$\phi(z) = (T,V,A,I).$$

Саме це відображення забезпечує перехід від очищених і нормалізованих

потокових ознак до високорівневих факторів, на яких надалі працюють нечітка система виведення, байєсівська мережа та DST-агрегація.

3.3 Параметризація факторного відображення $\phi(z)$

Після очищення та нормалізації експериментальних даних наступним етапом є побудова факторного відображення

$$\phi(z) = \big(T(z),V(z),A(z),I(z)\big),$$

яке переводить вектор потокових мережевих ознак $z \in \mathcal{D}$ у чотири нормовані фактори ризику. Це відображення є проміжною ланкою між низькорівневими характеристиками трафіку та математичною моделлю, сформульованою в розділі 2. Саме на його основі формуються вхідні дані для FIS- та BN-компонентів.

У базовому експериментальному сценарії фактори T, V та I будуються як зважені агрегати нормованих мережевих ознак. Фактор A, що відповідає критичності активу, задається сценарно, оскільки CIC-IDS2017 не містить інформації про бізнес-роль вузлів, цінність активів або організаційний контекст. Тому в цьому експерименті значення A фіксується як стала величина, що відповідає припущенню про помірну критичність цільового середовища.

Для кожного фактора $Q \in \{T,V,I\}$ задається множина індексів ознак $\mathcal{I}_Q$ та набір невід'ємних ваг w_{Qj}, які задовольняють умову нормування

$$w_{Qj} \geqslant 0, \qquad \sum_{j\in\mathcal{I}_Q} w_{Qj} = 1.$$

Відповідний фактор обчислюється за формулою

$$Q(z) = \sum_{j\in\mathcal{I}_Q} w_{Qj}\,\widetilde{z}_j, \qquad Q \in \{T,V,I\},$$

де $\widetilde{z}_j \in [0,1]$ – нормоване значення j-ї ознаки. За такої побудови фактори $T(z)$, $V(z)$ та $I(z)$ також належать інтервалу [0,1], що забезпечує узгодженість шкал між факторним поданням, нечіткою системою виведення та байєсівською мережею.

Фактор T інтерпретується як показник активності загрози. До його складу включено ознаки, що характеризують інтенсивність мережевого потоку, швидкість передавання даних, кількість пакетів, а також варіативність часових і розмірних характеристик трафіку. Такий вибір відображає припущення, що аномально інтенсивна або нерівномірна мережева активність може бути пов'язана з реалізацією загрози.

Фактор V у межах цього експерименту не слід ототожнювати з фактичною CVE-вразливістю. Оскільки датасет не містить даних про конфігурацію хостів, встановлені оновлення, результати сканування або відомі вразливості сервісів, V використовується як проксі-показник експозиції та підозрілої структури TCP-з'єднання. Отже, цей фактор відображає лише ту частину інформації про потенційну експозицію, яку можна наближено отримати з доступних потокових ознак.

Фактор I характеризує потенційний технічний вплив мережевого потоку. Для його побудови використано ознаки, пов'язані з обсягом переданих даних і розмірами пакетів, оскільки такі характеристики можуть опосередковано відображати масштаб впливу інциденту на доступність, цілісність або конфіденційність інформаційного ресурсу.

Окремо було виключено ознаку `Destination Port` зі складу фактора V. Номер порту є категоріальною або службовою характеристикою мережевого з'єднання і після числової нормалізації не має природної монотонної інтерпретації: більше значення номера порту не означає більшої вразливості чи експозиції. Тому в базовій реалізації для V залишено лише ознаки, пов'язані з TCP-прапорами та параметрами вікна з'єднання.

Використані ознаки та ваги наведено в табл. 3.3. Ваги задавалися як фіксована експертна параметризація базового сценарію до проведення

порівняльного аналізу моделей і не оптимізувалися на тестовій вибірці. Це дозволяє уникнути підлаштування факторного відображення під фінальні результати експерименту.

Таблиця 3.3 – Ознаки та ваги для побудови факторів ризику

Фактор	**Ознака**	**Вага / значення**
T	`Flow Packets/s`	0.25
	`Flow Bytes/s`	0.20
	`Packet Length Std`	0.20
	`Flow IAT Std`	0.15
	`Total Fwd Packets`	0.10
	`Total Backward Packets`	0.10
V	`SYN Flag Count`	0.25
	`RST Flag Count`	0.20
	`PSH Flag Count`	0.15
	`ACK Flag Count`	0.10
	`Init_Win_bytes_forward`	0.15
	`Init_Win_bytes_backward`	0.15
A	Сценарно задана критичність активу	0.60
I	`Total Length of Fwd Packets`	0.20
	`Total Length of Bwd Packets`	0.20
	`Flow Bytes/s`	0.20
	`Packet Length Mean`	0.15
	`Max Packet Length`	0.15
	`Average Packet Size`	0.10

Для подальшого використання в байєсівській мережі неперервні значення факторів було перетворено у дискретні стани

$$Low, \qquad Medium, \qquad High.$$

Для кожного фактора $Q \in \{T,V,A,I\}$ задавалися два пороги $\tau_{Q,1}$ та $\tau_{Q,2}$. Дискретний стан визначався правилом

$$q_Q(Q) = \begin{cases} Low, & 0 \leqslant Q \leqslant \tau_{Q,1}, \\ Medium, & \tau_{Q,1} < Q \leqslant \tau_{Q,2}, \\ High, & \tau_{Q,2} < Q \leqslant 1. \end{cases}$$

Пороги для факторів T, V та I визначалися лише на навчальній вибірці,

що запобігає використанню інформації з тестових даних під час параметризації моделі. Для фактора A використано сценарні межі, оскільки в базовому експерименті його значення є сталим. Значення порогів наведено в табл. 3.4.

Таблиця 3.4 – Пороги дискретизації факторів ризику

Фактор	**Поріг Low/Medium**	**Поріг Medium/High**
T	0.003643	0.055574
V	0.101115	0.216836
A	0.330000	0.660000
I	0.000378	0.026360

Отримане факторне подання використовується у двох формах. Неперервні значення T,V,A,I подаються на вхід нечіткої системи виведення, де вони фазифікуються за відповідними функціями належності. Дискретизовані стани цих самих факторів використовуються в байєсівській мережі для оцінювання апостеріорного розподілу $P(H \mid e)$.

Таким чином, відображення $\phi(z)$ забезпечує єдину основу для обох компонентів гібридної моделі. Воно зберігає узгодженість шкал і водночас не змішує нечітку та ймовірнісну інтерпретації ризику.

3.4 Реалізація нечіткої системи виведення

Нечітка система виведення використовується як експертно-інтерпретована компонента гібридної моделі. Її входами є неперервні фактори ризику

$$T(z), \quad V(z), \quad A(z), \quad I(z),$$

побудовані на етапі факторного відображення $\phi(z)$. Виходом системи є скалярна оцінка

$$R_{\mathrm{FIS}}(z) \in [0,1].$$

На відміну від байєсівської мережі, яка працює з дискретизованими станами факторів і формує апостеріорний розподіл на просторі гіпотез, нечітка система безпосередньо використовує неперервні значення T,V,A,I. Їх перетворення у числовий індекс ризику здійснюється через базу лінгвістичних правил.

Для кожного вхідного фактора $Q \in \{T,V,A,I\}$ використовувалися три лінгвістичні терми:

$$Low, \qquad Medium, \qquad High.$$

Для крайніх термів Low та $High$ застосовувалися плечові функції належності, а для терма $Medium$ – трикутна функція належності. Параметри функцій належності для факторів T, V та I узгоджувалися з порогами дискретизації, наведеними в табл. 3.4. Для фактора A використовувалися сценарні межі, оскільки в базовому експерименті його значення фіксовано як $A = 0.60$.

Позначимо через

$$\mu_{Q,l} : [0,1] \to [0,1], \qquad Q \in \{T,V,A,I\}, \quad l \in \{Low,Medium,High\},$$

функцію належності значення фактора Q до терма l. Фазифікація вхідного вектора полягає в обчисленні значень

$$\mu_{Q,l}(Q(z))$$

для всіх факторів і відповідних лінгвістичних термів. Ці величини не інтерпретуються як імовірності, а задають ступені належності факторів до нечітких категорій.

Вихідна змінна $Risk$ також задавалася на інтервалі [0,1] трьома термами Low, $Medium$ та $High$. У результаті роботи нечіткої системи для кожного спостереження формувалися:

$$R_{\text{FIS}}, \qquad FIS_{\text{level}},$$

а також ступені належності вихідної оцінки до трьох рівнів ризику. Надалі ці ступені використовуються під час формування базової функції маси для FIS-компоненти.

База правил містила всі $3^4 = 81$ комбінації лінгвістичних термів для чотирьох вхідних факторів. Щоб уникнути довільного ручного задання великої кількості правил, було використано монотонну схему їх автоматизованого формування. Для цього кожному терму ставилося у відповідність числове кодування

$$c(Low) = 0, \qquad c(Medium) = 0.5, \qquad c(High) = 1.$$

Для кожної комбінації термів

$$(l_T,l_V,l_A,l_I), \qquad l_T,l_V,l_A,l_I \in \{Low,Medium,High\},$$

обчислювався допоміжний експертний скор

$$s = 0.35\, c(l_T) + 0.25\, c(l_V) + 0.10\, c(l_A) + 0.30\, c(l_I).$$

На основі цього значення визначався вихідний терм правила:

$$Risk = \begin{cases} Low, & s < 0.30, \\ Medium, & 0.30 \leqslant s < 0.60, \\ High, & s \geqslant 0.60. \end{cases}$$

Отже, для кожної комбінації вхідних термів формувалося правило вигляду

IF T is l_T AND V is l_V AND A is l_A AND I is l_I

THEN $Risk$ is l_R,

де l_R визначається за правилом порогування скору s.

Допоміжний скор s не розглядається як окрема модель оцінювання ризику. Він використовується лише для узгодженого формування бази нечітких правил. Перевага такої схеми полягає в монотонності:

збільшення рівня загрози, експозиції, критичності активу або потенційного впливу не може зменшити вихідний лінгвістичний рівень ризику без спеціального предметного обґрунтування.

Для правила r_k з передумовою

$$(l_T^{(k)}, l_V^{(k)}, l_A^{(k)}, l_I^{(k)})$$

ступінь активації обчислювався за T-нормою мінімуму:

$$\omega_k = \min\{\mu_{T,l_T^{(k)}}(T(z)), \mu_{V,l_V^{(k)}}(V(z)), \mu_{A,l_A^{(k)}}(A(z)), \mu_{I,l_I^{(k)}}(I(z))\}.$$

Імплікація Мамдані реалізовувалася відсіканням вихідної функції належності відповідного терма:

$$\mu_R^{(k)}(r) = \min\{\omega_k, \mu_{R,l_R^{(k)}}(r)\}, \qquad r \in [0,1].$$

Після застосування всіх правил результуюча вихідна нечітка множина отримувалася агрегуванням за операцією максимуму:

$$\mu_R^{\text{agg}}(r) = \max_k \mu_R^{(k)}(r).$$

Для переходу від нечіткого висновку до числової оцінки використовувався метод центра ваги:

$$R_{\text{FIS}}(z) = \frac{\int_0^1 r\, \mu_R^{\text{agg}}(r)\, dr}{\int_0^1 \mu_R^{\text{agg}}(r)\, dr}.$$

За рахунок повного покриття комбінацій термів і перекриття функцій належності знаменник у цій формулі є додатним для допустимих вхідних значень.

У табл. 3.5 наведено фрагмент сформованої бази правил. Повна база містить усі 81 комбінацію термів і використовується під час обчислення R_{FIS} для кожного спостереження тестової вибірки.

Таблиця 3.5 – Приклади правил нечіткої системи виведення

T	V	A	I	$Risk$
Low	Low	Low	Low	Low
Low	Low	Medium	Medium	Low
Low	Low	Medium	High	Medium
Low	Medium	Medium	Medium	Medium
Medium	Low	Medium	Low	Low
Medium	Medium	Medium	Medium	Medium
Medium	High	Medium	High	High
High	Low	Medium	Medium	Medium
High	Medium	Medium	High	High
High	High	High	High	High

Таким чином, реалізована нечітка система задає відображення

$$f_{\text{FIS}} : [0,1]^4 \to [0,1], \qquad (T,V,A,I) \mapsto R_{\text{FIS}}.$$

Отримана оцінка відображає результат експертно-лінгвістичного аналізу факторів ризику. У наступному підрозділі розглядається ймовірнісна компонента моделі – байєсівська мережа, яка використовує дискретизовані стани тих самих факторів для отримання розподілу $P(H \mid e)$.

3.5 Реалізація байєсівської мережі

Байєсівська мережа використовується як імовірнісна компонента гібридної моделі. На відміну від нечіткої системи, яка працює з неперервними значеннями факторів T,V,A,I, байєсівська мережа використовує їх дискретизовані стани

$$T_d, \quad V_d, \quad A_d, \quad I_d$$

та формує апостеріорний розподіл на просторі гіпотез ризику

$$\Theta = \{H_{\text{low}}, H_{\text{med}}, H_{\text{high}}\}.$$

У практичній реалізації структура мережі відповідає схемі, введеній у другому розділі:

$$T_d,\, V_d,\, A_d,\, I_d \longrightarrow H,$$

де $H \in \Theta$ є цільовою змінною рівня ризику. Така структура означає, що рівень ризику визначається умовно відносно спільного стану чотирьох факторів, побудованих на етапі факторного відображення $\phi(z)$.

Для кожного спостереження z після дискретизації формується набір свідчень

$$e(z) = \{T_d = t,\, V_d = v,\, A_d = a,\, I_d = i\}.$$

У базовому експериментальному сценарії всі чотири фактори вважаються спостереженими. Тому апостеріорний розподіл $P(H \mid e(z))$ визначається відповідним рядком таблиці умовних ймовірностей

$$P(H \mid T_d,V_d,A_d,I_d).$$

Як навчальна ціль для оцінювання цієї таблиці використовувалася змінна `risk_level`, сформована на основі експериментального зіставлення класів трафіку з рівнями ризику. Початкова мітка `Label` не входила до множини вхідних вузлів байєсівської мережі та не використовувалася як ознака під час обчислення $P(H \mid e)$.

Оцінювання умовних імовірностей виконувалося лише на навчальній вибірці. Для кожної комбінації станів

$$e = (t,v,a,i)$$

обчислювалися кількість записів $N(e,h_k)$, що мають цю комбінацію факторів і рівень ризику $h_k \in \Theta$, а також загальна кількість записів $N(e)$ з відповідною комбінацією факторів.

Щоб уникнути нульових імовірностей для рідкісних або слабко представлених комбінацій, застосовувалося лапласівське згладжування з

параметром $\alpha = 1$:

$$\widehat{P}(H = h_k \mid e) = \frac{N(e,h_k) + \alpha}{N(e) + \alpha|\Theta|}, \qquad h_k \in \Theta.$$

Оскільки $|\Theta| = 3$, для комбінацій, які не спостерігалися в навчальній вибірці, така оцінка приводить до рівномірного розподілу між трьома рівнями ризику. Це запобігає появі невизначених або вироджених рядків у таблиці умовних імовірностей.

У базовому сценарії фактор A має стале значення $A = 0.60$. За порогами табл. 3.4 воно відповідає стану $A_d = Medium$. Отже, в експерименті фактично використовується зріз таблиці умовних імовірностей для середньої критичності активу. Це є наслідком відсутності в CIC-IDS2017 інформації про бізнес-критичність активів і враховується під час інтерпретації результатів.

Основним виходом байєсівської компоненти є не окрема категорія ризику, а повний апостеріорний розподіл

$$\widehat{P}(H_{\text{low}} \mid e), \quad \widehat{P}(H_{\text{med}} \mid e), \quad \widehat{P}(H_{\text{high}} \mid e).$$

Саме цей розподіл надалі використовується для формування базової функції маси m_{BN}.

Додатково, для порівняння з іншими компонентами та побудови baseline-моделей, вводиться числова оцінка ризику BN. Вона визначається як математичне сподівання за кодуванням рівнів ризику

$$u(H_{\text{low}}) = 0, \qquad u(H_{\text{med}}) = 0.5, \qquad u(H_{\text{high}}) = 1.$$

Тоді

$$R_{\text{BN}}(z) = \sum_{h_k \in \Theta} u(h_k) \widehat{P}(H = h_k \mid e(z)).$$

У розгорнутому вигляді це записується як

$$R_{\text{BN}} = 0 \cdot P_{\text{low}} + 0.5 \cdot P_{\text{med}} + 1 \cdot P_{\text{high}},$$

де P_{low}, P_{med} та P_{high} є відповідними компонентами апостеріорного розподілу.

Дискретна інтерпретація результату BN визначалася за найбільш імовірним рівнем ризику:

$$BN_{\text{level}} = \arg\max_{h_k \in \Theta} \widehat{P}(H = h_k \mid e).$$

Фрагмент отриманої таблиці умовних імовірностей наведено в табл. 3.6. Імовірності подано з округленням до трьох десяткових знаків, тому дуже малі значення можуть відображатися як 0.000.

Таблиця 3.6 – Фрагмент таблиці умовних імовірностей BN

T_d	V_d	A_d	I_d	N	P_{low}	P_{med}	P_{high}
Low	Low	Medium	Low	4 736	0.164	0.005	0.831
Low	Low	Medium	Medium	10 663	0.746	0.000	0.253
Low	Medium	Medium	Low	9 136	0.122	0.747	0.131
Medium	Medium	Medium	Low	6 091	0.082	0.598	0.320
Medium	High	Medium	Low	5 761	0.067	0.923	0.010
High	Medium	Medium	High	17 566	0.007	0.000	0.993
High	High	Medium	High	5 589	0.030	0.002	0.968

Наведені рядки демонструють, що байєсівська компонента не просто присвоює спостереженню один рівень ризику, а задає розподіл підтримки між трьома можливими гіпотезами. Це важливо для подальшої DST-агрегації, оскільки як імовірнісне джерело свідчень у гібридній моделі використовується саме розподіл $P(H \mid e)$, а не лише найбільш імовірний стан або числове значення R_{BN}.

3.6 Формування базових функцій маси та DST-агрегація

Після отримання виходів нечіткої системи та байєсівської мережі їх необхідно подати у спільному формалізмі, придатному для комбінування. Для цього результати FIS- та BN-компонентів перетворюються у базові

функції призначення маси на просторі гіпотез

$$\Theta = \{H_{\text{low}}, H_{\text{med}}, H_{\text{high}}\}.$$

Такий перехід є важливим, оскільки R_{FIS} є скалярною нечіткою оцінкою, тоді як байєсівська мережа формує апостеріорний розподіл $P(H \mid e)$. Після перетворення обидва джерела свідчень можна узгоджено агрегувати в межах теорії доказів.

Для FIS-компоненти маси визначалися через ступені належності значення R_{FIS} до вихідних термів ризику. Нехай

$$\mu_{R,k}(R_{\text{FIS}}), \qquad k \in \{\text{low,med,high}\},$$

позначає ступінь належності вихідної оцінки нечіткої системи до відповідного терма. Щоб отримати нормовані коефіцієнти підтримки гіпотез, використовувалося перетворення

$$\gamma_{\text{FIS},k} = \frac{\mu_{R,k}(R_{\text{FIS}})}{\sum\limits_{j \in \{\text{low,med,high}\}} \mu_{R,j}(R_{\text{FIS}})}.$$

За коректного покриття вихідного інтервалу функціями належності знаменник є додатним. Отримані значення $\gamma_{\text{FIS},k}$ задають відносну підтримку кожного рівня ризику з боку FIS-компоненти.

У базовому сценарії використовувалися однакові коефіцієнти надійності джерел:

$$\rho_{\text{FIS}} = 0.8, \qquad \rho_{\text{BN}} = 0.8.$$

Ці коефіцієнти не є рівнями ризику. Вони визначають частку маси, яку відповідне джерело призначає конкретним гіпотезам. Залишок маси переноситься на повну множину Θ, після чого трактується як невизначеність джерела.

Базова функція маси для FIS задавалася формулами

$$m_{\mathrm{FIS}}(\{H_k\}) = \rho_{\mathrm{FIS}}\gamma_{\mathrm{FIS},k}, \qquad H_k \in \Theta,$$

$$m_{\mathrm{FIS}}(\Theta) = 1 - \rho_{\mathrm{FIS}}.$$

Для всіх інших підмножин $B \subseteq \Theta$ покладалося

$$m_{\mathrm{FIS}}(B) = 0.$$

Для BN-компоненти джерелом підтримки гіпотез був апостеріорний розподіл, отриманий у попередньому підрозділі:

$$\widehat{P}(H_k \mid e), \qquad H_k \in \Theta.$$

Відповідна базова функція маси мала вигляд

$$m_{\mathrm{BN}}(\{H_k\}) = \rho_{\mathrm{BN}}\widehat{P}(H_k \mid e), \qquad H_k \in \Theta,$$

$$m_{\mathrm{BN}}(\Theta) = 1 - \rho_{\mathrm{BN}}.$$

Маси інших підмножин також вважалися нульовими. Отже, кожне джерело розподіляє частину маси між одноелементними гіпотезами H_{low}, H_{med}, H_{high}, а решту залишає на множині Θ.

Для оцінювання суперечності між FIS- та BN-компонентами обчислювався коефіцієнт конфлікту

$$K_{\mathrm{conflict}} = \sum_{\substack{H_i, H_j \in \Theta \\ i \neq j}} m_{\mathrm{FIS}}(\{H_i\}) m_{\mathrm{BN}}(\{H_j\}).$$

У цій реалізації конфлікт виникає тоді, коли два джерела призначають маси різним одноелементним гіпотезам. Маси, призначені множині Θ, не вважаються конфліктними, оскільки вони відображають невизначеність, а не підтримку альтернативного конкретного рівня ризику.

У базовому сценарії було використано поріг конфлікту

$$\tau = 0.5.$$

Якщо $K_{\text{conflict}} \leqslant \tau$, свідчення комбінувалися за правилом Демпстера. Оскільки фокальними елементами в реалізації є лише одноелементні гіпотези та множина Θ, для кожного $H_k \in \Theta$ спочатку обчислювалася узгоджена маса

$$\begin{aligned} S_k = {} & m_{\text{FIS}}(\{H_k\})m_{\text{BN}}(\{H_k\}) + m_{\text{FIS}}(\{H_k\})m_{\text{BN}}(\Theta) \\ & + m_{\text{FIS}}(\Theta)m_{\text{BN}}(\{H_k\}). \end{aligned}$$

Після цього агреговані маси визначалися як

$$m_D(\{H_k\}) = \frac{S_k}{1 - K_{\text{conflict}}}, \qquad H_k \in \Theta,$$

$$m_D(\Theta) = \frac{m_{\text{FIS}}(\Theta)m_{\text{BN}}(\Theta)}{1 - K_{\text{conflict}}}.$$

Таке нормування підсилює узгоджену частину свідчень і не переносить конфліктну масу до фінального розподілу.

Якщо $K_{\text{conflict}} > \tau$, застосовувалося правило Ягера. У цьому випадку конфліктна маса не перерозподіляється між окремими гіпотезами, а переноситься на множину невизначеності:

$$m_Y(\{H_k\}) = S_k, \qquad H_k \in \Theta,$$

$$m_Y(\Theta) = m_{\text{FIS}}(\Theta)m_{\text{BN}}(\Theta) + K_{\text{conflict}}.$$

Отже, за високого конфлікту модель не формує надмірно категоричний висновок, а збільшує масу, призначену Θ. Це дає змогу явно відобразити суперечливість свідчень.

Підсумкова агрегована функція маси визначалася за правилом

$$m = \begin{cases} m_D, & K_{\text{conflict}} \leqslant \tau, \\ m_Y, & K_{\text{conflict}} > \tau. \end{cases}$$

Після комбінування виконувалося pignistic-перетворення, що переводить агреговану функцію маси у розподіл на Θ. В реалізованій схемі ненульові маси можуть залишатися лише на одноелементних гіпотезах та на повній множині Θ, для кожного $H_k \in \Theta$ маємо

$$\text{BetP}(H_k) = m(\{H_k\}) + \frac{m(\Theta)}{|\Theta|}.$$

У розглянутій постановці $|\Theta| = 3$, тому маса невизначеності рівномірно розподіляється між трьома рівнями ризику.

Фінальний числовий індекс ризику обчислювався як математичне сподівання за тим самим кодуванням рівнів ризику, що використовувалося для інтерпретації результатів байєсівської мережі:

$$u(H_{\text{low}}) = 0, \qquad u(H_{\text{med}}) = 0.5, \qquad u(H_{\text{high}}) = 1.$$

Тоді

$$R_{\text{final}}(z) = \sum_{H_k \in \Theta} u(H_k)\,\text{BetP}(H_k).$$

У розгорнутому вигляді:

$$R_{\text{final}} = 0 \cdot \text{BetP}(H_{\text{low}}) + 0.5 \cdot \text{BetP}(H_{\text{med}}) + 1 \cdot \text{BetP}(H_{\text{high}}).$$

Отримане значення $R_{\text{final}} \in [0,1]$ є основним числовим результатом гібридної моделі. Дискретна інтерпретація фінального рішення визначалася як

$$Hybrid_{\text{level}} = \arg\max_{H_k \in \Theta} \text{BetP}(H_k),$$

і використовувалася для порівняння зі змінною `risk_level`.

Таким чином, DST-компонента забезпечує не лише фінальний

числовий індекс ризику, а й додаткові діагностичні величини: коефіцієнт конфлікту K_{conflict}, масу невизначеності $m(\Theta)$ та pignistic-розподіл між рівнями ризику.

3.7 Базовий експеримент і порівняння моделей

У цьому підрозділі наведено результати базового експерименту на тестовій вибірці. Порівняння виконувалося не лише для визначення найточнішої моделі за класифікаційними метриками, а й для оцінювання того, як запропонована DST-агрегація поводиться порівняно з окремими компонентами та простішими способами об'єднання їхніх числових оцінок.

Для аналізу розглядалися п'ять варіантів оцінювання: окрема нечітка система виведення FIS, окрема байєсівська мережа BN, просте середнє FIS-BN, зважене середнє та гібридна модель Hybrid DST. Для FIS використовувалася оцінка R_{FIS}, а для BN – числова інтерпретація апостеріорного розподілу R_{BN}, визначена в підрозділі 3.5.

Просте середнє задавалося формулою

$$R_{\text{avg}} = \frac{1}{2} R_{\text{FIS}} + \frac{1}{2} R_{\text{BN}}.$$

Зважене середнє обчислювалося як

$$R_{\text{wavg}} = 0.4 R_{\text{FIS}} + 0.6 R_{\text{BN}}.$$

Останній варіант використовується як простий baseline, що надає більшу вагу ймовірнісній компоненті. На відміну від Hybrid DST, такі усереднення не враховують ані конфлікту між джерелами свідчень, ані залишкової невизначеності. Тому вони є корисним орієнтиром для перевірки того, чи дає доказова агрегація перевагу порівняно зі звичайним числовим об'єднанням оцінок.

Якість моделей оцінювалася за відповідністю їхньої дискретної

інтерпретації експериментальній змінній `risk_level`. Для тестової вибірки

$$\{(z_i,y_i)\}_{i=1}^{N},$$

де $y_i \in \{Low,Medium,High\}$ – експериментальний рівень ризику, а $\widehat{y}_i$ – прогнозований рівень, обчислювалися метрики Accuracy та Macro F1:

$$Accuracy = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{\widehat{y}_i = y_i\},$$

$$Macro\,F1 = \frac{1}{3}\sum_{c\in\{Low,Medium,High\}} F1_c.$$

Метрика Accuracy відображає загальну частку правильно визначених рівнів ризику. Macro F1 усереднює якість за трьома класами та є більш чутливою до нерівномірності результатів між рівнями ризику.

Результати порівняння наведено в табл. 3.7 та на рис. 3.1. Найвищі значення Accuracy та Macro F1 отримано для байєсівської мережі. Це узгоджується з тим, що BN-компонента параметризується за навчальною вибіркою з використанням експериментальної змінної `risk_level`, тоді як FIS-компонента ґрунтується на фіксованій базі нечітких правил і не оптимізується за мітками класів.

Hybrid DST перевищує окрему FIS-компоненту та просте рівновагове середнє FIS–BN, однак поступається BN і зваженому середньому за обома класифікаційними метриками. Отже, у межах базового експерименту запропонована DST-агрегація не забезпечує максимального значення Accuracy або Macro F1. Це свідчить про те, що сама процедура доказового комбінування не гарантує автоматичного покращення класифікаційної якості порівняно з найсильнішим окремим джерелом.

Цей результат є важливим для коректної інтерпретації моделі. Перевага Hybrid DST полягає не в автоматичному підвищенні класифікаційної точності, а в можливості формально поєднувати

різнорідні джерела свідчень, оцінювати конфлікт між ними та зберігати інформацію про залишкову невизначеність.

Таблиця 3.7 – Порівняння моделей на тестовій вибірці

Модель	**Accuracy**	**Macro F1**
FIS	0.600	0.571
BN	0.823	0.787
Average FIS–BN	0.686	0.651
Weighted Average 0.4/0.6	0.713	0.685
Hybrid DST	0.695	0.661

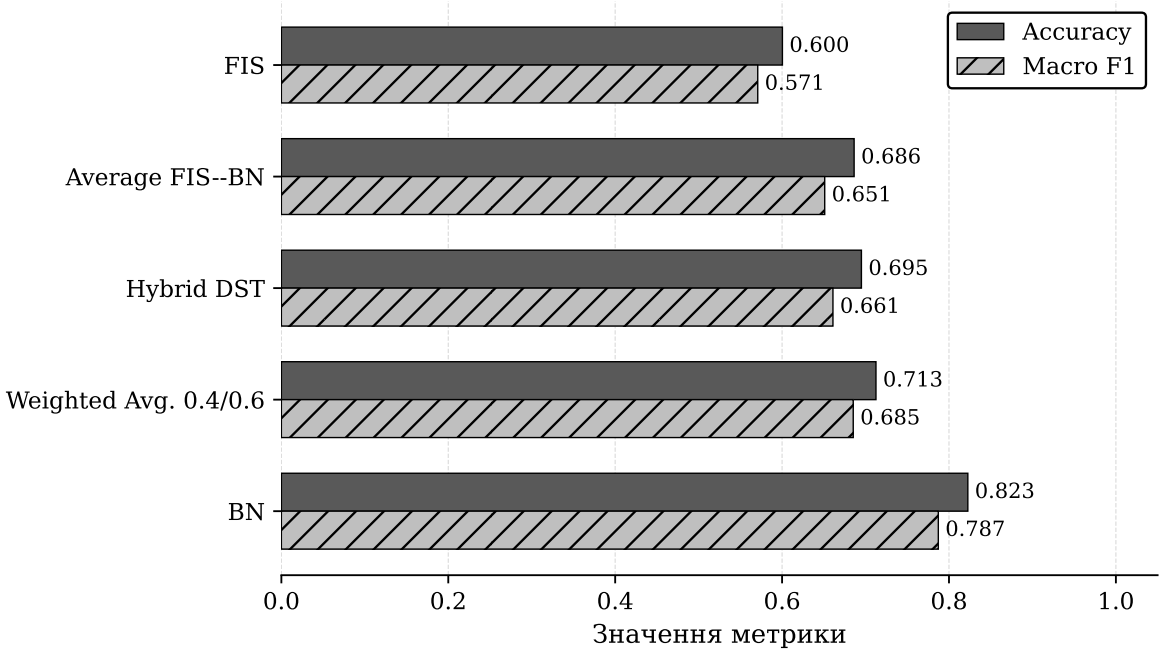

Рисунок 3.1 – Порівняння якості FIS, BN, average-baseline та Hybrid DST на тестовій вибірці

Таким чином, результати базового експерименту показують, що BN є найсильнішою компонентою з погляду класифікаційної відповідності змінній `risk_level`. Водночас Hybrid DST виконує іншу функцію в межах запропонованої моделі: він формує фінальний індекс ризику з урахуванням двох джерел свідчень, коефіцієнта конфлікту та маси невизначеності.

Подальший аналіз зосереджується не лише на класифікаційних

метриках, а й на поведінці R_{final}, розподілі результатів за класами атак, конфлікті між FIS і BN та стійкості моделі до неповноти факторної інформації.

3.8 Аналіз фінального індексу ризику

Класифікаційні метрики, розглянуті в попередньому підрозділі, оцінюють лише дискретну відповідність прогнозованого рівня ризику експериментальній змінній `risk_level`. Однак у запропонованій моделі важливим результатом є не тільки категорія Low, $Medium$ або $High$, а й фінальний нормований індекс ризику

$$R_{\text{final}}(z) \in [0,1],$$

отриманий після DST-агрегації та pignistic transformation. Тому окремо перевіряється, чи має цей індекс змістовну кількісну поведінку відносно експериментально заданих рівнів ризику.

Для кожного рівня $c \in \{Low, Medium, High\}$ обчислювалися середні значення оцінок

$$\overline{R}_{\text{FIS}}(c), \qquad \overline{R}_{\text{BN}}(c), \qquad \overline{R}_{\text{final}}(c).$$

Зокрема, для фінального індексу використовувалася формула

$$\overline{R}_{\text{final}}(c) = \frac{1}{N_c} \sum_{i:\, y_i = c} R_{\text{final}}(z_i),$$

де y_i – експериментальний рівень ризику для спостереження z_i, а N_c – кількість спостережень відповідного рівня. Аналогічно усереднювалися коефіцієнт конфлікту K_{conflict} та маса невизначеності $m(\Theta)$.

Результати наведено в табл. 3.8 та на рис. 3.2. Для фінального індексу отримано впорядкування

$$\overline{R}_{\text{final}}(Low) < \overline{R}_{\text{final}}(Medium) < \overline{R}_{\text{final}}(High).$$

Середнє значення R_{final} для низького рівня ризику становить 0.350, для середнього – 0.486, а для високого – 0.756. Отже, фінальний індекс узгоджується з очікуваною шкалою ризику: у середньому він зростає при переході від нормального трафіку та менш критичних сценаріїв до інтенсивніших і потенційно небезпечніших атак.

Таке впорядкування слід інтерпретувати саме як групову властивість середніх значень. Воно не гарантує безпомилкового ранжування кожного окремого спостереження. Значення R_{final} є агрегованим індексом, який залежить від виходів FIS і BN, коефіцієнта конфлікту та маси невизначеності. Тому для окремих потоків можливі ситуації, коли спостереження з нижчим експериментальним рівнем ризику отримує більший числовий індекс, ніж окреме спостереження з вищим рівнем. У цьому сенсі аналіз R_{final} доповнює, але не замінює класифікаційні метрики.

Порівняння окремих компонентів показує, що BN формує найбільш виражене розділення між рівнями ризику: середнє значення R_{BN} зростає від 0.412 для рівня Low до 0.848 для рівня $High$. Це узгоджується з результатами попереднього підрозділу, де BN мала найвищі значення Accuracy та Macro F1.

Нечітка система дає більш згладжені оцінки: $\overline{R}_{\text{FIS}}$ змінюється від 0.408 до 0.625. Така поведінка є очікуваною, оскільки FIS ґрунтується на фіксованих лінгвістичних правилах і не навчається безпосередньо на змінній `risk_level`.

Гібридний індекс R_{final} займає проміжне положення між цими двома компонентами. Він зберігає загальне зростання ризику від Low до $High$, але не повністю відтворює числову поведінку BN, оскільки під час DST-агрегації враховуються також нечітке джерело свідчень і залишкова невизначеність.

Зокрема, для рівня $Medium$ середня маса $m(\Theta)$ є найбільшою і становить 0.228. Це вказує на більшу невизначеність при віднесенні частини спостережень до проміжного рівня ризику. Для рівня $High$

середні значення конфлікту та маси невизначеності є нижчими, що відповідає більш узгодженій підтримці високого ризику з боку компонент моделі.

Таблиця 3.8 – Середні значення показників ризику за рівнями

Рівень ризику	**Кількість**	R_{FIS}	R_{BN}	R_{final}	K_{conflict}	$m(\Theta)$
Low	9 000	0.408	0.412	0.350	0.330	0.174
Medium	13 150	0.439	0.510	0.486	0.278	0.228
High	24 477	0.625	0.848	0.756	0.192	0.133

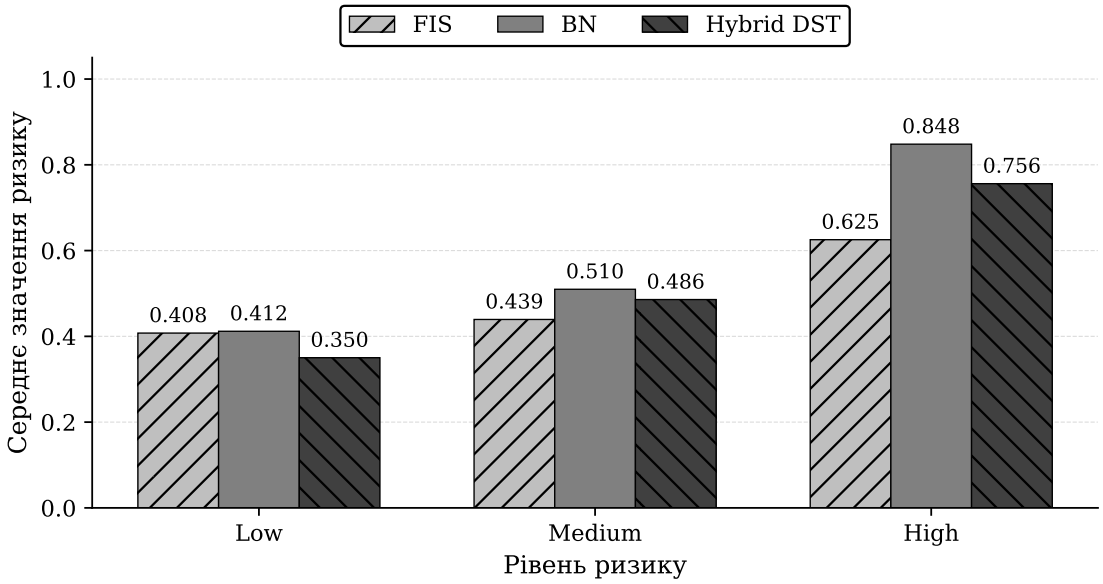

Рисунок 3.2 – Середні значення ризику за рівнями на тестовій вибірці

Отже, аналіз фінального індексу показує, що R_{final} є змістовною нормованою оцінкою ризику в межах прийнятого експериментального сценарію. Його середні значення узгоджені з рівнями Low, $Medium$ та $High$. Водночас цей індекс не слід трактувати як абсолютну ймовірність атаки або величину очікуваних збитків. У цій роботі він є інтегральним показником, який поєднує нечітку оцінку, байєсівський апостеріорний розподіл і результат доказової агрегації з урахуванням конфлікту та невизначеності.

3.9 Аналіз результатів за класами атак

Після аналізу результатів на рівні агрегованих категорій *Low*, *Medium* та *High* доцільно окремо розглянути поведінку моделі для конкретних класів CIC-IDS2017. Такий аналіз має діагностичний характер, оскільки дозволяє перевірити, чи узгоджується фінальний індекс ризику R_{final} не лише з укрупненою змінною `risk_level`, а й зі змістовною інтерпретацією окремих типів мережевої активності.

Нехай $\mathcal{L}_{\text{exp}}$ позначає множину класів, включених до базового експериментального сценарію, а $\ell \in \mathcal{L}_{\text{exp}}$ – один із таких класів. Для тестової вибірки введемо множину індексів

$$\mathcal{I}_\ell = \{i : \text{Label}(z_i) = \ell\},$$

де $\text{Label}(z_i)$ – початкова мітка класу для спостереження z_i. Ця мітка не використовується як вхідна ознака моделі, а застосовується лише для групування результатів під час діагностичного аналізу.

Для кожного класу ℓ обчислювалися середні значення виходів окремих компонентів, фінального індексу ризику, коефіцієнта конфлікту та маси невизначеності:

$$\overline{R}_{\text{FIS}}(\ell), \qquad \overline{R}_{\text{BN}}(\ell), \qquad \overline{R}_{\text{final}}(\ell), \qquad \overline{K}_{\text{conflict}}(\ell), \qquad \overline{m}_\Theta(\ell).$$

Зокрема, середнє значення фінального індексу для класу ℓ визначалося як

$$\overline{R}_{\text{final}}(\ell) = \frac{1}{|\mathcal{I}_\ell|} \sum_{i \in \mathcal{I}_\ell} R_{\text{final}}(z_i),$$

а середня маса невизначеності – як

$$\overline{m}_\Theta(\ell) = \frac{1}{|\mathcal{I}_\ell|} \sum_{i \in \mathcal{I}_\ell} m_i(\Theta),$$

де $m_i(\Theta)$ є масою, призначеною повній множині гіпотез після DST-агрегації

для спостереження z_i.

Результати подано на рис. 3.3 у вигляді діагностичної карти, яка відображає середні значення R_{FIS}, R_{BN}, R_{final}, K_{conflict} та $m(\Theta)$ для кожного класу. Такий формат дає змогу одночасно порівняти фінальний ризик, внесок окремих джерел свідчень і рівень їхньої узгодженості.

Отримані значення загалом відповідають очікуваній інтерпретації класів у межах прийнятого експериментального зіставлення з рівнями ризику. Найвищі значення $\overline{R}_{\text{final}}(\ell)$ спостерігаються для DoS-атак, зокрема `DoS Hulk`, `DoS GoldenEye`, `DoS Slowhttptest` та `DoS slowloris`, що узгоджується з їх проявом через інтенсивне або аномальне мережеве навантаження у факторах T та I.

Для класу `DDoS` фінальний індекс також залишається високим, хоча нижчим, ніж для частини DoS-класів. Це свідчить про неоднорідність атак категорії $High$ у просторі факторів (T,V,A,I) і показує, що модель частково зберігає відмінності між окремими типами високоризикової активності.

Класи `FTP-Patator`, `SSH-Patator` та `PortScan`, віднесені до рівня $Medium$, мають проміжні значення $\overline{R}_{\text{final}}(\ell)$. Нижчий ризик для `PortScan` порівняно з Patator-класами свідчить, що сканування портів менш виразно проявляється в ознаках технічного впливу та обсягу переданих даних. Клас `BENIGN` має найнижче середнє значення, що відповідає його інтерпретації як нормального трафіку.

Середні значення $\overline{K}_{\text{conflict}}(\ell)$ та $\overline{m}_{\Theta}(\ell)$ додатково характеризують узгодженість FIS- та BN-компонентів. Підвищені значення цих показників означають, що експертно-лінгвістична та ймовірнісна компоненти не повністю збігаються у підтримці рівнів ризику. У цьому підрозділі вони використовуються як діагностичні величини; детальніше їхню роль у поведінці Hybrid DST розглянуто далі.

Отже, аналіз за окремими класами показує, що R_{final} узгоджується не лише з агрегованими рівнями Low, $Medium$ та $High$, а також з відмінностями між конкретними типами мережевої активності. Водночас ці результати слід трактувати як внутрішню діагностику моделі на

вибраних класах CIC-IDS2017, а не як універсальне ранжування кіберзагроз поза межами використаного датасету.

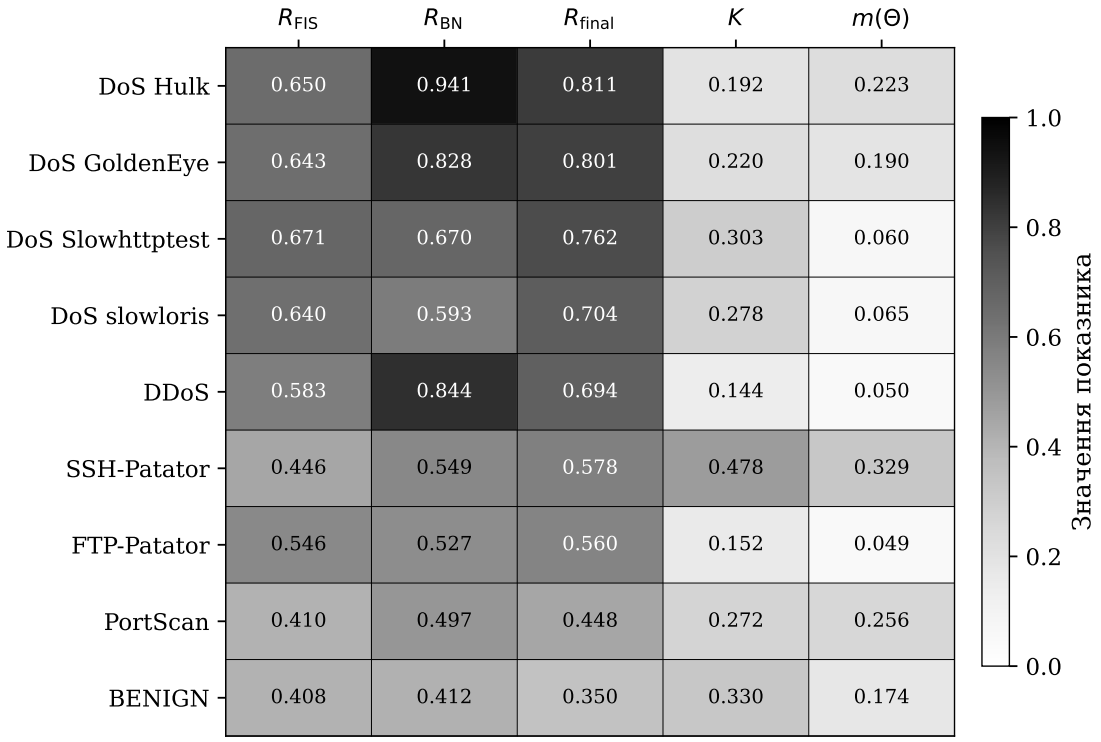

Рисунок 3.3 – Діагностична карта середніх значень ризику, конфлікту та невизначеності за класами

3.10 Аналіз конфлікту між джерелами свідчень

Однією з основних причин використання теорії доказів у запропонованій моделі є можливість явно фіксувати ситуації, у яких FIS- та BN-компоненти підтримують різні рівні ризику. На відміну від простого усереднення числових оцінок, DST-агрегація дозволяє не лише отримати фінальний індекс R_{final}, а й проаналізувати коефіцієнт конфлікту K_{conflict}, вибір правила комбінування та масу невизначеності $m(\Theta)$.

Для кожного тестового спостереження z_i обчислювалися значення

$$K_{\text{conflict}}(z_i), \qquad m_i(\Theta),$$

а також фіксувалося правило комбінування, застосоване відповідно до порога

$$\tau = 0.5.$$

Якщо $K_{\text{conflict}}(z_i) \leqslant \tau$, використовувалося правило Демпстера. Якщо $K_{\text{conflict}}(z_i) > \tau$, застосовувалося правило Ягера, за яким конфліктна маса переноситься на повну множину гіпотез Θ.

Підсумкові характеристики конфлікту на тестовій вибірці наведено в табл. 3.9. Середнє значення коефіцієнта конфлікту становить 0.243, медіанне – 0.170, а максимальне – 0.640. Отже, для більшості спостережень FIS і BN не перебувають у критичному конфлікті. Водночас у тестовій вибірці наявна помітна група випадків, у яких підтримка різних рівнів ризику з боку двох джерел істотно розходиться.

Таблиця 3.9 – Характеристики конфлікту та невизначеності

Показник	**Значення**
Кількість записів	46 627
Середнє значення K_{conflict}	0.243
Медіана K_{conflict}	0.170
Максимальне значення K_{conflict}	0.640
Кількість випадків із правилом Демпстера	36 564
Частка правила Демпстера	0.784
Кількість випадків із правилом Ягера	10 063
Частка правила Ягера	0.216
Середнє значення $m(\Theta)$	0.168
Медіана $m(\Theta)$	0.048
Максимальне значення $m(\Theta)$	0.680

Правило Демпстера було застосовано для 36 564 записів, тобто приблизно для 78.4% тестової вибірки. Правило Ягера застосовувалося для 10 063 записів, що становить близько 21.6%. Таким чином, перехід до консервативного правила комбінування не є поодиноким винятком: він спрацьовує для суттєвої частини спостережень із підвищеним конфліктом між джерелами свідчень. Це підтверджує доцільність використання адаптивної схеми DST-агрегації, у якій спосіб комбінування залежить від

рівня узгодженості між FIS- та BN-компонентами.

Середнє значення маси невизначеності $m(\Theta)$ дорівнює 0.168, тоді як медіана становить лише 0.048. Різниця між середнім і медіаною вказує на асиметричний характер розподілу: для значної кількості спостережень залишкова невизначеність є невеликою, але в окремих випадках вона суттєво зростає. Максимальне значення $m(\Theta) = 0.680$ відповідає ситуаціям, у яких конфлікт між джерелами є високим, а значна частина маси після застосування правила Ягера переноситься на Θ.

На рис. 3.4 подано розподіл коефіцієнта конфлікту та частки використання правил комбінування. Цей рисунок доповнює табл. 3.9, оскільки показує не лише узагальнені статистики, а й структуру переходу між режимами DST-агрегації. Значення K_{conflict}, що не перевищують поріг τ, відповідають області застосування правила Демпстера, тоді як правий хвіст розподілу формує область застосування правила Ягера.

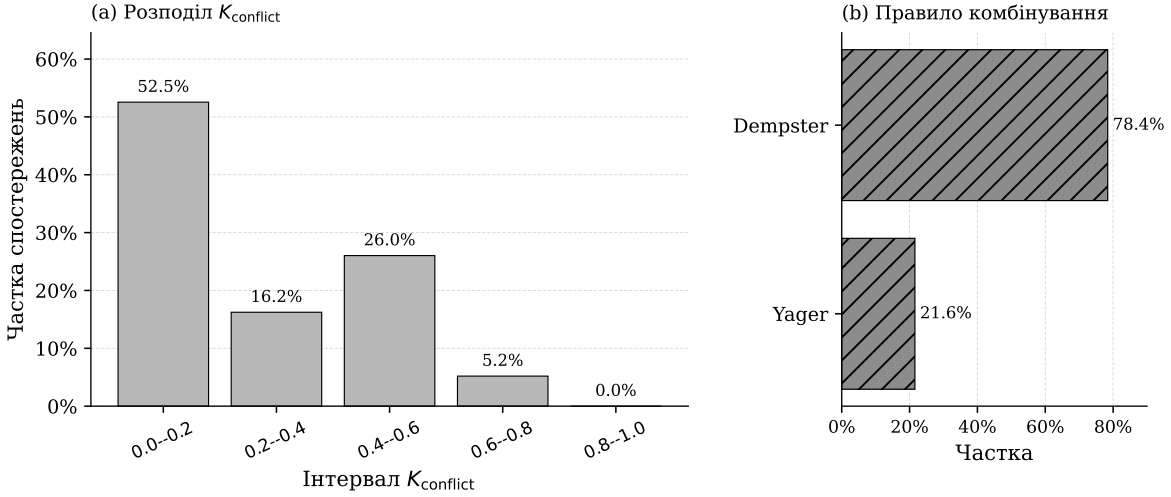

Рисунок 3.4 – Розподіл коефіцієнта конфлікту та вибір правила DST

Для змістовної інтерпретації висококонфліктних ситуацій у табл. 3.10 наведено репрезентативні приклади спостережень, для яких FIS- та BN-компоненти підтримують різні рівні ризику. У таблиці подано початковий клас трафіку, експериментальний рівень ризику, дискретні інтерпретації FIS і BN, значення коефіцієнта конфлікту та фінальний індекс R_{final}.

Таблиця 3.10 – Приклади конфлікту між FIS та BN

Тип випадку	**Клас**	**Рівень**	**FIS**	**BN**	K_{conflict}	R_{final}
Medium vs High	`DoS Hulk`	High	Medium	High	0.640	0.575
Medium vs High	`BENIGN`	Low	Medium	High	0.640	0.575
Medium vs Low	`DoS Hulk`	High	Medium	Low	0.639	0.475
High vs Low	`BENIGN`	Low	High	Low	0.637	0.478
High vs Medium	`BENIGN`	Low	High	Medium	0.634	0.579
Low vs High	`DoS Hulk`	High	Low	High	0.594	0.478
High vs Medium	`PortScan`	Medium	High	Medium	0.591	0.586

Наведені приклади демонструють типову поведінку Hybrid DST за умов суперечливих свідчень. Якщо FIS і BN підтримують різні гіпотези, наприклад $Medium$ проти $High$ або $High$ проти Low, коефіцієнт конфлікту наближається до верхніх значень, зафіксованих у тестовій вибірці.

У таких випадках застосування правила Ягера запобігає штучному посиленню однієї з конфліктних гіпотез і збільшує частку невизначеності. У результаті фінальний індекс часто переходить у проміжну область, що відображає обережну позицію моделі за наявності розбіжностей між джерелами.

Важливо, що конфлікт не слід інтерпретувати лише як помилку однієї з компонент. Він може виникати через різну природу FIS і BN. Нечітка система відображає експертно задану лінгвістичну логіку факторів T,V,A,I, тоді як байєсівська мережа спирається на статистичні зв'язки між дискретизованими факторами та експериментальною змінною `risk_level`. Тому розбіжність між їхніми висновками є інформативною характеристикою ситуації, а не лише недоліком моделі.

Таким чином, результати аналізу підтверджують практичну роль DST-шару в гібридній моделі. Він не лише формує фінальний індекс ризику, а й дозволяє явно виділити випадки, у яких джерела свідчень узгоджуються недостатньо. Саме наявність показників K_{conflict} та $m(\Theta)$ відрізняє Hybrid DST від простого числового усереднення FIS- і BN-оцінок та забезпечує додаткову інтерпретованість результату.

3.11 Аналіз чутливості та неповноти факторів ризику

Після базового порівняння моделей доцільно перевірити, наскільки результати Hybrid DST залежать від параметрів доказової агрегації та повноти факторного подання. У цьому підрозділі розглянуто два аспекти: чутливість до коефіцієнтів надійності джерел ρ_{FIS}, ρ_{BN} та поведінку моделі за умов вилучення окремих факторів ризику.

Спочатку проаналізовано вплив коефіцієнтів надійності. Вони визначають, яка частина маси кожного джерела призначається конкретним гіпотезам, а яка переноситься на множину невизначеності Θ. У базовому сценарії використовувалися однакові значення

$$\rho_{\text{FIS}} = \rho_{\text{BN}} = 0.8.$$

Додатково розглядалися конфігурації з більшою довірою до BN- або FIS-компоненти, а також сильніші зміщення на користь одного з джерел. Інші параметри моделі в усіх сценаріях залишалися незмінними.

Результати аналізу наведено в табл. 3.11. Найкращі значення Accuracy та Macro F1 отримано для базового сценарію з однаковими коефіцієнтами надійності. Збільшення довіри до BN не покращило загальні метрики, хоча BN-компонента окремо була найточнішою в базовому експерименті. Це пов'язано з тим, що зміна ρ_{FIS} і ρ_{BN} впливає не лише на вагу джерел, а й на перерозподіл маси між гіпотезами та множиною Θ, тобто змінює механізм DST-агрегації. Варіанти з пріоритетом FIS демонструють нижчі результати, що узгоджується з меншою точністю FIS-компоненти.

Окремо було проведено експеримент із неповнотою факторної інформації. Його мета полягає в перевірці того, як модель поводиться, якщо частина факторів T,V,A,I недоступна. Така ситуація є практично важливою, оскільки в реальних системах моніторингу частина ознак може бути відсутньою, пошкодженою або ненадійною.

Таблиця 3.11 – Чутливість Hybrid DST до надійності джерел

Сценарій	ρ_{FIS}	ρ_{BN}	**Accuracy**	**Macro F1**	$m(\Theta)$
`baseline_equal`	0.8	0.8	0.695	0.661	0.168
`bn_priority`	0.7	0.9	0.658	0.625	0.131
`fis_priority`	0.9	0.7	0.643	0.608	0.131
`strong_bn_priority`	0.6	0.9	0.670	0.634	0.080
`strong_fis_priority`	0.9	0.6	0.625	0.590	0.080

У межах експерименту розглядалися сценарії

`missing_T`, `missing_V`, `missing_I`, `missing_V_I`.

Фактор A не вилучався, оскільки в базовій постановці він задається сценарно і є сталим.

Для відсутніх факторів у FIS використовувалася нейтральна фазифікація: якщо фактор Q відсутній, то для всіх його термів покладалося

$$\mu_{Q,l}(Q) = 1, \qquad l \in \{Low, Medium, High\}.$$

За T-норми мінімуму це означає, що відсутній фактор не обмежує активацію правил. У BN відсутній фактор не фіксувався як evidence, а маргіналізувався за можливими станами відповідної дискретної змінної. Таким чином, обидві компоненти обробляли неповноту не через довільне заповнення значенням Low, $Medium$ або $High$, а через зменшення інформативності відповідного джерела.

Крім того, для сценаріїв неповноти використовувалися ефективні коефіцієнти надійності

$$\rho_{\text{eff}} = \rho_{\text{base}} \cdot \frac{d_{\text{avail}}}{4},$$

де d_{avail} – кількість доступних факторів із множини $\{T, V, A, I\}$, а $\rho_{\text{base}} = 0.8$. Такий підхід відображає припущення, що зменшення кількості доступних факторів має знижувати довіру до сформованих свідчень і збільшувати частку невизначеності.

Результати Hybrid DST за умов неповноти наведено в табл. 3.12. Вилучення одного фактора знижує якість моделі, але не руйнує її працездатність. Найбільше погіршення спостерігається за відсутності V та I: Macro F1 знижується до 0.442, а $m(\Theta)$ зростає до 0.389. Це підтверджує важливість експозиції та потенційного впливу для розрізнення рівнів ризику.

Таблиця 3.12 – Hybrid DST за умов неповноти факторів

Сценарій	**Відсутні фактори**	**Accuracy**	**Macro F1**	$m(\Theta)$
`full_data`	–	0.695	0.661	0.168
`missing_T`	T	0.634	0.565	0.201
`missing_V`	V	0.646	0.597	0.198
`missing_I`	I	0.652	0.612	0.198
`missing_V_I`	V,I	0.596	0.442	0.389

Для порівняння окремо оцінено FIS-, BN- та Hybrid DST-моделі в тих самих сценаріях неповноти; результати наведено в табл. 3.13.

Таблиця 3.13 – Порівняння моделей за умов неповноти факторів

Сценарій	**Модель**	**Accuracy**	**Macro F1**	**Середній ризик**
Повні дані	FIS	0.600	0.571	0.531
	BN	0.823	0.787	0.669
	Hybrid DST	0.695	0.661	0.601
Без фактора T	FIS	0.500	0.409	0.565
	BN	0.761	0.721	0.665
	Hybrid DST	0.634	0.565	0.618
Без фактора V	FIS	0.439	0.364	0.521
	BN	0.697	0.665	0.668
	Hybrid DST	0.646	0.597	0.573
Без фактора I	FIS	0.597	0.545	0.573
	BN	0.756	0.699	0.667
	Hybrid DST	0.652	0.612	0.626
Без факторів V,I	FIS	0.613	0.454	0.570
	BN	0.602	0.417	0.666
	Hybrid DST	0.596	0.442	0.601

За повних даних і за відсутності одного фактора BN здебільшого зберігає найвищі класифікаційні метрики. Натомість одночасне вилучення

V та I суттєво погіршує якість усіх моделей і зменшує різницю між ними, що підтверджує важливість цих факторів для розділення середнього та високого рівнів ризику.

На рис. 3.5 узагальнено вплив неповноти факторів на якість моделей і невизначеність фінального висновку.

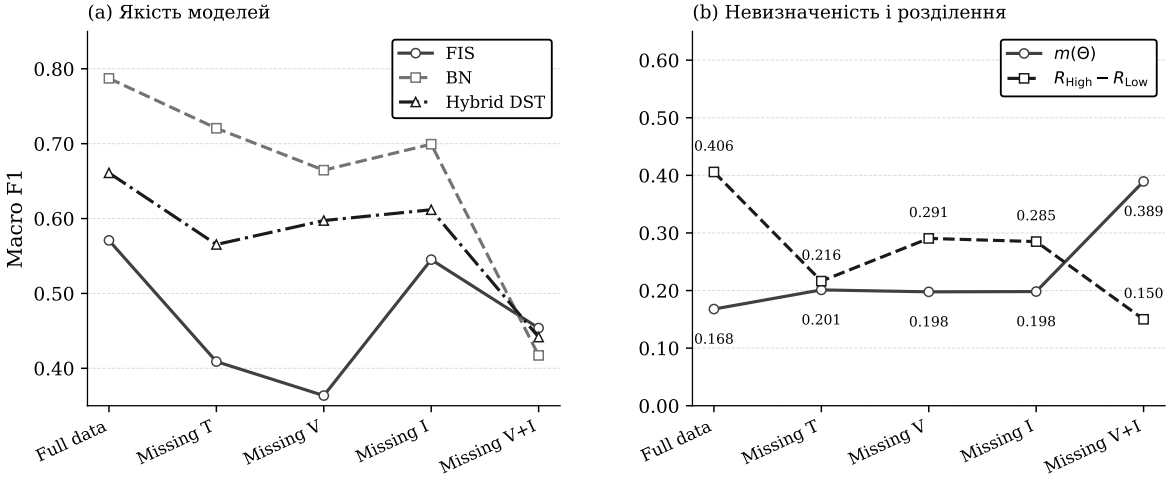

Рисунок 3.5 – Вплив неповноти факторів на результати моделей

Особливо показовим є сценарій `missing_V_I`, у якому одночасно знижується класифікаційна якість і зростає середня маса $m(\Theta)$. Отже, DST-агрегація не лише формує числовий компроміс між FIS і BN, а також відображає зменшення надійності висновку за умов недостатності факторної інформації.

Таким чином, аналіз чутливості показує, що базова конфігурація коефіцієнтів надійності є найкращою серед розглянутих варіантів за класифікаційними метриками. Експеримент із неповнотою факторів підтверджує, що запропонована модель реагує на втрату інформації не лише зниженням точності, а й збільшенням маси невизначеності. Це є важливою властивістю гібридної схеми, оскільки в задачах оцінювання кіберризику неповнота даних має відображатися не тільки у фінальному числовому індексі, а й у показниках довіри до отриманого висновку.

3.12 Обмеження реалізації та інтерпретація результатів

Отримані результати потрібно інтерпретувати з урахуванням того, що практична частина роботи є експериментальною перевіркою запропонованої математичної схеми, а не завершеною промисловою системою управління кіберризиками. Метою реалізації було перевірити, чи можна узгоджено поєднати FIS-, BN- та DST-компоненти в єдиний обчислювальний pipeline, а також дослідити поведінку фінального індексу ризику, конфлікту між джерелами свідчень і маси невизначеності на мережевих даних.

Насамперед слід зазначити, що Hybrid DST не перевищила байєсівську мережу за класифікаційними метриками. У базовому експерименті BN показала найвищі значення Accuracy та Macro F1, а зважене середнє FIS-BN також перевищило Hybrid DST за цими показниками. Отже, якщо розглядати задачу виключно як класифікацію мережевих потоків за експериментальними рівнями ризику, то BN-компонента є сильнішою моделлю в межах проведеного сценарію.

Тому Hybrid DST не слід інтерпретувати як підхід, що автоматично максимізує класифікаційну точність. Його роль полягає в іншому: у формальному поєднанні різнорідних джерел свідчень, явному врахуванні конфлікту між ними та збереженні інформації про залишкову невизначеність.

Окреме обмеження стосується природи цільової змінної `risk_level`. Вона була сформована на основі мітки класу `Label` шляхом експериментального зіставлення класів трафіку з рівнями Low, $Medium$ та $High$. Така змінна є зручною для навчання таблиць умовних імовірностей BN та оцінювання результатів, однак вона не є повною бізнес-оцінкою кіберризику. У датасеті CIC-IDS2017 відсутні дані про фінансові втрати, реальні наслідки інцидентів, критичність бізнес-процесів або контекст конкретної організації [18, 19]. Відповідно,

R_{final} у цій роботі слід трактувати як нормований експериментальний індекс ризику в межах заданого сценарію, а не як абсолютну оцінку збитків або універсальну міру критичності інциденту.

Ще одне обмеження пов'язане з інтерпретацією фактора V. У теоретичній постановці він відповідає вразливості або експозиції цільового середовища. Проте використаний датасет не містить інформації про CVE-ідентифікатори, версії програмного забезпечення, конфігурацію хостів, наявність оновлень або результати сканування вразливостей. Тому в експериментальній реалізації V не є безпосередньою оцінкою фактичної технічної вразливості активу. Він використовується як proxy-показник експозиції та TCP-поведінки з'єднання, побудований на основі доступних потокових ознак. Це обмежує предметну інтерпретацію результатів і потребує уточнення в разі перенесення моделі на реальну інфраструктуру.

Подібне застереження стосується фактора A, який описує критичність активу. Оскільки CIC-IDS2017 не містить відомостей про бізнес-роль вузлів або цінність активів, у базовому сценарії значення цього фактора було задано сценарно:

$$A = 0.60.$$

Це дозволило зберегти повну структуру факторного відображення

$$\phi(z) = (T,V,A,I),$$

узгоджену з математичною постановкою задачі. Водночас у проведеному експерименті фактор A не варіювався між спостереженнями, тому його вплив на диференціацію ризику не міг бути повноцінно перевірений на даних. У реальному застосуванні цей фактор має визначатися на основі інформації про роль активу, мережевий сегмент, важливість сервісу або інші контекстні характеристики.

Також слід враховувати сценарний характер параметризації моделі. Ваги у факторному відображенні ϕ, форма функцій належності, правила

FIS, пороги дискретизації, коефіцієнти надійності ρ_{FIS}, ρ_{BN} та поріг конфлікту τ були зафіксовані в межах базового експериментального сценарію. Частина цих параметрів має експертний або сценарний характер. Аналіз чутливості показав, що зміна коефіцієнтів надійності впливає на класифікаційні метрики та середню масу невизначеності, тому такі параметри не слід вважати універсальними. Для іншого датасету або іншого середовища моніторингу їх потрібно переоцінювати або додатково обґрунтовувати.

Ще одним обмеженням є те, що експериментальна перевірка виконувалася на одному наборі даних. CIC-IDS2017 є поширеним датасетом для дослідження мережевих атак, однак він не охоплює всіх можливих сценаріїв реальної інфраструктури, сучасних типів загроз, варіантів поведінки користувачів і конфігурацій захисних систем. Тому отримані числові результати не слід безпосередньо переносити на інші середовища без повторної параметризації, валідації та перевірки стійкості моделі.

Попри наведені обмеження, результати практичної частини підтверджують працездатність запропонованої математичної схеми. У межах експерименту було реалізовано послідовний перехід від мережевого спостереження до фінального індексу ризику:

$$\begin{aligned} z \;\longrightarrow\; \phi(z) \;\longrightarrow\; \big(R_{\text{FIS}}, P(H \mid e)\big) \\ \longrightarrow\; \big(m_{\text{FIS}}, m_{\text{BN}}\big) \;\longrightarrow\; m \;\longrightarrow\; \text{BetP} \;\longrightarrow\; R_{\text{final}}. \end{aligned}$$

Також показано, що фінальний індекс ризику має змістовну поведінку за експериментальними рівнями Low, $Medium$ та $High$. Аналіз конфлікту та неповноти факторів продемонстрував, що Hybrid DST дозволяє не лише отримати числовий результат, а й оцінити ступінь узгодженості джерел свідчень та рівень залишкової невизначеності.

Таким чином, практична реалізація підтверджує, що запропонована гібридна модель може бути використана як інтерпретований обчислювальний механізм оцінювання кіберризику в умовах неповних і

потенційно суперечливих даних. Водночас її результати мають розглядатися в межах прийнятого експериментального сценарію, а не як універсальна оцінка кіберризику для будь-якої інформаційної системи.

Висновки до розділу 3

У третьому розділі виконано програмну реалізацію та експериментальну перевірку гібридної моделі оцінювання кіберризику, формалізованої в попередньому розділі. На основі вибраних класів датасету CIC-IDS2017 було підготовлено експериментальний набір даних, сформовано змінну `risk_level`, побудовано факторне відображення

$$\phi(z) = (T,V,A,I),$$

а також реалізовано нечітку систему виведення, байєсівську мережу, процедури формування базових функцій маси та DST-агрегацію з переходом до фінального індексу R_{final}.

Результати базового експерименту показали, що байєсівська мережа є найсильнішою окремою компонентою за класифікаційними метриками Accuracy та Macro F1. Hybrid DST не перевищує BN за точністю, тому її не слід інтерпретувати як модель, що автоматично максимізує класифікаційну якість. Її основна роль полягає у формальному узгодженні різнорідних джерел свідчень, явному врахуванні конфлікту між ними та збереженні інформації про залишкову невизначеність.

Аналіз фінального індексу показав, що R_{final} має змістовну поведінку в межах прийнятого експериментального сценарію: його середні значення зростають при переході від рівня Low до $Medium$ і $High$. Додатковий аналіз за окремими класами атак підтвердив, що модель відображає не лише укрупнений поділ на три рівні ризику, а й певні відмінності між типами мережевої активності всередині цих рівнів.

Аналіз конфлікту між FIS- та BN-компонентами показав, що для

більшості спостережень конфлікт є помірним, однак для суттєвої частини тестової вибірки застосовується правило Ягера. Це підтверджує практичну роль DST-шару: він не лише формує фінальний індекс ризику, а й надає додаткові діагностичні характеристики у вигляді K_{conflict} та $m(\Theta)$.

Експеримент із неповнотою факторної інформації показав, що вилучення окремих факторів погіршує якість оцінювання, а одночасна відсутність факторів V та I призводить до найбільшого зниження Macro F1 і зростання середньої маси невизначеності. Це узгоджується з постановкою задачі роботи: у разі втрати важливої інформації модель має відображати не лише зміну фінальної оцінки ризику, а й зниження надійності висновку.

Водночас отримані результати слід розглядати з урахуванням обмежень експериментального сценарію. Змінна `risk_level` є експериментальною розміткою, сформованою на основі класів трафіку, а не повною бізнес-оцінкою кіберризику. Фактор V у межах CIC-IDS2017 є proxy-показником експозиції, а не реальною CVE-вразливістю, тоді як фактор A задається сценарно і не варіюється між спостереженнями. Тому числові результати не слід безпосередньо переносити на інші інформаційні системи без повторної параметризації та валідації.

Таким чином, у третьому розділі підтверджено працездатність запропонованої гібридної схеми. Практична реалізація показала, що модель може використовуватися як інтерпретований обчислювальний механізм оцінювання кіберризику в умовах неповноти та суперечливості даних, якщо її результати розглядати в межах чітко визначеного експериментального сценарію.

ВИСНОВКИ

У дипломній роботі розглянуто задачу побудови гібридної математичної моделі оцінювання ризиків кіберзагроз в умовах невизначеності даних. Мету роботи досягнуто: побудовано, формалізовано та експериментально перевірено модель, у якій нечітка система виведення та байєсівська мережа використовуються як різнорідні джерела свідчень, а їх результати агрегуються засобами теорії доказів Демпстера–Шейфера.

У ході дослідження проаналізовано сучасні стандарти, методології та математичні підходи до оцінювання ризиків інформаційної безпеки. Показано, що ISO/IEC 27005, NIST, Open FAIR та інші підходи задають процесну й термінологічну основу ризик-менеджменту, однак не визначають єдиного обчислювального механізму для поєднання експертних оцінок, ймовірнісних залежностей і конфліктних свідчень. Це обґрунтовує доцільність використання гібридної математичної моделі.

Виконано формалізацію простору гіпотез рівня ризику та множини високорівневих факторів кіберризику. Ризик подано через фактори активності загрози, вразливості або експозиції, критичності активу та потенційного впливу. Також визначено процедуру переходу від низькорівневих ознак мережевого трафіку до нормованих факторів ризику, що дозволило пов'язати технічні характеристики потоків із ризик-орієнтованою постановкою задачі.

Побудовано нечітку систему виведення, байєсівську мережу та процедуру перетворення їхніх виходів у базові функції призначення маси. На цій основі синтезовано алгоритм DST-агрегації з контролем конфлікту: правило Демпстера використовується для випадків помірної суперечності між джерелами, а правило Ягера – для ситуацій підвищеного конфлікту. Такий підхід дозволяє не лише отримувати фінальний індекс ризику, а й явно фіксувати конфлікт і залишкову невизначеність.

Виконано програмну реалізацію моделі мовою Python та проведено експериментальне дослідження на вибраних класах датасету CIC-IDS2017. Було підготовлено експериментальний набір даних обсягом 155 423 записи, сформовано навчальну та тестову вибірки, реалізовано факторне відображення, FIS-, BN- та DST-компоненти. Початкова мітка класу трафіку не використовувалась як вхідна ознака моделі, а застосовувалась лише для формування експериментальної змінної рівня ризику та оцінювання результатів.

Результати базового експерименту показали, що байєсівська мережа є найсильнішою окремою компонентою за класифікаційними метриками: для неї отримано Accuracy 0.823 та Macro F1 0.787. Hybrid DST показала Accuracy 0.695 та Macro F1 0.661, перевищивши окрему FIS-компоненту та просте середнє FIS-BN, але поступившись BN і зваженому середньому. Отже, запропонована гібридна модель не є найкращою саме як класифікатор рівнів ризику; її перевага полягає в інтерпретованій агрегації різнорідних свідчень, аналізі конфлікту та представленні невизначеності.

Аналіз фінального індексу ризику показав його змістовну кількісну поведінку в межах експериментального сценарію: середнє значення індексу становило 0.350 для низького рівня ризику, 0.486 для середнього та 0.756 для високого. Аналіз конфлікту між джерелами свідчень показав, що середнє значення коефіцієнта конфлікту становить 0.243, а правило Ягера застосовувалося приблизно для 21.6% спостережень. Це підтверджує практичну роль DST-шару як механізму виявлення недостатньої узгодженості між FIS- та BN-компонентами.

Експеримент із неповнотою факторної інформації показав, що втрата окремих факторів погіршує якість оцінювання та збільшує масу невизначеності. Найбільше погіршення спостерігалося при одночасній відсутності факторів експозиції та потенційного впливу: Macro F1 Hybrid DST знизилася до 0.442, а середня маса невизначеності зросла до 0.389. Отже, модель реагує на втрату важливої інформації не лише зміною

фінального індексу, а й зростанням показника невизначеності.

Отримані результати слід інтерпретувати з урахуванням обмежень експериментального сценарію. Змінна `risk_level` є експериментальною розміткою, сформованою на основі класів трафіку, а не повною бізнес-оцінкою кіберризику. Крім того, датасет CIC-IDS2017 не містить інформації про реальні CVE-вразливості, конфігурації хостів, фінансові втрати або бізнес-критичність активів. Тому фінальний індекс ризику слід розглядати як нормовану експериментальну оцінку в межах заданого сценарію, а не як універсальну міру кіберризику для будь-якої інформаційної системи.

Практичне значення роботи полягає в тому, що запропонована модель може бути використана як основа для обчислювального компонента систем підтримки прийняття рішень у сфері інформаційної безпеки. Вона дозволяє поєднувати мережеві ознаки, експертні правила, ймовірнісні залежності та механізм доказової агрегації, а також явно враховувати конфлікт між джерелами свідчень і залишкову невизначеність.

Подальші дослідження доцільно спрямувати на перевірку моделі на інших датасетах і в реальних мережевих середовищах, дослідження сценаріїв зашумлення ознак, уточнення параметрів моделі, а також заміну сценарної критичності активу та proxy-показника експозиції реальними даними про інфраструктуру, вразливості, конфігурації та стан захисних механізмів.

Таким чином, у роботі побудовано цілісну гібридну модель оцінювання кіберризику, що поєднує нечітке виведення, байєсівське моделювання та теорію доказів. Проведене дослідження підтвердило працездатність запропонованого підходу як інтерпретованого механізму оцінювання ризику в умовах неповноти та суперечливості даних за умови коректного врахування меж застосовності експериментальних результатів.

ПЕРЕЛІК ПОСИЛАНЬ

1. ISO/IEC 27005:2022. Information security, cybersecurity and privacy protection — Guidance on managing information security risks : Rep. / International Organization for Standardization ; executor: ISO/IEC. — Geneva, Switzerland : 2022. — Access mode: https://www.iso.org/standard/80585.html.

2. Guide for Conducting Risk Assessments : Rep. : NIST Special Publication 800-30 Revision 1 / National Institute of Standards and Technology ; executor: Joint Task Force Transformation Initiative. — Gaithersburg, MD : 2012. — Access mode: https://csrc.nist.gov/pubs/sp/800/30/r1/final.

3. The NIST Cybersecurity Framework (CSF) 2.0 : Rep. : NIST Cybersecurity White Paper 29 / National Institute of Standards and Technology ; executor: Pascoe Cherilyn, Quinn Stephen, Scarfone Karen. — Gaithersburg, MD : 2024. — Access mode: https://www.nist.gov/publications/nist-cybersecurity-framework-csf-20.

4. Risk Analysis (O-RA), Version 2.0.1 : Rep. : C20A / The Open Group ; executor: The Open Group. — Reading, UK : 2021. — Access mode: https://www.opengroup.org/open-fair.

5. Risk Taxonomy (O-RT), Version 3.0.1 : Rep. / The Open Group ; executor: The Open Group. — Reading, UK : 2021. — Access mode: https://www.opengroup.org/open-fair.

6. Zadeh Lotfi A. Fuzzy Sets // Information and Control. — 1965. — Vol. 8, no. 3. — P. 338–353.

7. Mamdani Ebrahim H., Assilian Sedrak. An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller // International Journal of Man-Machine Studies. — 1975. — Vol. 7, no. 1. — P. 1–13.

8. Ross Timothy J. Fuzzy Logic with Engineering Applications. — 4 ed. — Chichester, UK : John Wiley & Sons, 2017. — ISBN: 9781119235866.

9. Pearl Judea. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. — San Mateo, CA : Morgan Kaufmann, 1988. — ISBN: 9781558604797.

10. Koller Daphne, Friedman Nir. Probabilistic Graphical Models: Principles and Techniques. — Cambridge, MA : MIT Press, 2009. — ISBN: 9780262013192.

11. Wang Jiali, Neil Martin, Fenton Norman. A Bayesian Network Approach for Cybersecurity Risk Assessment Implementing and Extending the FAIR Model // Computers & Security. — 2020. — Vol. 89. — P. 101659.

12. Dempster Arthur P. Upper and Lower Probabilities Induced by a Multivalued Mapping // The Annals of Mathematical Statistics. — 1967. — Vol. 38, no. 2. — P. 325–339.

13. Shafer Glenn. A Mathematical Theory of Evidence. — Princeton, NJ : Princeton University Press, 1976. — ISBN: 9780691100425.

14. Yager Ronald R. On the Dempster-Shafer Framework and New Combination Rules // Information Sciences. — 1987. — Vol. 41, no. 2. — P. 93–137.

15. Combination of Evidence in Dempster-Shafer Theory : Rep. : SAND2002-0835 / Sandia National Laboratories ; executor: Sentz Kari, Ferson Scott. — Albuquerque, NM : 2002. — Access mode: https://www.osti.gov/biblio/800792.

16. Fenton Norman, Neil Martin. Risk Assessment and Decision Analysis with Bayesian Networks. — 2 ed. — Boca Raton, FL : CRC Press, 2019.

17. Han Jiawei, Pei Jian, Tong Hanghang. Data Mining: Concepts and Techniques. — 4 ed. — Cambridge, MA : Morgan Kaufmann, 2022. — ISBN: 9780128117606.

18. Sharafaldin Iman, Lashkari Arash Habibi, Ghorbani Ali A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization // Proceedings of the 4th International Conference on Information Systems Security and Privacy. — SciTePress. — 2018. — P. 108–116.

19. Sharafaldin Iman, Lashkari Arash Habibi, Ghorbani Ali A. A Detailed Analysis of the CICIDS2017 Data Set // Information Systems Security and Privacy. — Cham : Springer, 2019. — P. 172–188.

20. ISO 31000:2018. Risk management — Guidelines : Rep. / International Organization for Standardization ; executor: ISO. — Geneva, Switzerland : 2018. — Access mode: https://www.iso.org/standard/65694.html.

21. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection — Information security management systems — Requirements : Rep. / International Organization for Standardization ; executor: ISO/IEC. — Geneva, Switzerland : 2022. — Access mode: https://www.iso.org/standard/27001.

22. Integrating Cybersecurity and Enterprise Risk Management (ERM) : Rep. : NIST Interagency Report 8286 Revision 1 / National Institute of

Standards and Technology ; executor: Quinn Stephen, Chua Julie, Ivy Nahla et al. — Gaithersburg, MD : 2025. — Access mode: https://csrc.nist.gov/pubs/ir/8286/r1/final.

23. Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process : Rep. : CMU/SEI-2007-TR-012, ESC-TR-2007-012 / Carnegie Mellon University, Software Engineering Institute ; executor: Caralli Richard A., Stevens James F., Young Lisa R., Wilson William R. — Pittsburgh, PA : 2007. — Access mode: https://sei.cmu.edu/library/introducing-octave-allegro-improving-the-information-security-risk-assessment-process/.

24. ENISA. CRAMM: CCTA Risk Analysis and Management Method. — 2006. — Risk management inventory entry. Access mode: https://tools.enisa.europa.eu/topics/risk-management/current-risk/risk-management-inventory/rm-ra-methods/m_cramm.html (online; accessed: 2026-05-13).

25. Risk Management: Principles and Inventories for Risk Management / Risk Assessment Methods and Tools : Rep. / European Network and Information Security Agency ; executor: ENISA. — Heraklion, Greece : 2006. — Access mode: https://www.enisa.europa.eu/publications/risk-management-principles-and-inventories-for-risk-management-risk-assessment-methods-and-tools.

26. ENISA Threat Landscape 2025 : Rep. / European Union Agency for Cybersecurity ; executor: ENISA. — Athens, Greece : 2025. — Access mode: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025.

27. FIRST.Org, Inc. — Common Vulnerability Scoring System version 4.0: Specification Document : 2023. — Document Version 1.2. Access mode: https://www.first.org/cvss/v4.0/specification-document (online; accessed: 2026-05-13).

28. Improving Risk Assessment Model of Cyber Security Using Fuzzy Logic Inference System / Alali Mansour, Almogren Ahmad, Hassan Mohammad Mehedi, Rassan Iehab Al, and Bhuiyan Md. Zakirul Alam // Computers & Security. — 2018. — Vol. 74. — P. 323–339.

29. Jana Dipak Kumar, Ghosh Ramkrishna. Novel Interval Type-2 Fuzzy Logic Controller for Improving Risk Assessment Model of Cyber Security // Journal of Information Security and Applications. — 2018. — Vol. 40. — P. 173–182.

30. Bayesian Network Models in Cyber Security: A Systematic Review / Chockalingam Saba, Pieters Wolter, Teixeira André, and van Gelder Pieter // Secure IT Systems. — Cham : Springer. — 2017. — Vol. 10674 of Lecture Notes

in Computer Science. — P. 105–122.

31. Wang Jiali, Neil Martin. A Bayesian-Network-Based Cybersecurity Adversarial Risk Analysis Framework with Numerical Examples. — 2021. — 2106.00471.

32. Zadeh Lotfi A. A Simple View of the Dempster-Shafer Theory of Evidence and Its Implication for the Rule of Combination // AI Magazine. — 1986. — Vol. 7, no. 2. — P. 85–90.

33. Smets Philippe, Kennes Robert. The Transferable Belief Model // Artificial Intelligence. — 1994. — Vol. 66, no. 2. — P. 191–234.

34. A New Bayesian Network Model for Risk Assessment Based on Cloud Model, Interval Type-2 Fuzzy Sets and Improved D-S Evidence Theory / Xu Jintao, Ding Rui, Li Muye, Dai Tao, Zheng Mengyan, Yu Tao, and Sui Yang // Information Sciences. — 2022. — Vol. 618. — P. 336–355.

35. A New Fuzzy Bayesian Inference Approach for Risk Assessments / Xu Jintao, Sui Yang, Yu Tao, Ding Rui, Dai Tao, and Zheng Mengyan // Symmetry. — 2024. — Vol. 16, no. 7. — P. 786.

36. Quantifying Potential Cyber-Attack Risks in Maritime Transportation under Dempster-Shafer Theory FMECA and Rule-Based Bayesian Network Modelling / Uflaz Esma, Sezer Sukru Ilke, Tunçel Ahmet Lütfi, Aydin Muhammet, Akyuz Emre, and Arslan Ozcan // Reliability Engineering & System Safety. — 2024. — Vol. 243. — P. 109825.

37. Canadian Institute for Cybersecurity. CIC-IDS2017. — 2017. — Access mode: https://www.unb.ca/cic/datasets/ids-2017.html (online; accessed: 2026-05-13).

38. Canadian Institute for Cybersecurity. CICFlowMeter. — 2024. — Network traffic flow generator and analyzer. Access mode: https://github.com/ahlashkari/CICFlowMeter (online; accessed: 2026-05-13).

ДОДАТОК А ТЕКСТИ ПРОГРАМ

У додатку наведено ключові фрагменти програмної реалізації, використані для проведення експериментального дослідження. Повні notebook-файли містять також службові комірки для первинного огляду даних, проміжних перевірок, збереження результатів і побудови допоміжних графіків. У цьому додатку подано лише ті частини коду, які безпосередньо реалізують математичні компоненти запропонованої моделі та відображають основну логіку обчислювального експерименту.

Попередня обробка експериментальних даних включала завантаження CSV-файлів CIC-IDS2017, фільтрацію класів, формування змінної `risk_level`, поділ на навчальну та тестову вибірки, імпутацію пропущених значень і нормалізацію ознак. Основні етапи цієї процедури описано в підрозділі 3.2, тому в додатку основну увагу зосереджено на реалізації факторного відображення, FIS-, BN- та DST-компонентів.

А.1 Побудова факторного відображення

Факторне відображення $\varphi(z)$ реалізовано як зважене агрегування попередньо нормованих мережевих ознак. У програмі 1 наведено групи ознак і ваги, використані для побудови факторів T, V та I; критичність активу A у базовому сценарії задається сталою величиною.

```
factor_feature_weights = {
    "T": {
        "Flow Packets/s": 0.25,
        "Flow Bytes/s": 0.20,
        "Packet Length Std": 0.20,
        "Flow IAT Std": 0.15,
        "Total Fwd Packets": 0.10,
        "Total Backward Packets": 0.10,
    },

```

```
    "V": {
        "SYN Flag Count": 0.25,
        "RST Flag Count": 0.20,
        "PSH Flag Count": 0.15,
        "ACK Flag Count": 0.10,
        "Init_Win_bytes_forward": 0.15,
        "Init_Win_bytes_backward": 0.15,
    },

    "I": {
        "Total Length of Fwd Packets": 0.20,
        "Total Length of Bwd Packets": 0.20,
        "Flow Bytes/s": 0.20,
        "Packet Length Mean": 0.15,
        "Max Packet Length": 0.15,
        "Average Packet Size": 0.10,
    }
}

ASSET_CRITICALITY = 0.60
```

Лістинг 1: Ознаки та ваги факторного відображення

Після задання ваг кожен фактор ризику обчислюється як зважена сума відповідних нормованих ознак. У програмі 2 наведено основну функцію побудови вектора $\varphi(z) = (T,V,A,I)$.

```
def compute_weighted_factor(df, weights):
    result = np.zeros(len(df))
    for feature, weight in weights.items():
        result += weight * df[feature].astype(float).values
    return np.clip(result, 0, 1)

def build_phi_factors(df, asset_criticality=0.60):
    factors = pd.DataFrame(index=df.index)
    factors["T"] = compute_weighted_factor(df, factor_feature_weights["T"])
    factors["V"] = compute_weighted_factor(df, factor_feature_weights["V"])
    factors["A"] = asset_criticality
    factors["I"] = compute_weighted_factor(df, factor_feature_weights["I"])
    factors["Label"] = df["Label"].values
```

```
    factors["risk_level"] = df["risk_level"].values
    factors["source_file"] = df["source_file"].values
    return factors

train_factors = build_phi_factors(
    train_df,
    asset_criticality=ASSET_CRITICALITY
)

test_factors = build_phi_factors(
    test_df,
    asset_criticality=ASSET_CRITICALITY
)
```

Лістинг 2: Побудова факторів ризику T,V,A,I

Для використання факторів у байєсівській мережі неперервні значення T,V,A,I додатково переводяться у дискретні стани `Low`, `Medium` та `High`. Пороги для T, V та I визначаються за навчальною вибіркою, а для A задаються сценарно.

```
def get_quantile_thresholds(series):
    return {
        "low_medium": float(series.quantile(0.33)),
        "medium_high": float(series.quantile(0.66))
    }

factor_thresholds = {
    "T": get_quantile_thresholds(train_factors["T"]),
    "V": get_quantile_thresholds(train_factors["V"]),
    "I": get_quantile_thresholds(train_factors["I"]),
    "A": {
        "low_medium": 0.33,
        "medium_high": 0.66
    }
}

def discretize_value(value, low_medium, medium_high):
    if value <= low_medium:
```

```
19          return "Low"
20      elif value <= medium_high:
21          return "Medium"
22      else:
23          return "High"
24
25  def add_discrete_factors(df, thresholds):
26      df_discrete = df.copy()
27      for factor in ["T", "V", "A", "I"]:
28          low_medium = thresholds[factor]["low_medium"]
29          medium_high = thresholds[factor]["medium_high"]
30          df_discrete[f"{factor}_d"] = df_discrete[factor].apply(
31              lambda x: discretize_value(x, low_medium, medium_high)
32          )
33      return df_discrete
34
35  train_factors_bn = add_discrete_factors(train_factors, factor_thresholds)
36  test_factors_bn = add_discrete_factors(test_factors, factor_thresholds)
```

Лістинг 3: Дискретизація факторів ризику

A.2 Реалізація нечіткої системи виведення

Нечітка система виведення реалізована як система типу Мамдані. Спочатку задаються базові функції належності, які використовуються для фазифікації факторів ризику та вихідної змінної R_{FIS}.

```
1   def left_shoulder(x, a, b):
2       x = np.asarray(x, dtype=float)
3       y = np.zeros_like(x)
4       y[x <= a] = 1.0
5       y[x >= b] = 0.0
6       mask = (x > a) & (x < b)
7       y[mask] = (b - x[mask]) / (b - a)
8       return np.clip(y, 0, 1)
9
10  def right_shoulder(x, a, b):
```

```
    x = np.asarray(x, dtype=float)
    y = np.zeros_like(x)
    y[x <= a] = 0.0
    y[x >= b] = 1.0
    mask = (x > a) & (x < b)
    y[mask] = (x[mask] - a) / (b - a)
    return np.clip(y, 0, 1)

def triangular(x, a, b, c):
    x = np.asarray(x, dtype=float)
    y = np.zeros_like(x)
    left_mask = (x > a) & (x <= b)
    right_mask = (x > b) & (x < c)
    y[left_mask] = (x[left_mask] - a) / (b - a)
    y[right_mask] = (c - x[right_mask]) / (c - b)
    y[x == b] = 1.0
    return np.clip(y, 0, 1)
```

Лістинг 4: Функції належності нечіткої системи

Для факторів T, V та I функції належності будуються на основі порогів, визначених за навчальною вибіркою, тоді як для фактора A використовується окрема сценарна фазифікація.

```
TERMS = ["Low", "Medium", "High"]

def fuzzify_factor_by_thresholds(values, low_medium, medium_high):
    x = np.asarray(values, dtype=float)
    mid = (low_medium + medium_high) / 2
    return {
        "Low": left_shoulder(x, low_medium, medium_high),
        "Medium": triangular(x, low_medium, mid, medium_high),
        "High": right_shoulder(x, low_medium, medium_high),
    }

def fuzzify_asset_criticality(values):
    x = np.asarray(values, dtype=float)
    return {
        "Low": left_shoulder(x, 0.30, 0.45),
        "Medium": triangular(x, 0.35, 0.60, 0.75),
```

```
        "High": right_shoulder(x, 0.65, 0.80),
    }

def fuzzify_inputs(df):
    mu = {}
    for factor in ["T", "V", "I"]:
        low_medium = factor_thresholds[factor]["low_medium"]
        medium_high = factor_thresholds[factor]["medium_high"]
        mu[factor] = fuzzify_factor_by_thresholds(
            df[factor].values,
            low_medium,
            medium_high
        )

    mu["A"] = fuzzify_asset_criticality(df["A"].values)
    return mu
```

Лістинг 5: Фазифікація вхідних факторів ризику

База нечітких правил формується автоматично для всіх $3^4 = 81$ комбінацій термів. Вихідний терм правила визначається за монотонною зваженою схемою, що враховує внесок кожного фактора.

```
term_numeric = {
    "Low": 0.0,
    "Medium": 0.5,
    "High": 1.0,
}

rule_factor_weights = {
    "T": 0.35,
    "V": 0.25,
    "A": 0.10,
    "I": 0.30,
}

def choose_risk_term(score):
    if score < 0.30:
        return "Low"
    elif score < 0.60:
```

```
18          return "Medium"
19      else:
20          return "High"
21
22  fis_rules = []
23  for t_term, v_term, a_term, i_term in itertools.product(TERMS, TERMS, TERMS,
        TERMS):
24      score = (
25          rule_factor_weights["T"] * term_numeric[t_term]
26          + rule_factor_weights["V"] * term_numeric[v_term]
27          + rule_factor_weights["A"] * term_numeric[a_term]
28          + rule_factor_weights["I"] * term_numeric[i_term]
29      )
30      risk_term = choose_risk_term(score)
31      fis_rules.append({
32          "T": t_term,
33          "V": v_term,
34          "A": a_term,
35          "I": i_term,
36          "score": score,
37          "Risk": risk_term
38      })
39
40  fis_rules_df = pd.DataFrame(fis_rules)
```

Лістинг 6: Автоматизоване формування бази правил FIS

Наведений фрагмент реалізує Mamdani-виведення: активація кожного правила обчислюється за T-нормою мінімуму, після чого вихідні терми агрегуються та дефазифікуються методом центра ваги.

```
1  risk_grid = np.linspace(0, 1, 501)
2  risk_mf = {
3      "Low": left_shoulder(risk_grid, 0.30, 0.50),
4      "Medium": triangular(risk_grid, 0.25, 0.50, 0.75),
5      "High": right_shoulder(risk_grid, 0.50, 0.70),
6  }
7
8  risk_term_to_index = {
9      "Low": 0,
```

```
    "Medium": 1,
    "High": 2,
}

def compute_rule_output_heights(df):
    mu = fuzzify_inputs(df)
    n = len(df)
    output_heights = np.zeros((n, 3), dtype=float)
    for rule in fis_rules:
        activation = np.minimum.reduce([
            mu["T"][rule["T"]],
            mu["V"][rule["V"]],
            mu["A"][rule["A"]],
            mu["I"][rule["I"]],
        ])
        risk_idx = risk_term_to_index[rule["Risk"]]
        output_heights[:, risk_idx] = np.maximum(
            output_heights[:, risk_idx],
            activation
        )
    return output_heights

def defuzzify_centroid(output_heights, batch_size=5000):
    results = []
    low_mf = risk_mf["Low"]
    medium_mf = risk_mf["Medium"]
    high_mf = risk_mf["High"]
    for start in range(0, len(output_heights), batch_size):
        end = start + batch_size
        h = output_heights[start:end]
        clipped_low = np.minimum(h[:, [0]], low_mf[None, :])
        clipped_medium = np.minimum(h[:, [1]], medium_mf[None, :])
        clipped_high = np.minimum(h[:, [2]], high_mf[None, :])
        aggregated = np.maximum.reduce([
            clipped_low,
            clipped_medium,
            clipped_high
        ])
        denominator = aggregated.sum(axis=1)
```

```
        numerator = (aggregated * risk_grid[None, :]).sum(axis=1)
        batch_result = np.where(
            denominator > 0,
            numerator / denominator,
            0.5
        )
        results.append(batch_result)
    return np.concatenate(results)
```

Лістинг 7: Mamdani-виведення та дефазифікація

Після дефазифікації для кожного спостереження формується числова оцінка R_{FIS}, відповідний дискретний рівень ризику та нормовані ступені підтримки термів, які надалі використовуються під час формування BPA.

```
def classify_risk_score(r):
    if r < 0.33:
        return "Low"
    elif r < 0.66:
        return "Medium"
    else:
        return "High"

def compute_fis_gamma(r_values):
    r_values = np.asarray(r_values, dtype=float)
    mu_low = left_shoulder(r_values, 0.30, 0.50)
    mu_medium = triangular(r_values, 0.25, 0.50, 0.75)
    mu_high = right_shoulder(r_values, 0.50, 0.70)
    raw = np.vstack([mu_low, mu_medium, mu_high]).T
    denom = raw.sum(axis=1, keepdims=True)
    gamma = np.divide(
        raw,
        denom,
        out=np.zeros_like(raw),
        where=denom > 0
    )
    return gamma

```

```
def run_fis(df):
    output_heights = compute_rule_output_heights(df)
    r_fis = defuzzify_centroid(output_heights)
    gamma = compute_fis_gamma(r_fis)
    result = df.copy()
    result["R_FIS"] = r_fis
    result["FIS_level"] = result["R_FIS"].apply(classify_risk_score)
    result["gamma_FIS_Low"] = gamma[:, 0]
    result["gamma_FIS_Medium"] = gamma[:, 1]
    result["gamma_FIS_High"] = gamma[:, 2]
    result["fis_activation_low"] = output_heights[:, 0]
    result["fis_activation_medium"] = output_heights[:, 1]
    result["fis_activation_high"] = output_heights[:, 2]
    return result

train_fis = run_fis(train_factors)
test_fis = run_fis(test_factors)
```

Лістинг 8: Формування виходу FIS-компоненти

A.3 Реалізація байєсівської мережі

Байєсівська компонента використовує дискретизовані фактори T_d,V_d,A_d,I_d як батьківські змінні, а експериментальний рівень ризику `risk_level` – як цільову змінну H. У програмі 9 наведено основні стани та числове кодування рівнів ризику.

```
RISK_STATES = ["Low", "Medium", "High"]
PARENT_COLUMNS = ["T_d", "V_d", "A_d", "I_d"]
TARGET_COLUMN = "risk_level"
risk_numeric = {
    "Low": 0.0,
    "Medium": 0.5,
    "High": 1.0,
}
```

Лістинг 9: Задання станів байєсівської мережі

Таблиця умовних ймовірностей $P(H \mid T_d,V_d,A_d,I_d)$ оцінюється на навчальній вибірці за частотами комбінацій дискретних факторів. Для уникнення нульових ймовірностей використовується лапласівське згладжування.

```
def train_bn_cpt(train_df, parent_columns, target_column, states, alpha=1.0):
    parent_states = {col: states for col in parent_columns}
    all_parent_combinations = list(itertools.product(
        *[parent_states[col] for col in parent_columns]
    ))

    rows = []
    grouped = (
        train_df
        .groupby(parent_columns + [target_column])
        .size()
        .reset_index(name="count")
    )

    for combination in all_parent_combinations:
        condition = dict(zip(parent_columns, combination))
        subset = grouped.copy()
        for col, value in condition.items():
            subset = subset[subset[col] == value]
        counts = {state: 0 for state in states}
        for _, row in subset.iterrows():
            counts[row[target_column]] = int(row["count"])
        total = sum(counts.values())
        denominator = total + alpha * len(states)
        probabilities = {
            state: (counts[state] + alpha) / denominator
            for state in states
        }

        result_row = dict(condition)
        result_row["N"] = total
        for state in states:
            result_row[f"P_{state}"] = probabilities[state]
        rows.append(result_row)
```

```
35      cpt_df = pd.DataFrame(rows)
36      return cpt_df
37
38  ALPHA = 1.0
39  bn_cpt = train_bn_cpt(
40      train_df=train_bn_input,
41      parent_columns=PARENT_COLUMNS,
42      target_column=TARGET_COLUMN,
43      states=RISK_STATES,
44      alpha=ALPHA
45  )
```

Лістинг 10: Побудова таблиці умовних ймовірностей BN

Для швидкого отримання апостеріорного розподілу за заданим набором свідчень CPT перетворюється у словник, де ключем є комбінація станів (T_d,V_d,A_d,I_d).

```
1  def build_cpt_lookup(cpt_df, parent_columns, states):
2      lookup = {}
3      for _, row in cpt_df.iterrows():
4          key = tuple(row[col] for col in parent_columns)
5          probs = np.array([row[f"P_{state}"] for state in states],
       dtype=float)
6          lookup[key] = probs
7      return lookup
8  bn_cpt_lookup = build_cpt_lookup(bn_cpt, PARENT_COLUMNS, RISK_STATES)
```

Лістинг 11: Формування lookup-таблиці CPT

Оскільки в базовому експериментальному сценарії всі батьківські змінні є спостереженими, байєсівське виведення зводиться до вибору відповідного рядка CPT. Додатково обчислюється числова оцінка R_{BN} як математичне сподівання за кодуванням рівнів ризику.

```
1  def run_bn_inference(df, cpt_lookup, parent_columns, states):
2      probs = []
3      for _, row in df.iterrows():
4          key = tuple(row[col] for col in parent_columns)
5          if key in cpt_lookup:
```

```
            p = cpt_lookup[key]
        else:
            # fallback, should not normally happen
            p = np.ones(len(states)) / len(states)
        probs.append(p)
    probs = np.vstack(probs)
    result = df.copy()
    for i, state in enumerate(states):
        result[f"P_BN_{state}"] = probs[:, i]

    result["R_BN"] = (
        result["P_BN_Low"] * risk_numeric["Low"]
        + result["P_BN_Medium"] * risk_numeric["Medium"]
        + result["P_BN_High"] * risk_numeric["High"]
    )
    result["BN_level"] = result[[f"P_BN_{state}" for state in
    states]].idxmax(axis=1)
    result["BN_level"] = result["BN_level"].str.replace("P_BN_", "",
    regex=False)
    return result

train_bn = run_bn_inference(
    train_bn_input,
    bn_cpt_lookup,
    PARENT_COLUMNS,
    RISK_STATES
)

test_bn = run_bn_inference(
    test_bn_input,
    bn_cpt_lookup,
    PARENT_COLUMNS,
    RISK_STATES
)
```

Лістинг 12: Обчислення апостеріорного розподілу $P(H \mid e)$

А.4 Формування базових функцій маси та DST-агрегація

DST-компонента працює на спільному просторі гіпотез $\Theta = \{Low, Medium, High\}$. У базовому сценарії для FIS- та BN-компонентів використовуються однакові коефіцієнти надійності, а вибір між правилами Демпстера та Ягера здійснюється за порогом конфлікту.

```
STATES = ["Low", "Medium", "High"]
STATE_VALUES = {
    "Low": 0.0,
    "Medium": 0.5,
    "High": 1.0,
}
RHO_FIS = 0.80
RHO_BN = 0.80
CONFLICT_THRESHOLD = 0.50
```

Лістинг 13: Параметри DST-агрегації

Вихід FIS-компоненти перетворюється у BPA через нормовані ступені підтримки термів ризику, а вихід BN-компоненти – через апостеріорні ймовірності $P(H \mid e)$. Частина маси, що не призначається одноелементним гіпотезам, переноситься на повну множину Θ.

```
def normalize_rows(arr):
    arr = np.asarray(arr, dtype=float)
    denom = arr.sum(axis=1, keepdims=True)
    return np.divide(
        arr,
        denom,
        out=np.zeros_like(arr),
        where=denom > 0
    )

def extract_fis_gamma(df):
    gamma = df[
```

```
13          ["gamma_FIS_Low", "gamma_FIS_Medium", "gamma_FIS_High"]
14      ].values.astype(float)
15      return normalize_rows(gamma)
16
17  def extract_bn_probs(df):
18      probs = df[
19          ["P_BN_Low", "P_BN_Medium", "P_BN_High"]
20      ].values.astype(float)
21      return normalize_rows(probs)
22
23  def build_bpa_from_fis(df, rho_fis=0.80):
24      gamma = extract_fis_gamma(df)
25      singleton_masses = rho_fis * gamma
26      theta_mass = np.full(len(df), 1.0 - rho_fis)
27      return singleton_masses, theta_mass
28
29  def build_bpa_from_bn(df, rho_bn=0.80):
30      probs = extract_bn_probs(df)
31      singleton_masses = rho_bn * probs
32      theta_mass = np.full(len(df), 1.0 - rho_bn)
33      return singleton_masses, theta_mass
```

Лістинг 14: Формування BPA для FIS- та BN-компонентів

Наведена функція реалізує комбінування двох BPA. За помірного конфлікту використовується правило Демпстера, а за високого конфлікту – правило Ягера, у якому конфліктна маса переноситься на Θ.

```
1  def combine_bpa_dst(
2      m_fis_single,
3      m_fis_theta,
4      m_bn_single,
5      m_bn_theta,
6      conflict_threshold=0.50
7  ):
8      m_fis_single = np.asarray(m_fis_single, dtype=float)
9      m_bn_single = np.asarray(m_bn_single, dtype=float)
10     m_fis_theta = np.asarray(m_fis_theta, dtype=float)
11     m_bn_theta = np.asarray(m_bn_theta, dtype=float)
12     singleton_sum_fis = m_fis_single.sum(axis=1)
```

```
    singleton_sum_bn = m_bn_single.sum(axis=1)
    singleton_agreement = np.sum(m_fis_single * m_bn_single, axis=1)
    conflict_k = singleton_sum_fis * singleton_sum_bn - singleton_agreement
    unnormalized_single = (
        m_fis_single * m_bn_single
        + m_fis_single * m_bn_theta[:, None]
        + m_fis_theta[:, None] * m_bn_single
    )
    unnormalized_theta = m_fis_theta * m_bn_theta
    # Dempster rule
    denominator = 1.0 - conflict_k
    dempster_single = np.divide(
        unnormalized_single,
        denominator[:, None],
        out=np.zeros_like(unnormalized_single),
        where=denominator[:, None] > 1e-12
    )
    dempster_theta = np.divide(
        unnormalized_theta,
        denominator,
        out=np.zeros_like(unnormalized_theta),
        where=denominator > 1e-12
    )
    # Yager rule
    yager_single = unnormalized_single
    yager_theta = unnormalized_theta + conflict_k
    use_dempster = conflict_k <= conflict_threshold
    combined_single = np.where(
        use_dempster[:, None],
        dempster_single,
        yager_single
    )
    combined_theta = np.where(
        use_dempster,
        dempster_theta,
        yager_theta
    )
    used_rule = np.where(use_dempster, "Dempster", "Yager")
    total_mass = combined_single.sum(axis=1) + combined_theta
```

```
    combined_single = np.divide(
        combined_single,
        total_mass[:, None],
        out=np.zeros_like(combined_single),
        where=total_mass[:, None] > 0
    )
    combined_theta = np.divide(
        combined_theta,
        total_mass,
        out=np.zeros_like(combined_theta),
        where=total_mass > 0
    )
    return combined_single, combined_theta, conflict_k, used_rule
```

Лістинг 15: Обчислення конфлікту та комбінування BPA

Після комбінування агрегована функція маси перетворюється у ймовірнісний розподіл BetP. Оскільки в реалізації використовуються лише одноелементні гіпотези та повна множина Θ, маса $m(\Theta)$ рівномірно розподіляється між трьома станами.

```
def pignistic_transform(combined_single, combined_theta):
    n_states = combined_single.shape[1]
    betp = combined_single + combined_theta[:, None] / n_states
    betp = normalize_rows(betp)
    return betp

def classify_risk_by_score(r):
    if r < 0.33:
        return "Low"
    elif r < 0.66:
        return "Medium"
    else:
        return "High"

def classify_risk_by_betp(row):
    values = row[["BetP_Low", "BetP_Medium", "BetP_High"]].values
    return STATES[int(np.argmax(values))]
```

Лістинг 16: Пігністичне перетворення агрегованої функції маси

Фінальна функція об'єднує попередні етапи: формує BPA для FIS і BN, виконує DST-комбінування, обчислює BetP та отримує нормований індекс ризику R_{final}.

```
def run_dst_aggregation(
    fis_df,
    bn_df,
    rho_fis=0.80,
    rho_bn=0.80,
    conflict_threshold=0.50
):
    validate_alignment(fis_df, bn_df, name="DST input")
    m_fis_single, m_fis_theta = build_bpa_from_fis(fis_df, rho_fis=rho_fis)
    m_bn_single, m_bn_theta = build_bpa_from_bn(bn_df, rho_bn=rho_bn)
    combined_single, combined_theta, conflict_k, used_rule = combine_bpa_dst(
        m_fis_single,
        m_fis_theta,
        m_bn_single,
        m_bn_theta,
        conflict_threshold=conflict_threshold
    )
    betp = pignistic_transform(combined_single, combined_theta)
    state_values_array = np.array([
        STATE_VALUES["Low"],
        STATE_VALUES["Medium"],
        STATE_VALUES["High"]
    ])
    r_final = betp @ state_values_array
    result = pd.DataFrame(index=fis_df.index)
    # BPA FIS
    result["m_FIS_Low"] = m_fis_single[:, 0]
    result["m_FIS_Medium"] = m_fis_single[:, 1]
    result["m_FIS_High"] = m_fis_single[:, 2]
    result["m_FIS_Theta"] = m_fis_theta
    # BPA BN
    result["m_BN_Low"] = m_bn_single[:, 0]
    result["m_BN_Medium"] = m_bn_single[:, 1]
    result["m_BN_High"] = m_bn_single[:, 2]
    result["m_BN_Theta"] = m_bn_theta
```

```
    # DST combined mass
    result["m_DST_Low"] = combined_single[:, 0]
    result["m_DST_Medium"] = combined_single[:, 1]
    result["m_DST_High"] = combined_single[:, 2]
    result["m_DST_Theta"] = combined_theta
    # conflict and rule
    result["K_conflict"] = conflict_k
    result["DST_rule"] = used_rule
    # BetP
    result["BetP_Low"] = betp[:, 0]
    result["BetP_Medium"] = betp[:, 1]
    result["BetP_High"] = betp[:, 2]
    # final risk
    result["R_final"] = r_final
    result["R_final_level"] = result["R_final"].apply(classify_risk_by_score)
    result["DST_level"] = result.apply(classify_risk_by_betp, axis=1)
    return result

train_dst = run_dst_aggregation(
    train_fis,
    train_bn,
    rho_fis=RHO_FIS,
    rho_bn=RHO_BN,
    conflict_threshold=CONFLICT_THRESHOLD
)
test_dst = run_dst_aggregation(
    test_fis,
    test_bn,
    rho_fis=RHO_FIS,
    rho_bn=RHO_BN,
    conflict_threshold=CONFLICT_THRESHOLD
)
```

Лістинг 17: Формування фінального індексу ризику

A.5 Оцінювання результатів експерименту

Оцінювання результатів виконується за допомогою метрик `Accuracy` та `Macro F1`. У програмі 18 наведено функцію, яка формує один рядок порівняльної таблиці для заданої моделі.

```
from sklearn.metrics import accuracy_score, f1_score, confusion_matrix,
    classification_report
def model_metrics(df, pred_col, model_name):
    return {
        "model": model_name,
        "accuracy": accuracy_score(df["risk_level"], df[pred_col]),
        "macro_f1": f1_score(df["risk_level"], df[pred_col], average="macro")
    }
```

Лістинг 18: Обчислення основних метрик якості

Для порівняння окремих компонентів моделі використовується одна й та сама тестова вибірка. У програмі 19 наведено формування підсумкової таблиці для FIS, BN та Hybrid DST.

```
final_model_comparison_test = pd.DataFrame([
    model_metrics(test_dst, "FIS_level", "FIS"),
    model_metrics(test_dst, "BN_level", "BN"),
    model_metrics(test_dst, "R_final_level", "Hybrid DST")
])

final_model_comparison_test.to_csv(
    OUTPUT_TABLES_DIR / "final_model_comparison_test.csv",
    index=False
)
final_model_comparison_test
```

Лістинг 19: Порівняльна таблиця моделей на test-вибірці

Матриця помилок будується для фіксованого порядку класів `Low`, `Medium`, `High`. Наведений фрагмент показує обчислення confusion matrix для BN-компоненти на тестовій вибірці.

```
labels_order = ["Low", "Medium", "High"]
cm_bn_test = confusion_matrix(
    test_bn["risk_level"],
    test_bn["BN_level"],
    labels=labels_order
)
bn_test_cm_df = pd.DataFrame(
    cm_bn_test,
    index=[f"true_{x}" for x in labels_order],
    columns=[f"pred_{x}" for x in labels_order]
)
bn_test_cm_df
```

Лістинг 20: Матриця помилок для BN-компоненти

Для фінальної гібридної оцінки аналогічно обчислюється матриця помилок за рівнем ризику, отриманим із нормованого індексу R_{final}.

```
labels_order = ["Low", "Medium", "High"]
cm_dst_score = confusion_matrix(
    test_dst["risk_level"],
    test_dst["R_final_level"],
    labels=labels_order
)
dst_score_cm_df = pd.DataFrame(
    cm_dst_score,
    index=[f"true_{x}" for x in labels_order],
    columns=[f"pred_{x}" for x in labels_order]
)
display(dst_score_cm_df)
print("Hybrid by R_final_level Accuracy:",
      accuracy_score(test_dst["risk_level"], test_dst["R_final_level"]))
print("Hybrid by R_final_level Macro F1:",
      f1_score(test_dst["risk_level"], test_dst["R_final_level"],
    average="macro"))
```

Лістинг 21: Матриця помилок для фінальної гібридної оцінки

Додатково в експерименті порівнювалися фінальна DST-агрегація, просте середнє та зважене середнє виходів FIS і BN. Це дозволяє

перевірити, чи дає DST-компонента іншу поведінку порівняно з безпосереднім усередненням числових оцінок.

```
avg_df = test_dst.copy()
avg_df["R_avg"] = 0.5 * avg_df["R_FIS"] + 0.5 * avg_df["R_BN"]
avg_df["R_wavg"] = 0.4 * avg_df["R_FIS"] + 0.6 * avg_df["R_BN"]
def classify_score(r):
    if r < 0.33:
        return "Low"
    elif r < 0.66:
        return "Medium"
    else:
        return "High"
avg_df["R_avg_level"] = avg_df["R_avg"].apply(classify_score)
avg_df["R_wavg_level"] = avg_df["R_wavg"].apply(classify_score)

def compute_metrics(df, pred_col, model_name):
    return {
        "model": model_name,
        "accuracy": accuracy_score(df["risk_level"], df[pred_col]),
        "macro_f1": f1_score(df["risk_level"], df[pred_col], average="macro")
    }

baseline_comparison_test = pd.DataFrame([
    compute_metrics(avg_df, "FIS_level", "FIS"),
    compute_metrics(avg_df, "BN_level", "BN"),
    compute_metrics(avg_df, "R_avg_level", "Average FIS-BN"),
    compute_metrics(avg_df, "R_wavg_level", "Weighted Average 0.4/0.6"),
    compute_metrics(avg_df, "R_final_level", "Hybrid DST"),
])

baseline_comparison_test
```

Лістинг 22: Порівняння з простим та зваженим середнім