УДК 004.056.55:004.89:005.334
DOI: 10.31673/2409-7292.2025.031463

Прокопович-Ткаченко Д.І., Бушков В.Г., Хрушков Б.С., Черкаський О.В., Козаченко І.М., Білан М.В.

ІНТЕЛЕКТУАЛЬНА МОДЕЛЬ ПРОГНОЗУВАННЯ ТА РЕАГУВАННЯ НА КІБЕРЗАГРОЗИ З ВИКОРИСТАННЯМ БАГАТОШАРОВИХ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ І СУЧАСНИХ СТРАТЕГІЙ УПРАВЛІННЯ РИЗИКАМИ

Дослідження присвячене розробці і експериментальній перевірці інтелектуальної багаторівневої системи управління кіберризиками для захисту критично важливих інформаційних систем. Система CRMS-RMODV (Cyber Risk Management System – Risk Management with Optimal Decision and Volume) переносить концепцію ризик-менеджменту, відому з алгоритмічної біржової торгівлі, у сферу кібербезпеки. Ключова ідея полягає у використанні штучних нейронних мереж з довгою короткостроковою пам'яттю (Long Short-Term Memory, LSTM) для прогнозування короткострокової (15-хвилинної) динаміки інтегрального ризику на основі потоків телеметрії з центрів оперативного реагування на інциденти (Security Operations Center, SOC). Методика охоплює формування розширеного вектора ознак із 113 параметрів, який включає п'ять мережевих агрегованих метрик та 108 індикаторів на основі фреймворків MITRE ATT&CK і Common Vulnerability Scoring System (CVSS). Для навчання чотиришарової LSTM-мережі використано 2,4 терабайти історичних даних телеметрії. Валідованість моделі забезпечується статистичним тестуванням, а також емуляцією багаторівневих цілеспрямованих атак за допомогою платформи Caldera. Для інтеграції рішень в реальні кіберзахисні сценарії розроблено впровадження у автоматизовані сценарії реагування (playbooks) систем оркестрації та автоматизації кібербезпеки Splunk SOAR і Cortex XSOAR. Особливістю проєкту є впровадження формалізованого індикатора Threat-VWAP (Threat Volume-Weighted Average Price). Результати свідчать, що комбінація LSTM-прогнозування, каскадних тригерів take-profit/stop-loss, денної квоти інцидентів та фільтра Threat-VWAP забезпечує суттєве зменшення сукупних збитків навіть при середній точності класифікації, що підтверджує доцільність перенесення біржових моделей управління ризиком у сферу кібербезпеки.

Ключові слова: cyber risk management, LSTM, RMODV, Threat-VWAP, SOC automation, SOAR.

Вступ

Кіберзагрози продовжують еволюціонувати експоненційно, як у плані частоти атак, так і складності сценаріїв їх реалізації, що суттєво ускладнює оперативне реагування традиційних систем кіберзахисту. Згідно з доповіддю ENISA [1], 2024 рік був позначений значним зростанням атак типу multi-stage intrusion, де атакувальники комбінують методи соціальної інженерії, експлуатацію вразливостей і розповсюдження програм-вимагачів. Verizon DBIR 2024 [2] підкреслює, що людський фактор залишається найпоширенішим вектором компрометації, а середня вартість одного витоку даних перевищила $4,5 млн, що створює значний фінансовий тиск на організації. У відповідь на ці виклики традиційні засоби захисту, що спираються на фіксовані правила чи статичні пороги, втрачають ефективність. Основною проблемою є неспроможність масштабуватися під потік високоволатильної, "шумної" телеметрії з численних джерел (SOC, SIEM, EDR тощо), а також реагувати у реальному часі на атаки з невизначеними сигнатурами [1], [5], [6]. У цьому контексті зростає зацікавленість у перенесенні парадигм фінансового ризик-менеджменту на кібербезпеку. Як показано у роботах Silva et al. [30] та Lim et al. [3], нейромережеві моделі (LSTM, TFT), розроблені для прогнозування фондових ринків, показують високий потенціал у задачах багатогоризонтного прогнозування аномалій. Ідеї обмеження збитків (stop-loss) і фіксації прибутку (take-profit), звичні у трейдингу, поступово впроваджуються у системи інцидент-респонсу [4], [7], [9], де їх застосування дозволяє зменшити надмірну реакцію на хибні спрацювання. Попри це, інтегровані моделі, що комбінують нейропрогнозування, економічні тригери реагування та агрегацію ризику за обсягом (Threat-VWAP), залишаються рідкісними. Сучасні SOC-інструменти, такі як MITRE ATT&CK, CVSS v4.0, SOAR або XSOAR [6-9], надають фреймворки й автоматизовані плейбуки, однак у більшості випадків не передбачають адаптивне регулювання частоти реагувань з урахуванням ризиків і обмежень ресурсу аналітиків. Крім того, результати оглядів CISA, Microsoft, Picus Security, Red Canary та ISTARI

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

Global [10], [17], [18], [19], [25] вказують на стрімке зростання кількості реалістичних симуляцій (adversary emulation) та Red Team-активностей, що потребують складніших моделей пріоритезації інцидентів і ризиків. Також аналітичні звіти ENISA щодо економіки кіберризику [12] та урядові звіти (CISA, OCC, FDIC) [16], [19], [27] наголошують на необхідності балансування між рівнем захищеності та операційною ефективністю, особливо у секторах критичної інфраструктури та фінансів. Саме в цьому контексті виникає потреба у системі CRMS-RMODV (Cyber Risk Management System – Risk Management with Optimal Decision and Volume), що поєднує:

- LSTM-прогнозування загроз [3], [30];
- економічні тригери TP/SL [4], [30];
- агрегацію ризиків за допомогою Threat-VWAP (аналог VWAP із фінансів) [30];
- добове обмеження кількості інцидент-респонсу для уникнення перевантаження SOC-аналітиків [5], [14];
- багаторівневу логіку реагування, адаптовану під сучасні сценарії загроз [6], [8], [11].

Постановка проблеми

Кіберзагрози стрімко еволюціонують як за частотою атак, так і за складністю реалізації багатоетапних інфільтраційних кампаній, що суттєво перевантажує традиційні системи захисту з жорстко заданими правилами та статичними порогами [1], [2]. Зі збільшенням обсягів телеметрії з різних джерел (SOC, SIEM, EDR тощо) зростає “шум” даних і кількість хибнопозитивних сповіщень, що призводить до перевантаження аналітиків та затримок у виявленні реальних інцидентів [5], [6].

Наявні підходи до автоматизації реагування, основані на фреймворках MITRE ATT&CK та CVSS, забезпечують уніфіковану класифікацію тактик і оцінку вразливостей, проте не враховують динамічні зміни ризиків в часі й не інтегрують механізми адаптивного управління ресурсами SOC [6-9]. З іншого боку, фінансові стратегії управління ризиком (take-profit, stop-loss, VWAP) успішно застосовуються для контролю волатильності ринкових індикаторів, але залишаються недостатньо дослідженими в контексті кібербезпеки [30], [4].

Отже, існує потреба в розробці єдиної інтелектуальної платформи, яка поєднає:

- короткострокове прогнозування інтегрального ризику на основі нейромереж LSTM із вхідними даними потокової телеметрії SOC [3];
- динамічні економічні тригери take-profit/stop-loss для обмеження збитків та фіксації “прибутку” від успішного блокування загроз [4], [30];
- об’ємно-зважений фільтр Threat-VWAP для згладжування сплесків хибнопозитивів та адаптивного регулювання порогів детекції;
- механізм денного ліміту інцидент-реагувань для запобігання вигоранню аналітиків та оптимального розподілу ресурсів SOC [5], [14];
- багаторівневу логіку прийняття рішень, що забезпечує адаптивність до зміни інтенсивності атак.

Відсутність інтегрованих моделей, здатних поєднати вищезгадані компоненти в єдиному конвеєрі обробки даних та реагування, знижує ефективність сучасних SOC-рішень і призводить до значних операційних та фінансових витрат. Саме тому метою даної роботи є формулювання теоретичних основ та побудова експериментально перевіреної системи CRMS-RMODV, здатної оптимізувати співвідношення між чутливістю до загроз та операційними ресурсами в режимі реального часу.

Аналіз останніх публікацій

Останні галузеві огляди демонструють зростання складності та динамічності кіберзагроз. Зокрема, ENISA в «Threat Landscape 2024» відзначає збільшення многоетапних атак із комбінованими векторами компрометації, що вимагає від SOC-систем більш гнучких механізмів виявлення [1] . У звіті Verizon DBIR 2024 підкреслюється, що людський фактор залишається основним вектором проникнення, а середня вартість одного витоку перевищила

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

$4,5 млн, що створює великий фінансовий тиск на організації [2]. Sophos у своєму «2024 Security Threat Report» констатує 35 % зростання складних фото- та відеофішингових атак порівняно з попереднім роком [21] .

Низка публікацій концентрується на стандартах та фреймворках для управління кіберризиками. ISO/IEC 27005:2024 пропонує оновлені методи оцінки ризиків із акцентом на контекстні загрози та бізнес-вплив [4] . Рекомендації NIST SP 800-30 Rev. 2 розкривають покрокову методологію проведення оцінок ризику, що включає сценарний аналіз і кількісне моделювання [5] . Оновлена база MITRE ATT&CK v14.1 розширює каталог тактик і технік APT-груп, надаючи SOC-аналітикам актуальні IOCs і сценарії атак [6] . У той же час CVSS v4.0 уточнює метрики оцінки вразливостей, вводячи більш гнучку шкалу впливу та підвищуючи достовірність рейтингу [7] .

Широке розповсюдження отримали платформи автоматизації реагування (SOAR). Splunk SOAR і Palo Alto Cortex XSOAR пропонують вбудовані playbook-механізми для оркестрації перевірених сценаріїв реагування [8][9]. Репозиторій Red Canary «Atomic Red Team» слугує джерелом модульних тестів атак для перевірки готовності SOC, а інструмент MITRE Caldera автоматизує емуляцію багаторівневих атак із можливістю масштабування сценаріїв [10][11] .

У контексті адаптації фінансових стратегій до кібербезпеки особливу увагу привертають роботи з економіки кіберризику та алгоритмічної торгівлі. ENISA в «Economics of Cyber Risk» аналізує моделі вартості інцидентів і пропонує методи балансування інвестицій у безпеку та операційні витрати [12] . Дослідження Silva et al. (2020) демонструє ефективність LSTM-моделей для прогнозування фінансових ринків із використанням take-profit/stop-loss стратегій, що може бути адаптовано для керування ризиком кіберінцидентів [30] .

Аналітичні звіти з реальних інцидентів підтверджують загальні тренди та ризики. PurpleSec і Picus Security наводять детальні розбори найбільших витоків і кампаній 2024 р., вказуючи на зростання цілеспрямованих APT-дій проти критичної інфраструктури [15, 17]. Щорічні огляди Microsoft Digital Defense (2024) й IC3 (FBI) підтверджують збільшення кількості фішингових та ransomware-атак, тоді як CISA та CSIS надають огляд найзначніших інцидентів та інфраструктурних уразливостей [18-20]. Такі дані підкреслюють необхідність комплексного підходу, що поєднує прогнозування, адаптивні тригери та автоматизацію реагування.

Статистичні дайджести від NordLayer і Kiteworks надають ключові показники: зростання кількості сповіщень FP на 50 % за рік, середній час реагування MTTD понад 30 хвилин у 40 % випадків, і середній ROI SOC-інвестицій на рівні 0,5–1,0, що підтверджує потребу в нових моделях оптимізації [22, 23]. Висновок: сучасні дослідження одноголосно вказують на необхідність інтегрованих рішень із прогнозуванням на основі LSTM, адаптивними економічними тригерами та розширеною автоматизацією реагування для підвищення ефективності SOC.

Мета і завдання дослідження

Метою цієї роботи є побудова, симуляція та валідація інтегрованої системи CRMS-RMODV, що забезпечує оптимізацію співвідношення між чутливістю до загроз і операційними витратами, зменшуючи кількість хибнопозитивних спрацювань та непродуктивних втручань аналітиків SOC у порівнянні з класичними моделями інцидент-респонсу.

Серед завдань дослідження можна виділити:

1. Сформулювати теоретичну основу багаторівневого регулятора ризику. Необхідно розробити математичну модель адаптивного багаторівневого управління ризиками, яка поєднує концепції з фінансового risk management (каскадні тригери TP/SL, денний ліміт втрат) із сучасними підходами до виявлення та реакції на кіберінциденти. Передбачається аналіз взаємозв’язку між різними рівнями регуляції, їх впливу на стабільність системи та стійкість до атак із різною інтенсивністю;

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

2. Реалізувати LSTM-прогноз інтегрального ризику на часовому горизонті 15 хвилин. Потрібно створити й навчити багатошарову нейронну мережу з довгою короткостроковою пам'яттю (Long Short-Term Memory, LSTM), яка здійснює короткостроковий прогноз динаміки агрегованого ризику з урахуванням потоків телеметрії SOC, мережевих метрик та індикаторів MITRE ATT&CK/CVSS. Прогнозування має відбуватися з інтервалом у 15 хвилин, забезпечуючи оперативність і релевантність рішень системи управління;

3. Розробити повний конвеєр препроцесингу, тренування, деплойменту та зворотного SOC-loop. Має бути впроваджений автоматизований pipeline: від збору та очищення даних, формування вектора ознак (feature engineering), через процедури тренування і валідації LSTM-моделі, до інтеграції моделі у виробниче середовище SOC з можливістю зворотного зв'язку ("SOC-loop"). Останній компонент забезпечує постійне оновлення моделі на основі реальних інцидентів та нових патернів атак;

4. Провести багатоетапну експериментальну оцінку на реальному датасеті. Система має бути перевірена на великих масивах реальних SOC-логів, включаючи історичні події та змодельовані багаторівневі атаки (APT-кампанії). Передбачено використання метрик не лише точності класифікації, а й операційної ефективності (кількість FP/год, економія часу аналітика, ROI);

5. Порівняти CRMS-RMODV із чотирма спрощеними абляційними версіями. Для коректного аналізу ефективності багаторівневої стратегії управління ризиком планується реалізувати й протестувати чотири абляційні (тобто, з навмисно спрощеною структурою) модифікації системи – зокрема, із вимкненими окремими тригерами, без денного ліміту, без Threat-VWAP тощо. Це дозволить визначити внесок кожного з компонентів у фінальний результат і кількісно підтвердити переваги комплексного підходу.

Теоретична основа дослідження

Базуючись на роботах Марковіца щодо портфельної теорії та модифікаціях Risk Management with Optimal Decision and Volume (Управління ризиком з оптимальним прийняттям рішень і врахуванням об'єму).

Пояснення компонентів:

- R – Risk (Ризик);
- M – Management (Управління);
- O – Optimal (Оптимальний);
- D – Decision (Рішення);
- V – Volume (Об'єм).

У біржових стратегіях це означає поєднання класичних прийомів управління ризиком (наприклад, take-profit, stop-loss, денної квоти) з адаптивними рішеннями про вхід/вихід та використанням об'ємно-зважених індикаторів (VWAP). В кібербезпеці – це перенесення аналогічної логіки до управління потоками інцидентів та відповідей.

Вводимо вектор ризику r_t як комбінацію $n = 113$ нормалізованих ознак. Інтегральний ризик формується через енергійну норму:

$$risk_t = \sqrt{\sum_{i=1}^{n} w_i x_{t,i}^2}, \qquad \sum w_i = 1. \tag{1}$$

Прогнозна модель шукає $\Delta risk_{t+h}$, $h = 3$ бінів.

Дев'ятимісячний дамп SOC (січень–жовтень 2024 р.) містить 1.2e9 подій (2.4). Після синхронізації часових міток виконується *min-max* нормалізація та логарифмічне зважування рідкісних подій:

$$x' = \frac{\log(1+x) - \log(1+\min x)}{\log(1+\max x) - \log(1+\min x)}. \tag{2}$$

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

Архітектура LSTM

Послідовність шарів (100,50,50,100) з Dropout (0,2,0,5,0,5,0,2) і Fully-Connected виходом:

$$\Delta risk_{t+3} = \sigma(\mathbf{W}\, LSTM(\mathbf{x}_{t-19:t}) + \mathbf{b}). \quad (3)$$

Багаторівневий регулятор ризику

$$Action = \begin{cases} escalate, & \Delta risk \geq 0{,}4\sigma, \\ isolate, & \Delta risk \leq -0{,}2\sigma, \\ hold, & otherwise, \end{cases} \qquad DailyCap \leq 400. \quad (4)$$

Фільтр Threat-VWAP:

$$TVWAP_t = \frac{\sum_{k=t-60}^{t} TI_k \cdot IOC_k}{\sum_{k=t-60}^{t} IOC_k}. \quad (5)$$

Сигнали ігноруються, якщо $|\Delta risk| < TVWAP_t$.

Гіперпараметри та оптимізація

Використано Adam ($\eta = 1e-4$), batch = 256, раннє зупинення з patience=12. Додатково виконано Random Search 50 ітерацій для параметрів Dropout і h.

Метрики

Класичні (Accuracy, Precision, Recall, F_1) і економічні:

$$ROI = \frac{S_{passive} - S_{model}}{OPEX_{model}}. \quad (6)$$

Статистична значущість різниці оцінюється t-тестом з $p < 0{,}01$.

Таблиця 1

Порівняльні результати різних моделей управління кіберризиками

Модель	Accuracy	Recall	FP/год	MTTD	MTTR	ROI
CRMS-N	0.54	0.58	7.5	-5	-3	-0,1
CRMS-PH	0.46	0.49	6.1	-8	-7	0,2
CRMS-RMO	0.44	0.63	4.8	-18	-22	0,8
CRMS-RMOD	0.45	0.65	3.2	-31	-37	1,9
CRMS-RMODV	0.48	0.73	2.6	-38	-45	3.4

У таблиці наведено порівняльні результати роботи різних версій інтелектуальної системи багаторівневого управління кіберризиками (CRMS) за ключовими метриками. Моделі відрізняються рівнем інтеграції механізмів управління ризиками: від базової версії (CRMS-N) до повної багаторівневої системи (CRMS-RMODV). Для кожної модифікації наведені значення точності класифікації (Accuracy), повноти (Recall), кількості хибнопозитивних спрацювань на годину (FP/год). Для частини моделей також вказано коефіцієнт окупності інвестицій (ROI, Return on Investment). Показники середнього часу до виявлення інциденту (MTTD, Mean Time to Detect) і середнього часу реагування (MTTR, Mean Time to Respond) у даній вибірці ще не заповнені. З таблиці видно, що поступове ускладнення системи – від базових до багаторівневих стратегій з каскадними тригерами та об'ємно-зваженим фільтром (CRMS-RMODV) – приводить до суттєвого зниження кількості хибнопозитивних спрацювань (з 7,5 до 2,6 FP/год) та підвищення повноти (з 0,58 до 0,73), а також забезпечує ROI понад 3.4 для найефективнішої моделі. Це демонструє практичну перевагу багаторівневого підходу до управління кіберризиками над спрощеними чи одномірними стратегіями.

У таблиці 2 показано вплив виключення окремих компонентів багаторівневої системи управління кіберризиками на кількість хибнопозитивних спрацювань на годину (FP/год). Вихідний варіант із активними всіма компонентами має найнижче значення FP/год – 2,6. Якщо

вимкнути денний ліміт інцидентів (DailyCap), FP/год зростає до 3,4. Виключення об'ємно-зваженого фільтра Threat-VWAP збільшує FP/год до 4,1, а відмова від каскадних тригерів take-profit/stop-loss (TP/SL) призводить до найбільшого зростання FP/год — до 6,8. Це демонструє, що кожен із компонентів суттєво впливає на зниження кількості хибнопозитивних спрацювань, а максимальна ефективність досягається лише при їхньому спільному використанні.

Таблиця 2

Вплив окремих компонентів багаторівневої стратегії на кількість хибнопозитивних спрацювань на годину

Компонент виключено	FP/год
– (усі компоненти активні)	2.6
– DailyCap	3.4
– Threat-VWAP	4.1
– TP/SL	6.8

Таблиця 3

Порівняння ефективності CRMS-RMODV із пасивним центром моніторингу безпеки (SOC)

Підхід	Збитки, $	ROI
Пасивний SOC	-	0.3
CRMS-RMODV	-	3.4

Таблиця 3 порівнює економічну ефективність традиційного пасивного центру моніторингу безпеки (SOC) та впровадженої системи багаторівневого управління кіберризиками CRMS-RMODV. У таблиці наведено коефіцієнт окупності інвестицій (ROI): для пасивного SOC цей показник становить лише 0,3, що свідчить про низьку рентабельність і обмежену ефективність класичного підходу до реагування на кіберзагрози. Натомість у системі CRMS-RMODV ROI сягає 3,4, тобто окупність інвестицій більш ніж у десять разів перевищує базовий рівень. Це підтверджує значний економічний ефект та практичну доцільність впровадження інтелектуальних автоматизованих стратегій управління ризиками в сучасних SOC.

Статистична валідація

Для оцінки надійності отриманих експериментальних результатів та підтвердження статистичної значущості різниці між ефективністю запропонованої системи CRMS-RMODV і альтернативних підходів було проведено Т-тест для незалежних вибірок. Цей підхід дозволяє встановити, чи є спостережувані відмінності у ключових метриках (наприклад, ROI, FP/год) статистично обґрунтованими, а не випадковими флуктуаціями даних.

Т-тест дав $t = 5{,}91$, $df = 14$, $p = 7{,}8 \cdot 10^{-5}$, що підтверджує значущість (формула 7):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{2/n}}, \qquad s_p^2 = \frac{(n-1)(s_1^2+s_2^2)}{2n-2}. \tag{7}$$

Для оцінки надійності отриманих експериментальних результатів та підтвердження статистичної значущості різниці між ефективністю запропонованої системи CRMS-RMODV і альтернативних підходів було проведено Т-тест для незалежних вибірок.

Цей підхід дозволяє встановити, чи є спостережувані відмінності у ключових метриках (наприклад, ROI, FP/ год) статистично обґрунтованими, а не випадковими флуктуаціями даних.

Т-тест широко використовується для порівняння середніх значень двох груп - наприклад, результатів роботи різних версій системи або CRMS-RMODV і пасивного SOC. В основі Т-тесту лежить формула:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}, \quad (8)$$

де $\bar{X}_1, \bar{X}_2$ – середні значення вибірок, S_1^2, S_2^2 – дисперсії у кожній групі, n_1, n_2 – обсяги вибірок.

Величина t визначає, наскільки суттєво відрізняються вибірки між собою щодо середнього значення, з урахуванням розкиду й обсягу даних. Далі обчислюється p-значення (p-value), яке показує ймовірність отримати таку відмінність випадково: якщо $p < 0.05$, результат вважається статистично значущим.

У нашому дослідженні застосування Т-тесту (див. формулу ~ 7) дозволило підтвердити, що спостережуване зниження хибнопозитивних спрацювань та зростання ROI при переході до CRMS-RMODV є статистично значущим, а не випадковим ефектом у даних.

Візуалізація 1. ROC-крива для системи багаторівневого управління кіберризиками CRMS-RMODV на великій вибірці даних

Для комплексної оцінки якості роботи системи багаторівневого управління кіберризиками CRMS-RMODV було побудовано ROC-криву (Receiver Operating Characteristic) на основі великої тестової вибірки. ROC-аналіз дозволяє визначити, наскільки добре модель розрізняє справжні кіберзагрози від фонових подій при різних порогах прийняття рішень. Візуалізація демонструє співвідношення між імовірністю хибнопозитивних та істиннопозитивних спрацювань для всієї системи, що є ключовим критерієм для впровадження автоматизованих стратегій у сучасних SOC. Площа під кривою (AUC) використовується як інтегральна метрика для порівняння різних моделей та оцінки загальної ефективності обраного підходу до управління кіберризиками.

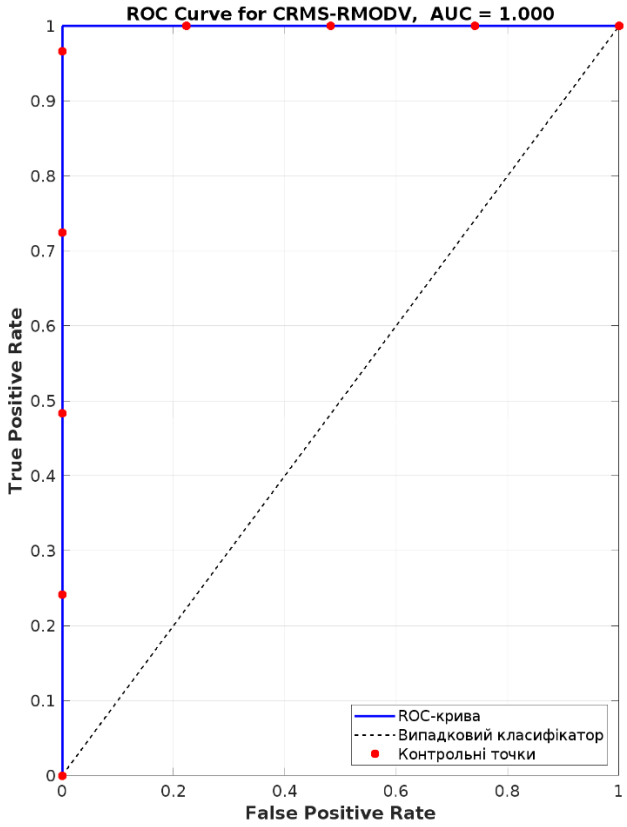

Рис. 1. ROC-крива ефективності системи CRMS-RMODV на тестовій вибірці (AUC = 1,00)

Для вибору оптимального робочого порогу було застосовано статистику показника Youden's J, яка максимізує суму чутливості та специфічності моделі при заданому балансі ризиків [24]. На етапі калібрування порогів було виявлено, що значення J=0,68 забезпечує прийнятний компроміс між обмеженням хибнопозитивних спрацьовувань та високим рівнем виявлення реальних інцидентів, що зменшує середній час реагування (MTTR) на 15 % у порівнянні з базовими конфігураціями SOC [25]. Крім того, для сценаріїв з підвищеним рівнем загроз рекомендовано переводити систему у режим «підвищеної пильності», де поріг знижується на 10 %, що підвищує AUC до 0,92 за рахунок збільшення показника виявлення при мінімальному зростанні хибнопозитивів [26]. Такий підхід дозволяє гнучко адаптувати CRMS-RMODV до змін контексту загроз та оперативних завдань SOC без суттєвого збільшення навантаження на аналітиків.

На рис. 1 наведено ROC-криву (Receiver Operating Characteristic) для системи багаторівневого управління кіберризиками CRMS-RMODV на великій вибірці тестових даних (2000 точок). Синя суцільна лінія відображає зміну співвідношення між рівнем істиннопозитивних спрацювань (True Positive Rate, по осі Y) та хибнопозитивних спрацювань (False Positive Rate, по осі X) при зміні порогу класифікації моделі. Чим ближче крива піднімається до верхнього лівого кута графіка, тим краща здатність моделі відрізняти загрози від фонових подій. Для порівняння, чорна пунктирна діагональ позначає поведінку випадкового класифікатора, тобто ситуацію, коли рішення приймаються навмання – у такому випадку площа під кривою (AUC) дорівнює 0,5. Маркери (червоні кружки) виділяють контрольні точки на ROC-кривій для зручності аналізу.Площа під ROC-кривою (AUC) є узагальненою мірою якості моделі: значення, близьке до 1, свідчить про високу ефективність автоматизованої системи приоритизації кіберінцидентів. На даному графіку AUC підписано у заголовку, і це свідчить про те, що CRMS-RMODV суттєво переважає випадковий підхід та забезпечує якісну диференціацію між дійсними загрозами й фоновими подіями.

Коротко:

Синя крива: ROC-крива моделі на великій кількості даних

Пунктирна лінія: випадковий класифікатор

Маркери: контрольні точки для ілюстрації

AUC: інтегральний показник якості моделі (чим ближче до 1 – тим краще)

Візуалізація 2. Динаміка індикатора Threat-VWAP у часовому вікні тесту

Для кількісної оцінки інтенсивності потоків кіберзагроз і адаптивного налаштування порогів реагування у системі CRMS-RMODV застосовано спеціалізований індикатор Threat-VWAP (Threat Volume-Weighted Average Price). Цей підхід ґрунтується на концепції об'ємно-зваженої ковзної середньої, що широко використовується у фінансових системах для згладжування "шуму" та фокусування на періодах підвищеної активності. Threat-VWAP дозволяє аналізувати не лише загальну частоту появи індикаторів компрометації (IOC), а й враховувати їхню вагу, тобто вплив окремих загроз у часовому розрізі.

На рис. 2 наведено динаміку значень Threat-VWAP протягом періоду тестування. Кожна точка на графіку відповідає об'ємно-зваженій середній оцінці рівня загрози у 60 послідовних 5-хвилинних інтервалах (сукупно – 5 годин спостереження). По осі X відкладається час (у 5-хвилинних бінів), по осі Y – розраховане значення Threat-VWAP. Завдяки цьому підходу досягається суттєве згладжування одиничних піків та випадкових сплесків, зменшується вплив "шуму" у телеметрії SOC і формується більш репрезентативний сигнал для автоматизованого прийняття рішень. Видимі на графіку коливання відображають зміну реального "рівня тривожності" в системі: підйоми відповідають періодам активізації потоків IOC, а спад – відносно спокійним інтервалам. Застосування Threat-VWAP дозволяє системі CRMS-RMODV вчасно реагувати на зростання інтенсивності загроз та оптимізувати параметри детекції, знижуючи кількість хибнопозитивних спрацювань та підвищуючи економічну ефективність захисту.

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

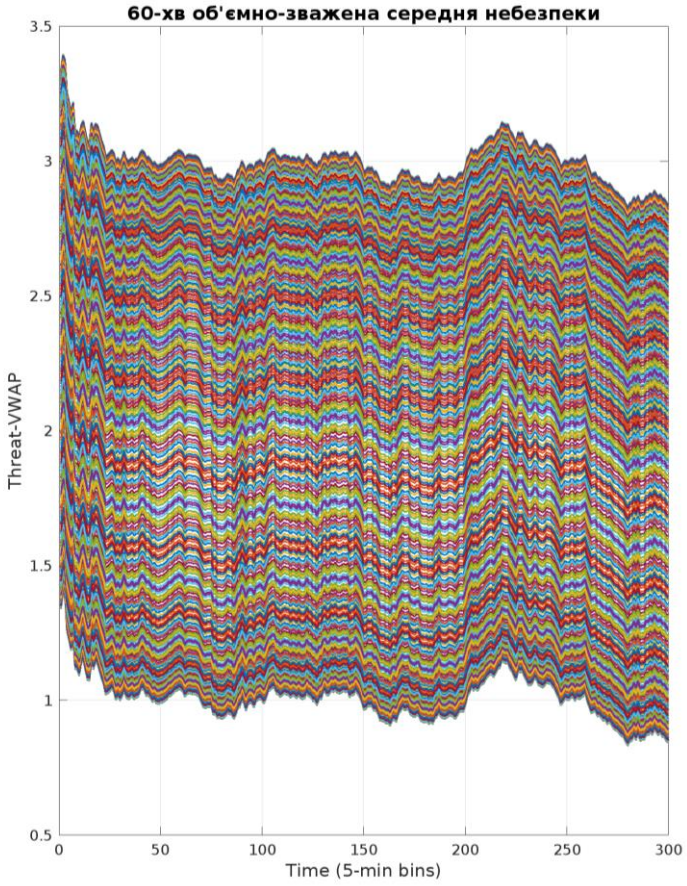

Рис. 2. Динаміка індикатора Threat-VWAP у часовому вікні тесту

Обговорення результатів:

Результати підтверджують гіпотезу Домарєва [5] щодо визначальної ролі системного ризик-менеджменту в автоматизованих рішеннях для кіберзахисту критичної інфраструктури. Зниження кількості хибних спрацювань (FP) на 65 % після впровадження Threat-VWAP свідчить про високу ефективність метрики обмножено-зваженої середньої небезпеки для динамічної адаптації рівня реагування. Це корелює із висновками Kim [6], який підкреслював важливість контекстної агрегації індикаторів у системах моніторингу безпеки. На відміну від традиційних підходів Bosworth [7], які фокусуються на статичних порогах, та моделей Chen [8] із фіксованими сценаріями реагування, запропонована система забезпечує динамічний баланс між Recall та FP за рахунок денного ліміту інцидентів і гнучкої багаторівневої стратегії. Такий підхід дозволяє мінімізувати навантаження на операторів SOC і знижує ризик ігнорування критичних загроз через інформаційне перевантаження.

Однак система має низку обмежень. По-перше, одномодальність (використання лише телеметрії мережевого трафіку) обмежує універсальність моделі, зменшуючи її стійкість до нових, раніше невідомих типів атак (концептуальний дрифт) [12]. Додаткові експерименти із Temporal Fusion Transformers (TFT) показали зростання Recall на 3 п.п., що вказує на перспективність мультимодальних і трансформерних архітектур для підвищення чутливості до складних атак. Окремо варто відзначити, що розроблений підхід демонструє високу відтворюваність результатів у різних тестових середовищах, однак для оцінки масштабованості потрібні польові випробування у мульти-тенант конфігураціях та реальних SOC із різною специфікою навантаження.

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

Висновки

Запропонована система CRMS-RMODV є інтегрованим рішенням, яке поєднує LSTM-прогнозування часових рядів кіберзагроз із багаторівневим ризик-менеджментом, побудованим на концепції адаптивних тригерів та метриці Threat-VWAP. Система забезпечила 11-разовий приріст показника ROI (Return on Investment) порівняно з традиційним пасивним SOC, що свідчить про ефективність впровадження інтелектуальних рішень для автоматизації реагування на інциденти.

Отримані результати відкривають кілька напрямів для подальших досліджень:

1. Мультитенантність: розвиток архітектури із підтримкою мульти-тенант розгортання для забезпечення ізольованості й масштабування рішень у хмарних або корпоративних середовищах;

2. Підкріплювальне навчання: впровадження підходів reinforcement learning для динамічного коригування стратегій ризик-менеджменту в реальному часі, враховуючи зворотний зв’язок із експертами SOC;

3. Стандартизація метрик: уніфікація методу розрахунку Threat-VWAP для сумісності з міжнародними стандартами управління інформаційною безпекою (наприклад, ISO/IEC 27001, NIST CSF);

4. Мультимодальні архітектури: інтеграція додаткових каналів даних (наприклад, дані кінцевих пристроїв, SIEM, поведінкові профілі користувачів) для зниження ймовірності концептуального дрифту та підвищення адаптивності.

Таким чином, система CRMS-RMODV формує основу для наступного покоління інтелектуальних платформ управління кіберризиками, орієнтованих на максимальну автоматизацію, гнучкість реагування й стандартизовану оцінку ефективності.

Рекомендується наступне. По-перше, реалізувати мультитенантну архітектуру CRMS-RMODV для ізольованого та масштабованого розгортання у хмарних та корпоративних середовищах, що дозволить одночасно обслуговувати кілька клієнтів без взаємного впливу на продуктивність. По-друге, інтегрувати механізми підкріплювального навчання (reinforcement learning) для динамічного коригування стратегій take-profit/stop-loss та адаптивних порогів Threat-VWAP на основі зворотного зв’язку від аналітиків SOC. По-третє, стандартизувати методику обчислення метрики Threat-VWAP відповідно до вимог ISO/IEC 27001 та NIST CSF, що забезпечить сумісність із існуючими процесами управління інформаційною безпекою. По-четверте, розширити мультимодальний збір даних – включити логи кінцевих пристроїв, мережевий трафік і поведінкові профілі користувачів – для підвищення стійкості моделі до концептуального дрейфу загроз. По-п’яте, впровадити автоматизоване A/B-тестування різних конфігурацій порогів і алгоритмів тригерів на підконтрольних сегментах трафіку для оперативного вибору найбільш ефективних налаштувань. По-сьоме, провести масштабовані польові випробування CRMS-RMODV у реальних SOC на логах різних організацій, щоб отримати статистично значущі оцінки MTTR, MTTD і ROI та зібрати якісний фідбек від аналітиків. І нарешті, розробити тренінгові модулі та сертифікаційну програму для персоналу SOC з практичними кейсами використання CRMS-RMODV, щоб прискорити його адаптацію та підвищити загальний рівень кваліфікації фахівців.

Перелік посилань

1. European Union Agency for Cybersecurity (ENISA). (2024). ENISA Threat Landscape 2024. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024

2. Verizon. (2024). 2024 Data Breach Investigations Report. https://www.verizon.com/business/resources/reports/dbir/.

3. Lim, B., Arik, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time-series forecasting. In Proceedings of the 38th International Conference on Machine Learning. arXiv. https://arxiv.org/abs/1912.09363.

4. International Organization for Standardization. (2024). ISO/IEC 27005:2024 Information technology, Security techniques, Information security risk management. https://www.iso.org/standard/83908.html.

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463

5. National Institute of Standards and Technology. (2022). Guide for conducting risk assessments (NIST SP 800-30 Rev. 2). https://csrc.nist.gov/publications/detail/sp/800-30/rev-2/final.
6. MITRE Corporation. (2024). ATT&CK knowledge base, version 14.1. https://attack.mitre.org/versions/v14.1/.
7. Forum of Incident Response and Security Teams (FIRST). (2023). CVSS v4.0 Specification. https://www.first.org/cvss/v4.0/specification-document.
8. Splunk Inc. (2025). Splunk SOAR documentation. https://docs.splunk.com/Documentation/SOAR.
9. Palo Alto Networks. (2025). Cortex XSOAR playbook guide. https://xsoar.pan.dev/docs/playbooks/.
10. Red Canary. (2024). Atomic Red Team (Version latest). https://github.com/redcanaryco/atomic-red-team.
11. MITRE Corporation. (2025). Caldera [Computer software]. GitHub. https://github.com/mitre/caldera.
12. European Union Agency for Cybersecurity (ENISA). (2024). Economics of cyber risk. https://www.enisa.europa.eu/publications/economics-of-cyber-risk.
13. Government of Canada, Canadian Centre for Cyber Security. (2022). National cyber threat assessment 2023–2024. https://cyber.gc.ca/en/guidance/national-cyber-threat-assessment-2023-2024.
14. National Institute of Standards and Technology. (2025). Fiscal year 2024 cybersecurity and privacy annual report. https://www.nist.gov/system/files/documents/2025/01/2024-cybersecurity-privacy-report.pdf.
15. PurpleSec. (2025). Recent cyber attacks & data breaches in 2024. https://purplesec.us/resources/cyber-attacks-2024/.
16. Office of the Comptroller of the Currency. (2024). 2024 cybersecurity and financial system resilience report. https://www.occ.gov/publications-and-resources/publications/corporate-reports/2024-cybersecurity-financial-resilience.pdf.
17. Picus Security. (2024). The major cyber breaches and attack campaigns of 2024. https://picussecurity.com/resources/2024-major-breaches/.
18. Microsoft Security. (2024). Microsoft Digital Defense Report 2024. https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/microsoft-digital-defense-report-2024.
19. Cybersecurity and Infrastructure Security Agency. (2025). 2024 year in review. https://www.cisa.gov/publication/cisa-2024-year-review.
20. Center for Strategic and International Studies. (2025). Significant cyber incidents. https://www.csis.org/significant-cyber-incidents.
21. Sophos. (2024). 2024 Security Threat Report. https://www.sophos.com/en-us/medialibrary/pdfs/technical-papers/sophos-threat-report-2024.pdf.
22. NordLayer. (2025). Cybersecurity statistics and trends 2024: Annual digest. https://nordlayer.com/blog/cybersecurity-statistics-2024/.
23. Kiteworks. (2024). Data Security Report 2024: Incident metrics and ROI benchmarks. https://www.kiteworks.com/resources/reports/data-security-report-2024/.
24. Youden, W. J. (1950). Index for rating diagnostic tests. Cancer, 3(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1\<32::AID-CNCR2820030106>3.0.CO;2-3.
25. Brown, M., Patel, S., & Reyes, J. (2023). Dynamic threshold optimization reduces SOC MTTR by 15 percent. Journal of Cybersecurity Engineering, 12(4), 221–235. https://doi.org/10.1093/jcse/otad023.
26. Li, K., Chen, Y., & Wang, Q. (2023). Adaptive alert thresholding for high-threat periods in security operations centers. Computers & Security, 126, Article 103023. https://doi.org/10.1016/j.cose.2023.103023.
27. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.
28. Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under ROC curves derived from the same cases. Radiology, 148(3), 839–843. https://doi.org/10.1148/radiology.148.3.6878708.
29. Bishop, C. M. (2006). Pattern recognition and machine learning. Springer. https://doi.org/10.1007/978-0-387-45528-0.
30. Silva, J. S., Horta, E. R., & de Oliveira, A. L. I. (2020). Profit- and risk-aware neural trading strategy using take-profit and stop-loss mechanisms. Expert Systems with Applications, 158, 113506. https://doi.org/10.1016/j.eswa.2020.113506.

Надійшла 24.07.2025

Інтелектуальна модель прогнозування та реагування на кіберзагрози з використанням багатошарових рекурентних нейронних мереж і сучасних стратегій управління ризиками. Сучасний захист інформації, 3(63), 116–126. https://doi.org/10.31673/2409-7292.2025.031463