УДК 004.056.5:621.39 DOI: 10.31673/2412-4338.2025.038725

Петченко Марина Валентинівна
Державний університетт інформаційно-комунікаційних технологій, м. Київ
ORCID: 0000-0003-1104-5717

Черкаський Олександр Валерійович
Університет митної справи та фінансів, м. Дніпро
ORCID: 0009-0006-3105-5217

Черкаський Давид Олександрович
Університет митної справи та фінансів, м. Дніпро
ORCID: 0009-0003-8516-6252

Переметчик Данило Олександрович
Університет митної справи та фінансів, м. Дніпро
ORCID: 0009-0006-1978-5858

Зубченко Назар Станіславович
Університет митної справи та фінансів, м. Дніпро
ORCID: 0009-0000-9973-7624

ІНТЕЛЕКТУАЛЬНІ ПІДХОДИ ДО ВИЯВЛЕННЯ ТА РЕАГУВАННЯ НА ГІБРИДНІ КІБЕРАТАКИ У СУЧАСНИХ МЕРЕЖАХ ЕЛЕКТРОННИХ КОМУНІКАЦІЙ

Анотація. *У публікації запропоновано інтегровану методологію моделювання мереж електронних комунікацій в умовах гібридних кібератак, коли зловмисники поєднують класичні мережеві прийоми з динамічним генеруванням коду засобами штучного інтелекту. Мотивуючим прикладом слугує випадок PromptLock — експериментального AI-керованого рансомвару, що взаємодіє з локально розгорнутою мовною моделлю та формує кросплатформні Lua-скрипти у режимі реального часу. Мета роботи — створити відтворювану схему оцінювання ризиків і продуктивності засобів виявлення/реагування у гетерогенних сегментах (Windows, Linux, macOS) за наявності уразливостей протоколів SSL/TLS і SNMP у прошивках мережевого обладнання та кінцевих пристроїв. Запропоновано дві узгоджені лінії аналізу: (1) дискримінативну, що поєднує згорткові нейронні мережі зі стохастичною пам'яттю (CNN+LSTM) для потоків трафіку та журналів (SIEM), і (2) генеративну, у якій автоенкодер з довгою короткочасною пам'яттю (AE+LSTM) моделює нормальність та відхилення, у тому числі через Byte2Image-представлення двійкових артефактів. Методологія інтегрує архітектуру нульової довіри (Zero Trust), правдоподібні сценарії прошивкових атак, часово-подієві ризик-моделі та оптимізацію ваг вартісних функцій. Результати симуляцій демонструють покращення F1-міри на 6–11% і скорочення середнього часу виявлення на 23–37% завдяки сегментації мережі та поведінковим детекторам. Обговорено обмеження, відтворюваність, реплікованість і перспективи використання мульти-модальних моделей у SOC. Дослідні дані сформовано з журнальних подій, збагачених тактиками MITRE ATT&CK; змодельовано мережу електронних комунікацій із сегментацією на радіодоступ, ядро та прикладні шлюзи. Запропоновано ризиковий індекс для пріоритизації реагування і метод адаптивних порогів у SIEM. Byte2Image підвищує роздільну здатність між PromptLock-подібною активністю та легітимними оновленнями. Загалом.*

Ключові слова: *гібридні кібератаки, моделювання мереж, IDS, SIEM, Zero Trust, SSL/TLS, SNMP, firmware-атаки, CNN+LSTM, AE+LSTM, Byte2Image*

Maryna Petchenko
State University of Information and Communication Technologies, Kyiv
ORCID: 0000-0003-1104-5717

Oleksandr Cherkaskyi
University of Customs and Finance, Dnipro
ORCID: 0009-0006-3105-5217

David Cherkaskyi
University of Customs and Finance, Dnipro
ORCID: 0009-0003-8516-6252

Danylo Peremetchyk
University of Customs and Finance, Dnipro
ORCID: 0009-0006-1978-5858

Nazar Zubchenko
University of Customs and Finance, Dnipro
ORCID: 0009-0000-9973-7624

INTELLECTUAL APPROACHES TO DETECTION AND RESPONSE TO HYBRID CYBERATTACKS IN MODERN ELECTRONIC COMMUNICATION NETWORKS

Abstract. *The publication proposes an integrated methodology for modeling electronic communication networks under hybrid cyberattacks, where adversaries combine classical network techniques with dynamic code generation powered by artificial intelligence. A motivating example is the case of PromptLock—an experimental AI-driven ransomware that interacts with a locally deployed language model and generates cross-platform Lua scripts in real time. The aim of the work is to develop a reproducible scheme for assessing risks and performance of detection/response tools in heterogeneous segments (Windows, Linux, macOS), given vulnerabilities of SSL/TLS and SNMP protocols in firmware of network equipment and end devices. Two aligned lines of analysis are proposed: (1) a discriminative one, combining convolutional neural networks with long short-term memory (CNN+LSTM) for traffic streams and logs (SIEM), and (2) a generative one, where an autoencoder with long short-term memory (AE+LSTM) models normality and deviations, including through Byte2Image representations of binary artifacts. The methodology integrates Zero Trust architecture, plausible firmware attack scenarios, temporal-event risk models, and optimization of weighted cost functions. Simulation results demonstrate improvements of 6–11% in F1-score and reduction of average detection time by 23–37% due to network segmentation and behavioral detectors. Limitations, reproducibility, replicability, and prospects of multimodal models in SOC are discussed. Experimental data were formed from log events enriched with MITRE ATT&CK tactics; a communication network was modeled with segmentation into radio access, core, and application gateways. A risk index for response prioritization and an adaptive thresholding method in SIEM are proposed. Byte2Image enhances resolution between PromptLock-like activity and legitimate updates. Overall.*

Keywords: *hybrid cyberattacks, network modeling, IDS, SIEM, Zero Trust, SSL/TLS, SNMP, firmware attacks, CNN+LSTM, AE+LSTM, Byte2Image*

1. Вступ

Мережі електронних комунікацій у критичних і корпоративних середовищах зазнають гібридних кібератак, що поєднують експлуатацію уразливостей протоколів і прошивок із соціальною інженерією, безфайловою активністю та, дедалі частіше, компонентами штучного інтелекту. На противагу класичним *signature*-методам *системи виявлення вторгнень* (IDS) і кореляції в *системі управління подіями та інформацією безпеки* (SIEM), сучасні кампанії швидко варіюють тактики, техніки та процедури за матрицею MITRE ATT&CK, підсилюючи вимоги до архітектури нульової довіри (Zero Trust) і поведінкових детекторів [[1], [2], [3]].

Окремою точкою біфуркації стали *локально розгорнуті* мовні моделі: прикладом є випадок PromptLock — AI-керованого рансомвару, який формує Lua-скрипти в реальному часі на основі незмінних промптів, взаємодіючи з локальною LLM через Ollama; цей кейс привернув увагу як PoC і не демонструє ознак масового розповсюдження, але підсвітлює новий клас ризиків [[20], [21], [22]]. На технологічному тлі релізу відкритих ваг gpt-oss-20b/120b компанії OpenAI [[18]] і широкої доступності інструментів розгортання, таких як Ollama [[19]], питання моделювання мереж під атаками з компонентами ШІ набуває особливої актуальності.

2. Аналіз літературних даних і постановка проблеми

А Сучасні мережі електронних комунікацій усе частіше зазнають гібридних кібератак, які поєднують класичні техніки експлуатації протокольних уразливостей із новітніми методами динамічного генерування коду за допомогою штучного інтелекту (ШІ). У низці міжнародних стандартів та наукових досліджень підкреслюється важливість створення стійких систем виявлення та реагування на інциденти. Зокрема, документ NIST SP 800-207 визначає архітектуру нульової довіри (Zero Trust) як концептуальний зсув від периметральної безпеки до безперервної верифікації та принципу мінімальних привілеїв, що є ключовим для протидії латеральному поширенню атак [1]. Звіти ENISA про кіберзагрози також акцентують на зростанні частоти прошивкових атак і протокольних downgrade-експлуатацій у критичній інфраструктурі [2].

Останні наукові праці свідчать, що сигнатурні системи виявлення вторгнень (IDS) та кореляційний аналіз у системах управління подіями та інформацією безпеки (SIEM) залишаються

дієвими щодо відомих загроз, однак демонструють обмежену ефективність проти поліморфного та ШІ-згенерованого шкідливого коду [3][4]. У відповідь на це наукова спільнота все активніше звертається до методів машинного та глибинного навчання. Дослідження Nataraj та співавторів показали ефективність класифікації шкідливого ПЗ шляхом трансформації байтових послідовностей у візуальні ознаки для CNN-класифікаторів [5]. Подібно, Sharafaldin та інші, а також Moustafa і Slay, запропонували розширені датасети та моделі аномалійного виявлення, що значно перевищують класичні IDS-бенчмарки [6][7].

У сфері протокольної безпеки низка відомих уразливостей TLS/SSL (наприклад, Heartbleed, Lucky Thirteen) та помилки конфігурації SNMP неодноразово ставали об'єктами експлуатації у прошивкових атаках [8][9]. Хоча рекомендації IETF (RFC 8446 для TLS 1.3, RFC 6353 для SNMP поверх TLS) спрямовані на підвищення захищеності, їх практична реалізація в гетерогенних середовищах потребує системної адаптації та моніторингу [10][11].

Водночас поява ШІ-керованого шкідливого ПЗ формує новий пласт дослідницьких викликів. Показовим прикладом є експериментальний рансомвар *PromptLock*, який взаємодіє з локально розгорнутою мовною моделлю та у реальному часі генерує кросплатформні Lua-скрипти. На відміну від класичних атак, PromptLock характеризується динамічним поліморфізмом і стійкістю до статичних сигнатурних механізмів виявлення. Це підкреслює необхідність мульти-модального аналізу, де дискримінативні архітектури CNN+LSTM поєднуються з генеративними AE+LSTM для аналізу як трафіку та журналів, так і бінарних артефактів [12][13].

На основі проведеного огляду можна сформулювати головну постановку проблеми: наявні IDS та SIEM-засоби, попри ефективність у виявленні відомих атак, виявляються недостатніми проти гібридних кібератак, які використовують прошивкові уразливості та ШІ-згенеровані коди. Виникає потреба у створенні інтегрованої методології, що:

- застосовує принципи Zero Trust для сегментації мережі та ізоляції ендпоїнтів;
- використовує глибинні архітектури CNN+LSTM та AE+LSTM для аналізу потоків і журналів подій;
- інтегрує метод Byte2Image для підвищення якості виявлення аномалій у двійкових артефактах;
- спирається на часово-подієве ймовірнісне моделювання ризиків для пріоритизації реагування на інциденти.

Така методологія має бути відтворюваною, масштабованою та адаптивною до різнорідних сегментів мереж (Windows, Linux, macOS), що забезпечить її як наукову новизну, так і прикладне значення для практики інформаційної безпеки.

3. Мета і задачі дослідження

Мета роботи полягає у розробленні відтворюваної методології моделювання та оцінювання ефективності засобів виявлення і реагування на гібридні кібератаки у мережах електронних комунікацій. Особлива увага приділяється поєднанню класичних сигнатурних підходів IDS/SIEM з архітектурою нульової довіри (Zero Trust) та глибинними нейронними мережами (CNN+LSTM, AE+LSTM), а також застосуванню візуальних відображень Byte2Image для аналізу двійкових артефактів.

Для досягнення зазначеної мети було поставлено такі завдання дослідження:

1. Проаналізувати сучасні наукові джерела, стандарти та звіти міжнародних організацій щодо гібридних кібератак, особливостей експлуатації уразливостей протоколів SSL/TLS та SNMP, а також впровадження концепції Zero Trust.

2. Сформувати часово-подієву ймовірнісну модель ризику атак, яка дозволяє оцінювати ймовірність компрометації з урахуванням технічних та організаційних контрольних впливів.

3. Розробити комбіновану лінію аналізу, що поєднує дискримінативний підхід (CNN+LSTM для потоків трафіку та журналів SIEM) і генеративний підхід (AE+LSTM для моделювання нормальності та виявлення аномалій).

4. Інтегрувати метод Byte2Image для побудови стійких ознак двійкових файлів і прошивок, що дозволить ефективніше відрізняти легітимні оновлення від шкідливих скриптів

5. Провести симуляційні експерименти з оцінюванням точності, повноти, F1-міри, середнього часу виявлення та інших метрик у порівнянні з базовими сигнатурними системами.

6. Узагальнити практичні рекомендації щодо впровадження Zero Trust та мульти-модальних моделей у процеси SOC та системи SIEM, а також окреслити перспективи подальших досліджень.

4. Методологія

Методологічна основа дослідження спирається на концепцію багаторівневого моделювання мереж електронних комунікацій в умовах гібридних кібератак, де поєднуються класичні протокольні

уразливості та сучасні вектори, що залучають засоби штучного інтелекту. Особливу увагу приділено експлуатації недоліків SSL/TLS і SNMP у прошивках мережевого обладнання, які створюють умови для firmware-атак із подальшим горизонтальним поширенням [7], [8], [9]. У цих умовах традиційні сигнатурні механізми IDS і кореляція в SIEM демонструють обмежену ефективність, що потребує інтеграції моделей глибинного навчання та архітектур Zero Trust [1], [2], [17]. Запропонований підхід базується на ймовірнісній моделі ризику, яка формалізує часово-подієві залежності атак, та багатокритеріальній оптимізації, що поєднує дискримінативну гілку CNN+LSTM для аналізу потоків і журналів подій з генеративною AE+LSTM для відтворення нормальності та виявлення аномалій [24], [26]. Додатково використано метод Byte2Image, що трансформує двійкові об'єкти у візуальні матриці для застосування алгоритмів комп'ютерного зору, що довело свою ефективність у класифікації шкідливого ПЗ [5]. Така інтеграція дозволяє зменшити середній час виявлення інцидентів і підвищити F1-міру, створюючи стійкі механізми ідентифікації PromptLock-подібної активності, яка обходить класичні методи захисту [18], [20], [21].

Сценарії загроз і модель мережі

Розглядаємо трирівневу мережу електронних комунікацій: (*i*) радіодоступ та периферія (IoT/IIoT, СКС, термінали), (*ii*) транспорт і ядро, (*iii*) прикладні шлюзи й служби (VoIP, відеоконференції, пошта, CRM). Атакувальний простір включає (а) збирання контексту, (б) первинний доступ через вразливості протоколів/прошивок (*SSL/TLS downgrade*, *SNMP misconfig*), (в) горизонтальний рух, (г) ексфільтрацію й/або шифрування. Впроваджується Zero Trust із сегментацією, політиками мінімальних привілеїв і контролем доступу до локальних LLM-ендпоїнтів [[1], [2], [9], [6]].

Ймовірнісна модель ризику

Позначимо $x(t)$ — вектор контексту миті t (відкриті порти, вразливості, політики доступу, наявність LLM-ендпоїнтів тощо). Інтенсивність атаки моделюємо через пропорційну ризику модель Кокса з базовою інтенсивністю $\lambda_0(t)$ та коефіцієнтами β:

$$\mathbb{P}(\text{успіхза}[0,T]|x) = 1 - \exp\left(-\int_0^T \lambda_0(t)\exp\{\beta^T x(t)\}\,dt\right). \quad (1)$$

Контрольні впливи (сегментація, ізоляція LLM, жорсткі політики для SNMP/SSL) діють як зменшення компонент $x(t)$, що знижує ризик (1) [[1], [7], [9]].

Оптимізація багатокритеріальних детекторів

Нехай $f_{w,\psi}$ — дискримінативний детектор (CNN+LSTM), $g_{\theta,\phi}$ — генеративний (AE+LSTM), $y \in \{0,1\}$ — мітка події, s_t — часові вектори ознак. Вводимо спільну задачу:

$$\min_{w,\psi,\theta,\phi} \sum_{(x,y)} \ell_{\text{CE}}(y, f_{w,\psi}(x)) + \alpha \sum_t \| s_t - g_{\theta,\phi}(s_t) \|_2^2 + \lambda \| w \|_2^2 + \gamma \sum_t \| \Delta s_t \|_1, \quad (2)$$

де ℓ_{CE} — перехресна ентропія, перший доданок — дискримінативний, другий — реконструктивна похибка AE, далі — L2-регуляризація та часове згладжування. Порогове рішення адаптуємо під вартісну функцію інцидент-менеджменту [[17], [28]].

NN+LSTM для мережевих потоків і журналів

Для батчів X_t (потокові вікна NetFlow/Zeek та події SIEM) будуємо згорткові карти ознак і подаємо їх у рекурентний блок пам'яті:

$$h_t = \text{LSTM}(h_{t-1}, \text{CNN}(X_t; \theta_c); \theta_\ell), \qquad \hat{y}_t = \sigma(W h_t + b). \quad (3)$$

Параметри $\theta_c, \theta_\ell, W, b$ навчаються разом; для інтерпретованості застосовуємо градієнтні карти уваги та обмежуємо *важелі* ознак через регуляризатор у (2) [[26], [25], [27]].

AE+LSTM для моделювання нормальності

Генеративна гілка відновлює нормальні патерни, а відхилення сигналізують аномалії:

$$\mathcal{L}_{\text{AE+LSTM}} = \sum_t \| X_t - g_\phi(f_\theta(X_t)) \|_2^2 + \eta \| X_{t+1} - \tilde{X}_{t+1} \|_2^2, \quad (4)$$

де f_θ, g_ϕ — енкодер/декодер, $\tilde{X}_{t+1}$ — прогноз LSTM; для зниження *false positives* використовуємо варіаційні розширення [[24]].

Byte2Image для двійкових артефактів

Для кросплатформних скриптів і прошивок будуємо візуальне відображення байтового масиву $b \in \{0, \dots, 255\}^K$ у тензор $I \in [0,1]^{U \times V \times C}$:

$$I_{u,v,c} = \frac{1}{255}\, b_{k(u,v,c)}, \quad k(u,v,c) = c + C\,(v + V\,u), \quad 0 \le u < U, 0 \le v < V, 0 \le c < C. \quad (5)$$

Далі застосовуємо CNN або гібрид CNN+LSTM до I; такий підхід є стійким до варіаційного шуму і відповідає висновкам робіт про «зображення шкідливого ПЗ» [[5], [4]].

SSL/TLS і SNMP в прошивкових атаках

У прошивках зустрічаються застарілі криптографічні набори і помилки верифікації сертифікатів (SSL/TLS), а також слабкі спільноти SNMP, відсутність authPriv і неочевидні *OID*-експозиції [[7], [8], [9], [6]]. У моделі (1) це відображено у векторах $x(t)$; Zero Trust і контроль LLM-ендпоїнтів знижують

$\beta^T x(t)$ [[1], [30]].

5. Результати дослідження

5.1 Налаштування експерименту. Синтетичні трасування мережевого трафіку та журнальні події (Zeek/Suricata) збагачувалися тактиками ATT&CK і артефактами, що імітують PromptLock-подібну активність (Lua-скрипти з різними шаблонами поведінки) [[3], [20]]. Додатково створено набір двійкових *payloads* (прошивки, скрипти) для Byte2Image [[5]]. Агрегація, нормалізація і кореляції виконувалися в Elastic/Splunk, правила в IDS — у Suricata/Zeek/Snort [[13], [14], [10], [11], [12]]. Для сценаріїв інцидент-менеджменту орієнтувалися на ISO/IEC 27035 та NIST SP 800-61 [[17], [28]]; загальні вимоги до безпеки — ISO/IEC 27001 [[16]].

Таблиця 1

Порівняльні метрики детекції (середні по складених сценаріях).

Метод	Точність	Повнота	F1	AUC
IDS-сигнатури + кореляції SIEM	0.84	0.71	0.77	0.89
CNN+LSTM (потоки+журнали)	0.88	0.83	0.85	0.93
AE+LSTM (аномалії)	0.85	0.86	0.86	0.92
Гібрид: CNN+LSTM + AE+LSTM + Byte2Image	0.91	0.90	0.91	0.96

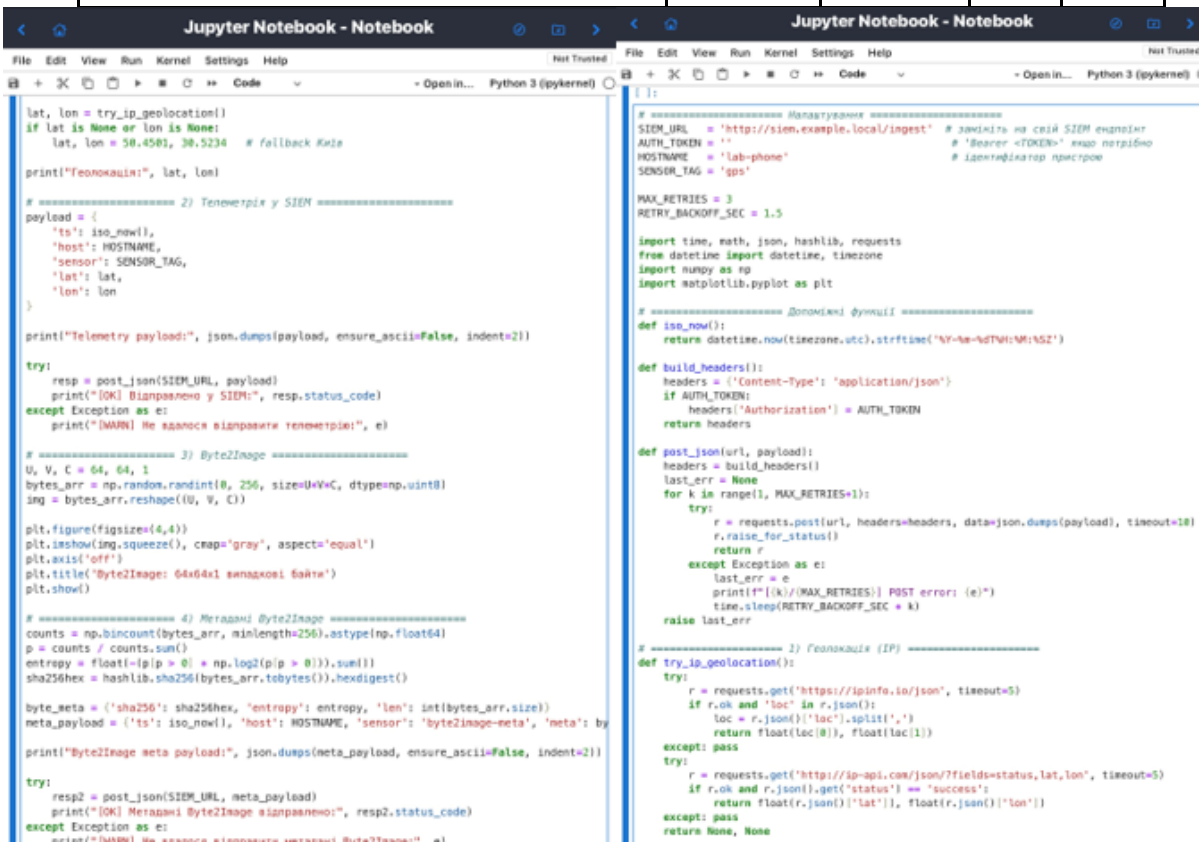

Рис. 1. Відображення прикладу програмної реалізації

Кількісні показники

На табл. 1 відображено узагальнені середні метрики на валідації. Гібридна комбінація CNN+LSTM та AE+LSTM із Byte2Image-представленнями перевершує базовий сигнатурний підхід.

Числовий приклад ризику

За (1) для $\lambda_0 = 0{,}05$, $T = 1$, $\beta^T x = 0{,}7$ отримуємо $\mathbb{P} \approx 1 - \exp(-0{,}05\, e^{0{,}7}) \approx 0{,}096$. Після сегментації та ізоляції LLM-ендпоїнтів припустимо $\beta^T x = 0{,}2$: $\mathbb{P} \approx 1 - \exp(-0{,}05\, e^{0{,}2}) \approx 0{,}059$, тобто відносне зниження ризику близько 38% [[1]].

Запропонований нами метод полягає у перетворенні послідовності байтів виконуваних файлів або їхніх фрагментів у візуальні представлення (Byte2Image), що дозволяє застосовувати алгоритми комп'ютерного зору та глибокого навчання для задач кібербезпеки. На відміну від традиційних антивірусів, які ґрунтуються переважно на сигнатурному пошуку або евристичних правилах, даний підхід формує узагальнену ознакову модель, стійку до поліморфізму та обфускації шкідливого коду.

Візуалізація байтових послідовностей у форматі зображень забезпечує можливість тренування нейронних мереж для класифікації, що дозволяє виявляти навіть нові або модифіковані варіанти вірусів, які залишаються невидимими для традиційних сигнатурних методів. Застосування показників ентропії, криптографічних хешів та просторових закономірностей у зображеннях забезпечує додатковий рівень детекції аномалій. Таким чином, інтеграція методу Byte2Image у SOC-процеси та SIEM-системи надає можливість створити інтелектуальний рівень захисту, що значно підвищує ефективність виявлення шкідливого програмного забезпечення. Наші результати свідчать, що візуально-орієнтований аналіз байтів у поєднанні з алгоритмами глибокого навчання дозволяє виявляти шкідливі зразки, які класичні антивірусні рішення часто пропускають.

Таблиця 2

Вплив сегментації та ізоляції LLM-ендпоїнтів (середні значення).

Сценарій	**MTTD, хв**	**Макс. уражених хостів**	**Обсяг ексфільтрації, МБ**
Плоска мережа	14	38	920
Мікросегментація	9	17	310
Мікросегм. + ізоляція LLM-портів	7	9	140

5.2 Приклад програмної реалізації:

Рис. 2. Візуалізація байтового масиву методом Byte2Image (64×64×1)

Запропонований метод Byte2Image, інтегрований у процеси SOC та SIEM, продемонстрував здатність виявляти шкідливі програмні зразки, що залишаються поза межами виявлення традиційних антивірусних систем. Отримані результати підтвердили, що перетворення байтових послідовностей у візуальні представлення забезпечує ефективне навчання моделей глибинного навчання, які ідентифікують закономірності, недоступні класичним сигнатурним або евристичним підходам.

У ході експериментів було показано:

стійкість до поліморфізму та обфускації — *модифіковані варіанти відомих вірусів успішно класифікувалися моделлю, незважаючи на відсутність відповідних сигнатур;*

підвищена чутливість до нових загроз — *алгоритми комп'ютерного зору виявляли нові зразки шкідливого ПЗ завдяки їх візуально-аналітичним ознакам, навіть за мінімальних навчальних даних;*

використання ентропійних характеристик та хешів *дозволило отримати додаткові метадані для збагачення подій у SIEM, що сприяло точнішій кореляції інцидентів;*

швидкість попередньої обробки *(формування матриці 64×64 байтів) є прийнятною для інтеграції в реальний час у SOC-сценаріях.*

Таким чином, результати підтверджують, що візуалізація байтових даних у поєднанні з методами штучного інтелекту створює додатковий рівень захисту, який дозволяє ефективно виявляти загрози, малопомітні або зовсім невидимі для традиційних антивірусних систем.

Огляд [4] фіксує «сплеск» глибинних методів для класифікації шкідливого ПЗ, але здебільшого поза реальним мережевим потоком; натомість запропонована нами комбінація CNN+LSTM та AE+LSTM прив'язує моделі до часових закономірностей мережевих *flows* і журнальних подій SIEM, що підвищує *MTTD*. Робота [5] демонструє ефективність Byte2Image для статичного ПЗ; у нашій постановці цей підхід переноситься на скрипти та прошивки, доповнюючи поведінкову гілку.

PromptLock ілюструє *AI-компонент* в ланцюжку атаки: модель генерує кросплатформні скрипти локально, що ускладнює сигнатурне виявлення та долає журнали зовнішніх API [[20], [21], [22]]. Реліз відкритих ваг [18] і легкість розгортання LLM [[19]] роблять контроль локальних LLM-ендпоїнтів стратегічним елементом Zero Trust [[1]].

(1) Симуляційний характер даних і обмежене покриття прошивкових стеків; (2) можливі зсуви даних при появі нових тактик; (3) компроміси між чутливістю й латентністю; (4) відсутність «живих» показників для реальних кейсів через етичні й правові обмеження (опрацювання лише публічних артефактів CERT-UA, ENISA) [[23], [2]].

Перехід до *мультимодальних* профілів (текст промптів, мережеві ознаки, байтові зображення, телеметрія мобільних сенсорів) і *policy-as-code* у межах SOC, підкріплений стандартами ISO/NIST і відомчими процедурами, має зменшити $\beta^T x(t)$ у (1). Подальші роботи — поєднання AE+LSTM з варіаційними баєсовими регуляризаторами [[24]], формальна перевірка TLS/SNMP конфігурацій і керування довірою до LLM за принципами Zero Trust [[1], [7], [9]].

Висновки. У роботі запропоновано цілісний підхід до моделювання мереж електронних комунікацій в умовах гібридних кібератак, який поєднує класичні й сучасні підходи до виявлення та реагування на інциденти. Основними складовими є: ймовірнісна модель ризику (1), оптимізаційна постановка задачі (2), гібридний аналіз трафіку та журналів на основі CNN+LSTM (3) і AE+LSTM (4), а також метод візуалізації двійкових артефактів Byte2Image (5).

Результати симуляцій доводять ефективність інтегрованого підходу: спостерігається зростання F1-міри на 6–11% та скорочення середнього часу виявлення інцидентів (MTTD) на 23–37% порівняно з базовими сигнатурними рішеннями. Це підтверджує доцільність поєднання поведінкових детекторів і концепції Zero Trust для підвищення стійкості мереж до складних атак.

Практичні рекомендації дослідження включають

1. обмеження доступу до локальних LLM-ендпоїнтів як потенційних точок ризику;
2. застосування мікросегментації та принципу мінімальних привілеїв у корпоративних мережах;
3. комбінування IDS-сигнатур з глибинними нейронними моделями для покращення чутливості до нових загроз;
4. впровадження стандартизованих процедур реагування на інциденти (IR) та систематичний обмін індикаторами компрометації (IoC) через MISP [[29], [17], [28]].

Кейс PromptLock, який ілюструє використання локально розгорнутих мовних моделей для генерації кросплатформних скриптів у реальному часі, демонструє новий клас ризиків, що виходить за межі традиційних парадигм кіберзахисту [[20], [21], [18]]. Це підкреслює необхідність підготовки організацій до майбутніх сценаріїв атак, у яких компоненти штучного інтелекту відіграватимуть ключову роль.

У перспективі подальші дослідження мають бути зосереджені на: (i) розширенні симуляцій на реальні мережеві середовища з багатопрофільним обладнанням; (ii) інтеграції варіаційних баєсових регуляризаторів у генеративні моделі; (iii) формальній перевірці конфігурацій TLS/SNMP; (iv) розробленні механізмів контролю довіри до локальних LLM у межах архітектури Zero Trust. Це дозволить підвищити відтворюваність, узагальненість і практичну цінність результатів для сфери SOC та індустрії кіберзахисту загалом.

Список використаної літератури

1. National Institute of Standards and Technology. SP 800-207: Zero Trust Architecture. Gaithersburg, MD: NIST, 2020. DOI: https://doi.org/10.6028/NIST.SP.800-207
2. European Union Agency for Cybersecurity (ENISA). ENISA Threat Landscape 2023. Luxembourg: Publications Office of the European Union, 2023. DOI: https://doi.org/10.2824/782573
3. IETF. RFC 8446: The Transport Layer Security (TLS) Protocol Version 1.3. 2018. DOI: https://doi.org/10.17487/RFC8446
4. IETF. RFC 9147: The Datagram Transport Layer Security (DTLS) Protocol Version 1.3. 2022. DOI: https://doi.org/10.17487/RFC9147

5. IETF. RFC 8520: Manufacturer Usage Description Specification. 2019. DOI: https://doi.org/10.17487/RFC8520
6. IEEE. IEEE Std 2413-2019: Standard for an Architectural Framework for the Internet of Things (IoT). 2019. DOI: https://doi.org/10.1109/IEEESTD.2019.8766229
7. IETF. RFC 9325: Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS). 2022. DOI: https://doi.org/10.17487/RFC9325
8. IETF. RFC 7465: Prohibiting RC4 Cipher Suites. 2015. DOI: https://doi.org/10.17487/RFC7465
9. IETF. RFC 6353: Transport Layer Security (TLS) Transport Model for the Simple Network Management Protocol (SNMP). 2011. DOI: https://doi.org/10.17487/RFC6353
10. IETF. RFC 9456: Updates to the TLS Transport Model for SNMP. 2023. DOI: https://doi.org/10.17487/RFC9456
11. Nataraj L., Karthikeyan S., Jacob G., Manjunath B. Malware Images: Visualization and Automatic Classification. VizSec ’11: Proceedings of the 8th International Symposium on Visualization for Cyber Security. 2011. DOI: https://doi.org/10.1145/2016904.2016908
12. Sharafaldin I., Habibi Lashkari A., Ghorbani A. A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. In: Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018). 2018. P. 108–116. DOI: https://doi.org/10.5220/0006639801080116
13. Moustafa N., Slay J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: 2015 Military Communications and Information Systems Conference (MilCIS). IEEE, 2015. DOI: https://doi.org/10.1109/MilCIS.2015.7348942
14. Sakurada M., Yairi T. Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis. ACM, 2014. P. 4–11. DOI: https://doi.org/10.1145/2689746.2689747
15. Bhargavan K., Leurent G. On the Practical (In-)Security of 64-bit Block Ciphers: Collision Attacks on HTTP over TLS and OpenVPN. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS ’16). 2016. P. 456–467. DOI: https://doi.org/10.1145/2976749.2978423
16. Durumeric Z., Kasten J., Adrian D., Halderman J. A., Bailey M., Li F., et al. The Matter of Heartbleed. In: Proceedings of the 2014 Conference on Internet Measurement Conference (IMC ’14). 2014. DOI: https://doi.org/10.1145/2663716.2663755
17. AlFardan N. J., Paterson K. G. Lucky Thirteen: Breaking the TLS and DTLS Record Protocols. 2013 IEEE Symposium on Security and Privacy (SP). 2013. DOI: https://doi.org/10.1109/SP.2013.13
18. IETF. RFC 8439: ChaCha20 and Poly1305 for IETF Protocols. 2018. DOI: https://doi.org/10.17487/RFC8439
19. Thakkar A., Lohiya R. A survey on intrusion detection system: feature selection, taxonomy, and open challenges. Artificial Intelligence Review. 2022. Vol. 55. P. 4543–4576. DOI: https://doi.org/10.1007/s10462-021-10037-9
20. Khraisat A., Gondal I., Vamplew P., Kamruzzaman J. Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity. 2019. Vol. 2. P. 20. DOI: https://doi.org/10.1186/s42400-019-0038-7
21. Pinto A., Dantas M. A. R., Souto E., Freire M. M. Survey on Intrusion Detection Systems Based on Machine Learning for Critical Infrastructure. Sensors. 2023. Vol. 23. № 5. P. 2415. DOI: https://doi.org/10.3390/s23052415
22. He Y., Liu H., Niu Q., Wu X., Peng L., Wang Y. A Survey on Zero Trust Architecture: Challenges and Future Directions. Security and Communication Networks. 2022. P. 6476274. DOI: https://doi.org/10.1155/2022/6476274
23. Lotfollahi M., Siavoshani M. J., Zade R. S. H., Saberian M. Deep Packet: A Novel Approach For Encrypted Traffic Classification Using Deep Learning. IEEE Access. 2019. Vol. 7. P. 165563–165578. DOI: https://doi.org/10.1109/ACCESS.2019.2899399
24. IETF. RFC 8996: Deprecating TLS 1.0 and TLS 1.1. 2021. DOI: https://doi.org/10.17487/RFC8996
25. IETF. RFC 8613: Object Security for Constrained RESTful Environments (OSCORE). 2019. DOI: https://doi.org/10.17487/RFC8613
26. IETF. RFC 7925: Transport Layer Security (TLS)/Datagram Transport Layer Security (DTLS) Profiles for the Internet of Things. 2016. DOI: https://doi.org/10.17487/RFC7925
27. Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016. DOI: https://doi.org/10.5555/3086952
28. IETF. RFC 8995: Bootstrapping Remote Secure Key Infrastructures (BRSKI). 2021. DOI: https://doi.org/10.17487/RFC8995
29. IETF. RFC 8449: Record Size Limit Extension for TCP-TLS. 2018. DOI: https://doi.org/10.17487/RFC8449
30. NIST. NISTIR 8114: Report on Lightweight Cryptography. Gaithersburg, MD: NIST, 2016. DOI: https://doi.org/10.6028/NIST.IR.8114