ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ ІНФОРМАЦІЇ
КАФЕДРА СИСТЕМ ТА ТЕХНОЛОГІЙ КІБЕРБЕЗПЕКИ

ВСЕУКРАЇНСЬКА НАУКОВО-ПРАКТИЧНА КОНФЕРЕНЦІЯ

«ЦИФРОВА ТРАНСФОРМАЦІЯ КІБЕРБЕЗПЕКИ»

Тези доповідей

29 квітня
2026

м. Київ

Рекомендовано до друку кафедрою Сиситем та технологій Державного університету інформаційно-комунікаційних технологійй (протокол № 8/2 від 30.04.2026 р.)

Цифрова трансформація кібербезпеки: Матеріали Всеукраїнської науково-практичної конференції (м. Київ, 29 квітня 2026 року). Навчально-науковий інститут кібербезпеки та захисту інформації, Державний університет інформаційно-комунікаційних технологій. Київ, 2026. 131 с. Збірник призначений для науковців, викладачів, докторантів, аспірантів і студентів закладів вищої освіти, фахівців з кібербезпеки та захисту інформації, працівників органів державної влади та місцевого самоврядування. Редакційна колегія не несе відповідальності за зміст матеріалів, що опубліковані у збірнику. Тези подані в авторській редакції та відображають персональну позицію учасників конференції.

Зміст

Bobyk Y.V.
PhD student, ICT, LPNU
Lviv, Ukraine

MONITORING OF SERVICE QUALITY IN CLOUD AND IOT SYSTEMS: APPROACHES AND CHALLENGES

The rapid evolution of cloud computing combined with Internet of Things (IoT) technologies has produced highly intricate large-scale distributed architectures. These infrastructures demand reliable mechanisms for service quality monitoring to sustain system dependability, operational efficiency, and user satisfaction. The current work analyzes up-to-date techniques for evaluating service quality in Cloud-IoT settings, outlines the primary obstacles encountered, and proposes promising directions for continued research. Special attention is given to the challenges associated with scaling and handling data with minimal latency, and integrating advanced monitoring strategies based on AI.

Keywords: Cloud, IoT, Service Quality, Monitoring, Edge Computing, Performance.

The convergence of cloud computing and IoT has produced broad distributed systems implemented across fields such as intelligent urban systems, medical monitoring, industrial automation, and other sectors. Guaranteeing consistent service quality is vital given the heterogeneous devices, variable processing demands, and substantial volumes of data these systems handle. Service quality is generally assessed according to core indicators such as latency, availability, throughput, and reliability [1, p. 1]. Mathematically, the service quality of a system can be formulated as a function of these variables:

$$Q_s = f(A_v, L_t, T_h, R_l) \quad (1)$$

where A_v – availability, L_t – latency, T_h - throughput and R_l – reliability. This formulation outlines the primary variables influencing service quality without defining their exact quantitative contribution. In practice, a more detailed evaluation can be performed by assigning weights to each metric and by also considering additional parameters, such as the energy consumption of edge devices. Overall service quality may then be expressed through a weighted summation:

$$Q_s = w_1 A_v + w_2 \frac{1}{L_t} + w_3 T_h + w_4 R_l + w_5 \frac{1}{E_c} \quad (2)$$

where w_i denotes the relative importance of each parameter. In this formulation, increased latency or higher energy consumption tends to reduce QoS, whereas greater availability, throughput, and reliability improve it. This quantitative model allows a more precise assessment of system performance and supports comparative analysis between different monitoring approaches.

Additionally, the weights may be dynamically adjusted in consideration of operational status and application-specific requirements applications, enabling more flexible and context-aware QoS evaluation.

Cloud-IoT platforms confront several critical difficulties, notably the swift proliferation of connected devices, technological heterogeneity, massive data flows, and rigorous real-time constraints. For this reason, continuous QoS monitoring is indispensable for stable and protected functioning. Assessment strategies include agent-based, agentless, edge, cloud, and hybrid variants. Agent-based methods collect real-time data via software agents, while agentless techniques rely on APIs or network monitoring. Edge solutions are suitable for latency-sensitive tasks, whereas cloud-based approaches support centralized analysis and visualization, while hybrid configurations optimally combine the strengths of both [2, p. 3].

Current QoS systems additionally integrate security and privacy dimensions, covering data integrity, confidentiality, regulatory adherence, and safeguards against cyber attacks. These factors are measured through indicators such as incident frequency, authentication success rates, and encryption reliability.

Recent advancements highlight the role of AI in predicting failures, detecting anomalies, and enabling proactive resource management. In addition, self-adaptive systems improve resilience by automatically reconfiguring components and maintaining service continuity. Combining edge and cloud resources facilitates local processing of data at or near its source, reducing latency while still enabling centralized monitoring and control [3, p. 15]. This shift toward intelligent monitoring reflects the growing need for autonomous decision-making in complex distributed environments

Real-world implementations highlight the essential role of uniting service quality monitoring with security protocols. In smart cities, it enhances traffic regulation and public protection by promptly detecting irregularities and possible cyber threats. In healthcare, wearable IoT devices support continuous monitoring, where protecting sensitive information and ensuring its accuracy is crucial, while AI helps detect early irregularities.

Despite recent progress, cloud-IoT systems still face key challenges, including scalability problems stemming from the ongoing rise in device numbers, hardware and protocol diversity, and the necessity of managing high-volume data streams in real time. Further complications emerge from security and privacy matters, including internal threats, unidentified weaknesses, and regulatory obligations.

These challenges are further intensified by the decentralized nature of IoT environments and the lack of unified standards [4, p. 12].

Advancing QoS in Cloud-IoT systems requires prioritizing the establishment and adoption of standardized monitoring frameworks, such as ETSI MANO (Management and Orchestration), ISO/IEC 25010 quality models, and OpenTelemetry-based monitoring architectures, which enable unified metric collection and interoperability in heterogeneous Cloud-IoT environments. These frameworks can facilitate consistent QoS evaluation and improve integration between edge and cloud layers.

In addition, there is a growing need to explore AI-driven adaptive approaches in Cloud-IoT environments. ML approaches are able to recognize anomalies and uncovering patterns in data. Deep Learning methods, including recurrent neural networks (RNNs) and long short-term memory networks (LSTMs), are employed to model temporal dynamics, while Reinforcement Learning supports adaptive allocation of resources and autonomous optimization of distributed systems. These techniques greatly enhance scalability, enable predictive maintenance, and support autonomous decision-making under dynamic workloads. Furthermore, hybrid AI approaches that combine multiple techniques are gaining increasing attention. For instance, integrating deep learning models with reinforcement learning enables both accurate prediction and adaptive decision-making. Such combinations are particularly effective in complex Cloud-IoT environments, where system conditions change dynamically and require both short-term responsiveness and long-term optimization.

For example, LSTM models are especially suitable for capturing temporal dependencies in IoT data streams, as reinforcement learning supports flexible decision-making in systems experiencing constantly changing operational demands.

Another promising direction involves the integration of federated learning, which allows distributed IoT devices to collaboratively train models while preserving data privacy, addressing both security and regulatory challenges.

An additional important aspect is energy efficiency, especially for edge and IoT devices that operate with restricted resources. Balancing energy consumption with QoS requirements requires intelligent workload distribution and energy-aware monitoring strategies.

Moreover, employing real-time data analytics and stream processing frameworks plays a crucial role in modern QoS monitoring systems. Utilizing technologies like Apache Kafka and Apache Flink, systems can continuously ingest and analyze data, quickly recognizing deviations in performance or unusual behavior. This is particularly important in latency-sensitive Cloud-IoT applications, where even minor delays can significantly impact system reliability and user experience.

Furthermore, the integration of digital twin technology is emerging as a promising approach for QoS monitoring. Digital twins represent physical IoT components in a virtual environment,

allowing performance trends to be estimated and system operations to be refined. This approach allows for a more accurate evaluation of service quality and supports proactive system management.

Future work should explore the deployment of standardized QoS frameworks, such as ETSI MANO, ISO/IEC 25010, and OpenTelemetry, in real Cloud-IoT environments. AI techniques, including Machine Learning, Deep Learning, and Reinforcement Learning, should undergo experimental testing to evaluate their performance in detecting anomalies, forecasting trends, and managing resources adaptively. Effective QoS monitoring in cloud-IoT environments requires integrating performance, security, and intelligent analytics into a unified framework. Increasing complexity in distributed environments requires monitoring frameworks that are both adaptive and scalable, ensuring consistent performance even in dynamic scenarios. Simultaneously, the growth of edge computing and decentralized systems highlights the importance of developing efficient solutions capable of operating under limited resource conditions.

Список використаних джерел:

1. Younas, M. I., Iqbal, M. J., Aziz, A., & Sodhro, A. H. (2023). Toward QoS Monitoring in IoT Edge Devices Driven Healthcare – A Systematic Literature Review. Sensors, 23(21), 8885. https://doi.org/10.3390/s23218885

2. Hasan, M. K., Hasan, M. M., Al-Qirim, N., Abdullah, S. N. H. S., Islam, S., & Razzaque, M. A. (2026). Review on data privacy and security for IoT-based multifunctional layers of cyber-physical systems in smart grids. *Journal on Information Security*. https://doi.org/10.1186/s13635-026-00225-x

3. Ahmad, W., Rasool, A., Javed, A. R., Baker, T., & Jalil, Z. (2022). Cyber Security in IoT-Based Cloud Computing: A Comprehensive Survey. *Electronics, 11*(1), 16. https://doi.org/10.3390/electronics11010016

4. Chang, K.-C., Niu, H., Kim, B., et al. (2024). IoT Privacy Risks Revealed. *Entropy, 26*(7), 561. https://doi.org/10.3390/e26070561

Абібулаєв А.Р.
аспр. гр. КБа-21, каф.ЗІ, ІКТА, НУ «Львівська політехніка», Львів, Україна
Піскозуб Андріян Збігнєвич
к.т.н.,доц., каф.ЗІ, ІКТА, НУ «Львівська політехніка», Львів, Україна

ML-ОРІЄНТОВАНЕ ОЦІНЮВАННЯ РИЗИКУ НЕБЕЗПЕЧНИХ ПОДІЙ В API/HTTP-ТРАФІКУ ХМАРНИХ СЕРВІСІВ

У тезі розглянуто підхід до оцінювання ризику небезпечних подій в API/HTTP-трафіку хмарних сервісів на основі методів машинного навчання. Запропоновано поєднання двох числових каналів, а саме керованого каналу на структурованих ознаках події та каналу нетипової поведінки для нових або рідкісних сценаріїв. Оцінки моделей приводяться до спільної шкали ризику, а робочий поріг вибирається за FPR-бюджетом. Результати експериментів підтверджують високу якість підходу та його придатність до практичного використання у хмарному середовищі.

Ключові слова: машинне навчання; оцінювання ризику; захист API/HTTP; визначення аномалій; OWASP Top 10; ризикоорієнтовані сповіщення; хмарні мікросервіси.

Сучасні хмарні сервіси функціонують у середовищі високої динамічності, розподіленої мікросервісної архітектури та інтенсивної API-взаємодії. За таких умов класичні сигнатурні засоби виявлення загроз часто втрачають ефективність, оскільки слабко враховують контекст активу, зміни поведінкових профілів сервісів і потенційний вплив події на інфраструктуру. Для практичної експлуатації систем безпеки цього недостатньо, оскільки важливо не лише фіксувати підозрілу активність, а й оцінювати її ризик у конкретних умовах функціонування хмарного сервісу. У сучасних дослідженнях з безпеки хмарних середовищ методи машинного навчання розглядаються як засіб підвищення адаптивності виявлення, однак їх практична придатність безпосередньо залежить від якості даних, стабільності рішень, контролю хибно-позитивних спрацювань і зрозумілості результатів для операційних процесів [1-4].

Особливо актуальною ця задача є для API/HTTP-трафіку, через який реалізується значна частина атак на хмарні системи, зокрема ін’єкції, BruteForce, Credential Stuffing, зловживання бізнес-логікою та спроби несанкціонованого доступу. Одна й та сама подія може мати різний рівень небезпеки залежно від середовища виконання, типу ресурсу, характеру даних і зовнішньої доступності сервісу. Саме тому доцільно переходити від бінарного рішення «атака або не атака» до ризик-орієнтованого оцінювання подій.

У межах запропонованого підходу використано два взаємодоповнювальні ML-канали. Перший канал є керованим і навчається на розмічених даних для стабільного розділення нормальних подій та атак. Другий канал орієнтований на виявлення нетипової поведінки та формує додатковий сигнал відхилення від норми, що є корисним для нових або слабко представлених у навчальній вибірці сценаріїв. Така побудова відповідає загальній логіці реалізованої системи, у якій два ML-канали надалі поєднуються модулем об’єднання рішень (Рис.1).

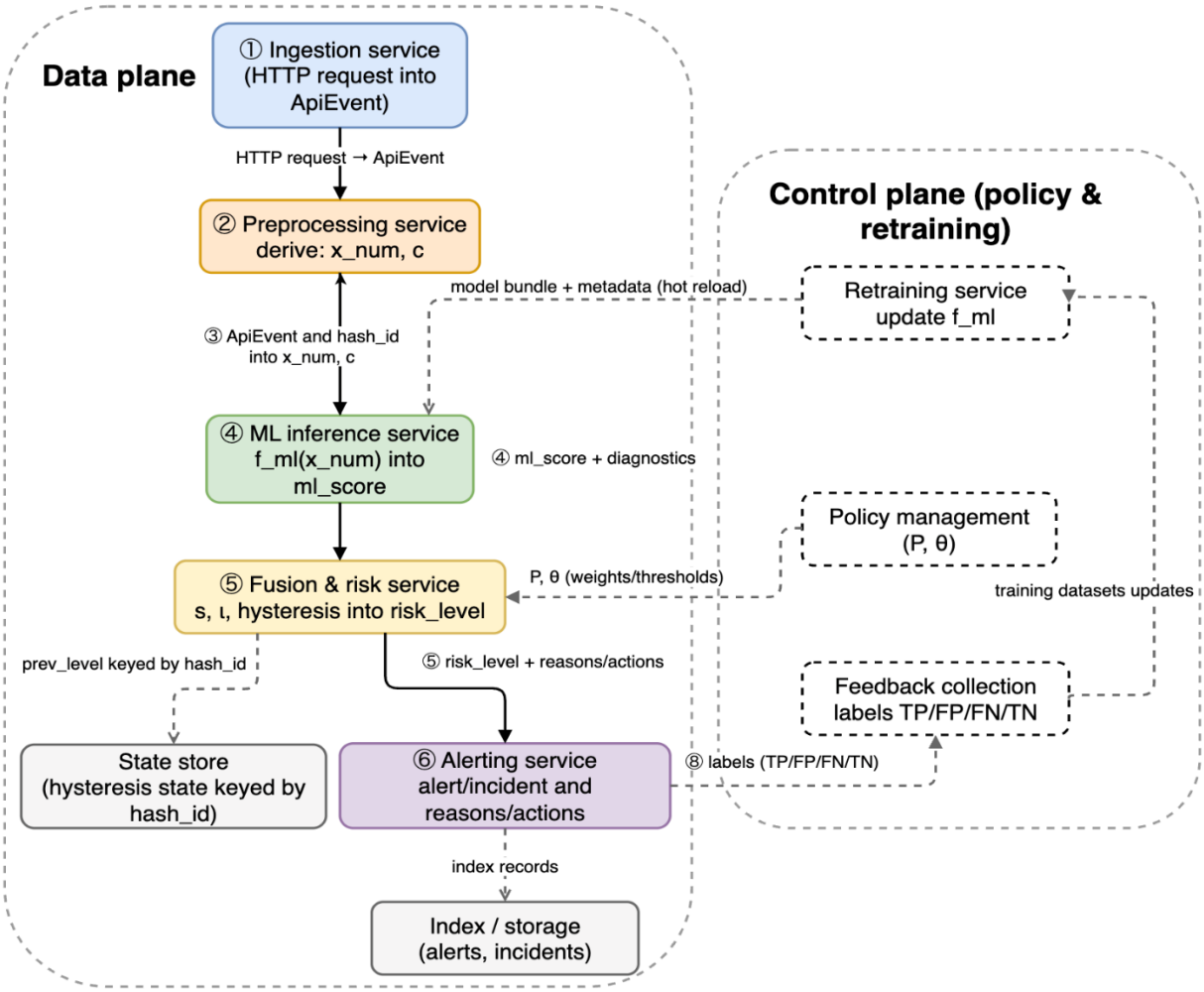

Рис. 1. Схема ML-конвеєра оцінювання ризику подій в API/HTTP-трафіку хмарних сервісів

Для кожної події API/HTTP-трафіку після попередньої обробки формується вектор структурованих числових ознак

$$x = (x_1, x_2, \ldots, x_d) \quad (1),$$

де x є поданням події у просторі ознак, а d є кількістю числових характеристик, що використовуються моделями.

Ключова методична ідея підходу полягає у приведенні різнорідних виходів керованих та допоміжних каналів до спільної шкали ризику. Це дає змогу порівнювати оцінки моделей різної природи й використовувати їх у межах єдиного системного рішення. Для практичного застосування у потоковому режимі фінальне спрацювання визначається за пороговою політикою з контролем рівня хибно-позитивних спрацювань

$$\hat{y}(x) = 1\left[p_{final}(x) \geq \tau_{fusion}\right], \qquad \tau_\alpha = Q_{1-\alpha}(\{s_i : y_i = 0\}), 0 < \alpha < 1 \quad (2),$$

де $p_{final}(x)$ є фінальною оцінкою ризику, (τ_α є порогом, вибраним за допустимим бюджетом FPR, а $Q_{1-\alpha}$ є емпіричним квантилем для нормальних подій.

Така постановка є важливою перевагою підходу, оскільки робить систему придатною не лише до точного виявлення, а й до керованої експлуатації в умовах реального потоку подій. У потоковому застосуванні стабільність рішень додатково підсилюється політиками керованості, зокрема, гістерезисом, який зменшує коливання біля робочого порогу.

Експериментальна перевірка підходу показала, що керований числовий канал є основним джерелом стабільного розділення нормальних подій та атак, тоді як канал нетипової поведінки виконує допоміжну роль і формує корисний сигнал новизни. У дослідженні якість

аналізувалася за ROC-AUC, PR-AUC, F1, точністю та повнотою, а вибір порога здійснювався з урахуванням FPR-бюджету. Для основного ML-ансамблю на тестовій вибірці отримано ROC-AUC = 0.9843, PR-AUC = 0.9511 та F1 = 0.8400 при FPR близько 0.051. Це свідчить про високу якість базового розділення нормальних подій і атак. Водночас допоміжний канал нетипової поведінки є практично корисним у ситуаціях, коли подія суттєво відхиляється від нормального режиму, але недостатньо добре описується наявними розміченими прикладами. Отже, цей канал не замінює керований, а доповнює його в межах спільної ризикової оцінки.

Таким чином, запропонований підхід доцільно розглядати як основу для побудови ризик-орієнтованого конвеєра моніторингу API/HTTP-трафіку в хмарному середовищі. Його перевагою є поєднання високої якості керованого розділення подій, здатності реагувати на нетипову поведінку та керованого контролю хибних сповіщень про небезпеку. Отримані результати підтверджують придатність підходу до інтеграції в інфраструктуру моніторингу та реагування, де важливими є стабільність рішень, відтворюваність політики порогів і можливість подальшого оновлення моделей.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Belal M. M., Sundaram D. M. Comprehensive Review on Intelligent Security Defences in Cloud: Taxonomy, Security Issues, ML/DL Techniques, Challenges and Future Trends. *Journal of King Saud University - Computer and Information Sciences*. 2022. Vol. 34.
2. Alzoubi Y. I., Mishra A., Topcu A. E. Research Trends in Deep Learning and Machine Learning for Cloud Computing Security. *Artificial Intelligence Review*. 2024. Vol. 57.
3. Mamidi S. R. The Role of AI and Machine Learning in Enhancing Cloud Security. *Journal of Artificial Intelligence and General Science*. 2024. Vol. 3.
4. Rakgoale D. M., Kobo H. I., Mapundu Z. Z., Khosa T. N. A Review of AI/ML Algorithms for Security Enhancement in Cloud Computing with Emphasis on Artificial Neural Networks. In: *Proceedings of the 2024 4th International Multidisciplinary Information Technology and Engineering Conference*. 2024.

Ахмедов А.Ф.
студент групи БСД-22, ННІКБЗІ ДУІКТ,
Київ, Україна

ДЕЦЕНТРАЛІЗОВАНА БІОМЕТРИЧНА ІДЕНТИФІКАЦІЯ ЯК МЕХАНІЗМ ПІДВИЩЕННЯ СТІЙКОСТІ IoT-ІНФРАСТРУКТУРИ

Розглянуто фундаментальні вразливості централізованих систем автентифікації в IoT-мережах. Запропоновано епістемологічно обґрунтовану гіпотезу використання децентралізованих ідентифікаторів (DID) на основі моделі identity-as-existence у поєднанні з логічним принципом Бритви Оккама. Обґрунтовано, що за умови реалізації біометричного протоколу з ефективною ентропією, еквівалентною рівню постквантової стійкості NIST Level 5, потреби в розподілених реєстрах та PQC-алгоритмах можуть розглядатися як умовно надлишкові.

Ключові слова: IoT, децентралізовані ідентифікатори (DID), самосуверенна ідентичність (SSI), постквантова криптографія, Бритва Оккама, біометрія, кібербезпека.

Сучасні стратегії розгортання IoT-мереж характеризуються системною вразливістю через домінування централізованих моделей ідентифікації (Active Directory, OAuth, RFID, а також хмарних сервісів автентифікації). Означена архітектура детермінує три ключові вектори ризику:

- наявність «єдиної точки відмови» (Single Point of Failure) призводить до втрати функціональності всієї периферійної IoT-інфраструктури у разі недоступності центрального сервера автентифікації через кібератаки або технічні аварії;
- відсутність інтероперабельності обмежує можливість використання уніфікованих ідентифікаційних даних у різнорідних IoT-екосистемах;
- концентрація персональних даних у централізованих сховищах створює "медові пастки" (honeypots) - високовартісні цілі, компрометація яких спричиняє незворотні наслідки для приватності користувача.

Технологічною альтернативою є стандартизована консорціумом W3C концепція децентралізованих ідентифікаторів (Decentralized Identifiers, DID) та самосуверенної ідентичності (Self-Sovereign Identity, SSI) [1, розд. 1]. Оскільки суб'єкт створює та контролює такий ідентифікатор самостійно без залучення централізованих реєстраторів, відповідний DID-документ містить лише публічні ключі та сервісні точки для безпечної автентифікації. Ключоваа перевага полягає у можливості локальної верифікації, коли IoT-пристрій самостійно перевіряє цифровий підпис користувача через бездротові інтерфейси за допомогою легковагових криптографічних протоколів. Такий підхід гарантує повну функціональну автономність інфраструктури навіть за умов відсутності зв'язку із зовнішніми мережами.

У даній роботі використано підхід передбачуваної гіпотези у розумінні Карла Поппера [2]. Замість твердження про безумовне існування конкретного технічного рішення висунуто припущення про можливість біометричного протоколу із наперед визначеними властивостями. До таких властивостей належать детермінована генерація стабільного коду та достатня ентропія разом із криптографічною незворотністю. Така методологічна рамка

забезпечує фальсифікованість гіпотези та її відкритість до емпіричної перевірки згідно з науковим критерієм демаркації.

Для досягнення постквантової стійкості застосовано логічний принцип Бритви Оккама («не слід вводити зайві сутності без потреби») [3, p. 12] як евристичний інструмент оптимізації архітектури, а не як доведений аргумент. Якщо біометричний протокол детермінованого перетворення здатен генерувати стабільний ідентифікаційний код з ентропією 512 біт, що забезпечує ефективну 256-бітну постквантову стійкість з урахуванням квадратичного прискорення за алгоритмом Гровера, то виникає підстава переглянути необхідність традиційних компонентів системи. Зокрема, розподілені реєстри (блокчейн) як сховища ідентифікаторів, криптографічні пари ключів та PQC-алгоритми можуть розглядатися як умовно надлишкові за умови, що біометричний код забезпечує еквівалентний рівень захисту. Згідно з вимогами Національного бюро стандартів США (National Institute of Standards and Technology, NIST), зазначений рівень ентропії відповідає найвищому рівню постквантової стійкості Level 5 [4, sec. 3]. Такий підхід концептуально наближається до моделі «особистісне шифрування» (Identity-Based Encryption, IBE), де біометричний ідентифікатор безпосередньо виконує функцію криптографічного ключа, реалізуючи пряму схему «Біометрія → Доступ» замість багаторівневої «Біометрія → Ключ → Реєстр → Доступ».

Для досягнення філософської глибини аналізу необхідно експлікувати три парадигми цифрової ідентифікації:

identity-as-record (ідентичність як запис) - класична модель DID, де легітимність суб’єкта визначається наявністю відповідного запису в розподіленому реєстрі;

identity-as-attribute (ідентичність як набір ознак) - традиційна біометрія, що базується на зовнішньому порівнянні пред’явлених даних із еталоном;

identity-as-existence (ідентичність як факт існування) - пропонована модель, у якій ідентифікатор є похідним від біометричної сутності суб’єкта, а «джерелом істини» (Source of Truth) виступає сама людина.

Реалізація моделі identity-as-existence передбачає три ключові умови. До них належать висока відтворюваність через детерміновану генерацію коду без зберігання еталонів та ентропійна масштабованість згідно з вимогами NIST разом із криптографічною незворотністю перетворення. У межах даного дослідження ці умови розглядаються як гіпотетичні та не є емпірично доведеними. Остання властивість вважається критичною за умови відсутності відомих ефективних атак на простір біометричних представлень. Потенційна технічна реалізація може базуватися на хеш-функціях або нечітких екстракторах, проте детальний аналіз конкретних алгоритмів перетворення не входить до основних завдань цієї праці.

Таким чином, запропонований підхід дозволяє інтерпретувати біометричну ідентифікацію не лише як суто технічний механізм, а як концептуальний перехід до парадигми «доказу замість довіри» (proof-based trust), де джерелом легітимності стає безпосереднє фізичне існування суб'єкта. IoT-інфраструктура трансформується з централізованої та вразливої системи на автономну, безпечну та інтероперабельну екосистему, звільняючись від ризиків зберігання чутливих даних та тягаря адміністрування централізованих реєстрів. Подальші дослідження мають бути спрямовані на емпіричну верифікацію висунутої гіпотези, формалізацію вимог до біометричних протоколів та технічну розробку прототипу біометричного DID-протоколу.

Перелік посилань:

1. W3C, "Decentralized Identifiers (DIDs) v1.0", World Wide Web Consortium, 2022.

URL: https://www.w3.org/TR/did-core/

2. Popper K. The Logic of Scientific Discovery. London: Routledge, 1959.

3. Stanford Encyclopedia of Philosophy, "William of Ockham", 2022.

URL: https://plato.stanford.edu/entries/ockham/

4. NIST, "Post-Quantum Cryptography Standardization", National Institute of Standards and Technology, 2022. URL: https://csrc.nist.gov/Projects/post-quantum-cryptography

Барилюк В.В.
студет групи БСДМ-61, ННІКБЗІ, ДУІКТ, Київ, Україна

АРХІТЕКТУРА ZERO TRUST ДЛЯ СУЧАСНИХ ОРГАНІЗАЦІЙ

У роботі досліджується перехід від традиційної моделі захисту периметра до архітектури Zero Trust в умовах використання хмарних середовищ та віддаленої роботи. Обґрунтовується необхідність відмови від довіри за замовчуванням. Розглядаються ключові шляхи захисту: багатофакторна автентифікація, мікросегментація мережі, дотримання принципу найменших привілеїв та безперервний моніторинг подій. Доведено, що впровадження Zero Trust мінімізує площу атаки і сприяє цифровій трансформації бізнесу.

Ключові слова: Zero Trust, кібербезпека, захист периметра, управління доступом, багатофакторна автентифікація, принцип найменших привілеїв, мікросегментація мережі, безперервний моніторинг.

Протягом тривалого часу основою корпоративної кібербезпеки багатьох організацій була модель захисту периметра. Це означає що усі користувачі, пристрої та додатки, що знаходяться всередині корпоративної мережі, вважалися довіреними, тоді як усе, що знаходилося зовні вважається потенційною загрозою. Такий захист будувався з використанням брандмауерів, системи виявлення вторгнень та VPN, які повинні захистити цей зовнішній периметр.

Однак у сучасних умовах, коли відбулася трансформація робочих процесів в організаціях і як наслідок, відбулось розмиття або повне зникнення традиційного корпоративного периметра [1-2]. Можна виділити декілька ключових факторів, які вплинули на цей процес:

1. Перехід на хмарні середовища. Дані та додатки більше не зберігаються виключно на локальних серверах організації, а розподілені між публічними, приватними та гібридними хмарами (SaaS, PaaS, IaaS).

2. Віддалена та гібридна робота. Співробітники підключаються до корпоративних ресурсів з різних точок доступу, які не контролюються IT-відділом компанії.

3. *Концепція BYOD.* Багато організацій дозволило використання особистих смартфонів та ноутбуків для доступу до корпоративних даних, що створює додаткові ризики та зони, які не контролюються службою безпеки.

2. Шляхи вирішення: Впровадження моделі Zero Trust

Шляхом вирішення даної проблеми для побудови сучасної архітектури організації є концепція безпеки *Zero Trust*. Гаслом цієї концепції є наступний вираз: *«Ніколи не довіряй, завжди перевіряй»* [3]. Основою концепції є те, що довіра не повинна надаватися за замовчуванням нікому і нічому, незалежно від того, де знаходиться користувач чи пристрій — всередині корпоративної мережі чи за її межами.

Для вирішення проблеми вразливості традиційних мереж організації повинні впровадити архітектуру Zero Trust (ZTA), яка базується на кількох фундаментальних принципах та технологічних рішеннях (рис.1).

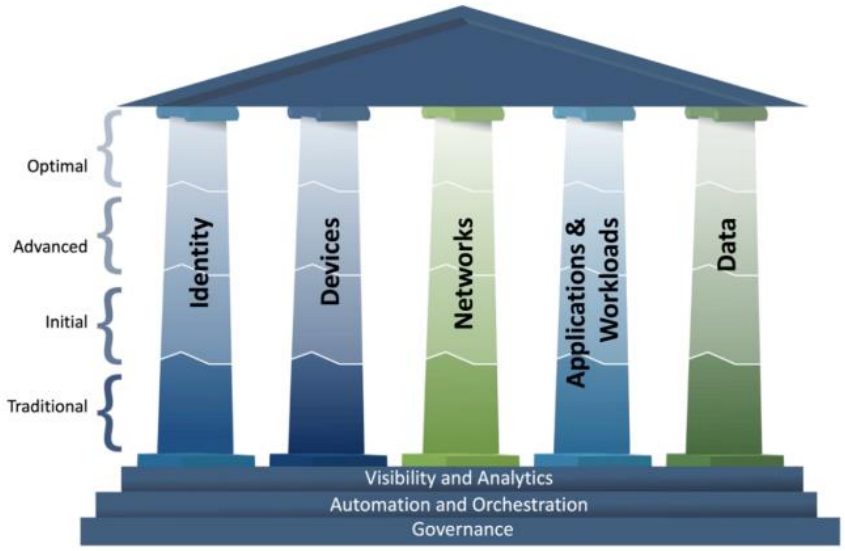

Рис.1. Модель Zero Trust [1]

Основним принципом даної моделі є сувора ідентифікація та управління доступом. Це надає є розуміння того, *хто* і *з чим* намагається з'єднатися і одних паролів тут вже недостатньо. Організації мають впроваджувати багаторівневу перевірку особи за допомогою додаткових механізмів:

Багатофакторна автентифікація;

Єдиний вхід (Single Sign-On, SSO);

Оцінка контексту.

Багатофакторна автентифікація є обов'язковою умовою для доступу до будь-якого ресурсу системи організації.

SSO дозволяє централізовано управляти сесіями доступу користувачів системи організації.

Система повинна оцінювати контекст, тобто аналізувати не лише логін і пароль, але й час доби, геолокацію, тип пристрою та його стан безпеки (чи оновлений антивірус, чи встановлені патчі ОС).

Крім цього велику увагу повинно приділятися принципу найменших привілеїв. Це означає, що кожному користувачеві, програмі чи системному процесу повинен надаватися лише той мінімальний рівень доступу, який суворо необхідний для виконання їхніх поточних робочих завдань, і лише на необхідний проміжок часу. Такий підхід гарантує, що у разі компрометації одного облікового запису, зловмисник не зможе отримати доступ до критичних даних або системи управління організації [4].

Для усунення горизонтального поширення кібератак, організації повинні впроваджувати мікросегментацію корпоративної мережі, тобто архітектура мережі організації повинна складатися з логічно ізольованих зон від рівня окремих додатків до віртуальних машин. Це забезпечує входження в систему заново за рахунок повторної автентифікації та авторизації в систему [1-2].

З урахуванням того, що автентифікація в моделі Zero Trust не є одноразовою подією, під час входу в систему організації, а це є безперервним процесом, організаціям необхідно впроваджувати у свою систему рішення класу SIEM та SOAR для безперервного моніторингу усіх систем у режимі реального часу. Аналіз подій в системі, а саме поведінка користувачів, яка стає підозрілою, повинна автоматично знижувати рівень довіри і вимагати додаткової перевірки або блокувати доступ до системи [1-2].

Кожен пристрій, який підключається до корпоративної мережі (смартфон, ноутбук, сервер), в концепції Zero Trust розглядається як потенційний вектор атаки. Для своєчасного виявлення шкідливої активності необхідно використовувати рішення, які можуть виявляти таку активність на пристроях і ізолювати їх до того, як вони зможуть нашкодити іншим

компонентам корпоративної мережі. До класу таких рішень відносяться системи EDR або XDR [4].

Отже, архітектура Zero Trust це не просто черговий набір рішень з кібербезпеки, а стратегічний зсув у підходах трансформації корпоративних IT-процесів організацій до сучасних методів та заходів кібербезпеки. Впровадження архітектури Zero Trust дозволить організаціям вирішити основну проблему сучасних кіберзагроз, а саме мінімізувати площу атаки та унеможливити вільне переміщення зловмисників всередині мережі.

У кінцевому підсумку, перехід до Zero Trust не лише знижує ризики катастрофічних витоків даних, але й сприяє цифровій трансформації бізнесу. Це дозволяє співробітникам безпечно і продуктивно працювати з будь-якої точки світу та з будь-якого пристрою, а бізнесу — зберігати довіру клієнтів та партнерів у висококонкурентному цифровому світі.

Перелік посилань:

1. Zero Trust Architecture : NIST Special Publication 800-207. National Institute of Standards and Technology (NIST); U.S. Department of Commerce, 2020. URL: https://csrc.nist.gov/publications/detail/sp/800-207/final (дата звернення: 14.04.2026).
2. Zero Trust Maturity Model. Cybersecurity and Infrastructure Security Agency (CISA), 2023. URL: https://www.cisa.gov/zero-trust-maturity-model (дата звернення: 14.04.2026).
3. Kindervag J. Build Security Into Your Network's DNA: The Zero Trust Network Architecture. Forrester Research, 2010. URL: https://www.forrester.com/report/build-security-into-your-networks-dna-the-zero-trust-network-architecture/RES56682 (дата звернення: 14.04.2026).
4. Market Guide for Zero Trust Network Access / Gartner Research. Gartner, 2025. URL: https://www.gartner.com/en/documents/4011400 (дата звернення: 14.04.2026).

Бовкун І.В.
студент групи ПД-43, ННІІТ ДУІКТ, Київ,
Україна

Методи захисту інформації у системах безконтактного ресторанного сервісу

Дана робота присвячена дослідженню та розробці системи захисту інформації у WEB-застосунку для безконтактного оформлення замовлень у ресторанах. Проаналізовано побудову системи на основі мікросервісної архітектури з використанням Java Spring, PostgreSQL та сервісу Keycloak. Висвітлено ключові механізми безпеки: ізоляцію компонентів у Docker-контейнерах, централізовану автентифікацію через протоколи OAuth2/OpenID Connect та управління доступом на основі ролей (RBAC). Описано заходи захисту від SQL-ін'єкцій, CSRF-атак, а також валідацію JWT-токенів, вхідних даних і використання HTTPS. Запропонований комплексний підхід забезпечує надійний захист фінансових транзакцій і персональних даних.

Ключові слова: захист інформації, web-застосунок, безконтактні замовлення, мікросервісна архітектура, Java Spring, Keycloak, Docker-контейнеризація, автентифікація, авторизація, RBAC, JWT-токени, HTTPS.

У сучасному ресторанному бізнесі впровадження систем безконтактної оплати є стратегічним кроком для підвищення якості обслуговування та швидкості розрахунків. Проте обробка фінансових транзакцій та персональних даних клієнтів у вебсередовищі створює значні ризики, що потребують комплексного підходу до інформаційної безпеки [1].

Дане дослідження присвячене розробці системи безконтактної оплати, побудованої на мікросервісній архітектурі з використанням фреймворку Java Spring, СУБД PostgreSQL та сервісу ідентифікації Keycloak. Оскільки застосунок оперує конфіденційними даними (історія замовлень, платіжні токени, профілі користувачів), захист інформації на кожному рівні взаємодії є критично важливим.

Архітектура системи передбачає розгортання сервісів застосунку в ізольованих Docker-контейнерах, що дозволяє відокремити сервер додатків, базу даних та сервіс авторизації, мінімізуючи ризик «ефекту доміно» при компрометації одного з вузлів.

Для забезпечення належного рівня безпеки в проєкті реалізовано наступні механізми:

1. Централізована автентифікація та авторизація.

Використання Keycloak на основі протоколів OAuth2 та OpenID Connect дозволяє винести логіку керування користувачами за межі основного коду. Це забезпечує надійне збереження паролів (хешування) та підтримку багатофакторної автентифікації [2].

2. Управління доступом на основі ролей (RBAC).

За допомогою Spring Security реалізовано чіткий розподіл прав доступу [3]. Наприклад, роль CUSTOMER має доступ лише до своїх чеків, тоді як CHEF або ADMINISTRATOR можуть переглядати статус замовлень та керувати меню. Приклад реалізації наведено на рис.1.

```
public class OrderController {

    private final OrderService orderService;

    @PostMapping
    @PreAuthorize("hasRole('ROLE_CUSTOMER')")
    public ResponseEntity<OrderResponseDto> createOrder(@RequestBody OrderRequestDto dto) {
        return ResponseEntity.ok(orderService.createOrder(KeycloakUtil.getCurrentUser().getId(), dto));
    }
```

Рис 1. Приклад реалізації доступу на основі ролей (RBAC)

3. Безпека даних у БД.

Використання Spring Data JPA з параметризованими запитами автоматично нівелює ризики SQL-ін’єкцій. Для збереження чутливих налаштувань (ключі API платіжних систем, паролі БД) використовуються змінні оточення (Environment Variables) всередині Docker, що виключає потрапляння секретів у вихідний код.

4. Захист мережевого рівня.

Взаємодія між клієнтом та сервером здійснюється виключно через протокол HTTPS. На рівні Spring Boot налаштовано фільтри для захисту від Cross-Site Request Forgery (CSRF) та Cross-Origin Resource Sharing (CORS) політик.

5. Валідація JWT-токенів.

Сервер Java Spring перевіряє підпис кожного токена, виданого Keycloak, що гарантує цілісність даних користувача та неможливість підробки сесії.

6. Валідація вхідних даних.

Застосування Jakarta Validation та Hibernate Validator дозволяє забезпечити цілісність даних на рівні DTO та сутностей. Використання анотацій, таких як @NotNull, @Size та @Pattern, гарантує, що до обробки бізнес-логікою потрапляють лише дані, які відповідають встановленим форматам. Це створює додатковий рівень захисту від некоректного введення та атак, спрямованих на маніпуляцію структурою запитів.

Впровадження даного комплексу заходів дозволяє створити стійку до атак систему, що відповідає сучасним вимогам безпеки електронних платежів. Подальші дослідження можуть бути спрямовані на впровадження систем моніторингу підозрілої активності та використання інструментів автоматичного сканування уразливостей у контейнерах.

Перелік посилань:

1. Бондаренко О., Ушкаленко І. (2017). Безпека web-додатків: актуальні проблеми та їх аналіз. *Формування ринкової економіки в Україні*. Вип. 38., 28-36.

2. Keycloak: Server Administration Guide [Електронний ресурс]. — Режим доступу: URL https://www.keycloak.org/docs/latest/server_admin/

3. Spring Security Reference [Електронний ресурс]. — Режим доступу: URL https://docs.spring.io/spring-security/reference/index.html

Будзинський Олександр Володимирович
аспірант кафедри УКБЗІ, ННІЗІ ДУІКТ, Київ, Україна

ГРАФОВА МОДЕЛЬ ОЦІНЮВАННЯ РИЗИКІВ ДОСТУПУ ДО КОРПОРАТИВНИХ БАЗ ДАНИХ

У роботі запропоновано графову модель оцінювання ризиків доступу до корпоративних баз даних, яка враховує структурні особливості мережі та можливі шляхи поширення атак. Мережа подається у вигляді орієнтованого графа, де вузли характеризуються рівнем аномальності, визначеним за допомогою моделей машинного навчання (Isolation Forest, LSTM, автоенкодер). Запропоновано підхід до обчислення ймовірності переходу між вузлами та інтегрального ризику шляху до бази даних із використанням експоненційної функції критичності. Розроблений метод дозволяє ідентифікувати найбільш небезпечні маршрути доступу та забезпечує основу для динамічної сегментації мережі й автоматичного застосування політик безпеки в системах класу SOC.

Ключові слова: захист корпоративних баз даних, модель оцінювання ризиків, моделі машинного навчання, графові моделі доступу.

Сучасні корпоративні інформаційні системи характеризуються високим рівнем складності мережевої інфраструктури, значною кількістю взаємопов’язаних вузлів та багаторівневими механізмами доступу до критичних ресурсів, зокрема баз даних. У таких умовах традиційні підходи до забезпечення безпеки, що базуються на аналізі окремих подій або пакетів, не дозволяють адекватно оцінити ризик поширення атак у мережі. Особливої актуальності набуває підхід, у якому доступ до бази даних розглядається як проходження потенційної атаки через множину вузлів мережі, що формують різні шляхи доступу.

Одним із перспективних напрямів є використання графових моделей для представлення процесів доступу до баз даних, що дозволяє враховувати не лише окремі події, а й цілісні шляхи взаємодії в інформаційній системі [1-3]. Такий підхід створює передумови для переходу від подієво-орієнтованого до шляхово-орієнтованого аналізу ризиків, що є більш адекватним сучасним сценаріям атак.

В узагальненій моделі інфраструктури (рис. 1) у вигляді орієнтованого графа G=(V,E) позначені: V – множина вузлів (логічні елементи інфраструктури: хости, сервери, сервіси); E – множина дозволених з'єднань між ними.

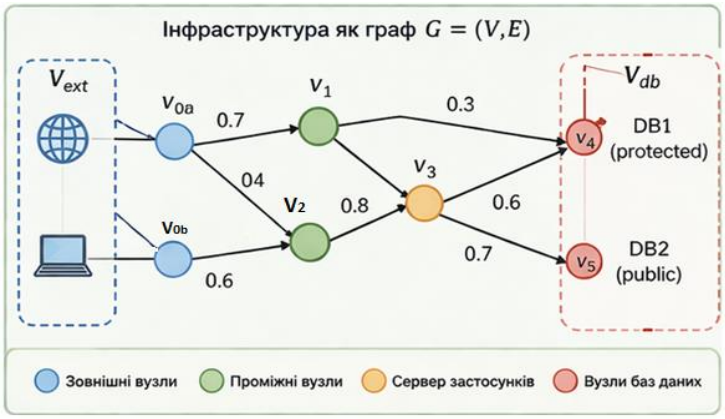

Рис. 1. Принципова графова інфраструктура

Залежно від рівня абстракції, вузли можуть відповідати фізичним об'єктам (робоча станція користувача, сервер застосунку, сервер бази даних), логічним компонентам (сегмент мережі (VLAN), сервіс або мікросервіс, рівень доступу). Набір шляхів від зовнішніх вузлів V_{ext} до вузлів баз даних V_{db} задає поверхню атаки. Мінімізація поверхні атаки – це мінімізація набору ребер/зв'язків, через які можливий доступ. Ребра графа характеризуються ймовірностями переходів P(π,t), що можуть змінюватися в часі та відображають динаміку доступу. Для кожного вузла визначається рівень аномальності як сума застосованих моделей машинного навчання ($s_{IF}(v,t)$ – Isolation Forest, $s_{LSTM}(v,t)$ – LSTM, $s_{AE}(v,t)$ – автоенкодер):

$$q_V(t) = \sigma(\alpha s_{IF}(v,t) + \beta s_{LSTM}(v,t) + \gamma s_{AE}(v,t)), \quad (1)$$

де:s_{IF}, s_{LSTM}, s_{AE}– оцінки моделей машинного навчання; $\sigma(\cdot)$– сигмоїдна функція нормалізації; α,β,γ – вагові коефіцієнти моделей, визначаються на основі метрики якості (наприклад, F1-score), α+β+γ=1, $q_v(t) \in [0,1]$.

Для переходу до рівня міжвузлових взаємодій використовується агрегуюча функція, яка формує оцінку для ребра графа на основі скорів суміжних вузлів і забезпечує узгодженість між вузловим та сегментним представленням інфраструктури:

$$p_{uv}(t) = \sqrt{q_u(t) \cdot q_v(t)} \quad (2)$$

Ймовірність успішного проходження атаки по всьому шляху π (рис. 2) в момент часу t визначається як добуток:

$$P(\pi,t) = \prod_{(u,v)\in\pi} p_{uv}(t) \quad (3)$$

Рис. 2. Визначення ймовірності поширення атаки на кожному шляху

Ризик компрометації шляху визначається як добуток імовірності проходження атаки уздовж шляху та сумарної критичності ресурсів, що входять до цього шляху, при цьому аномальність вузлів враховується опосередковано через імовірності переходів між вузлами, що дозволяє уникнути подвійного врахування та забезпечує коректність. Для врахування нелінійного ефекту розвитку атаки запропоновано модифіковану експоненційну модель (рис. 3), у якій ймовірність компрометації шляху входить до показника експоненти:

$$R(\pi,t) = I(DB)\cdot\exp(\lambda\cdot P(\pi,t)\cdot\sum_{v\in\pi} I(v)) \quad (4)$$

де: $I(v)$– критичність вузла; I(DB)– критичність бази даних; λ – **коефіцієнт ваги (масштабування)**, пояснення)

Коефіцієнт λ визначає ступінь впливу критичності проміжних вузлів на загальний ризик шляху. При λ=0 модель зводиться до оцінювання ризику лише на основі цільового ресурсу, тоді як при збільшенні λ зростає значущість внутрішньої структури мережі та ролі окремих вузлів у поширенні атаки.

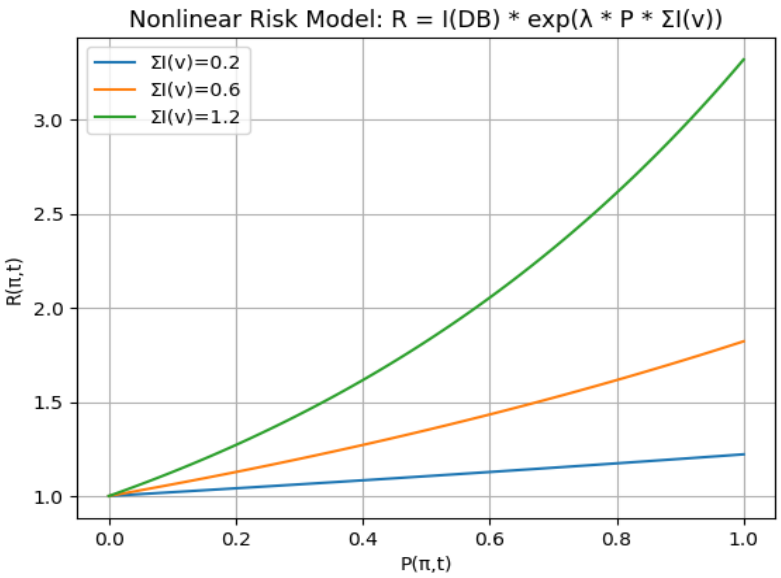

Рис. 3. Графіки залежності ризику шляху від критичності вузлів

Це дозволяє моделювати прискорене зростання ризику при збільшенні ймовірності атаки, що відповідає реальним сценаріям ескалації привілеїв та поширення атак у мережі, враховує чутливість до критичних вузлів та забезпечує ефективність виявлення небезпечних шляхів. Запропонована модель враховує як ймовірність атаки, так і критичність вузлів, дозволяє оцінювати ризик не окремих подій, а цілих шляхів, при цьому експоненційна залежність відображає ефект накопичення вразливостей.

Графова модель може бути використана у системах SOC для автоматичного управління політиками безпеки.

Перелік посилань:

1. Pujol-Perich, D., Suarez-Varela, J., Ferriol, M., Xiao, S., Wu, B., Cabellos-Aparicio, A., & Barlet-Ros, P. (2021). IGNNITION: Bridging the Gap between Graph Neural Networks and Networking Systems. IEEE Network, 35(6), 171–177. https://doi.org/10.1109/mnet.001.2100266
2. Zhong, M., Lin, M., Zhang, C., & Xu, Z. (2024). A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges. Computers & Security, 141, 103821. https://doi.org/10.1016/j.cose.2024.103821
3. Caville, E., Lo, W. W., Layeghy, S., & Portmann, M. (2022). Anomal-E: A self-supervised network intrusion detection system based on graph neural networks. Knowledge-Based Systems, 110030. https://doi.org/10.1016/j.knosys.2022.110030

Буханець Ілля Сергійович
студент групи БСДМ-51, ННІЗІ ДУІКТ,
Київ, Україна

ЗАХИСТ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

Штучний інтелект у розробці ПЗ слід розглядати як фундаментальний інструмент та новий вектор вразливостей, що потребує інтеграції безпеки безпосередньо в процес створення програмних продуктів (DevSecOps). Запропоновано практичний підхід до впровадження ШІ в розробку, що поєднує безпечне керування ланцюгом постачання AI-моделей (AI Supply Chain), валідацію коду, згенерованого ШІ, на наявність логічних вразливостей, захист середовищ розробки від prompt injection та data poisoning, а також забезпечення прозорості та відтворюваності побудови моделей. Важливим аспектом є впровадження автоматизованих інструментів контролю якості та безпеки (AI-enhanced security testing) на кожному етапі життєвого циклу розробки ПЗ. Очікуваний ефект — прискорення темпів розробки без шкоди для безпеки, мінімізація ризиків впровадження вразливого коду та забезпечення високого рівня надійності програмних продуктів, створених за допомогою інтелектуальних систем. **Ключові слова**: штучний інтелект, розробка ПЗ, DevSecOps, безпека коду, життєвий цикл розробки, AI-assisted development, безпека моделей.

Штучний інтелект стає невід’ємною складовою сучасного процесу розробки ПЗ (SDLC), трансформуючи підходи до написання коду, тестування та архітектурного проектування. Однак інтеграція ШІ в корпоративні продукти створює нову площину ризиків, де поряд із класичними вразливостями ПЗ виникають специфічні загрози, пов’язані з використанням моделей, даних та сторонніх інтеграцій. У таких умовах безпека розробки (DevSecOps) повинна охоплювати весь життєвий цикл ШІ-компонентів: від навчання моделей до їхньої експлуатації в продуктовому середовищі. Першим викликом для розробників є безпека даних та навчання. Якщо модель навчається на неперевірених або спотворених даних, це ставить під загрозу надійність усього програмного продукту. На відміну від традиційного коду, помилки в ШІ-моделях можуть бути неявними, що призводить до некоректної логіки прийняття рішень. Тому процес розробки має передбачати суворий контроль цілісності даних, автоматизований аудит навчальних вибірок та валідацію поведінки моделі на кожному етапі ітеративного циклу розробки. Другим критичним аспектом є захист від ін'єкцій у запити (prompt injection) на рівні коду застосунку. При розробці систем, що інтегрують LLM, розробники часто стикаються з проблемою непередбачуваної поведінки моделі при взаємодії з зовнішніми інструментами та API. Захист таких систем вимагає впровадження багаторівневої архітектури: від фільтрації вхідних даних та ізоляції середовищ виконання (sandbox) до жорсткого обмеження контексту, в якому працює ШІ. Не менш важливим напрямом є управління безпекою ланцюга постачання (AI Supply Chain). Сучасні проєкти з використанням ШІ часто базуються на сторонніх бібліотеках, готових архітектурах моделей та хмарних сервісах. Скомпрометована залежність або вразлива модель можуть стати «бекдором» для всього застосунку. Процес розробки має включати сканування залежностей, контроль версій моделей, мінімізацію привілеїв сервісних акаунтів та регулярне оновлення компонентів. Практичний підхід до розробки безпечних ШІ-рішень доцільно будувати за п’ятьма напрямами: Безпека даних (Data SecOps): автентифікація джерел, контроль цілісності та простежуваність даних протягом усього циклу розробки. Захист моделей: забезпечення цілісності вагових коефіцієнтів, захист інтелектуальної власності та конфігурацій середовища виконання. Безпечне розгортання: сегментація інфраструктури, використання контейнеризації та валідація вихідних даних на рівні API-шлюзів. Моніторинг продуктової поведінки: впровадження систем журналювання та виявлення аномалій, що дозволяють оперативно

реагувати на девіації в роботі ШІ-компонентів. AI Governance у розробці: інтеграція політик безпеки, аудит на відповідність стандартам якості та регулярна оцінка ризиків, пов'язаних із впровадженням ШІ-інструментів у програмний код.

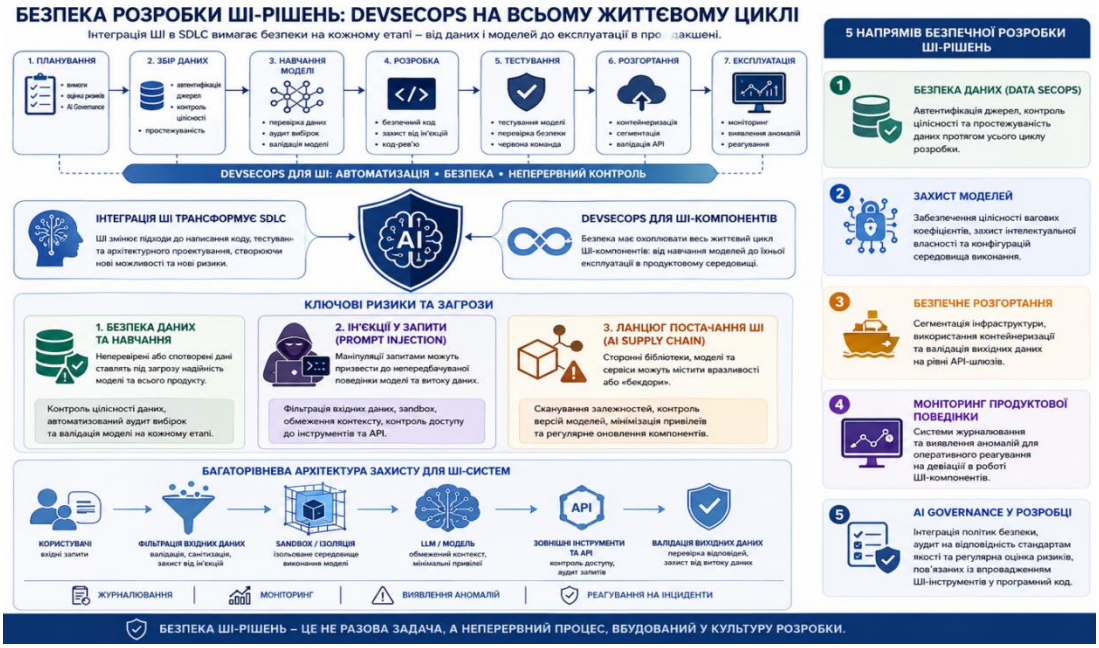

Рис. 1. Безпека розробки ШІ-РІШЕНЬ

Зазначені напрями захисту мають застосовуватися в розробці ПЗ як цілісна екосистема, де безпека даних, стійкість моделей, захист інфраструктури розгортання та політики governance взаємопідсилюють один одного. Наприклад, навіть найсучасніша архітектура моделі не гарантує безпеки без контролю якості навчальних даних, а автоматизований моніторинг стає неефективним без чітко визначених процедур реагування на інциденти в межах CI/CD-процесів. Саме комплексне поєднання технічних засобів (Security-as-Code) та організаційних правил дозволяє мінімізувати ризики компрометації програмного продукту та підвищити його стійкість до специфічних атак на ШІ. Особливе значення для розробника має забезпечення надійності після виходу продукту в експлуатацію. Навіть ретельно протестована модель може демонструвати непередбачувані реакції у реальних умовах, стаючи об'єктом маніпуляцій. Тому життєвий цикл розробки ШІ-рішень обов'язково має включати етапи threat modeling (моделювання загроз) та AI red teaming, які дозволяють ще на стадії проєктування виявити потенційні вектори атак на логіку моделі. Не менш важливим є впровадження принципів Zero Trust при інтеграції ШІ в програмні екосистеми. У сучасних застосунках модель часто взаємодіє з користувацькими інтерфейсами та зовнішніми API, що створює широку поверхню для несанкціонованого доступу чи витоку даних. Розробники повинні впроваджувати сувору автентифікацію, фільтрацію вхідних промптів, обмеження швидкості запитів (rate limiting) та контроль вихідної інформації, не довіряючи жодному запиту за замовчуванням. Таким чином, інтеграція ШІ в розробку ПЗ вимагає переходу до моделі Secure AI Development. Ефективний підхід охоплює всі етапи життєвого циклу ПЗ: від безпечного вибору джерел даних та навчання моделей до їх безпечного розгортання, постійного моніторингу та управління ризиками. Поєднання DevSecOps-практик, регулярного аудиту безпеки та AI governance дозволяє не лише мінімізувати площу атаки, а й забезпечити стабільну, передбачувану та етичну роботу інтелектуальних систем. З огляду на стрімке поширення ШІ-рішень, питання їх захисту набуває стратегічного значення для компаній-розробників. Надійне програмне забезпечення, що базується на ШІ, повинне проєктуватися за принципом «Security by Design» на кожному етапі — від створення коду до оновлення моделей. Лише системний підхід, що об'єднує технічні засоби захисту, постійний контроль та

чіткі організаційні правила, гарантує успішне та безпечне впровадження штучного інтелекту в сучасні цифрові продукти.

Перелік посилань:

1 Artificial Intelligence Risk Management Framework (AI RMF 1.0) [Електронний ресурс] // National Institute of Standards and Technology (NIST). − Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (дата звернення: 12.04.2026).
2 Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile [Електронний ресурс] // National Institute of Standards and Technology (NIST). − Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf (дата звернення: 14.04.2026).
3 OWASP Top 10 for Large Language Model Applications 2025 [Електронний ресурс] // OWASP Foundation. − Режим доступу: https://genai.owasp.org/llm-top-10/ (дата звернення: 13.04.2026).
4 Guidelines for Secure AI System Development [Електронний ресурс] // National Cyber Security Centre. − Режим доступу: https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development (дата звернення: 15.04.2026).).

Вдовенко К. В.
курсант, ЖВІ,
Житомир, Україна

Жуков А. О.
старший викладач, ЖВІ,
Житомир, Україна

АНАЛІЗ ТА МЕТОДИ ЗАХИСТУ АВТОМАТИЗОВАНИХ РОБОЧИХ МІСЦЬ ВІД ВИТОКУ ДАНИХ ЧЕРЕЗ ЗНІМНІ НОСІЇ ІНФОРМАЦІЇ

У роботі розглянуто основні загрози витоку інформації через знімні носії даних та визначено їх вплив на безпеку автоматизованих робочих місць. Проаналізовано сучасні організаційні, програмні та технічні методи захисту, а також принципи контролю використання USB-пристроїв. Особливу увагу приділено процесам моніторингу підключень, управлінню доступом до зовнішніх носіїв та запобіганню несанкціонованому копіюванню даних. Розглянуто інтеграцію засобів контролю периферійних пристроїв із системами запобігання витоку даних (DLP) та централізованими системами управління інформаційною безпекою. Визначено ключові переваги комплексного підходу до захисту, основні виклики впровадження та перспективи розвитку таких систем з урахуванням автоматизації та інтелектуального аналізу подій безпеки.

Ключові слова: інформаційна безпека, автоматизовані робочі місця, знімні носії, USB-пристрої, витік даних, DLP, контроль доступу, моніторинг подій, кіберзахист.

Автоматизовані робочі місця (АРМ) є невід'ємною складовою сучасної інформаційної інфраструктури організацій. Саме на них здійснюється обробка, зберігання та передавання значних обсягів службової, комерційної та конфіденційної інформації. У зв'язку з цим забезпечення належного рівня інформаційної безпеки АРМ є одним із ключових завдань кіберзахисту.

Одним із найбільш поширених каналів витоку даних залишаються знімні носії інформації, зокрема USB-накопичувачі. Їх доступність, портативність та простота використання створюють значні ризики несанкціонованого копіювання інформації, а також занесення шкідливого програмного забезпечення в інформаційні системи організації.

Основною проблемою є відсутність належного контролю за використанням зовнішніх носіїв інформації. Це може призводити до:

- несанкціонованого копіювання конфіденційних даних;
- витоку службової інформації;
- зараження комп'ютерів шкідливим програмним забезпеченням;
- порушення цілісності та доступності інформаційних ресурсів.

З огляду на зростання кількості кіберінцидентів, питання захисту АРМ від витоку даних через USB-пристрої набуває особливої актуальності. [1].

Метою роботи є аналіз основних загроз витоку інформації через знімні носії даних та визначення ефективних методів захисту АРМ.

У ході дослідження було проаналізовано основні підходи до захисту інформаційних систем від витоку даних через зовнішні носії. Виокремлено такі основні методи:

• Організаційні заходи безпеки: впровадження політик використання зовнішніх носіїв, регламентація доступу до інформаційних ресурсів та підвищення обізнаності персоналу з питань кібербезпеки.

• Програмні засоби контролю: використання спеціалізованого програмного забезпечення для моніторингу підключення USB-пристроїв, ведення журналу подій та блокування несанкціонованих носіїв інформації.

• Метод авторизації пристроїв: формування списку дозволених USB-носіїв та надання доступу лише до зареєстрованих пристроїв із фіксацією подій безпеки у системному журналі.

Додатковим ефективним механізмом підвищення рівня захисту є використання систем запобігання витоку даних (Data Loss Prevention – DLP) [2]. Такі системи дозволяють контролювати операції копіювання інформації на зовнішні носії, аналізувати типи файлів та обмежувати передачу конфіденційних даних. Застосування DLP-систем у поєднанні з механізмами контролю USB-пристроїв забезпечує комплексний підхід до захисту інформаційних ресурсів.

Проведений аналіз показав, що використання окремих методів захисту не забезпечує достатнього рівня безпеки інформаційних систем. Найбільш ефективним підходом є комплексне поєднання організаційних та програмних механізмів контролю доступу до зовнішніх носіїв інформації. У сучасних умовах особливу роль відіграють системи контролю периферійних пристроїв, які дозволяють здійснювати моніторинг підключення USB-носіїв, ідентифікувати пристрої за унікальними параметрами та обмежувати доступ до них відповідно до встановлених політик безпеки [3].

Важливим напрямом розвитку таких систем є інтеграція засобів контролю USB-пристроїв із системами управління інформаційною безпекою організації. Це дозволяє централізовано управляти політиками доступу, здійснювати журналювання подій безпеки та своєчасно виявляти спроби несанкціонованого використання зовнішніх носіїв даних. Крім того, використання сучасних систем моніторингу дозволяє автоматично повідомляти адміністратора безпеки про підключення невідомих пристроїв та потенційні інциденти.

Підсумовуючи результати дослідження, можна зазначити, що ефективний захист автоматизованих робочих місць від витоку інформації через знімні носії можливий лише за умови комплексного застосування організаційних політик безпеки, технічних засобів контролю та регулярного моніторингу використання периферійних пристроїв. Реалізація механізмів авторизації USB-носіїв та систем моніторингу подій безпеки дозволяє значно знизити ризик витоку конфіденційної інформації та підвищити загальний рівень захищеності інформаційної інфраструктури організації.

Перелік посилань:

1. Scarfone K., Padgette J. Guide to Enterprise Telework, Remote Access, and BYOD Security. – NIST Special Publication 800-46, 2020.

2. Whitman M., Mattord H. Principles of Information Security. – Cengage Learning, 2018.

3. ДСТУ ISO/IEC 27001:2018 Системи управління інформаційною безпекою. Вимоги. – Український національний стандарт, 2018.

4. НД ТЗІ 3.1.020-99 Засоби криптографічного захисту інформації. – Кабінет Міністрів України, 1999.

Гавриленко Д.П.
студент групи БСДМ-51, ННІКБЗІ ДУІКТ,
Київ, Україна

ПОБУДОВА SOC

Сучасний етап розвитку інформаційного суспільства характеризується стрімким зростанням кількості та складності кіберзагроз, що особливо актуально в умовах цифрової трансформації та євроінтеграційного курсу України. Організації різних галузей дедалі частіше стикаються з необхідністю забезпечення безперервного контролю за станом інформаційної безпеки та оперативного реагування на інциденти. Одним із ключових рішень у цій сфері є створення центрів моніторингу безпеки — Security Operations Center (SOC), які забезпечують централізований збір, аналіз та обробку подій безпеки. SOC дозволяє підвищити рівень захищеності інформаційних систем, своєчасно виявляти кіберзагрози та мінімізувати їх вплив на діяльність організації. Актуальність теми зумовлена необхідністю побудови ефективних систем кіберзахисту, що відповідають сучасним стандартам і вимогам. Метою роботи є аналіз принципів побудови SOC, його структури, основних компонентів та етапів впровадження.

Ключові слова: SOC, кібербезпека, SIEM, моніторинг безпеки, інциденти інформаційної безпеки, аналіз логів, реагування на інциденти, загрози, захист інформації.

Зі зростанням кількості кібератак та курсом України на євроінтеграцію особлива увага приділяється кібербезпеці. Центр моніторингу безпеки (SOC) виступає ключовим елементом захисту, забезпечуючи безперервний контроль, аналіз інцидентів і реагування на загрози. Його впровадження дозволяє своєчасно виявляти атаки та мінімізувати їх наслідки, що відповідає базовим принципам захисту інформаційних систем [1].

SOC визначається як централізована структура, що об’єднує людей, процеси та технології для забезпечення кібербезпеки. Основні функції включають моніторинг подій, аналіз загроз і реагування на інциденти із застосуванням підходів моделювання загроз [2].

Архітектура SOC включає такі компоненти, як SIEM, EDR, SOAR та інші інструменти аналізу. Вони забезпечують збір і кореляцію логів, виявлення аномалій та автоматизацію реагування, що відповідає сучасним підходам до побудови захищених систем [3]. Процес побудови SOC складається з етапів планування, проєктування, вибору технологій, формування команди та запуску системи. Важливу роль відіграє дотримання міжнародних стандартів і нормативних вимог, що підвищує ефективність управління інцидентами та загальний рівень кіберзахисту організації [4].

Таким чином, SOC є комплексним рішенням, яке поєднує технічні засоби, організаційні процеси та кваліфікований персонал для забезпечення ефективної протидії кіберзагрозам.

Процес створення SOC включає кілька ключових етапів:

- На етапі планування виконується аналіз ризиків, визначаються вимоги безпеки, цілі та модель функціонування SOC (власний або MSSP), із урахуванням міжнародних стандартів, зокрема ISO/IEC 27001 та NIST CSF [5].
- Далі формується архітектура SOC, яка включає інтеграцію з корпоративними системами, розробку процедур реагування, ескалації та звітності. Важливим елементом є створення стандартних інструкцій (runbooks), що забезпечують уніфіковане реагування на інциденти [6].
- На етапі впровадження обираються технології: SIEM для кореляції подій, EDR для моніторингу кінцевих пристроїв, SOAR для автоматизації процесів та платформи Threat Intelligence для аналізу загроз. Їх комплексне використання дозволяє ефективно виявляти та обробляти інциденти [7].

- Формування SOC-команди передбачає визначення ролей (аналітики L1–L3, інцидент-респондери, інженери безпеки) та організацію цілодобової роботи. Перед запуском проводиться тестування шляхом імітації атак (red team), що відповідає сучасним практикам перевірки захисту [8].
- Після запуску SOC працює в режимі 24/7, здійснюючи постійний моніторинг та аналіз подій. Важливим є регулярне оцінювання ефективності роботи через показники, такі як MTTD, MTTR та рівень хибних спрацювань, що дозволяє оптимізувати процеси безпеки [9].
- Організаційно SOC повинен відповідати законодавчим та нормативним вимогам. В Україні це включає рекомендації Держспецзв'язку, які передбачають моделювання атак, аналіз журналів подій та взаємодію з державними структурами (CERT-UA, НКЦКІ) [10].

Побудова SOC не стосується лише технічних засобів – вона тісно пов'язана з організаційними процедурами та нормативною базою. По-перше, необхідно визначити відповідальних керівників і забезпечити формалізацію політик (інформаційна безпека, реагування на інциденти, конфіденційність). SOC повинен координуватися з іншими підрозділами (IT, аудиту, кадровим, юридичним), а також узгоджувати дії з зовнішніми структурами.

Таблиця 1 - Інструменти та технології SOC

Категорія інструменту	**Опис та функції**	**Приклади (вендори)**
SIEM (Security Information and Event Management)	Збирає та аналізує події з різних систем, забезпечує кореляцію логів і виявлення загроз із автоматичним реагуванням	Splunk ES, IBM QRadar, Microsoft Sentinel
EDR (Endpoint Detection and Response)	Моніторить кінцеві пристрої, виявляє та ізолює шкідливу активність, збирає дані для розслідування	CrowdStrike Falcon, Microsoft Defender for Endpoint, SentinelOne
SOAR (Security Orchestration, Automation and Response)	Автоматизує процеси реагування, об'єднує інструменти та реалізує сценарії (playbook'и) обробки інцидентів	Palo Alto XSOAR, Splunk Phantom, IBM Resilient
Threat Intelligence Platform	Агрегує дані про загрози, надає актуальну інформацію та оновлює правила виявлення	Recorded Future, MISP, Anomali, Cisco Talos

Ефективність SOC оцінюється за ключовими показниками (KPI), серед яких основними є: MTTD (час виявлення інциденту), MTTR (час реагування), кількість інцидентів, рівень хибних спрацювань (FPR) та пропущених загроз (FNR), а також витрати на обробку інцидентів. Ці метрики дозволяють оцінити швидкість і якість реагування SOC, виявити проблеми в процесах або інструментах. Регулярний моніторинг KPI дає змогу підвищувати ефективність роботи SOC шляхом оптимізації технологій, процедур і підготовки персоналу [9].

Отже, було розглянуто основні принципи побудови Security Operations Center (SOC), включаючи архітектуру, функціональні компоненти та етапи впровадження. Встановлено, що ефективність SOC залежить від правильної організації процесів моніторингу, аналізу подій безпеки та реагування на інциденти. Проаналізовано ключові технології, зокрема SIEM,

системи управління інцидентами та інструменти збору логів, які забезпечують централізований контроль та виявлення загроз. Визначено, що впровадження SOC дозволяє значно підвищити рівень захищеності інформаційних систем, своєчасно виявляти атаки та мінімізувати їх наслідки. Таким чином, побудова SOC є важливим елементом сучасної системи кібербезпеки організації та відповідає вимогам цифрової трансформації та захисту інформації.

Перелік посилань:

1. Saltzer J. H., Schroeder M. D. The Protection of Information in Computer Systems // Proceedings of the IEEE. 1975. Vol. 63, No. 9. P. 1278–1308. URL: https://web.mit.edu/6.857/OldStuff/Fall03/ref/saltzer-schroeder.pdf
2. Threat Modeling Process // OWASP. [Електронний ресурс]. Режим доступу: https://owasp.org/www-community/Threat_Modeling_Process
3. Android security architecture and features // Source.android.com Docs. [Електронний ресурс]. Режим доступу: https://source.android.com/docs/security/features
4. Behl A., Behl K. Cybersecurity and Cyberwar: What Everyone Needs to Know. Oxford University Press, 2017.
5. ISO/IEC 27001:2022 Information Security Management Systems — Requirements. [Електронний ресурс]. Режим доступу: https://www.iso.org/standard/27001
6. Computer Security Incident Handling Guide (NIST SP 800-61 Rev. 2) // NIST. [Електронний ресурс]. Режим доступу: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf
7. Behl A., Behl K. Cybersecurity and Cyberwar: What Everyone Needs to Know. Oxford University Press, 2017.
8. MITRE ATT&CK Framework // MITRE. [Електронний ресурс]. Режим доступу: https://attack.mitre.org/
9. Measuring SOC Performance: Key Metrics // Splunk. [Електронний ресурс]. Режим доступу: https://www.splunk.com/
10. Методичні рекомендації з кіберзахисту (Наказ №570 від 03.07.2023) // Держспецзв’язку України. [Електронний ресурс].

Dmytro Hamza, Halyna Haidur, Serhii Zybin
State University of Information and Communication Technologies, Department of Information and Cyber Security, Kyiv, Ukraine

MULTI-CRITERIA SELECTION OF OPTIMAL HYBRID STACKING NSEMBLE MODEL FOR INTRUSION DETECTION YSTEMS: PARETO FRONT AND ABOVE-AVERAGE RULE FILTERING

This study proposes a weight-free, two-stage multi-criteria selection procedure for optimizing hybrid stacking ensemble models in Intrusion Detection Systems. By sequentially applying Pareto front identification and above-average rule filtering, the method objectively balances classification accuracy, F1-score, and inference time, reliably isolating the single optimal configuration on the CSE-CIC-IDS2018 dataset.

Keywords: Intrusion Detection Systems, stacking ensemble, multi-criteria optimization, Pareto front, above-average rule, model selection, cybersecurity, CSE-CIC-IDS2018, classification accuracy, inference latency.

Choosing the right configuration for a hybrid stacking ensemble is rarely a matter of picking the most accurate model. In Intrusion Detection Systems (IDS), raising classification accuracy tends to push inference time upward, while efforts to suppress false positives typically inflate model complexity [1]. These conflicts turn the selection problem into a genuine multi-criteria optimization task - one that cannot be resolved by ordering models on a single metric.

Let

$$M = \{m_1, \dots, m_n\},$$

denote the space of candidate stacking configurations, and let

$$f(m_i) = (\text{Accuracy}, \text{F1} - \text{score}, \text{Latency}).$$

be the vector of efficiency criteria associated with configuration m_i. The goal is to identify $m^* \in M$ that is jointly optimal across all three criteria - a goal for which no single right answer exists once the criteria conflict.

The MCDM literature offers several approaches, each with its own trade-offs.

The simplest route - **single-criterion ranking** - collapses the problem to maximizing one metric:

$$m^* = \text{argmax} f_1(m),$$

By design it ignores everything else, which is precisely why it often returns a model with unacceptable latency.

Weighted sum (scalarization) sidesteps this by combining the criteria into a single scalar,

$$Q = w_1 \cdot \text{Acc} + w_2 \cdot \text{F1} - w_3 \cdot T,$$

but the weights w_i must be supplied by the researcher - introducing subjectivity - and the method cannot reach points on concave regions of the Pareto front [2].

The ε-constraint method optimizes one criterion while holding the others within user-defined thresholds, yet forfeits access to compromise alternatives along the front.

Lexicographic ordering works when criteria can be ranked by strict importance - a condition rarely met in IDS, where accuracy and latency carry comparable operational weight [3].

Two approaches sidestep these limitations. The first - the **Pareto front** - identifies the set of non-dominated solutions. A configuration m_a dominates m_b,

$$m_a \succ m_b,$$

when it is no worse across every criterion and strictly better in at least one. The Pareto-optimal set follows naturally as

$$P = \{m \in M \mid \nexists\, m' \in M : m' \succ m\}.$$

This set captures every defensible compromise without demanding a single weight [4, 5]. The cost is that |P| can be large, leaving the analyst to choose among several equally valid candidates.

Above-average rule filtering (also known in the literature as mean-based thresholding or satisficing selection) takes the opposite approach: it keeps only those configurations that beat the sample mean across all three criteria,

$$f_1 > \overline{f}_1 \;\wedge\; f_2 > \overline{f}_2 \;\wedge\; f_3 < \overline{f}_3.$$

It is computationally trivial and weight-free. Its weakness is the mirror image of Pareto's: when applied to the full set M, the rule can silently discard non-dominated configurations whose values dip below the mean on one axis while excelling on another.

These observations suggest combining the two methods rather than choosing between them. We propose a two-stage selection procedure in which the Pareto front and above-average rule filtering each handle the task for which they are best suited.

Stage 1 - Pareto front. The full set M is filtered to its non-dominated subset P. The operation runs in $O(n^2)$ for naive pairwise comparison (or $O(k \cdot n^2)$ generalized to k criteria) and produces a compact set of mutually incomparable candidates. Dominated configurations drop out entirely: each is provably inferior to at least one alternative on every axis.

Stage 2 - above-average rule filtering. The Pareto set P is then narrowed to those candidates that also exceed the sample means computed over the full M,

$$\overline{f}_1 = 0.9775;\ \overline{f}_2 = 0.9621;\ \overline{f}_3 = 7.28\,ms,$$

where $f_3(m)$ is minimized (shorter inference is preferred) while Accuracy and F1 are maximized. Because thresholds are derived from the full sample rather than from P alone, they remain objectively fixed and are not distorted by the subset's size. The step costs $O(|P|)$ and yields a single configuration - one that is simultaneously non-dominated (Stage 1) and demonstrably above-average on every criterion (Stage 2).

The sequence matters. Pareto front answers **"Which configurations have no obvious shortcomings?"**; above-average filtering then answers **"Which of those is practically the best?"** Reversing the order is unsafe - a configuration passing only the above-average filter may still be dominated by another and is not, in general, guaranteed to be Pareto-optimal.

We evaluate the procedure on 10 stacking ensemble configurations trained on CSE-CIC-IDS2018 [6]. Preprocessing follows a standard pipeline: SMOTE oversampling for class imbalance [7], Min-Max normalization, and PCA reduction from 80 to 18 components (95% variance retained). Results appear in Table 1.

Table 1.

Stacking model metrics and two-stage selection results

Base models	Meta-classifier	Acc.	F1	T, ms (↓)	Pareto (Stage 1)	Above-avg. (Stage 2)
XGBoost + CatBoost + LightGBM	XGBoost	**0.9807**	**0.9657**	**7.16**	✓	✓
XGBoost + CatBoost + Random Forest	XGBoost	0.9801	0.9648	7.57	-	-
XGBoost + CatBoost + LightGBM	Gradient Boosting	0.9796	0.9639	7.40	-	-

XGBoost + CatBoost + Random Forest	Gradient Boosting	0.9791	0.9631	7.83	-	-
XGBoost + CatBoost + LightGBM	Random Forest	0.9784	0.9643	7.48	-	-
XGBoost + CatBoost + Random Forest	Random Forest	0.9779	0.9628	7.91	-	-
XGBoost + CatBoost + LightGBM	Logistic Regression	**0.9762**	**0.9611**	**6.91**	✓	-
XGBoost + CatBoost + Random Forest	Logistic Regression	0.9755	0.9598	7.31	-	-
CatBoost + LightGBM + Extra Trees	XGBoost	**0.9741**	**0.9587**	**6.51**	✓	-
CatBoost + LightGBM + Extra Trees	Gradient Boosting	0.9733	0.9572	6.73	-	-
Sample mean (thresholds)	**-**	**0.9775**	**0.9621**	**7.28**	**-**	**-**

Note. The downward arrow (↓) in the T, ms column indicates that f_3 (prediction time) is minimized; Accuracy and F1 are maximized.

Stage 1 (Pareto front) isolates three non-dominated configurations (rows 1, 7, 9 of Table 1), leaving seven dominated candidates excluded. Each of the three survivors wins on a different trade-off: row 1 attains the highest Accuracy (0.9807) at a latency cost of 7.16 ms; row 7 sits in the middle (0.9762 at 6.91 ms); row 9 achieves the lowest latency (6.51 ms) at the cost of Accuracy (0.9741). At this stage no configuration is objectively superior - the preferred one depends on the deployment profile of the IDS.

Stage 2 (above-average rule filtering) applies the thresholds computed over the full 10-configuration sample,

$$\overline{f}_1 = 0.9775;\ \overline{f}_2 = 0.9621;\ \overline{f}_3 = 7.28\,ms.$$

The evaluation demonstrates that Configuration 1 meets all three thresholds, while Configurations 7 and 9 are excluded due to accuracy below the sample mean (0.9762 and 0.9741 against the threshold of 0.9775, respectively). Configuration 1 is thus selected as the final recommendation.

The recommended configuration is **XGBoost + CatBoost + LightGBM with XGBoost as meta-classifier**. It delivers Accuracy 0.9807, F1-score 0.9657, and 7.16 ms per network flow - comfortably within the latency budget expected of real-time IDS [8].

Conclusions

1. A two-stage procedure is proposed that combines the Pareto front and above-average rule filtering. Stage 1 extracts all non-dominated candidates; Stage 2 selects among them the configuration that also exceeds the sample mean for each criterion. The procedure is fully formalized, reproducible, and does not contain user-entered weights.

3. Experimental evaluation on CSE-CIC-IDS2018 confirmed the procedure's effectiveness: Stage 1 reduced the 10-configuration search space to 3 Pareto-optimal candidates, and Stage 2 uniquely identified XGBoost + CatBoost + LightGBM / XGBoost (Accuracy 0.9807, F1-score 0.9657, latency 7.16 ms) as the final selection.

4. The order of application is not interchangeable. Pareto front answers a structural question - which configurations are Pareto-optimal; above-average filtering answers a practical one - which among them performs above the sample norm. Reversing the steps may yield a configuration that passes the above-average test yet is itself dominated.

References

1. Pinto A., Herrera L.-C., Donoso Y., Gutierrez J. A. Survey on Intrusion Detection Systems Based on Machine Learning Techniques for the Protection of Critical Infrastructure. Sensors. 2023. Vol. 23, No. 5. Art. 2415. DOI: https://doi.org/10.3390/s23052415

2. Marler R. T., Arora J. S. Survey of multi-objective optimization methods for engineering. Structural and Multidisciplinary Optimization. 2004. Vol. 26, No. 6. P. 369–395. DOI: https://doi.org/10.1007/s00158-003-0368-6

3. Miettinen K. Nonlinear Multiobjective Optimization. Boston: Kluwer Academic Publishers, 1999. 298 p. DOI: https://doi.org/10.1007/978-1-4615-5563-6

4. Ehrgott M. Multicriteria Optimization. 2nd ed. Berlin: Springer, 2005. 323 p. DOI: https://doi.org/10.1007/3-540-27659-9

5. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation. 2002. Vol. 6, No. 2. P. 182–197. DOI: https://doi.org/10.1109/4235.996017

6. Sharafaldin I., Lashkari A. H., Ghorbani A. A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP). 2018. P. 108–116. DOI: https://doi.org/10.5220/0006639801080116

7. Chawla N. V., Bowyer K. W., Hall L. O., Kegelmeyer W. P. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research. 2002. Vol. 16. P. 321–357. DOI: https://doi.org/10.1613/jair.953

8. Haidur H. I., Hamza D. Ye. Hybrid method for malicious activity detection based on stacking ensemble of classifiers. Modern Information Protection. 2025. No. 3(63). P. 20–26. DOI: https://doi.org/10.31673/2409-7292.2025.030315

9. Hamza D. Ye. Impact of CSE-CIC-IDS2018 dataset optimization on the efficiency of the hybrid stacking model for network intrusion detection. Cybersecurity: Education, Science, Technique. 2025. No. 2(30). P. 766–777. DOI: https://doi.org/10.28925/2663-4023.2025.30.963

Голота В.В., Дмітрієв В.Є.
Студент групи БСД-42, ННІКБЗІ ДУІКТ, Київ,
Україна

СИСТЕМА ЗАХИСТУ ІНФОРМАЦІЙНОЇ СИСТЕМИ ОРГАНІЗАЦІЇ ВІД ВНУТРІШНІХ ЗАГРОЗ НА БАЗІ TERAMIND

Захист інформаційної системи (ІС) сучасної організації сьогодні виходить за межі простого встановлення антивірусів чи налаштування мережевих екранів. Основна небезпека для стабільності ІС часто знаходиться всередині самої структури. Внутрішні загрози, що виникають через дії легітимних користувачів, є найбільш складними для ідентифікації, оскільки вони не потребують зламу системи ззовні. Створення ефективної моделі захисту на базі рішення Teramind дозволяє забезпечити цілісний контроль над робочим середовищем та мінімізувати ймовірність компрометації критично важливих ресурсів.

Ключові слова: Інформаційна система, внутрішні загрози, кіберзахист, Teramind, моніторинг активності, поведінковий аналіз, захист периметра, інсайдерські ризики.

Внутрішня загроза в інформаційній системі — це потенційна можливість порушення безпеки через використання наданих прав доступу. Проблема полягає в тому, що класичні засоби захисту фокусуються на подіях (наприклад, спроба входу в систему), але не на контексті дій користувача. Рішення Teramind пропонує концепцію «прозорості» дій, що перетворює кожного суб'єкта системи на об'єкт моніторингу в реальному часі.

Побудова системи захисту на базі Teramind дозволяє вирішити комплекс завдань, спрямованих на нейтралізацію внутрішніх ризиків:

1. Контроль привілейованих користувачів. Адміністратори систем та топ-менеджмент мають найширші права доступу. Teramind дозволяє фіксувати кожен їхній крок, що виключає можливість безконтрольної зміни налаштувань безпеки або видалення системних логів.
2. Виявлення аномальної поведінки (UEBA). Система автоматично створює базову лінію «нормальної» активності для кожного працівника. Будь-яке відхилення (наприклад, масове завантаження файлів у неробочий час або нетипове використання командного рядка) миттєво ініціює тривожний сигнал для служби безпеки.
3. Візуальна верифікація інцидентів. На відміну від лог-файлів, які важко аналізувати, Teramind надає відеозапис екрана та можливість перегляду дій у реальному часі. Це є незаперечним доказом при проведенні внутрішніх розслідувань.

Практична реалізація захисту інформаційної системи забезпечується через централізовану консоль управління Teramind (рис. 1). Візуалізація даних у реальному часі дозволяє офіцеру безпеки миттєво отримувати зріз активності по всій організації.

На відміну від традиційних систем логування, дашборд Teramind інтегрує в єдиному вікні показники використання програм, відвідані вебресурси та рівень інтенсивності роботи. Це дозволяє ідентифікувати аномалії не лише за технічними параметрами, а й за відхиленнями від типового робочого профілю працівника. Зокрема, функція відстеження "живих" сесій надає можливість негайного втручання у разі виявлення підозрілих операцій з конфіденційними файлами.

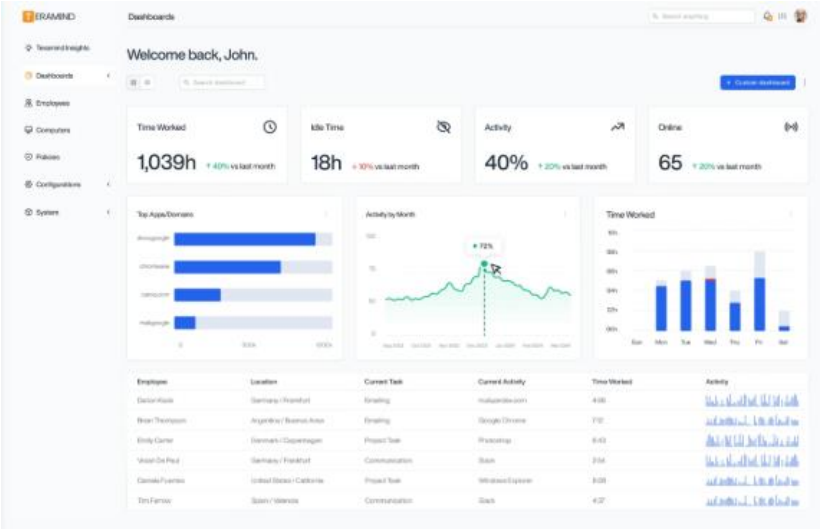

Рис. 1 інтерфейс панелі керування Teramind

Важливим елементом системи є механізм «реактивного керування». На основі отриманих даних адміністратор може створювати автоматизовані правила: наприклад, блокування доступу до певного сегмента мережі, якщо система зафіксувала підозрілу активність у браузері або месенджері. Це створює додатковий рівень захисту всередині ІС, який працює паралельно з правами доступу, наданими операційною системою.

Окрім технічного аспекту, впровадження Teramind формує середовище відповідальності. Співробітники, усвідомлюючи наявність системи моніторингу, рідше вдаються до порушень регламентів безпеки, що значно знижує ризик «халатних» загроз, таких як використання незахищених зовнішніх носіїв або відвідування шкідливих вебресурсів.

Побудова системи захисту на базі Teramind забезпечує глибоку ешелоновану оборону інформаційної системи організації. Поєднання функцій моніторингу, автоматичного аналізу ризиків та можливості негайного реагування дозволяє нейтралізувати внутрішні загрози на етапі їх виникнення. Це робить інформаційну систему стійкою не лише до технічних збоїв, а й до непередбачуваного людського фактора.

Перелік посилань:

1. Teramind UAM | Програмне забезпечення для моніторингу працівників. *Softprom – IT Distributor | Cyber Security, Cloud, IT Systems, CCTV, CAD*. URL: https://softprom.com/ua/vendor/teramind/product/teramind-uam-programne-zabezpechennya-dlya-monitoringu-pratsivnikiv (дата звернення: 22.04.2026).

2. Privileged Access Management (PAM) - контроль привілейованих користувачів. *Softprom – IT Distributor | Cyber Security, Cloud, IT Systems, CCTV, CAD*. URL: https://softprom.com/ua/vendor/cyberark/product/privileged-access-management-pam-kontrol-privileyovanih-koristuvachiv (дата звернення: 22.04.2026).

3. Workforce Analytics & Insider Risk Platform | Teramind. *Teramind*. URL: https://www.teramind.co/ (дата звернення: 22.04.2026).

Дасюк Юрій Євгенійович
студент групи БСДМ-52, ННІКБЗІ ДУІКТ,
Київ,Україна

СУЧАСНІ DEEPFAKE-ТЕХНОЛОГІЇ В РАМКАХ СОЦІАЛЬНОЇ ІНЖЕНЕРІЇ

Дана теза має на меті розглянути технології Deepfake як один з найбільш небезпечних інструментів сучасної соціальної інженерії, що загрожує як приватним особам, так і корпоративній та національній безпеці.

В умовах стрімкого розвитку штучного інтелекту, створення високоякісного синтетичного контенту (аудіо та відео) стало доступним навіть для користувачів без глибоких технічних знань. Завдяки цьому з'явились абсолютно нові вектори атаки для зловмисників, які використовують згенеровані штучним інтелектом матеріали в рамках соціальної інженерії для незаконного заволодіння коштами, шантажу, інформаційної боротьби та сіяння паніки.

Ключові слова: Deepfake, соціальна інженерія, кібербезпека, штучний інтелект.

Основними загрозами, що виникають через доступність методів генерації дипфейків, є: фінансове шахрайство, автоматизований вішинг (голосовий фішинг), обхід біометричної автентифікації, викрадення цифрової особистості, шантаж та поширення глибокої дезінформації.

Найбільш критичними для кібербезпеки є зличинні способи використання ШІ для створення діпфейків, зокрема:

- **Атаки з використанням синтетичного голосу.** Зловмисники імітують голос вищого керівництва компаній для того щоб «терміново» погодити та провести фіктивні фінансові операції. [4, с. 37]
- **Масштабований вішинг:** використання ШІ для одночасної генерації сотень, або навіть тисяч персоналізованих голосових дзвінків від імені банків або державних установ на телефони жертв.
- **Обхід систем віддаленої ідентифікації:** використання штучного інтелекту для створення «масок» на відео та фото в режимі реального часу для обману систем безпеки, що підтверджують особу у банківських установах та сервісах перевірки особистості.
- **Шантаж та дискредитація:** створення компрометуючих відео та фото матеріалів жінок з обличчям жертви для вимагання коштів, послуг, шантажу або ж інтернет травлі. [4, с. 38]
- **Політична дезінформація**, створення фейкових відеозвернень політиків чи військового керівництва для того щоб сіяти паніку та маніпулювати думкою суспільства в умовах збройного конфлікту. [4, с. 36]
- **Підрив довіри до медіа:** Поширення фейкових новин, імітування подій, що робить справжні відеоматеріали сумнівними, суттєво ускладнюючи розрізнення правди та брехні.

Технології Deepfake вже давно вийшли за межі експериментів та розваг, ставши потужною зброєю в руках кіберзлочинців. Шахраї використовують синтетичне аудіо та відео для крадіжки мільйонів доларів, корпоративного шпигунства та проведення психологічних операцій. [2]

Основними джерелами даних для створення якісних підробок та тренування генеративного штучного інтелекту є:

- **Відкриті профілі в соціальних мережах:** велика кількість фотографій та відео з різних ракурсів дозволяє якісно навчити нейромережу міміці людини.
- **Публічні виступи, подкасти та записи інтерв'ю** - ідеальне джерело чистого аудіо без фонового шуму для створення високоточних копій голосу. Створити якісну підробку можна всього завдяки кільком секундам запису. [4, с. 36]
- **Витоки корпоративних даних** можуть містити внутрішні відеозаписи нарад, що дає не лише біометричні дані, а й контекст для побудови правдоподібного сценарію атаки на інших співробітників компанії. [3]
- **Тіньові сервіси,** платформи, форуми та магазини в Darknet, що надають послуги зі створення діпфейків на замовлення, що знижує поріг входу для потенційних кіберзлочинців. [3]

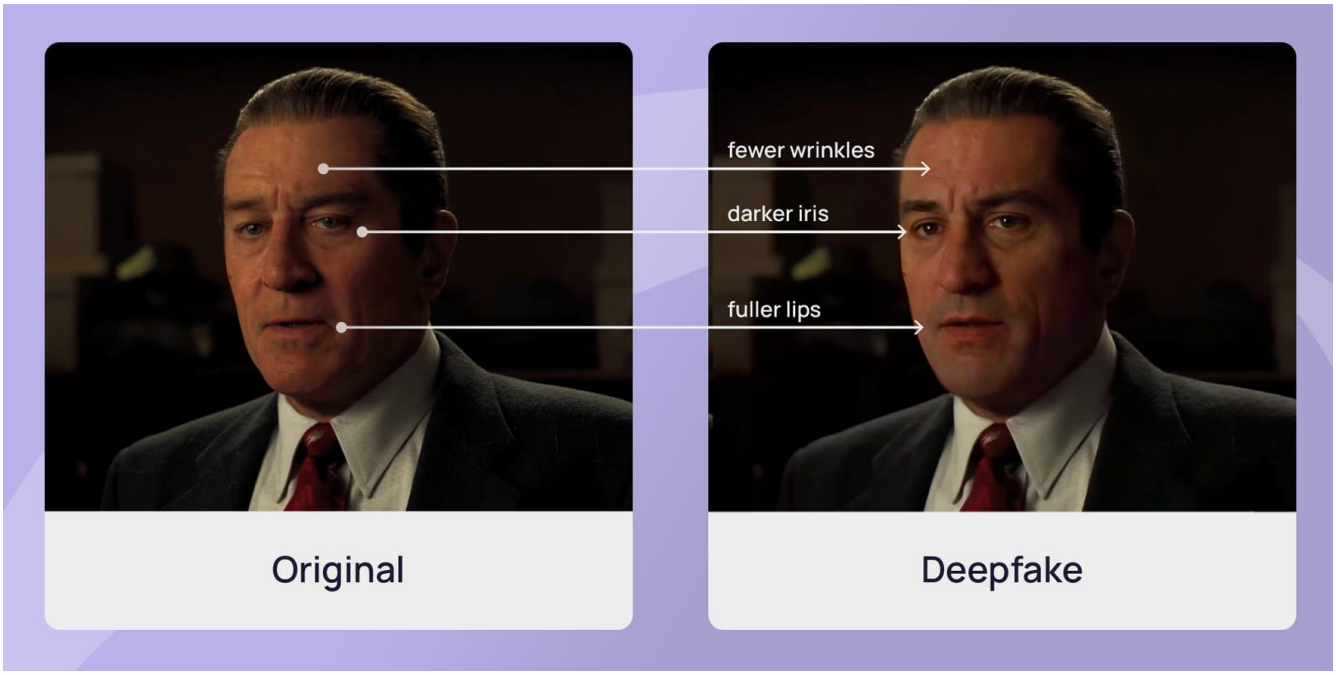

Рис. 1. Приклад використання Deepfake для імітації обличчя жертви

Можливість зловмисників створювати переконливі підробки голосу та зображення керівників або партнерів відкриває нові вектори для атак на бізнес та державний сектор України. [2]

Для ефективної протидії сучасним Deepfake-інструментам на державному та корпоративному рівнях необхідний комплексний підхід:

- **Впровадження автоматичних алгоритмів виявлення фейків** за допомогою інтеграції рішень на основі ШІ для автоматичного виявлення аномалій моргання, розсинхронізації губ, неприродних тіней та об’єктів в кадрах у файлах та медіапотоках. [1]
- **Перехід до архітектури Zero Trust:** запровадження політик, за якими жоден запит чи наказ, навіть голосовий або на відео від керівника не виконується без додаткової верифікації по іншому каналу зв'язку.
- **Фізична багатофакторна автентифікація:** відмова від виключно біометричних методів на користь фізичних апаратних ключів безпеки. [1]
- **Навчання персоналу:** проведення регулярних тренінгів та симуляцій фішингових атак з використанням діпфейків для підвищення рівня обізнаності співробітників. [1]

Робота по виявленню, захисту та зменшенню збитків, завданих за допомогою діпфейк атак – це комплексна задача, яка вимагає правильне об'єднання зусилль на технологічному, організаційному та освітньому рівнях.

Технології Deepfake з часом стали потужною зброєю в руках кіберзлочинців та несуть пряму загрозу як приватним особами, так і корпоративній та навіть національній безпеці. Протидія їм просто не може обмежуватися лише програмними засобами.

Ефективна стратегія кіберзахисту має обов'язково проєднувати не лише використання інтрументів ШІ для виявлення фейків та подвійну верифікацію особи за допомогою фізичних методів, а й постійне навчання персоналу та роботу з населенням для створення суспільного імунітету до діпфейків.

Разом з тим людям варто частіше використовувати сервіси для багаторазової автентифікації для захисту своїх акаунтів, менше постити себе, свої відео, фото та аудіо в соціальні мережі та використовувати кодові слова під час розмови з друзями, родиною або колегами що можна запитати у разі підозрілих «термінових» прохань.

Перелік посилань:

1. Deepfake Technologies and Social Engineering Trends 2025. Cyber Defense Magazine. https://www.cyberdefensemagazine.com/cybersecurity-trends-for-2025/ (дата звернення 15.04.2026).

2. Analysis of AI-driven synthetic media in modern conflict. Institute for Strategic Studies (2025). https://niss.gov.ua/en/news/statti/ai-frontiers-innovations-breakthroughs-challenges-november-2025 (дата звернення 14.04.2026).

3. Dark Net Recruitment is Turning Employees into Malicious Insiders For-Hire. Infosecurity Magazine (2019). https://www.infosecurity-magazine.com/opinions/dark-net-recruitment-malicious/ (дата звернення 16.04.2026).

4. Юртаєва, К. В. (2021). Кримінологічний аналіз використання технології Deepfake: коли фейк стає злочином. Вісник Кримінологічної асоціації України. https://dspace.univd.edu.ua/server/api/core/bitstreams/f94ec02c-eb36-4adc-8cf5-c121c5580020/content

Журавель Олександр Олегович,
студент групи БСДМ-51, ННІКБЗІ ДУІКТ, Київ,
Україна

ВИЯВЛЕННЯ ТА РЕАГУВАННЯ НА ІНЦИДЕНТИ БЕЗПЕКИ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

Ефективне виявлення та реагування на інциденти безпеки є важливою складовою забезпечення захисту інформаційних систем. У сучасних умовах постійного зростання кількості кіберзагроз організації повинні забезпечувати безперервний моніторинг подій та своєчасне реагування на потенційні інциденти. Від якості цих процесів залежить здатність системи протистояти атакам і мінімізувати їх негативні наслідки.

Ключові слова: інциденти безпеки, виявлення та реагування на інциденти безпеки в інформаційних системах, кібербезпека, моніторинг.

У сучасних інформаційних системах процес обробки інцидентів безпеки базується на комплексному підході, що включає етапи виявлення (**Detect**) та реагування (**Respond**). Відповідно до рекомендацій NIST, ці етапи є взаємопов'язаними та спрямовані на своєчасне виявлення загроз і обмеження впливу на систему.

Процес виявлення інцидентів безпеки

Виявлення інцидентів безпеки передбачає постійний моніторинг (**Continuous Monitoring**) інформаційних систем з метою ідентифікації підозрілої активності. Основною метою цього етапу є своєчасне виявлення потенційних інцидентів та визначення їх характеру [1, с.15].

Для підвищення ефективності виявлення інцидентів застосовуються сигна-турний аналіз та виявлення аномалій (рис.1).

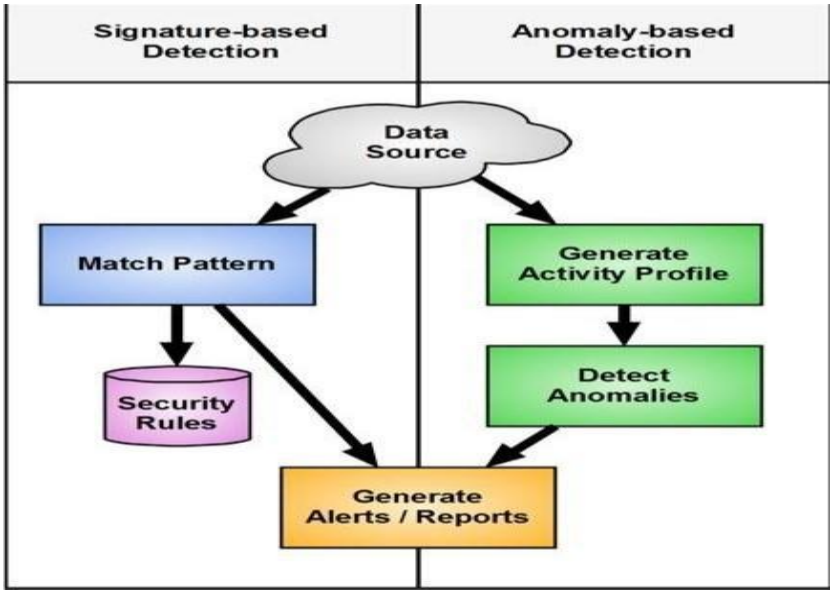

Рис. 1 – Схема сигнатурного аналізу та аналізу виявлення аномалій в інформаційній системі

Сигнатурний аналіз базується на використанні відомих атак і забезпечує швидку ідентифікацію вже відомих загроз. Підхід, заснований на виявленні аномалій, передбачає аналіз типової поведінки системи з подальшим визначенням відхилень від неї. Такий підхід дозволяє виявляти раніше невідомі загрози, але може супроводжуватися підвищеним рівнем помилкових спрацювань.

Реагування на інциденти безпеки

Реагування на інциденти безпеки включає комплекс заходів, спрямованих на обмеження їх впливу та відновлення функціонування системи. Згідно NIST, процес управління інцидентами має циклічний характер і включає кілька етапів реагування (рис. 2) [2, с. 21]: підготовка (**Preparation**), виявлення та аналіз (**Detection and Analysis**), стримування, ліквідація та відновлення (**Containment, Eradication and Recovery**), діяльність після інциденту (**Post-Incident Activity**).

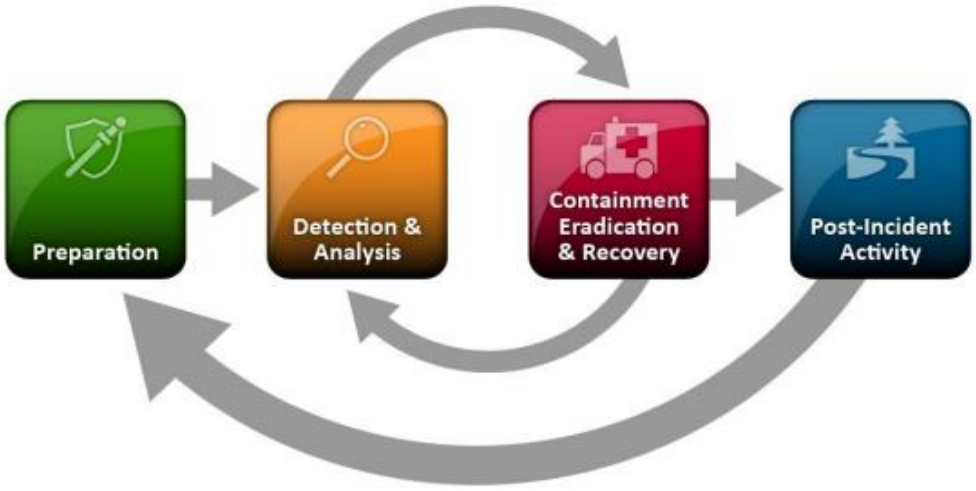

Рис. 2 – Основні етапи реагування на інциденти безпеки

Етап підготовки передбачає створення необхідної організаційної та технічної бази для реагування, включаючи розробку політик безпеки та процедур обробки інцидентів. Виялення та аналіз спрямовані на ідентифікацію інциденту, визначення його характеру та оцінку можливих наслідків.

Безпосереднє реагування включає стримування інциденту безпеки, ліквідацію загрози та відновлення працездатності системи. Стримування спрямоване на обмеження поширення інциденту, ліквідація – усунення його причин, відновлення – повернення системи до нормального функціонування.

Кінцевим етапом реагування є діяльність після інциденту, яка передбачає аналіз отриманого досвіду та вдосконалення заходів безпеки. Циклічний характер процесу забезпечує безперервне вдосконалення механізмів виявлення та реагування на інциденти безпеки [2, с. 21].

Виклики щодо виявлення та реагування на інциденти безпеки

Процеси виявлення та реагування на інциденти безпеки в інформаційних системах пов’язані з рядом суттєвих труднощів. Однією з ключових проблем є ускладнення сучасних кібератак, які часто мають багаторівневу структуру та здатні тривалилий час залишатися непоміченими засобами захисту.

Значним викликом виступають великі обсяги даних, що ускладнюють їх аналіз і своєчасне виялення потенційних загроз, що вимагає суттєвих обчислювальних і людських ресурсів та підвищує навантаження на фахівців з інформаційної безпеки. Варто ще відзначити дефіцит кваліфікованих спеціалістів і складність організації ефективної координації процесів реагування, що може спричинити затримки у процесі виявлення та стримування інцидентів безпеки [3]. Таким чином, ефективність процесів виялення та реагування визначається здатністю своєчасно ідентифікувати інциденти та забезпечити узгоджене виконання заходів щодо стримування, ліквідації та відновлення функціонування інформаційних систем.

Перелік посилань:

1. NIST. The NIST Cybersecurity Framework (CSF) 2.0. National Institute of Standards and Technology

This publication is available free of charge from: https://doi.org/10.6028/NIST.CSWP. 29 February 26, 2024. URL: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf (дата звернення 20.04.2026).

2. NIST. Computer Security Incident Handling Guide. Recommendations of the National Institute of Standards and

Technology. NIST Special Publication 800-61 Revision 2. Paul R. Cichonski, Thomas Millar, Timothy Grance, Karen Scarfone URL: https://www.nist.gov/publications/computer-security-incident-handling-guide (дата звернення 20.04.2026).

3. SentinelOne. Cyber Security Incident Response: Definition and Best Practices. URL: https://www.sentinelone.com/cybersecurity-101/cybersecurity/cyber-security-incident-response/ (дата звернення 21.04.2026).

Зелінський М.С.
студент групи ПД-43, ННІІТ ДУІКТ,
Київ, Україна

ЗАХИСТ ДАНИХ В ЕЛЕКТРОННІЙ СИСТЕМІ АГРОЗАМОВЛЕНЬ

Цифрові технології швидко змінюють українське сільське господарство. Замість старих способів доставки товарів тепер працюють онлайн-системи. На цих платформах зберігається критично важлива інформація - починаючи від грошей, що проходять через них, закінчуючи особистими даними користувачів. Чим дорожчі дані, тим частіше хакери обирають їх метою, особливо коли сайти мають ненадійний захист або уразливі API. Коли перевірка користувачів, контроль доступу чи безпека сховищ ігноруються, компанії опиняються перед серйозним ризиком зриву роботи. Отже створення сталого та багатошарового захисту даних у спеціалізованих системах посідає провідне місце під час їхньої побудови.

Ключові слова: кібербезпека, захист даних, аграрні ресурси, автентифікація, OAuth 2.0, Keycloak, JWT, вебзастосунок.

Замість простої торгівлі тепер все частіше діють онлайн-платформи, де зберігають важливу інформацію про поставки, грошові операції, запаси й дані людей. Ці цифрові майданчики приваблюють хакерів, тому що там легко знайти слабкі місця. Наприклад, коли система погано перевіряє користувачів, або дає всім однаковий доступ без обмежень. Ще одна проблема - відкриті програмні інтерфейси, якими можуть скористатися не за призначенням. У результаті компанії ризикують втратити великі суми чи конфіденційну інформацію [1].

Це дослідження оглядає захист даних у системі для замовлення й доставки сільськогосподарських матеріалів. Замість класичного підходу, система працює на Java з допомогою Spring Framework - тут задіяні Spring Boot, Spring Security, Spring Data JPA і Spring MVC. PostgreSQL виступає основним сховищем інформації, до якого отримують доступ завдяки Hibernate як ORM-рішенню. Систему запускають через Docker Compose, одночасно з цим специфікацію REST API. На фронтенді працює фреймворк Next.js на базі React. Через безпечний REST API блоки спілкуються один з одним, ніби живуть окремо. Такий поділ обмежує пряме торкання до серверної частини. Spring Security додає ще один шар, коли справа доходить до типових атак (наприклад, налаштування CORS та захист від CSRF при взаємодії між Next.js та Java).

Три ролі діють у системі - менеджер, адміністратор і постачальник, причому кожна з ролей отримала свій чіткий набір правил для роботи. Через це безпечний процес входу став головною справою при побудові всього задуму. Щоб перевіряти й допускати користувачів, взяли OAuth 2.0 разом із Keycloak. Цей протокол дає клієнту токени доступу з чітко визначеними правами, не передаючи паролі між частинами системи. Сам Keycloak працює окремо, всередині контейнера, стаючи єдиним пунктом контролю над користувачами: реєструє людей, видає і оновлює токени, веде облік ролей, працює з OpenID Connect та SAML. Коли потрібно - підключається до сторонніх систем, наприклад Google, Microsoft або внутрішній LDAP, через що користувачі можуть увійти один раз і мати доступ одразу до багатьох сервісів [2].

Додатково у системі передбачено захист на рівні взаємодії з базою даних. Замість простих запитів, Spring Data JPA разом із Hibernate використовує тільки параметризовані команди для PostgreSQL - це блокує можливість SQL-ін’єкцій, які часто використовують у атаках на веб-додатки. Паролі адміністраторів зберігаються у криптографічно стійких хешах. Дані між користувачем і сервером проходять через захищене з’єднання за допомогою HTTPS протоколу. Через такий підхід закриваються слабкі місця не тільки при рівні автентифікації, а й там, де інформацію тримають чи пересилають.

Опісля вдалої перевірки особи, Keycloak надсилає клієнтові JWT-токен у форматі JSON, зашифрований методом RS256. Замість одного спільного секрету, у RS256 - два ключі:

один закритий, інший відкритий. Той, що прихований, ставить підпис, а доступний кожному перевіряє його. Всередині цього токена визначено, хто користувач, яку має посаду, і чому йому можна довіряти; ніякої додаткової пам'яті на сервері не треба. Кожен новий запит проходить крізь спеціальні бар'єри в Spring Security, через ланцюжок фільтрів. Перевірка підпису й строку дії токену виконується на місці, без звертання до сховища даних. Час обробки відповідей скоротився значно, а система легко розтягується на десятки вузлів, бо жоден з них не залежить від спільних сесій. Окремий доступ у системі працює так: менеджер керує замовленнями, хоч і не бачить фінансові дані. Постачальник отримує призначені йому замовлення. Адміністратор тримає під контролем кожен елемент. У Keycloak реєструються всі невдалі входи, й облікові записи можуть в такому випадку тимчасово блокувати - це сильно знижує ризик атак перебором [2].

Особисту інформацію учасників системи обробляють строго з урахуванням вимог чинного законодавства. Відповідно до Закону України "Про захист персональних даних", збір і обробка персональних даних допускаються лише у чітко визначених цілях та з обов'язковим забезпеченням їх захисту від несанкціонованого доступу [3]. У межах системи контроль над доступом працює через ролі: до даних постачальників і менеджерів має доступ тільки адміністратор. Запропоновані рішення OAuth 2.0, JWT-токенів з підписом RS256, а також управління ролями через Spring Security разом із Keycloak, створюють комплексний бар'єр проти загроз. Цей захист не перевантажує основну структуру програми, водночас дозволяючи їй можливість подальшого масштабування. Можливо, в майбутньому стане доречною заміна refresh-токенів через регулярні інтервали, однаковий спосіб записування подій безпеки в єдиному журналі та налаштовування SSO для корпоративних клієнтів.

Перелік посилань:

1. Боскін О.О., Корніловська Н.В., Поліщук В.М., Сарафаннікова Н.В. (2023). Безпека веб-додатків та хакерські атаки. Вісник Херсонського національного технічного університету, № 3(86), 83-02. https://doi.org/10.35546/knutu2078-4481.2023.3.11

2. Spilca L. Spring Security in Action. – Shelter Island: Manning Publications, 2020. – 558 p.

3. Закон України «Про захист персональних даних» від 01.06.2010 №2297-VI. [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/2297-17

Кравченко Є.Ю.
студент групи ПД – 44, ННІІТ, ДУІКТ
Київ, Україна

ЗАХИСТ КРИТИЧНОЇ ІНФРАСТРУКТУРИ У WEB-ЗАСТОСУНКУ ДЛЯ ВІЗУАЛІЗАЦІЇ ПОВІТРЯНИХ ЗАГРОЗ

У роботі досліджено комплексний підхід до забезпечення безпеки web-застосунку для візуалізації повітряних загроз у реальному часі. Основна увага приділена мінімізації поверхні атаки шляхом використання технології SignalR замість традиційних REST-ендпоінтів. Реалізовано багаторівневий захист: автентифікація через Telegram Mini App, обмеження лімітів з'єднань та приховування адміністративних інтерфейсів (код 404). Для захисту мережевої інфраструктури впроваджено сервіси Cloudflare, що забезпечують приховування реальної IP-адреси сервера, фільтрацію трафіку (WAF) та ефективну протидію DDoS-атакам. Запропоновані заходи забезпечують стабільну роботу та конфіденційність даних у системах моніторингу.

Ключові слова: інформаційна безпека, SignalR, Cloudflare, DDoS-захист, Telegram Mini App, візуалізація загроз, Web-безпека.

У сучасних web-застосунках, що працюють у режимі реального часу та взаємодіють із великою кількістю користувачів, питання захисту серверної інфраструктури від несанкціонованого доступу та атак типу відмови в обслуговуванні (DDoS) є критично важливим. Особливо це актуально для систем, які надають дані про повітряні загрози в реальному часі, оскільки вони є привабливою ціллю для перевантаження серверів та спроб обходу механізмів доступу. Для забезпечення стабільної роботи таких систем необхідно впроваджувати комплексний підхід, що поєднує захист на рівні мережі, протоколів передачі даних та архітектури самого застосунку.

У рамках даного проєкту реалізовано web-застосунок, у якому передача актуальних даних про цілі відбувається за допомогою технології SignalR. Ця бібліотека для ASP.NET Core дозволяє серверному коду миттєво надсилати сповіщення підключеним клієнтам, використовуючи двосторонній зв'язок [3]. Такий підхід дозволяє повністю відмовитися від класичних REST-ендпоінтів для отримання даних клієнтом, що значно знижує ризик їх прямого використання зловмисниками для генерації великої кількості запитів та спроб автоматизованого збору інформації. Відсутність відкритих HTTP-ендпоінтів для отримання критичних даних суттєво ускладнює проведення DDoS-атак, спрямованих на виснаження ресурсів API [2].

На серверній частині системи реалізовано набір адміністративних ендпоінтів для керування конфігурацією парсера. Доступ до цих елементів керування обмежується за допомогою спеціального ключа доступу, який передається у запиті та перевіряється на рівні спеціального програмного забезпечення середнього шару (middleware). Унікальною особливістю реалізації є те, що у разі відсутності або некоректності ключа сервер повертає відповідь з кодом 404 (Not Found) замість 401 або 403. Це дозволяє повністю приховати сам факт існування адміністративних інтерфейсів від автоматизованих сканерів та зловмисників.

Додатково впроваджено обмеження на кількість перепідключень до SignalR-хабу, що запобігає атакам на виснаження ресурсів через масове створення нових з'єднань.

Для забезпечення суворого контролю доступу лише цільовою аудиторією впроваджено багаторівневу систему автентифікації на основі Telegram Mini App та стандартів JWT (JSON Web Token). Процес ініціалізації взаємодії починається із надсилання клієнтом запиту на спеціалізований ендпоінт авторизації, який містить пакет даних initData. Серверна частина виконує криптографічну валідацію отриманого хеш-підпису, звіряючи його із секретним ключем бота, що дозволяє однозначно підтвердити автентичність запиту та цілісність переданих параметрів. У разі успішної валідації сервер повертає клієнту короткоживучий

JWT-токен, який стає єдиним легітимним ідентифікатором для всіх подальших операцій. Важливою архітектурною особливістю є те, що всі наступні взаємодії, включаючи встановлення дуплексного каналу через SignalR-хаб, відбуваються виключно за умови наявності валідного токена у заголовках запиту. Для мінімізації ризиків, пов'язаних із потенційною компрометацією сесії, час життя токена (TTL) обмежено однією годиною. Це змушує клієнтську частину регулярно оновлювати авторизаційні дані, що суттєво підвищує стійкість системи до атак типу "повторне відтворення" та забезпечує динамічне керування правами доступу в режимі реального часу.

Передача даних у системі здійснюється виключно через захищений протокол HTTPS, що базується на використанні TLS для шифрування трафіку. Це гарантує конфіденційність та цілісність даних під час їх транзиту, захищаючи користувачів від атак типу «Man-in-the-Middle» (перехоплення даних). Крім того, для ускладнення аналізу клієнтської частини застосовано обфускацію JavaScript-коду. Цей процес перетворює вихідний код у важкочитабельну форму, що значно знижує ризик реверс-інжинірингу та пошуку логічних вразливостей у роботі фронтенд-частини застосунку.

Для захисту мережевої інфраструктури та приховування реального походження сервера використано платформу Cloudflare. Cloudflare діє як реверс-проксі, що дозволяє фільтрувати шкідливий трафік на рівні edge-вузлів ще до того, як він досягне цільового сервера. Використання Web Application Firewall (WAF) забезпечує захист від поширених вразливостей, описаних у списках OWASP Top 10 [1]. Приховування реальної IP-адреси сервера унеможливлює проведення прямих атак на хостинг-провайдера (рис. 1).

Запропонований комплекс заходів, що включає перехід на SignalR, інтеграцію з Telegram Mini App, використання Cloudflare та багаторівневу перевірку прав доступу, дозволяє значно підвищити рівень безпеки web-застосунку. Мінімізація поверхні атаки та приховування критичних вузлів системи забезпечують стабільну роботу та захист даних у системах моніторингу повітряних загроз.

Перелік посилань:

1. OWASP (2025). OWASP Top 10: The Ten Most Critical Web Application Security Risks [Електронний ресурс]. — Режим доступу: URL https://owasp.org/Top10/2025/

2. OWASP (2023). OWASP API Security Top 10 [Електронний ресурс]. — Режим доступу: URL https://owasp.org/API-Security/editions/2023/en/0x00-header/

3. Microsoft (2024). ASP.NET Core SignalR documentation [Електронний ресурс]. — Режим доступу: URL https://learn.microsoft.com/aspnet/core/signalr/introduction

Кутовий Д.С.
Студент групи БСДМ-52, ННІКБЗІ ДУІКТ,
Київ, Україна

МЕНЕДЖМЕНТ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Менеджмент інформаційної безпеки (МІБ) є критично важливою складовою стратегічного управління сучасною організацією. В умовах стрімкого зростання кількості кіберзагроз та складності IT-інфраструктури, захист активів перестає бути суто технічним завданням і переходить у площину системного управління ризиками, процесами та людським фактором. Впровадження системи управління інформаційною безпекою (СУІБ) дозволяє організаціям не лише мінімізувати ймовірність інцидентів, а й забезпечити безперервність бізнес-процесів та відповідність міжнародним стандартам.

Ключові слова: менеджмент інформаційної безпеки, СУІБ, управління ризиками, ISO/IEC 27001, кіберстійкість.

Ефективний менеджмент інформаційної безпеки ґрунтується на розумінні того, що безпека — це не стан, а безперервний процес. Згідно зі звітами ENISA (2024) та Gartner (2024), понад 70 % успішних кібератак пов’язані з прорахунками в управлінських процесах або людським фактором, що підкреслює пріоритетність організаційних заходів над суто технічними рішеннями [2; 3]. Основною метою МІБ є забезпечення тріади CIA: конфіденційності, цілісності та доступності інформації.

Фундаментом для побудови системного менеджменту є міжнародні стандарти, серед яких провідну роль займає ISO/IEC 27001:2022. Він пропонує процесну модель, що базується на циклі Дефакторумінга (PDCA: Plan-Do-Check-Act), що дозволяє організації постійно вдосконалювати рівень захисту. Окрім ISO, широкого розповсюдження набув фреймворк NIST Cybersecurity Framework (2.0), який у 2024 році був оновлений з акцентом на функцію «Governance» (Управління), що підкреслює роль керівництва у прийнятті рішень щодо безпеки [1, с. 12–15].

Центральним елементом менеджменту є управління ризиками. Процес включає ідентифікацію інформаційних активів, оцінку вразливостей та загроз, а також вибір стратегії обробки ризику (прийняття, зниження, передача або уникнення). Використання кількісних та якісних методів аналізу ризиків (наприклад, за методиками ISO/IEC 27005) дозволяє раціонально розподіляти бюджет на безпеку, інвестуючи в найбільш критичні напрямки. Дослідження підтверджують, що компанії з впровадженим ризик-орієнтованим підходом зазнають у середньому на 40 % менше фінансових збитків від кіберінцидентів [4, с. 202].

Важливою складовою МІБ є управління людським капіталом. Оскільки соціальна інженерія залишається одним із головних векторів атак, менеджмент має фокусуватися на створенні культури кібербезпеки (Security Awareness). Це включає не лише періодичне навчання, а й розробку політик прийнятного використання ресурсів (AUP) та чітких процедур реагування на інциденти. Системний підхід до навчання персоналу дозволяє знизити ризик успішного фішингу та інших атак, спрямованих на людину, на 60–70 % [1, с. 58].

Сучасний менеджмент також охоплює питання забезпечення безперервності бізнесу (BCP) та відновлення після катастроф (DRP). У цифровій економіці час простою критичних систем безпосередньо корелює з репутаційними та фінансовими втратами. Менеджмент безпеки координує створення резервних копій, розробку планів відновлення та проведення регулярних кібернавчань для перевірки готовності персоналу до кризових ситуацій [3].

Окремим викликом для менеджменту є комплаєнс — відповідність законодавчим та галузевим вимогам (наприклад, GDPR, PCI DSS, або національне законодавство у сфері захисту критичної інфраструктури). Координація технічних засобів захисту, таких як DLP-системи, MFA та засоби криптографії, повинна повністю відповідати юридичним зобов'язанням організації перед клієнтами та державою.

Отже, менеджмент інформаційної безпеки є багаторівневою дисципліною, яка поєднує стратегічне планування, технічний контроль та управління персоналом. Тільки за умови інтеграції безпеки в усі бізнес-процеси організація може досягти стану кіберстійкості — здатності не лише протистояти атакам, а й швидко відновлюватися після них у динамічному цифровому середовищі.

Перелік посилань:

1. Humphrey C., Williams J. Information Security Management Principles. BCS Learning & Development Limited, 2021. 320 p.
2. ENISA Threat Landscape 2024. European Union Agency for Cybersecurity. (https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024).
3. ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements. ISO, 2022.
4. Tipton H. F., Krause M. Official (ISC)2 Guide to the CISSP CBK. CRC Press, 2023. 1200 p.

Леонов А.В.
студент кафедри захисту інформації,
НУ «Львівська політехніка»
Львів, Україна

Івкова В.С.
асистент кафедри захисту інформації,
НУ «Львівська політехніка»
Львів, Україна

СИСТЕМИ АВТЕНТИФІКАЦІЇ КОНТЕНТУ ЯК ЗАСІБ КІБЕРЗАХИСТУ: ПРОТИДІЯ ЗАГРОЗАМ НА ОСНОВІ ШІ В OSINT-РОЗВІДЦІ

Швидкий розвиток генеративного штучного інтелекту створює критичну вразливість для OSINT-розвідки, засмічуючи інформаційний простір синтетичними медіафайлами та підриваючи довіру до джерел аналізу. Пропонується проактивний інструмент кіберзахисту - система автентифікації контенту, основою якого є стандарт C2PA, що дозволяє аналітикам швидко верифікувати походження та цілісність матеріалів. Імплементація такої системи забезпечує вищу надійність висновків OSINT-аналітиків та підвищує ефективність протидії інформаційним операціям різного роду.

Ключові слова: OSINT-розвідка, генеративний штучний інтелект, автентифікація контенту, стандарт C2PA.

Стрімкий розвиток генеративного штучного інтелекту створює новий вектор кіберзагроз, особливо для OSINT-розвідки, більшою мірою шляхом «отруєння» даних синтетичним контентом. Значною мірою ця тенденція ускладнює встановлення факту достовірності аналізу і створює ризики для персональної, корпоративної та національної безпеки.

Метою є дослідження концептуального фреймворку системи автентифікації контенту як проактивного засобу кіберзахисту що є актуальним й для покращення стійкості OSINT-процесів.

Синтетичні медіазагрози реалізуються на усіх рівнях, їх класифікація дозволить зрозуміти вектор потенційних атак. Так, наприклад на державному рівні вони спрямовані на підрив національної безпеки через ІПСО (наприклад, поширення deepfake-новин особливо в період воєнного стану), дискредитацію міжнародних відносин (сфабриковані аудіозаписи державних діячів). Щодо корпоративного рівня то це промисловий шпіонаж (створення фейкових профілів експертів за для виманювання даних) та ринкові маніпуляції (обвал акцій конкурента через компроментуюче deepfake-відео). На персональному рівні [1, с. 1772] загрози здебільшого охоплюють шантаж та шахрайство. Не можна оминути й технічний рівень реалізації загроз такого типу, а саме «отруєння» даних, що призводить до катастрофічно неправильних аналітичних висновків при автоматизованому та мануальному OSINT-дослідженню.

Для ефективної протидії широкому колу загроз на зразок описаних вище необхідний перехід від реактивної до проактивної архітектури кіберзахисту. Основою такої архітектури може стати система автентифікації контенту, що дозволяє верифікувати контент в момент його аналізу.

Пропонується концепція, що базується на технології підтвердження походження контенту (Content Provenance). Фундаментальний принцип полягає у вбудовуванні в медіафайл невидимого для людського сприйняття, але машинозчитуваного, криптографічно стійкого маркера, що є по суті методом стеганографії [2, с. 150]. Для OSINT-досліджень такий

маркер стає фундаментальним інструментом верифікації, що супроводжуватиме контент протягом усього його життєвого циклу, та дозволяє оцінити його надійність.

Ефективність системи забезпечується безперервністю процесу на трьох ключових етапах:

1. Етап створення: маркування повинно відбуватися безпосередньо в момент генерації контенту. Розробники генеративних моделей мають інтегрувати у свої системи функціонал, який автоматично вбудовує маркер. Цей маркер міститиме критично важливу оцінювальну інформацію, до прикладу джерело, модель генеративного засобу та часову мітку.
2. Етап поширення: вимогою до маркера є його стійкість до обробки файлу. Він має зберігатись після стиснення, зміни розміру, завантаження на онлайн-платформи. Це гарантуватиме, що «сертифікат походження» не буде втрачено в процесі поширення інформації.
3. Етап верифікації: на цьому етапі потенційний OSINT-аналітик маючи відповідні інструменти може миттєво перевірити контент. Таким інструментом може бути браузерний плагін, що позначає синтетичний контент; API для верифікації в професійних платформах як от Maltego; портативна версія аналізатора для перевірки файлів завантажених з ненадійних джерел.

Запропонована концепція спирається на реальні індустріальні стандарти, зокрема на ініціативу C2PA,які підтримують провідні технологічні компанії. Згідно з офіційною специфікацією, головною метою стандарту є розв'язання проблеми довіри в цифровому просторі: «Встановлення походження медіа є критично важливим для забезпечення прозорості, розуміння і, зрештою, довіри» [3, с. 6].

Для OSINT, цей стандарт надає єдиний технічний фреймворк для автоматизованої верифікації походження та цілісності контенту. Це дозволяє аналітикам швидко відфільтровувати синтетичні дані, перетворюючи C2PA з інструменту для захисту авторських прав на фундаментальний елемент архітектури кіберзахисту.

Інтеграція системи автентифікації контенту в робочі процеси OSINT-розвідки потребує переходу від просто поділу інформації на «правдиву» або «фейкову» до більш гнучкої моделі, яка оцінює рівень довіри до різноманітних джерел.

Такий підхід дозволить аналітикам робити більш зважені та обґрунтовані висновки в умовах широкого застосування генеративного медіаконтенту.

Ключовим елементом такої системи стає впровадження «шкали довіри» в OSINT-інструментах. Ця шкала допомагає аналітику розрізняти контент з високою довірою (що має криптографічний підпис від перевіреного джерела), неверифікований (що вимагає класичних підходів OSINT) та з підтвердженням генеративності (об'єкт для аналізу дезінформаційних кампаній).

Впровадження глобальної системи автентифікації контенту стикається з трьома ключовими викликами:

а) Технічна протидія: зловмисники постійно розроблятимуть методи затирання маркерів.

Варіант вирішення: забезпечення безперервного розвитку криптографічно стійких алгоритмів маркування для унеможливлення фальсифікації «сертифікату довіри»

б) Масове впровадження: успіх стандарту залежить від його прийняття широким колом розробників, включаючи спільноту відкритого коду.

Варіант вирішення: глобалізацію даного стандарту може забезпечити розробка міжнародних нормативно-правових актів, що зобов'язують маркувати згенерований контент, а також тиск з боку великих платформ.

в) До стандартний контент: система не зможе автоматично автентифікувати мільярди файлів, створених до її впровадження.

Варіант вирішення: слід розуміти, що ця система доповнює, а не замінює існуючі методики OSINT-верифікації та OSINT-дослідження, фокусуючись на новому та найбільш небезпечному контенті.

Таким чином, впровадження подібних систем автентифікації контенту, що спираються на C2PA, є не просто технічним нововведенням, а стратегічною відповіддю на загрози, створені генеративним ШІ. Для OSINT-розвідки це означає перехід від реактивних методів боротьби з дезінформацією до проактивної моделі верифікації, що значно підвищує надійність та швидкість аналізу Це дозволяє протидіяти так званому «дивіденду брехуна», коли через поширення фейків суспільство починає сумніватися навіть у достовірності справжніх доказів [1, с. 1785]. В кінцевому підсумку, такий підхід є основою для нової архітектури довіри, де підтвердження походження та цілісність даних стають головним критерієм цінності в цифровому середовищі.

Список використаних джерел:

1. Чесней Р., Сітрон Д. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security // Lawfare Institute, 2022. [Електронний ресурс] – Режим доступу: https://scholarship.law.bu.edu/faculty_scholarship/640/

2. Прокопович-Ткаченко Д. Глибокі автоенкодери для приховування інформації: сучасні підходи та перспективи розвитку // Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка». – 2025. – № 4(28). – С. 150–161. [Електронний ресурс] – Режим доступу: https://doi.org/10.28925/2663-4023.2025.28.765

3. Коаліція за походження та автентичність контенту (C2PA) // C2PA Technical Specification // 2024. URL: https://spec.c2pa.org/specifications/specifications/1.0/specs/_attachments/C2PA_Specification.pdf

Лисенко Д.В.
студент групи ПД-43, ННІІТ ДУІКТ,
Київ, Україна

ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ REST API У WEB-ЗАСТОСУНКУ ДЛЯ ВЕДЕННЯ ТРЕНУВАЛЬНОГО ЩОДЕННИКА

У роботі досліджено підходи до гарантування безпеки REST API для web-застосунку з відстеження тренувального прогресу. Проаналізовано застосування Spring Security для побудови безсесійної архітектури на основі JSON Web Token (JWT). Описано механізми захисту конфіденційних даних за допомогою алгоритму BCrypt для хешування паролів та налаштування політики CORS. Розглянуто використання Spring Data JPA для захисту від SQL-ін'єкцій. Запропоноване рішення дозволяє створити надійну архітектуру захисту персональних даних користувачів у Health & Fitness застосунках та забезпечити масштабованість системи.

Ключові слова: інформаційна безпека, REST API, Spring Security, JSON Web Token (JWT), BCrypt, CORS, SQL-ін'єкції, Health & Fitness застосунки.

Стрімка цифровізація сфери Health & Fitness докорінно змінила підхід до моніторингу фізичної активності, перевівши його з площини ручного документування у формат високотехнологічного аналізу. Традиційні методи фіксації результатів, як-от паперові нотатки, поступово втрачають актуальність через неможливість оперативного опрацювання великих масивів інформації. Розробка спеціалізованого web-застосунку дозволяє автоматизувати розрахунок критичних показників, зокрема сумарного об'єму навантаження, інтенсивності та індивідуальних максимумів. Це не лише спрощує процес ведення щоденника, а й забезпечує глибоку візуалізацію прогресу через інтерактивні дашборди.

Разом з цим, такі платформи акумулюють детальні персональні дані (фізіологічні показники та параметри тренувального процесу), тому особлива увага приділяється захисту серверної інфраструктури [1]. Високий рівень безпеки backend-частини є необхідною умовою для запобігання компрометації даних та збереження іміджу продукту.

У цій роботі описано підхід до побудови архітектури безпеки для REST API web-застосунку з відстеження тренувального прогресу, з акцентом на конкретних технічних рішеннях, прийнятих у процесі реалізації. Програмне рішення реалізовано мовою Java з використанням екосистеми Spring Boot. Оскільки взаємодія між клієнтською частиною та сервером базується на архітектурному стилі REST, це вимагає впровадження надійних stateless-механізмів автентифікації та контролю доступу [2]. Для гарантування конфіденційності та цілісності інформації інтегровано модуль Spring Security, який надає гнучкі інструменти для налаштування правил безпеки на рівні окремих ендпоінтів.

Основою системи захисту розробленого застосунку є токенізація доступу за допомогою JSON Web Token (JWT). Структурно JSON Web Token складається з трьох частин: заголовка (header), набору заяв (payload) та криптографічного підпису (signature). У заголовку вказується тип токена та алгоритм шифрування (наприклад, HS256). Корисне навантаження містить ідентифікатор користувача, його ролі та термін дії токена. Використання підпису гарантує, що дані не були змінені третіми сторонами під час передачі, оскільки для його перевірки сервер використовує секретний ключ, відомий лише backend-частині. Такий підхід забезпечує повну автономність системи, оскільки серверу не потрібно зберігати стан сесії в оперативній пам'яті, що значно підвищує масштабованість додатку. Замість класичних серверних сесій реалізовано механізм, за якого після успішної авторизації клієнт отримує криптографічно підписаний

токен, що передається у HTTP-заголовку при кожному наступному запиті [3]. Це дозволяє безпечно ідентифікувати користувача та розмежовувати доступ до ресурсів, ізолюючи функціонал запису підходів таким чином, щоб клієнт міг взаємодіяти лише з власною історією тренувань. Водночас зберігання облікових даних у базі PostgreSQL у відкритому вигляді є неприпустимим, тому для нейтралізації цієї загрози застосовано алгоритм незворотного хешування BCrypt [1, 4]. BCrypt автоматично генерує унікальну сіль для кожного пароля, тому навіть якщо зловмисник отримає доступ до бази даних, використання райдужних таблиць або прямого перебору стає практично неможливим — кожен хеш доведеться зламувати окремо.

Додатковим рівнем безпеки є налаштування політики Cross-Origin Resource Sharing (CORS), яка суворо регламентує, які зовнішні домени мають право звертатися до API, відсікаючи несанкціоновані запити зі сторонніх ресурсів. Ланцюг фільтрів безпеки перехоплює всі вхідні запити, виконуючи валідацію JWT до того, як дані досягнуть контролерів бізнес-логіки [3]. Крім того, завдяки використанню технологій Spring Data JPA та Hibernate на рівні доступу до даних, система автоматично застосовує параметризовані запити, що ефективно нівелює ризики SQL-ін'єкцій та захищає базу даних від несанкціонованих маніпуляцій [2, 4].

Таким чином, поєднання JWT-автентифікації, BCrypt-хешування, CORS-політики та захисту від SQL-ін'єкцій через JPA дозволило сформувати достатньо надійну архітектуру безпеки для тренувального онлайн щоденника.

Перелік посилань:

1. Боскін О.О., Корніловська Н.В., Поліщук В.М., Сарафаннікова Н.В. (2023). Безпека веб-додатків та хакерські атаки. *Вісник Херсонського національного технічного університету, № 3(86),* 83-92. https://doi.org/10.35546/kntu2078-4481.2023.3.11

2. Офіційна документація Spring Security [Електронний ресурс]. – Режим доступу: https://docs.spring.io/spring-security/reference/

3. Специфікація JSON Web Token (RFC 7519) [Електронний ресурс]. – Режим доступу: https://datatracker.ietf.org/doc/html/rfc7519

4. Walls C. (2022). Spring in Action. Shelter Island: Manning Publications. 544 p. URL: https://www.manning.com/books/spring-in-action-sixth-edition

Михайленко Юлія Іванівна
Студентка групи БСДМ-52, ННІКБЗІ ДУІКТ,
Київ, Україна

АВТОМАТИЗАЦІЯ КІБЕРРОЗВІДКИ: OSINT ЗА ДОПОМОГОЮ ШІ

У сучасному цифровому середовищі обсяги інформації стрімко зростають, що ускладнює її аналіз та використання у сфері кібербезпеки. Особливо це стосується відкритих джерел інформації, які активно використовуються в процесах кіберрозвідки. Одним із ключових підходів до збору таких даних є OSINT (Open Source Intelligence), що передбачає отримання та аналіз інформації з публічно доступних ресурсів.

До основних джерел OSINT належать соціальні мережі, новинні ресурси, форуми, веб-сайти та відкриті реєстри. Водночас зростання обсягів цих даних створює проблему їх ефективної обробки, що зумовлює необхідність автоматизації відповідних процесів. У цьому контексті особливого значення набуває використання технологій штучного інтелекту, які дозволяють значно підвищити ефективність аналізу інформації та оптимізувати процеси кіберрозвідки.

Ключові слова: кіберрозвідка, OSINT, відкриті джерела інформації, штучний інтелект, автоматизація, кібербезпека, аналіз даних, машинне навчання.

OSINT є важливим інструментом у кіберрозвідці, оскільки дозволяє отримувати актуальну інформацію без використання закритих або конфіденційних джерел. Основною перевагою цього підходу є легальність і доступність даних. Процес OSINT включає кілька етапів: збір інформації, її фільтрацію, аналіз та формування висновків. У традиційному підході ці етапи виконуються людиною вручну, що потребує значних часових і когнітивних ресурсів. Приклад реалізації OSINT без застосування штучного інтелекту наведено на рисунку 1 [1].

Наприклад, аналітик може здійснювати пошук інформації про організацію, її співробітників, соціальні мережі, форуми та інші відкриті джерела. Такий підхід має обмеження, пов'язане з людським фактором, оскільки швидкість або обмеження можливості обробки великих обсягів даних.

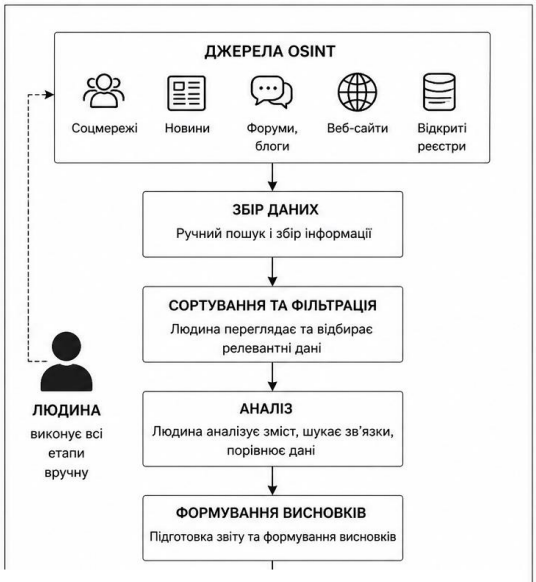

Рис.1. OSINT без застосування штучного інтелекту

Як результат, аналіз без використання штучного інтелекту передбачає значні витрати часу на збір та аналіз даних, що знижує ефективність і підвищує ризик пропуску важливої інформації.

Застосування штучного інтелекту дозволяє автоматизувати ключові етапи роботи з OSINT, зокрема збір, обробку та аналіз даних. Алгоритми машинного навчання здатні обробляти великі масиви інформації, виявляти закономірності та знаходити зв'язки між об'єктами. Подібно до сучасних пошукових систем, які здатні знаходити релевантну та схожу інформацію за запитом користувача, системи на основі штучного інтелекту можуть аналізувати великі обсяги відкритих даних, визначати ключові теми та виявляти потенційні загрози [3].

Наприклад, у процесі кіберрозвідки штучний інтелект може автоматично аналізувати повідомлення в соціальних мережах або новинних ресурсах, виділяти ключові слова, групувати інформацію та сигналізувати про підозрілі активності. Приклад автоматизованого процесу OSINT із застосуванням штучного інтелекту наведено на рисунку 2.

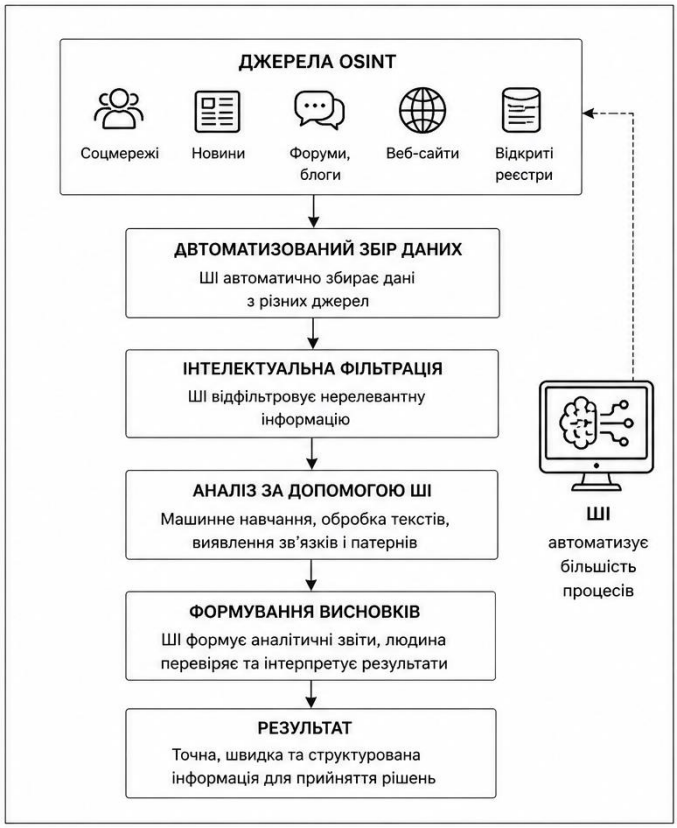

Рис.2. OSINT із застосуванням штучного інтелекту

Як показано на рисунку 2, використання штучного інтелекту дозволяє автоматизувати обробку інформації, прискорити аналіз даних та підвищити ефективність кіберрозвідки. Таким чином, використання штучного інтелекту значно підвищує швидкість та ефективність обробки інформації у порівнянні з ручним аналізом [2,4].

Використання штучного інтелекту в кіберрозвідці має як переваги, так і обмеження. До переваг належать висока швидкість обробки даних, можливість аналізу великих обсягів інформації та автоматизація процесів. Штучний інтелект також здатний виявляти закономірності та зв'язки, які складно визначити вручну. Водночас існують і певні недоліки. Зокрема, штучний інтелект не завжди здатний повноцінно інтерпретувати зміст інформації, оскільки його робота базується на пошуку шаблонів і ключових слів. Це може призводити до неточностей або помилкових висновків.

Крім того, результати аналізу значною мірою залежать від якості вхідних даних. У разі використання недостовірної або неповної інформації існує ризик отримання некоректних результатів. Також важливо враховувати відсутність критичного мислення у штучного інтелекту, що обмежує його здатність до глибокого аналізу.

Таким чином, інтеграція штучного інтелекту в OSINT відкриває нові можливості для розвитку кіберрозвідки. Використання інтелектуальних алгоритмів дозволяє автоматизувати обробку великих обсягів даних, підвищити швидкість аналізу та ефективність виявлення загроз, проте потребує відповідального та комплексного підходу до його використання.

Перелік посилань:

1. Навчальний посібник: Open Source Intelligence Tools and Resources Handbook. – Режим доступу: https://docslib.org/doc/11283797/open-source-intelligence-tools-and-resources-handbook-2020
2. Стаття: An Introduction to Artificial Intelligence. – Доступ: https://arxiv.org/abs/1911.05755
3. Стаття: Comprehensive Overview of Artificial Intelligence Applications. – Доступ: https://arxiv.org/abs/2409.13059
4. Стаття: Explainable Artificial Intelligence (XAI). – Доступ: https://arxiv.org/abs/1910.10045

М'якота І.А.
студент групи ПД-44, ННІІТ ДУІКТ,
Київ, Україна

ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ДАНИХ У WEB-ЗАСТОСУНКУ MINI TRELLO ДЛЯ УПРАВЛІННЯ ЗАВДАННЯМИ ТА КОМАНДНОЇ ВЗАЄМОДІЇ

У сучасних web-застосунках питання захисту інформації є одним із ключових, оскільки програмні системи працюють із персональними даними користувачів, обліковими записами, файлами, ролями доступу та внутрішньою службовою інформацією. Особливо це стосується систем керування завданнями, у яких одночасно обробляються дані про користувачів, дошки, задачі, вкладення, повідомлення та сповіщення. Тому забезпечення конфіденційності, цілісності та доступності даних є обов'язковою умовою якісної реалізації програмного продукту [1, 2, 3].

У межах дипломної роботи розробляється web-застосунок Mini Trello, призначений для управління задачами, дошками, спринтами, повідомленнями та взаємодією між користувачами. Така система містить чутливі дані: електронні адреси користувачів, паролі, ролі доступу, налаштування профілю, файли-вкладення, а також інформацію про робочі процеси команди. У зв'язку з цим під час реалізації програмного продукту було приділено увагу основним механізмам інформаційної безпеки.

Одним із базових рівнів захисту є безпечне зберігання паролів. У моделі користувача Mini Trello пароль не зберігається у відкритому вигляді, а обробляється за допомогою функції generate_password_hash, а перевірка виконується через check_password_hash. Такий підхід дозволяє зменшити ризик компрометації облікових даних у разі несанкціонованого доступу до бази даних. Крім того, у системі передбачено механізм генерації токена для скидання пароля з обмеженим часом дії, що є додатковим елементом захисту облікового запису.

Другим важливим компонентом є захист сесії користувача. У застосунку використовується Flask-Login із сесійною автентифікацією. Для cookie сесії налаштовано параметри SESSION_COOKIE_HTTPONLY та SESSION_COOKIE_SAMESITE=Lax, а також аналогічні параметри для cookie "remember me". Атрибут HttpOnly зменшує ризик викрадення cookie через клієнтські сценарії, а SameSite частково обмежує міжсайтову передачу сесійних даних. У production-середовищі також доцільно обов'язково використовувати параметр SESSION_COOKIE_SECURE=True, що забезпечує передавання cookie лише через HTTPS-з'єднання [2, 3].

Ще одним механізмом захисту є контроль доступу на основі ролей. У Mini Trello реалізовано два рівні авторизації: глобальні ролі користувача (viewer, manager, admin) та ролі всередині конкретної дошки (viewer, member, admin, owner). Для цього у програмному коді використано окремий сервіс RBAC, який виконує перевірку прав доступу перед роботою з API. Завдяки цьому користувачі з роллю viewer мають доступ лише на читання, а редагування задач, вкладень, учасників дошки та інших сутностей дозволяється лише тим ролям, яким це передбачено бізнес-логікою системи. Такий підхід знижує ризик несанкціонованої зміни або видалення даних.

Важливим елементом безпеки є також захист HTTP-відповідей за допомогою Content Security Policy (CSP). У застосунку після кожного запиту встановлюється заголовок Content-Security-Policy, який обмежує джерела завантаження скриптів, стилів, зображень, шрифтів та зовнішніх підключень. Це зменшує ймовірність виконання небажаного коду та ускладнює реалізацію частини XSS-атак. Хоча у поточній конфігурації для сумісності залишено unsafe-inline та unsafe-eval, сам факт використання CSP уже є важливим кроком до підвищення рівня захищеності системи. Надалі доцільно поступово відмовлятися від цих винятків і переходити до більш жорсткої політики безпеки.

Окрему увагу у Mini Trello приділено завантаженню файлів. Система підтримує додавання вкладень до задач, а також завантаження аватарів користувачів. Для цього

використовується перевірка розширень файлів, обмеження максимального розміру, генерація безпечних імен файлів та збереження їх у визначених каталогах. Для аватарів дозволені лише конкретні графічні формати (png, jpg, jpeg, webp, gif), а розмір файлу обмежено 5 МБ. Для загальних вкладень застосовується список дозволених розширень із конфігурації системи. Такий підхід зменшує ризик завантаження небезпечного вмісту та відповідає базовим рекомендаціям щодо безпечної обробки файлів [1].

Крім того, у системі реалізовано валідацію вхідних даних на рівні API. Наприклад, під час редагування профілю перевіряється довжина імені користувача, біографії, часової зони, мови інтерфейсу; під час створення задач перевіряються назва, пріоритет, виконавець, батьківська задача та інші параметри. Аналогічні перевірки застосовуються для зміни пароля, додавання коментарів, вкладень, повідомлень і запрошень. Це дозволяє знизити ризик логічних помилок, спотворення даних та частини атак, пов’язаних із передачею некоректних параметрів.

Слід зазначити, що для підвищення рівня захисту в подальшому доцільно розширити реалізовані механізми безпеки. Зокрема, варто впровадити повноцінний CSRF-захист для форм і state-changing запитів, обов’язкове використання HTTPS у production-середовищі, суворішу політику CSP без unsafe-inline, а також додаткову перевірку MIME-типів і вмісту файлів під час завантаження. Також перспективним є аудит WebSocket-з’єднань та додавання централізованого журналювання подій безпеки.

Отже, у web-застосунку Mini Trello вже реалізовано низку важливих механізмів захисту інформації: хешування паролів, токени скидання пароля, захищені cookie-параметри, рольову модель доступу, Content Security Policy, контроль завантаження файлів і валідацію вхідних даних. У сукупності це формує базовий рівень захисту персональних і службових даних користувачів. Забезпечення безпеки інформації є невід’ємною складовою якісної розробки web-застосунків, тому в подальшому така система потребує постійного вдосконалення відповідно до сучасних загроз та практик безпечного програмування.

ДЖЕРЕЛА

1. OWASP Foundation. File Upload Cheat Sheet [https://cheatsheetseries.owasp.org/cheatsheets/File_Upload_Cheat_Sheet.html].
2. OWASP Foundation. Session Management Cheat Sheet [https://cheatsheetseries.owasp.org/cheatsheets/Session_Management_Cheat_Sheet.html].
3. Pallets. Security Considerations — Flask Documentation [https://flask.palletsprojects.com/en/stable/web-security/].

Осадчий М.В.
студент групи ПД-42, ННІІТ, ДУІКТ,
Київ, Україна

КІБЕРЗАХИСТ WEB-СЕРВІСУ ДЛЯ ЗБОРУ ПРОПОЗИЦІЙ З САЙТІВ ПРОДАЖУ АВТО

Досліджено питання забезпечення захисту інформації у web-агрегаторі оголошень вторинного авторинку. Проведено аналіз потенційних загроз та уразливостей, характерних для сервісів збору даних (scraping) та платформ з великим обсягом динамічного контенту. Запропоновано архітектурні рішення для безпечного зберігання бази даних агрегатора та впровадження протоколів шифрування, що дозволяє мінімізувати ризики витоку персональних даних та отримання недостовірної інформації про транспортні засоби.

Ключові слова: кіберзахист, web-агрегатор, вторинний авторинок, моніторинг оголошень, цілісність даних.

Сучасний ринок авто з пробігом розширює свої кордони, виходячи з країн, що розвиваються до країн розвинених. Згідно звіту [1] розмір ринку вживаних автомобілів у 2025 році становив 1,9 трильйона доларів США і, як очікується, перевищить 3,47 трильйона доларів США до 2035 року, зростаючи на понад 6,2% у середньому протягом прогнозованого періоду. Така тенденція сприяє зростанню популярності цифрових платформ для купівлі/продажу вживаних авто. А це, в свою чергу, вимагає інноваційних підходів до розробки програмних продуктів, зокрема web-агрегаторів.

У сучасних web-системах, що збирають та обробляють великі обсяги даних із різних джерел, питання захисту інформації набуває особливого значення. Програмний продукт для агрегації пропозицій з сайтів продажу автомобілів з пробігом накопичує не лише технічні характеристики транспортних засобів, а й дані про користувачів, їхні пошукові запити, обрані фільтри, історію переглядів, контактні дані продавців та іншу інформацію, пов’язану з використанням сервісу. Зловмисники використовують різноманітні тактики, щоб експлуатувати уразливості та завдати шкоди. Серед поширених загроз безпеці web-сайтів агрегаторів виділяють DDoS-атаки, фішинг, SQL-ін'єкції, XSS та інше зловмисне програмне забезпечення [2]. Наслідками потенційної атаки можуть стати недоступність ресурсу, виведення з ладу, витік конфіденційних даних, перекручення або втрата інформації, що потягне за собою як репутаційні, так і матеріальні збитки.

У даній роботі подано аналіз реалізованих підходів до забезпечення безпеки інформації у web-сайті для агрегації пропозицій з сайтів продажу автомобілів з пробігом. Для такого ресурсу забезпечено базові вимоги конфіденційності, цілісності й доступності даних, оскільки порушення безпеки може призвести до витоку персональної інформації, підміни оголошень, поширення шкідливого контенту або несанкціонованого доступу до службових функцій [3].

Одним із базових напрямів захисту в розробленому web-сайті є правильна організація автентифікації та контролю доступу. У системі передбачено захист облікових записів користувачів, особистого кабінету, збережених оголошень, підписок на нові пропозиції та взаємодії з адміністративною панеллю. Паролі користувачів хешуються за допомогою bcrypt перед збереженням у базі даних, а доступ до адміністративних функцій розмежовано відповідно до ролей користувачів. Для пароля задається 8 раундів хешування - чим більше раундів, тим безпечніше, хоча довше триває обчислення.

Іншим важливим аспектом є захист від типових web-загроз. Для сайту такого типу враховано ризики порушення контролю доступу, ін’єкцій, помилок конфігурації безпеки, недоліків автентифікації, а також загрози, пов’язані з некоректною обробкою зовнішніх даних, що надходять із інших ресурсів.

Оскільки агрегатор автоматично отримує або оновлює інформацію з різних джерел, у системі реалізовано перевірку вхідних даних, безпечну обробку параметрів запитів,

екранування вмісту перед виведенням у браузері та обмеження можливості виконання шкідливих сценаріїв. Для доступу до БД використано підготовлені SQL-запити через Hibernate ORM, а сервер налаштовано для роботи через HTTPS із використанням TLS-сертифіката [4].

Окрему увагу під час розробки програмного продукту приділено безпеці самого процесу створення web-сайту. Вимоги безпеки враховано ще на етапі проєктування, виконується перевірка коду, контроль залежностей, своєчасне оновлення компонентів та журналювання подій безпеки. Такий підхід зменшує ризик появи уразливостей у готовій системі та підвищує її стійкість до атак.

Отже, у web-сайті для агрегації пропозицій з сайтів продажу автомобілів з пробігом реалізовано основні механізми захисту інформації, необхідні для його надійної роботи. Використання правових вимог щодо захисту персональних даних, сучасних підходів до захисту web-застосунків та безпечних практик розробки дало змогу знизити ризики витоку інформації, несанкціонованого доступу та компрометації сервісу.

Перелік посилань:

1. Used Car Market Overview— Режим доступу: https://www.researchnester.com/reports/used-car-market/6271
2. Website security and cyberthreats – Режим доступу: https://www.one.com/en-gb/website-security/what-is-website-security/
3. Firebase Security Rules and Firebase Authentication — Google Developers, 2026. — [Електронний ресурс]. — Режим доступу: https://firebase.google.com/docs/rules/rules-and-auth
4. Пірог О.В. Безпека вебдодатків : навч. посібн. / О.В. Пірог. – Електронні дані. – Житомир : Житомирська політехніка, 2025. – 290 с. http://eztuir.ztu.edu.ua/123456789/8813

Піскунов Костянтин Валерійович
студент групи БСДМ-51, ННІЗІ ДУІКТ,
Київ, Україна

ЗАХИСТ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

Системи штучного інтелекту слід розглядати як окремий об’єкт кіберзахисту в су часних інформаційних системах. Запропоновано практичний підхід до їх захисту, що поєднує контроль цілісності й походження даних, безпечну розробку та розгортан ня моделей, обмеження доступу до AI-компонентів і зовнішніх інтеграцій, протидію атакам типу prompt injection і data poisoning, постійний моніторинг поведінки моде лей, AI red teaming, а також впровадження механізмів AI governance на всіх етапах життєвого циклу системи. Очікуваний ефект − зменшення площі атаки, підвищення стійкості AI-систем до маніпуляцій і витоків даних, зниження ризику компрометації моделей та забезпечення контрольованого використання штучного інтелекту в інфор маційних системах.

Ключові слова: штучний інтелект, кібербезпека, AI-системи, безпека моделей, prompt injection, data poisoning, AI governance.

Штучний інтелект швидко інтегрується в корпоративні сервіси, аналітичні платформи, системи підтримки рішень і цифрові продукти. У таких умовах захисту потребують не лише мережі, сервери та застосунки, а й самі AI-системи: моделі, навчальні дані, промпти, API, векторні бази, зовнішні інструменти та середовище розгортання. Це пов’язано з тим, що ШІ створює нову поверхню атаки, де поряд із класичними кіберзагрозами виникають специфічні ризики, притаманні саме моделям і даним.

Однією з головних загроз є компрометація даних. Якщо модель навчається або працює на недостовірних, спотворених чи навмисно отруєних даних, це безпосередньо впливає на її результати, стійкість і безпечність. На відміну від звичайних IT-систем, у ШІ навіть часткова зміна набору даних може поступово змінити поведінку моделі без прямого втручання в програмний код. У результаті система може формувати помилкові висновки, надавати некоректні рекомендації або приймати рішення, що не відповідають реальним умовам. Тому захист AI-систем має включати перевірку джерел даних, контроль цілісності, аудит змін, обмеження несанкціонованого доступу та валідацію перед використанням у навчанні чи інференсі.

Другою критичною загрозою є prompt injection. Йдеться про ситуацію, коли зловмисник через спеціально сформований запит або вміст зовнішнього дже рела змушує модель ігнорувати системні обмеження, видавати чутливу інфор мацію або виконувати небажані дії. Ризик особливо високий у тих системах, де модель взаємодіє з файлами, поштою, пошуковими сервісами чи зовнішні ми API. У такому випадку захист повинен базуватися не лише на фільтрації запитів, а й на розмежуванні прав, перевірці контексту, ізоляції інструментів і жорсткому контролі дозволених дій.

Важливим напрямом є також захист самої моделі та її оточення. Сучасні AI-рішення часто залежать від готових бібліотек, зовнішніх моделей, сторонніх датасетів і хмарних сервісів. Це створює ризики ланцюга постачання: скомпро метована залежність або ненадійний компонент можуть вплинути на всю си стему. Тому безпечне середовище AI має передбачати перевірку залежностей, контроль версій, ізоляцію середовищ навчання і розгортання, а також мінімі зацію привілеїв для сервісних облікових записів.

Практичний підхід до захисту AI-систем доцільно будувати за п’ятьма на прямами. Перший − безпечні дані: перевірка джерел, контроль цілісності, про стежуваність походження. Другий − захист моделі: обмеження доступу до ваг, конфігурацій і середовища виконання. Третій − безпечне розгортання: сегмен тація, контроль інтеграцій, мінімізація привілеїв, валідація вхідних і вихідних даних. Четвертий − постійний моніторинг: журналювання,

виявлення анома лій, аналіз запитів і відповідей, реагування на інциденти. П’ятий − AI governance: визначення відповідальних осіб, аудит, політики використання та оцінка ризиків.

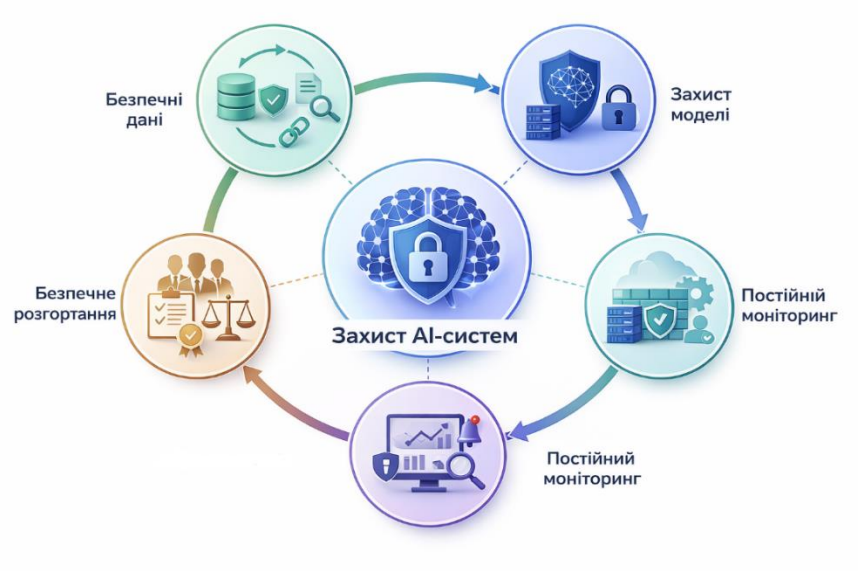

Рис. 1. П’ять ключових напрямів захисту AI-систем

Запропоновані напрями захисту мають застосовуватися не ізольовано, а як взаємопов’язана система заходів, у якій безпечні дані, захист моделі, безпечне розгортання, моніторинг і AI governance підсилюють один одного. Наприклад, навіть належний захист моделі не гарантує безпеки без контролю джерел да них, а ефективний моніторинг не дає повного результату без визначених політик реагування та відповідальності. Саме комплексне поєднання технічних і орга нізаційних механізмів дозволяє зменшити ризик компрометації AI-системи та підвищити її стійкість до сучасних атак.

Окреме значення має моніторинг уже після запуску системи. Навіть проте стована модель може демонструвати небажані реакції в нових умовах, помиля тися на нетипових запитах або ставати об’єктом зловживань. Саме тому захист ШІ не завершується на етапі розробки. Потрібні постійне спостереження за по ведінкою моделі, фіксація подій, контроль змін і готовність до реагування на інциденти. Для цього доцільно застосовувати threat modeling та AI red teaming, які дозволяють моделювати реальні сценарії атак і перевіряти стійкість системина практиці.

Неменш важливим є організаційний рівень захисту. Упровадження AIgovernance означає, що організація має чітко визначити, хто відповідає за вибір моделі, джерела даних, безпечне розгортання, аудит, оновлення та реагування на інци денти. Без такого розподілу відповідальності навіть технічно захищена система залишається вразливою через слабкі процедури контролю.

Додатковим важливим аспектом є забезпечення безпеки взаємодії AI-систем із зовнішніми користувачами та сервісами. У сучасних умовах моделі штучного інтелекту часто інтегруються в складні екосистеми, де вони обробляють запити з різних джерел, включаючи вебінтерфейси, мобільні застосунки та API. Це підвищує ризик несанкціонованого доступу, витоку інформації та зловживання функціональністю системи. Для мінімізації таких ризиків необхідно впроваджувати механізми автентифікації та авторизації, обмеження швидкості запитів (rate limiting), фільтрацію вхідних даних і контроль вихідної інформації. Окрім цього, важливо застосовувати принципи zero trust, за яких жоден запит не вважається безпечним за замовчуванням, а кожна взаємодія перевіряється на відповідність політикам безпеки. Такий підхід дозволяє значно знизити ризик експлуатації вразливостей та підвищити загальний рівень захищеності AI-систем.

Отже, захист самих систем штучного інтелекту є окремим і важливим на прямом сучасної кібербезпеки. Ефективний підхід до нього повинен охоплювати весь життєвий цикл

системи: дані, модель, середовище розгортання, моніторинг і управління ризиками. Поєднання безпечної розробки, контролю доступу, пе ревірки інтеграцій, AI red teaming і механізмів governance дає змогу зменшити площу атаки, підвищити стійкість моделей і забезпечити надійне використання ШІ в інформаційних системах.

З урахуванням швидкого поширення AI-рішень у різних сферах, питання їх захисту набуває не лише технічного, а й стратегічного значення. Надійна AI-система повинна проєктуватися з урахуванням принципів безпеки на всіх етапах життєвого циклу: від підготовки даних і навчання моделі до її впровадження, супроводу та оновлення. Лише за умови системного підходу, який поєднує технічні засоби захисту, постійний контроль, аудит і чіткі організаційні правила, можна забезпечити стабільну, передбачувану та безпечну роботу штучного інтелекту в сучасних інформаційних системах та зменшити ризики несанкціонованого впливу на їх функціонування.

Перелік посилань:

1. Artificial Intelligence Risk Management Framework (AI RMF 1.0) [Електронний ресурс] // National Institute of Standards and Technology (NIST). − Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (дата звернення: 12.04.2026).
2. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile [Електронний ресурс] // National Institute of Standards and Technology (NIST). − Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf (дата звернення: 14.04.2026).
3. OWASP Top 10 for Large Language Model Applications 2025 [Електронний ресурс] // OWASP Foundation. − Режим доступу: https://genai.owasp.org/llm-top-10/ (дата звернення: 13.04.2026).
4. Guidelines for Secure AI System Development [Електронний ресурс] // National Cyber Security Centre. − Режим доступу: https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development (дата звернення: 15.04.2026).

Поремський Я.С.
студент групи БСДМ-51, ННІКБЗІ ДУІКТ,
Київ, Україна

ВИКОРИСТАННЯ AI-АГЕНТІВ У СИСТЕМАХ SOC ДЛЯ ВИЯВЛЕННЯ ТА РЕАГУВАННЯ НА КІБЕРЗАГРОЗИ

Ефективність SOC визначається здатністю своєчасно виявляти інциденти та ініціювати обґрунтоване реагування в умовах високої щільності подій і дефіциту людських ресурсів. AI-агенти забезпечують новий рівень автоматизації, поєднуючи багатокроковий аналіз, контекстне збагачення та виконання дій через інструменти SOC із контролем доступів і журналюванням. Цей підхід дозволяє зменшити навантаження від алерт-шуму, підвищити якість тріажу й кореляції та скоротити час до стримування загрози, за умови належного управління ризиками автономності.

Ключові слова: SOC, AI-агенти, виявлення загроз, реагування на інциденти, SIEM, EDR.

Цифрова трансформація підприємств супроводжується зростанням обсягів даних, ускладненням інфраструктури та підвищенням вимог до швидкості реагування на кіберзагрози. За результатами дослідження Vectra AI [1], у 2026 році команди SOC отримують у середньому близько 3000 алертів щоденно, з яких приблизно 63% залишаються необробленими через перевантаження та обмежені ресурси. Це призводить до високого операційного навантаження, ознак вигорання фахівців і зниження ефективності виявлення

загроз. У таких умовах зростає потреба у впровадженні інтелектуальних систем, здатних автоматизувати обробку подій і підтримувати прийняття рішень на рівні SOC.

Класичний цикл SOC умовно охоплює збір подій, нормалізацію, кореляцію, тріаж, розслідування, реагування та післяінцидентний аналіз. У традиційних реалізаціях автоматизація концентрується навколо правил (use-cases), статичних кореляцій і плейбуків, які запускаються за формальними умовами. Такий підхід добре працює для повторюваних сценаріїв, але погано масштабується у ситуаціях, де потрібні контекст, гіпотезо-орієнтоване розслідування та гнучка послідовність дій.

На відміну від простої автоматизації на основі правил, характерної для традиційних систем безпеки, agentic AI здатний оркеструвати роботу кількох інструментів, інтегрувати контекстну інформацію з різних джерел і підтримувати процес прийняття рішень шляхом обробки неструктурованих даних [2]. Водночас такі системи зазвичай функціонують під наглядом людини або в межах попередньо налаштованих політик, а не як повністю автономні механізми навчання та управління у виробничому середовищі.

Agentic AI використовує здатність динамічно навчатися на основі середовища. Він підсилює процеси кібербезпеки за рахунок безперервного моніторингу та реагування на загрози в режимі реального часу, автоматизації повторюваних завдань SOC із мінімальним втручанням людини та надання контекстно обґрунтованої підтримки прийняття рішень. На рис. 1 представлена архітектурна схема інтеграції AI-агентів із SOC (Model Registry, AI Inference Service, процеси детекції та реагування).

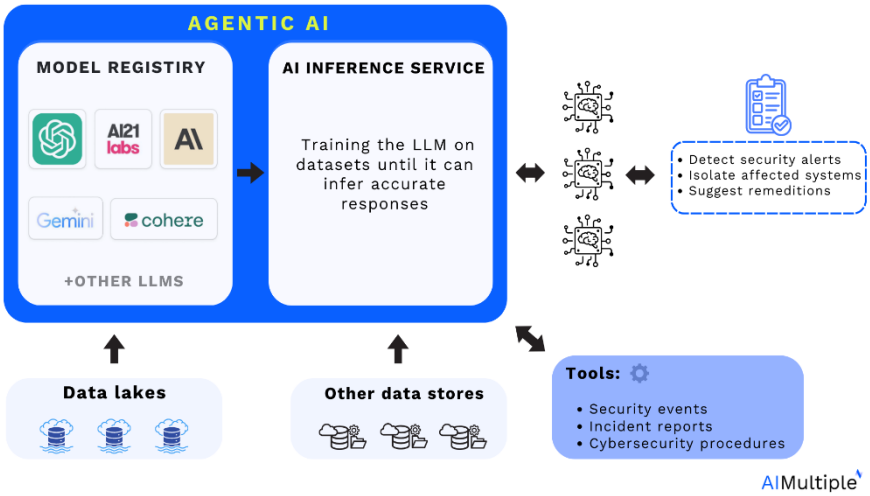

Рис. 1. Архітектура інтеграції AI-агентів із SOC

З точки зору операцій SOC, функціональні можливості agentic AI доцільно розглядати через основні етапи обробки інцидентів, незалежно від класичної моделі поділу на рівні аналітиків (Tier 1–Tier 3). У цьому контексті умовно можна виділити три ключові напрями застосування агентів.

Перший напрям пов’язаний із зменшенням навантаження від алертів і підвищення якості тріажу: AI-агенти виконують класифікацію подій, дедуплікацію, групування пов’язаних алертів і їх збагачення контекстною інформацією, що дозволяє зменшити шум і підвищити точність пріоритизації загроз.

Другий напрям охоплює автоматизацію процесів розслідування, зокрема збір і кореляцію даних із різних джерел (SIEM, EDR, систем управління ідентифікацією), формування контексту інциденту та підтримку прийняття рішень аналітиками).

Третій напрям пов’язаний із реагуванням на інциденти, де AI-агенти можуть ініціювати або виконувати визначені дії відповідно до політик безпеки, зокрема ізоляцію скомпрометованих систем, блокування доступу або запуск сценаріїв стримування, зазвичай під контролем людини.

Основні можливості Agentic AI:

Інтелектуальний тріаж і збагачення алертів. Agentic-системи здатні класифікувати та пріоритизувати алерти, зменшуючи рівень шуму й допомагаючи аналітикам SOC зосередитися на дійсно значущих загрозах.

Автоматизована підтримка розслідувань. Такі системи можуть збирати контекстну інформацію (зокрема дані кіберрозвідки та результати кореляції журналів) і узагальнювати отримані результати для подальшого аналізу людиною.

Стримування загроз і виконання плейбуків. Agentic AI здатний виконувати дії з реагування, наприклад ізоляцію скомпрометованих вузлів або застосування обмежень доступу відповідно до автоматизованих сценаріїв, із дотриманням політик управління та під контролем людини.

Підтримка Threat Hunting. Системи допомагають аналітикам шляхом кореляції індикаторів компрометації (IOC) між різними джерелами даних і формування гіпотез для подальшого розслідування, хоча остаточна інтерпретація залишається за фахівцем.

Аналіз і пріоритизація вразливостей. AI-системи дозволяють масштабовано аналізувати вразливості та оцінювати їх критичність для ефективного розподілу ресурсів.

Використання agentic AI у кібербезпеці супроводжується низкою суттєвих викликів [3], пов’язаних із автономністю таких систем та їх інтеграцією у корпоративну інфраструктуру. Однією з ключових проблем є розширення поверхні атаки, оскільки кожен AI-агент виступає окремою точкою доступу, яку може бути використано для компрометації систем . Додаткові ризики виникають через доступ агентів до великого обсягу даних і можливість прийняття рішень без повного контексту, що може призводити до помилкових або небезпечних дій. Окрему загрозу становлять атаки на самі механізми функціонування агентів, зокрема prompt injection та отруєння даних, які дозволяють маніпулювати поведінкою системи. Крім того, складність багатокомпонентних агентних систем ускладнює контроль і аудит, що підвищує ризик втрати прозорості та керованості процесів безпеки.

З метою мінімізації зазначених ризиків впровадження agentic AI у SOC доцільно здійснювати поетапно, починаючи з режиму «read-only», у якому агенти виконують аналіз і збагачення даних без впливу на інфраструктуру. Наступним етапом є використання режиму рекомендацій із поясненням прийнятих рішень, після чого можливий перехід до обмеженої автоматизації низькоризикових дій. Ключовим принципом залишається збереження «людини в контурі» для критичних операцій, що дозволяє контролювати автономність систем. Також необхідно впроваджувати принцип найменших привілеїв, забезпечувати журналювання дій агентів і контроль доступу до даних, а також приділяти особливу увагу якості телеметрії та інтеграцій, оскільки саме вони визначають ефективність і безпечність роботи AI-агентів. Такий підхід дозволяє досягти балансу між автоматизацією та керованістю, зменшуючи ризики та підвищуючи ефективність SOC.

Перелік посилань:

1. Vectra AI. New Vectra AI Research Finds Cyber Resilience Lagging in the AI Era. URL: https://www.vectra.ai/about/news/new-vectra-ai-research-finds-cyber-resilience-lagging-in-the-ai-era (дата звернення: 08.04.2026).
2. AIMultiple. Agentic AI in Cybersecurity. URL: https://aimultiple.com/agentic-ai-cybersecurity (дата звернення: 08.04.2026).
3. N-iX. Exploring Agentic AI Cybersecurity: Top Use Cases and Challenges. URL: https://www.n-ix.com/agentic-ai-cybersecurity/ (дата звернення: 08.04.2026).

Савчук О.С.
Студент групи БСД-42, ННІКБЗІ ДУІКТ,
Київ, Україна

СИСТЕМА ПРОТИДІЇ ВИТОКУ КОНФІДЕНЦІЙНОЇ ІНФОРМАЦІЇ В ОРГАНІЗАЦІЇ НА БАЗІ TERAMIND

У сучасних умовах цифровізації бізнес-процесів захист інформаційних активів став основою економічної стабільності будь-якої організації. Традиційно основна увага фахівців з кібербезпеки приділялася зовнішнім загрозам — хакерським атакам, вірусам та мережевим зломів. Проте статистика останніх років свідчить, що найбільш руйнівні наслідки мають саме внутрішні загрози, ініційовані персоналом. Людський фактор залишається "найслабшою ланкою", оскільки працівники мають легітимний доступ до систем, що дозволяє їм обходити класичні бар'єри захисту. Тому розробка та впровадження комплексних систем моніторингу, таких як Teramind, є критично необхідним кроком для сучасного підприємства.

Ключові слова: Кібербезпека, внутрішні загрози, інсайдерські ризики, витік конфіденційної інформації, система моніторингу, Teramind, DLP**.**

Питання внутрішніх загроз розглядаються в межах концепції DLP (*Data Loss Prevention* — запобігання витоку даних) та розширеного підходу UAM (*User Activity Monitoring* — моніторинг активності користувачів). Платформа Teramind виділяється серед конкурентів завдяки поєднанню технічного контролю за передачею файлів та інтелектуального аналізу поведінки працівників.

Внутрішня загроза в контексті кібербезпеки — це не лише навмисна крадіжка даних. Ми класифікуємо її за трьома основними напрямками [2]:

1. Халатність: помилкові дії працівників (відправка листа не тому адресату, використання слабких паролів).
2. Зловмисність: свідоме копіювання бази клієнтів або інтелектуальної власності для перепродажу.
3. Компрометація: використання хакерами облікових даних реального співробітника.

Головна складність виявлення таких загроз полягає в тому, що дії порушника на перший погляд не відрізняються від повсякденної робочої діяльності. Саме цю проблему вирішує впровадження системи Teramind [1].

Функціональні можливості Teramind дозволяють організації здійснювати повний візуальний та технічний контроль. Система автоматично фіксує всі дії користувача: від натискання клавіш та рухів миші до використання вебсайтів та месенджерів. Важливою перевагою є функція відеозапису екрана в реальному часі, що дозволяє фахівцю з безпеки побачити контекст події. Якщо працівник намагається скопіювати конфіденційний документ на особисту флеш-накопичувач або завантажити його у хмарне сховище, система негайно блокує дію та сповіщає адміністратора [3].

Логіка функціонування платформи Teramind базується на замкненому циклі управління безпекою (рис. 1), який охоплює чотири критичні стадії: безперервний моніторинг (*Monitor*), інтелектуальне виявлення аномалій (*Detect*), адміністрування прав та доступів (*Manage*) та оперативне реагування на інциденти (*Respond*). Такий підхід дозволяє не лише фіксувати факт порушення, а й проводити ретроспективний аналіз для розслідування інцидентів (*Audit & Forensics*) та оптимізувати робочі процеси всередині організації.

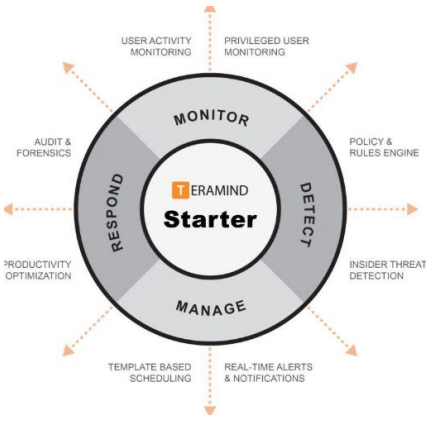

Рис. 1 Функціональна модель циклу протидії внутрішнім загрозам у рішенні Teramind

Окрім прямого блокування витоків (DLP), Teramind використовує алгоритми аналізу поведінки. Система створює "цифровий портрет" кожного працівника. Наприклад, якщо бухгалтер зазвичай працює лише з програмою 1С та Excel з 9:00 до 18:00, а раптом починає активність о 2-й ночі та заходить у системні налаштування мережі, система ідентифікує це як аномалію. Такий проактивний підхід дозволяє зупинити інцидент ще до того, як дані фактично залишать межі організації.

Впровадження такої системи також позитивно впливає на загальну дисципліну та продуктивність. Аналітичні звіти Teramind допомагають керівництву виявити неефективне використання робочого часу, що є додатковим аргументом для бізнесу при розрахунку окупності (ROI) інструментів безпеки.

Система Teramind є ефективним інструментом протидії внутрішнім загрозам, оскільки вона закриває "сліпі зони", які ігнорують звичайні антивіруси чи мережеві екрани. Завдяки автоматизації моніторингу та швидкому реагуванню на аномалії, організація значно знижує ризик втрати конфіденційної інформації. Проте варто пам'ятати, що технологія є лише частиною стратегії захисту; вона повинна працювати разом із чіткими політиками безпеки та регулярним навчанням персоналу правилам цифрової гігієни.

Перелік посилань:

1. Home | Teramind Knowledge Base. *Home | Teramind Knowledge Base*. URL: https://kb.teramind.co/en/ (дата звернення: 20.04.2026).

2. Що таке Внутрішня Загроза - Терміни та Визначення Кібербезпеки. *Keepsolid VPN Unlimited*. URL: https://www.vpnunlimited.com/ua/help/cybersecurity/insider-threat?srsltid=AfmBOopSGrSSlaisTBR5kmZVgsRe2h-SLhAly1xtQV4MTmNfzsmQKWjA (дата звернення: 20.04.2026).

3. Teramind UAM | Програмне забезпечення для моніторингу працівників. *Softprom – IT Distributor | Cyber Security, Cloud, IT Systems, CCTV, CAD*. URL: https://softprom.com/ua/vendor/teramind/product/teramind-uam-programne-zabezpechennya-dlya-monitoringu-pratsivnikiv (дата звернення: 20.04.2026).

Севертока О.А.
студент групи БСДМ-51, ННІКБЗІ ДУІКТ, Київ, Україна

БЕЗПЕКА ВПРОВАДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ В SOC: РИЗИКИ, КОНТРОЛІ ТА ПРАКТИЧНА МОДЕЛЬ ЗАХИСТУ

Анотація. У тезах розглянуто безпечне впровадження інструментів штучного інтелекту в роботу центру операцій безпеки. Визначено ключові ризики для даних, моделей та процесів реагування на інциденти, а також запропоновано практичну модель контролів, що поєднує управління ризиками, людську верифікацію, журналювання, оцінювання якості відповідей і захист від маніпуляцій. Підхід орієнтований на підвищення швидкості аналізу подій без втрати довіри до рішень аналітика.

Ключові слова: штучний інтелект, SOC, кібербезпека, інцидент, модель загроз, управління ризиками.

Центр операцій безпеки (SOC) поступово переходить від класичного моніторингу журналів до інтелектуального аналізу великих потоків подій. У такому середовищі штучний інтелект може виконувати попередню класифікацію сповіщень, групувати схожі інциденти, пояснювати ланцюги атаки природною мовою та допомагати аналітику швидше сформувати гіпотезу. Проте автоматизація не усуває відповідальність людини: вона переносить частину ризику з ручної обробки подій на якість даних, модель, політики доступу та процедури перевірки результатів.

Актуальність теми посилюється тим, що сучасний ландшафт загроз характеризується масовістю інцидентів, повторним використанням інструментів атакувальників і залежністю організацій від цифрових сервісів. За даними ENISA Threat Landscape 2025, аналіз охопив 4875 інцидентів за період з 1 липня 2024 року до 30 червня 2025 року, що підтверджує потребу в більш швидкому та контекстному прийнятті рішень у кіберзахисті [4]. SOC у таких умовах має не лише виявляти події, а й підтримувати безперервне управління ризиком.

Основна проблема впровадження ШІ в SOC полягає у суперечності між швидкістю автоматизованої допомоги та необхідністю довіряти результату. Модель може помилково знизити пріоритет критичного сповіщення, некоректно узагальнити контекст інциденту або відтворити конфіденційні фрагменти журналів у відповіді. Крім того, атакувальник може навмисно формувати події так, щоб вплинути на пояснення моделі, приховати ознаки компрометації або спровокувати аналітика на неправильну дію. Тому ШІ у SOC треба розглядати не як самостійного суб'єкта реагування, а як контрольований інструмент підтримки рішень.

Доцільно виділити три групи ризиків. Перша група пов'язана з даними: витік журналів, персональних даних, індикаторів компрометації та внутрішньої інформації про інфраструктуру. Друга група стосується моделі: отруєння даних, підміна контексту, prompt injection, помилкові висновки та залежність від зовнішніх постачальників. Третя група охоплює операційний процес: надмірну довіру аналітика до відповіді, відсутність трасування рішень і нечіткі правила ескалації. Саме поєднання цих груп створює найбільшу небезпеку для корпоративних інформаційних систем.

Практична модель безпечного впровадження має будуватися за логікою управління ризиками. NIST Cybersecurity Framework 2.0 пропонує функції Govern, Identify, Protect, Detect, Respond і Recover, які зручно застосувати до життєвого циклу інструментів ШІ в SOC [1]. На рівні Govern визначаються власники процесу, межі використання ШІ та критерії прийнятності ризику. На рівні Identify описуються джерела даних, типи подій, залежності від сервісів і сценарії можливих зловживань. Функція Protect охоплює контроль доступу, маскування даних, ізоляцію середовища та заборону передавання критичних секретів у модель.

На рівні Detect модель використовується для кореляції подій, але її результат має перевірятися технічними правилами SIEM, EDR або SOAR. На рівні Respond потрібно встановити принцип human-in-the-loop: модель може запропонувати план дій, однак блокування облікового запису, ізоляція вузла або повідомлення керівництва виконуються лише після підтвердження відповідальним аналітиком. Рекомендації NIST SP 800-61 Rev. 3 підкреслюють, що реагування на інциденти має бути інтегроване в загальне управління кіберризиками, а не існувати окремою технічною процедурою [2].

Окрему увагу слід приділити захисту самої системи ШІ. NIST AI RMF 1.0 розглядає довіреність ШІ через керування, картування, вимірювання та управління ризиками протягом життєвого циклу [3]. Для SOC це означає потребу регулярно тестувати модель на стійкість до маніпулятивних запитів, вимірювати частку помилкових рекомендацій, зберігати версії промптів і політик, а також фіксувати джерела даних, на основі яких сформовано відповідь. Якщо модель не може пояснити підстави висновку, такий результат повинен мати нижчий рівень довіри.

Для практичного застосування можна запропонувати мінімальний набір контролів: псевдонімізація чутливих полів перед передаванням у модель; розмежування ролей між аналітиком, адміністратором платформи та власником ризику; журналювання всіх запитів і відповідей; тестові набори інцидентів для перевірки якості; фільтри проти prompt injection; заборона автоматичного виконання руйнівних дій; регулярний перегляд помилкових спрацювань і пропусків. Такий набір не ускладнює роботу SOC, але створює прозору основу для аудиту.

Важливим є також організаційний аспект. Якщо аналітики не розуміють обмежень ШІ, вони можуть сприймати відповідь моделі як остаточний висновок. Тому навчання персоналу має включати приклади помилкових пояснень, атак на контекст, витоків даних і сценаріїв неправильної ескалації. Корисною практикою є правило подвійного підтвердження для інцидентів високого рівня критичності, а також післяінцидентний аналіз, у якому оцінюється не лише дія людини, а й якість рекомендацій ШІ.

Отже, штучний інтелект може суттєво підвищити ефективність SOC, але лише за умови контрольованого впровадження. Найкращий результат досягається тоді, коли модель допомагає структурувати інформацію, а не замінює професійне судження аналітика. Безпека такого підходу залежить від поєднання технічних контролів, управління ризиками, прозорого журналювання і регулярного тестування. У перспективі саме захищені та перевірювані ШІ-рішення можуть стати основою більш зрілих, швидких і стійких центрів операцій безпеки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Pascoe C., Quinn S., Scarfone K. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. 2024. URL: https://doi.org/10.6028/NIST.CSWP.29 (дата звернення: 22.04.2026).
2. Nelson A., Rekhi S., Souppaya M., Scarfone K. Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile. NIST SP 800-61 Rev. 3. 2025. URL: https://doi.org/10.6028/NIST.SP.800-61r3 (дата звернення: 22.04.2026).
3. Tabassi E. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. 2023. URL: https://doi.org/10.6028/NIST.AI.100-1 (дата звернення: 22.04.2026).
4. ENISA Threat Landscape 2025. European Union Agency for Cybersecurity. 2025. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025 (дата звернення: 22.04.2026).

Снітка Н.В.
студент групи КНД-42, ННІІТ ДУІКТ,
Київ, Україна

ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ВЕБ-ЗАСТОСУНКУ НА БАЗІ LARAVEL

У тезах розглянуто ключові аспекти забезпечення безпеки веб-застосунку, розробленого на базі фреймворку Laravel. Визначено основні кіберзагрози для сучасних веб-рішень, зокрема порушення контролю доступу, ін'єкційні атаки, XSS та помилкові налаштування безпеки. Окреслено вбудовані механізми Laravel, які сприяють захисту даних, безпечній автентифікації, авторизації, валідації запитів і стійкій експлуатації веб-систем.

Ключові слова: веб-застосунок, Laravel, кібербезпека, автентифікація, авторизація, захист даних

Умови цифрової трансформації сприяють стрімкому поширенню веб-застосунків у бізнесі, освіті, державному управлінні та сфері послуг. Такі системи забезпечують взаємодію з користувачами, оброблення персональних даних, керування контентом і виконання критичних бізнес-операцій. Водночас саме веб-застосунки залишаються однією з найпоширеніших цілей кіберзловмисників, оскільки помилки в логіці доступу, обробленні даних або конфігурації середовища можуть призводити до витоку інформації, несанкціонованих змін і порушення доступності сервісу.

Одними з найбільш актуальних ризиків для сучасних веб-рішень є порушення контролю доступу, ін'єкційні атаки, міжсайтовий скриптинг, помилкові налаштування безпеки та використання вразливих компонентів. Для веб-застосунку ці загрози мають практичний характер: зловмисник може намагатися отримати доступ до службових розділів, змінити дані через підроблені запити, ввести шкідливий код у форми або використати незахищені API-методи. Тому питання безпеки має враховуватися не після завершення розробки, а на всіх етапах життєвого циклу програмного продукту.

Фреймворк Laravel є зручною основою для створення захищених веб-застосунків, оскільки містить вбудовані механізми, що підтримують принципи secure development. Насамперед ідеться про засоби автентифікації та авторизації, які дають змогу реалізувати рольову модель доступу й чітко розмежувати права користувачів. Для перевірки дозволів доцільно використовувати policies і gates, що зменшує ризик ескалації привілеїв і несанкціонованого виконання дій у системі.

Важливим елементом захисту є також коректна обробка вхідних даних. Laravel підтримує серверну валідацію запитів, роботу з ORM Eloquent і автоматичне екранування в шаблонах Blade. У поєднанні з CSRF-захистом це дає змогу знизити ймовірність реалізації типових атак, пов'язаних із підміною запитів, SQL-ін'єкціями та XSS. Додатково доцільно застосовувати безпечне хешування паролів, обмеження кількості спроб входу, використання HTTPS, а для API - токенову автентифікацію з мінімально необхідними правами доступу.

Окрему увагу слід приділяти безпечній експлуатації застосунку. Навіть за наявності вбудованих засобів захисту фреймворку ризики зростають у разі неправильного налаштування серверного оточення, використання застарілих пакетів або відсутності моніторингу подій безпеки. Тому під час розробки доцільно впроваджувати регулярне оновлення залежностей, аудит логів, резервне копіювання, тестування ролей доступу та контроль конфігурації

середовища. Такий підхід забезпечує не лише захист даних, а й підвищує загальну стійкість веб-системи до сучасних кіберзагроз.

Крім базових механізмів захисту, важливо враховувати безпеку взаємодії веб-застосунку з зовнішніми сервісами та адміністративною частиною системи. У практичних проєктах це стосується інтеграції API, завантаження файлів, надсилання службових повідомлень і керування обліковими записами. Для зменшення ризиків доцільно використовувати принцип найменших привілеїв, перевірку типів і розмірів файлів, журналювання критичних дій, а також розділення прав звичайних користувачів і адміністраторів. Такий підхід дозволяє своєчасно виявляти підозрілу активність і знижує ймовірність компрометації системи через допоміжні модулі.

Отже, безпека веб-застосунку на базі Laravel повинна розглядатися як комплекс технічних і організаційних заходів, інтегрованих у весь процес розробки. Використання вбудованих механізмів фреймворку в поєднанні з дотриманням принципів безпечного програмування дає змогу істотно зменшити ризики компрометації системи. Це підтверджує доцільність застосування Laravel для створення сучасних веб-рішень, у яких вимоги до функціональності мають поєднуватися з належним рівнем кібербезпеки.

Список використаних джерел

1. OWASP Foundation. OWASP Top Ten Web Application Security Risks. URL: https://owasp.org/www-project-top-ten/ .

2. Laravel. Authentication. URL: https://laravel.com/docs/13.x/authentication .

3. Laravel. Authorization. URL: https://laravel.com/docs/13.x/authorization .

4. Laravel. CSRF Protection. URL: https://laravel.com/docs/13.x/csrf .

5. Laravel. Validation. URL: https://laravel.com/docs/13.x/validation .

6. Laravel. Hashing. URL: https://laravel.com/docs/13.x/hashing .

Талан Анна Владиславівна
Студентка групи БСДМ-52, ННІКБЗІ, Київ,
Україна

ПОБУДОВА СУЧАСНОГО SOC ТА МЕНЕДЖМЕНТ ІНЦИДЕНТІВ: АДАПТАЦІЯ ДО ФРЕЙМВОРКУ NIST CSF 2.0

Що ми розуміємо під ефективним Security Operations Center (SOC)? Це не просто набір інструментів моніторингу, а комплексна екосистема, яка об'єднує кваліфікований персонал, формалізовані процеси та передові технології для безперервного виявлення, аналізу та реагування на кіберзагрози в корпоративних інформаційних системах.

Зі зростанням складності кібератак традиційні реактивні підходи до захисту інфраструктури втрачають свою ефективність, вимагаючи від бізнесу переходу до проактивного менеджменту інформаційної безпеки.

Ключові слова: Security Operations Center (SOC), кібербезпека, NIST CSF 2.0, менеджмент інцидентів, SIEM, SOAR, управління ризиками, комплаєнс, захист даних.

У 2024 році Національний інститут стандартів і технологій США (NIST) випустив оновлену версію фреймворку кібербезпеки – NIST CSF 2.0. Головною інновацією стало додавання функції «Govern» (Управління), яка наголошує, що кібербезпека є не лише технічною проблемою, а й ключовим елементом корпоративного управління ризиками. Інтеграція цієї функції в архітектуру SOC дозволяє узгодити процеси розслідування інцидентів із загальною бізнес-стратегією компанії. [1]

Основою технологічного стеку сучасного SOC є системи класу SIEM (Security Information and Event Management) та SOAR (Security Orchestration, Automation, and Response). Вони дозволяють автоматизувати збір логів, агрегацію артефактів та запуск стандартизованих плейбуків реагування (Playbooks). Проте навіть найкращі інструменти неефективні без правильної категоризації векторів атак та оцінки впливу на бізнес. [2]

Для пріоритизації інцидентів на першій лінії (L1) SOC доцільно впроваджувати метрики кількісної оцінки ризику.

Рис. 1 – Архітектура реагування на кіберінциденти в екосистемі SOC

Особливої уваги під час розслідування інцидентів вимагає правовий аспект, зокрема дотримання вимог щодо захисту персональних даних. Узгодження внутрішніх політик SOC із європейським регламентом GDPR та актуальними вимогами українського законодавства (зокрема, щодо підзвітності та принципів обробки даних) є критично важливим для уникнення регуляторних штрафів після витоку інформації. [4]

Отже, побудова резильєнтного SOC вимагає системного підходу. Використання міжнародних фреймворків, таких як NIST CSF 2.0, у поєднанні з автоматизацією процесів реагування та чітким дотриманням законодавчих норм, забезпечує надійний захист корпоративних інформаційних систем від сучасних викликів.

Перелік посилань:

1. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. URL: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf.
2. Muniz J. (2021). The Modern Security Operations Center. Cisco Press.
3. Alqahtani A., Al-Makhadmeh Z. Automated Incident Response and Threat Intelligence in SOC. ResearchGate. URL: https://www.researchgate.net/publication/351234567_Automated_Incident_Response
4. GDPR and Incident Response: A comprehensive guide to compliance and reporting. IT Governance. URL: https://www.itgovernance.eu/blog/en/gdpr-incident-response-guide.

Тарасовська Владислава Валентинівна
студент факультету автоматизації і інформаційних технологій
Київський національний університет будівництва і архітектури
м. Київ, Україна
E-mail: hurianova.vlada@gmail.com

Науковий керівник:
Шабала Євгенія Євгенівна
кандидат технічних наук, доцент
доцент кафедри кібербезпеки та комп'ютерної інженерії
Київський національний університет будівництва і архітектури
м. Київ, Україна

ВИКОРИСТАННЯ OSINT У РОЗСЛІДУВАННІ КІБЕРЗЛОЧИНІВ: МЕТОДИ ТА ІНСТРУМЕНТИ

У тезі розглянуто технічні аспекти застосування OSINT (Open Source Intelligence) у розслідуванні кіберзлочинів, описано методи мережевої розвідки, пасивного та активного сканування, аналізу метаданих і моніторингу інфраструктури, а також охарактеризовано ключові інструменти та підходи до автоматизації збору й обробки даних з відкритих джерел. Особливу увагу приділено питанням ефективності використання OSINT-технологій для виявлення цифрових слідів, ідентифікації потенційних загроз та встановлення взаємозв'язків між об'єктами розслідування. Також проаналізовано переваги й обмеження застосування відкритих розвідувальних даних у контексті сучасної кібербезпеки.

Ключові слова: OSINT, мережева розвідка, кіберрозслідування, Shodan, Maltego, метадані, DNS, пасивна розвідка, автоматизація.

Вступ. Зростання кіберзлочинів (від APT-атак і ransomware до фішингових кампаній та зламів критичної інфраструктури) вимагає ефективних технічних засобів розслідування [1, розділ 1]. Одним із найпотужніших підходів є OSINT: систематичний збір, нормалізація та кореляція даних з публічно доступних джерел: масових реєстрів, соціальних платформ, пошукових індексів, без вразливостей та блокчейн-мереж.

Класифікація методів OSINT у контексті кіберрозслідувань

За характером взаємодії з цільовою системою OSINT-методи поділяються на пасивні та напів-пасивні. Пасивна розвідка не передбачає жодного звернення до інфраструктури суб'єкта: дослідник оперує виключно даними, вже проіндексованими третіми сторонами (пошуковиками, реєстрами, архівами) [1], однак напів-пасивна розвідка допускає мінімально помітні запити до публічних сервісів (WHOIS та DNS), що за інтенсивністю не відрізняються від звичайного мережевого трафіку [1].

З технічного погляду виокремлюють такі напрями, що дозволяють ідентифікувати цифрові сліди зловмисників:

- Мережева розвідка (Network Intelligence): аналіз IP-просторів, ASN-блоків, BGP-маршрутів, відкритих портів і банерів сервісів [3];
- Аналіз доменної інфраструктури: дослідження DNS-записів (A, MX, TXT, SPF, DKIM), сертифікатів TLS (Certificate Transparency Logs), WHOIS-історії та пов'язаних доменів [3];
- Аналіз метаданих: вилучення EXIF-даних із зображень, OLE-метаданих із документів Office, геолокаційних тегів та часових міток [3];

– Моніторинг соціальних мереж і форумів (SOCMINT): відстеження акаунтів, публікацій і цифрових слідів підозрюваних на відкритих платформах [3];
– Аналіз блокчейн-транзакцій: трасування потоків криптовалют через публічні реєстри для виявлення гаманців, пов'язаних із злочинною діяльністю [3].

Технічний арсенал OSINT-розслідувань

Ключовим інструментом мережевої розвідки є Shodan (пошукова система, що індексує банери відкритих портів мільярдів інтернет-вузлів), за допомогою якого аналітик може виявляти серверну інфраструктуру зловмисника, ідентифікувати версії програмного забезпечення та відомі CVE-вразливості, а також встановлювати зв'язки між окремими IP-адресами через спільні SSL-сертифікати або унікальні HTTP-заголовки. Разом із можливостями пошукових систем мережевої розвідки виникає потреба у застосуванні додаткових інструментів глибшого технічного аналізу.

Для комплексного аналізу мережевої інфраструктури застосовуються:

– DNSdumpster та SecurityTrails для побудови повної карти DNS-зон і виявлення субдоменів [4];
– Censys для глибокого аналізу TLS-сертифікатів і відкритих портів [4];
– VirusTotal – для перевірки репутації IP-адрес, доменів і файлових хешів через агрегацію даних понад 70 антивірусних рушіїв [4].

Центральне місце у графовому аналізі зв'язків займає Maltego - платформа, що дозволяє будувати візуальні графи залежностей між доменами, IP-адресами, e-mail-акаунтами, іменами осіб та організаціями, автоматично збагачуючи ці об'єкти даними через «трансформи» (API-інтеграції з десятками джерел) [2].

Наступним етапом OSINT-аналізу є автоматизований збір даних із відкритих джерел.

Автоматизацію процесу збору даних забезпечують спеціалізовані фреймворки:

– Recon-ng - модульна платформа на Python для систематичного збору OSINT-даних з підтримкою баз даних і звітності;
– SpiderFoot - автономний сканер, здатний агрегувати понад 200 джерел даних за мінімального втручання оператора;
– TheHarvester - утиліта для збору email-адрес, піддоменів і відкритих портів через пошукові системи і DNS-бруте.

Автоматизація та обробка даних

Окрему увагу приділено практичному OSINT-розслідуванню, що передбачає обробку великих масивів неструктурованих даних, що вимагає застосування автоматизованих конвеєрів та стандартів обміну даними. Саме тому інтеграція OSINT-даних у платформи threat intelligence дозволяє автоматично збагачувати індикатори компрометації (IoC) контекстом та будувати ідентифікацію кампаній за технікою MITRE ATT&CK.

Технічні виклики та обмеження

Ефективна протидія технікам анонімізації та швидкій ротації інфраструктури (fast-flux DNS) вимагає крос-платформного аналізу та виявлення спільних інфраструктурних артефактів [1]. Сучасні розслідування часто стикаються з проблемою «bulletproof-хостингу», що вимагає ще глибшої технічної експертизи та використання розширених пошукових можливостей [4].

Висновки

Проведений аналіз свідчить, що OSINT є невід'ємним компонентом сучасного технічного арсеналу кіберрозслідувань, ефективність процесу якого суттєво підвищується за умови комплексного використання спеціалізованих інструментів, таких як Shodan, Censys [4] та Maltego [2], а також через стандартизацію процесів у рамках фреймворку MITRE ATT&CK.

Список використаних джерел

1. Bazzell M. Open Source Intelligence Techniques. 9th ed. IntelTechniques Media, 2022. 512 p. URL: https://dokumen.pub/osint-techniques-resources-for-uncovering-online-information-1nbsped-9798345969250.html;

2. Maltego Technologies. Maltego Documentation. URL: https://docs.maltego.com/en/support/solutions/articles/15000019166-what-is-maltego-;
3. MITRE ATT&CK Framework. URL: https://attack.mitre.org;
4. Censys. Attack Surface Management. URL: https://50081908.fs1.hubspotusercontent-na1.net/hubfs/50081908/1-Pagers/CensysASMDatasheet.pdf.

Твердохліб Я.М.
студентка групи БСДМ-52, ННІКБЗІ ДУІКТ,
Київ, Україна

РОЛЬ SIEM-СИСТЕМ У ФУНКЦІОНУВАННІ SOC

Впровадження систем управління інформацією та подіями безпеки (SIEM) є невід'ємним елементом побудови ефективного Центру операцій безпеки (SOC), спрямованим на забезпечення безперервного моніторингу, своєчасне виявлення інцидентів та мінімізацію ризиків кіберзагроз. Концепція функціонування SOC базується на централізованому зборі, нормалізації та кореляції різнорідних подій безпеки, що неможливо реалізувати без технологічного ядра у вигляді SIEM. Такий підхід забезпечує аналітиків необхідним інструментарієм для розслідувань, проактивного пошуку загроз та часткової автоматизації реагування, формуючи основу сучасної корпоративної кібербезпеки.

Ключові слова: SOC, SIEM, кібербезпека, кореляція подій, виявлення інцидентів, реагування на інциденти, SOAR

Центр операцій безпеки (SOC) є функціональною одиницею кіберзахисту, що забезпечує безперервний моніторинг, виявлення та реагування на інциденти інформаційної безпеки. У інформаційних системах, які характеризуються високою складністю та значними обсягами розрізнених даних, що генеруються IT-інфраструктурою, критичну роль відіграють системи класу SIEM (Security Information and Event Management), які, власне, і є основою будь-якого SOC. Це набуває особливої актуальності в умовах постійного зростання складності та прихованості кіберзагроз.

Першочергова функція SIEM полягає у збиранні журналів подій із максимально широкого кола джерел: міжмережевих екранів, IDS/IPS, кінцевих точок, серверів, хмарних середовищ, баз даних та застосунків. Оскільки кожен вендор має власний формат логів, SIEM нормалізує всі вхідні дані до єдиної схеми. Це дозволяє аналітикам порівнювати події з абсолютно різних систем в єдиному інтерфейсі, не витрачаючи час на розбір форматів.

Друга, і найбільш критична функція SIEM у функціонуванні SOC полягає у здатності до перехресної кореляції журналів подій. Використовуючи спільні атрибути, система пов'язує розрізнені логи у цілісні сценарії атак, що дозволяє аналітикам миттєво отримувати сповіщення про критичні інциденти. У структурі SOC інструментарій SIEM діє як радар, що забезпечує раннє виявлення загроз у цифровому просторі [1, ст.36].

Жоден окремий засіб захисту не бачить повної картини атаки. Найнебезпечніші кібератаки – це не поодинокі події, а ланцюжки дій, розтягнуті в часі. Зловмисник може спочатку провести сканування портів, через кілька днів підібрати пароль до облікового запису, потім використати легітимний інструмент для бокового переміщення мережею, а потім запустити шифрувальник. Окремо взятий фаєрвол не побачить, що за скануванням послідував успішний логін. Окремо взята система автентифікації не знає, що після успішного логіну з'явилася підозріла активність на сусідньому сервері. SIEM аналізує всі події в єдиному потоці, застосовує задані правила кореляції і склеює розрізнені події в осмислений інцидент безпеки. Потужність кореляції подій полягає в тому, що вона може працювати з тисячами журналів з різних пристроїв, вибирати цю конкретну послідовність подій з вашого брандмауера, сервера Windows та сервера бази даних і сповіщати вас протягом кількох секунд. [2, ст.11] Без кореляції SOC перетворюється на чергову службу, яка реагує лише на очевидні сповіщення, але не може виявити складну цілеспрямовану атаку.

Третє призначення SIEM полягає у забезпеченні можливостей для розслідування інцидентів та проактивного пошуку загроз. У процесі реагування на інциденти SIEM

забезпечує аналітиків необхідною доказовою базою та інструментами для проведення глибинного розслідування (forensics). Завдяки накопиченню та збереженню історичних даних, а також можливостям швидкого пошуку й фільтрації за різними атрибутами, команда SOC здатна відтворити повну хронологію атаки, ідентифікувати початкову точку проникнення та оцінити масштаб і наслідки компрометації інформаційних систем.

Четверта функція, яка посилюється в останні роки, – це автоматизація реагування та інтеграція з SOAR. SIEM не лише виявляє загрози, але й може стати тригером для автоматичних дій. Після спрацювання кореляційного правила SIEM здатен надіслати команду на фаєрвол для блокування зловмисного IP, запустити скрипт для збору доказів на скомпрометованому сервері тощо. Це значно скорочує час між виявленням атаки і початком реагування, що критично важливо, оскільки сучасні атаки розвиваються лічені хвилини. Автоматизація, яку запускає SIEM, перетворює SOC із пасивного спостерігача на активну захисну систему. Водночас, інтеграція із SOAR-платформами дозволяє стандартизувати процеси реагування за допомогою плейбуків (playbooks), що зменшує ймовірність людської помилки. У результаті підвищується загальна ефективність роботи аналітиків і забезпечується більш швидке та узгоджене реагування на інциденти.

Таким чином, роль SIEM у функціонуванні SOC є критичною. Без неї центр кібербезпеки втрачає можливість ефективно агрегувати та аналізувати події, виявляти складні атаки, проводити розслідування й забезпечувати своєчасне реагування. Фактично, SOC без SIEM залишається лише формальною структурою без належної технологічної спроможності. Натомість належно впроваджена та налаштована SIEM-система формує основу дієздатного SOC, саме рівень її налаштування, повнота покриття джерел даних і якість кореляційних правил безпосередньо визначають здатність SOC протидіяти кіберзагрозам.

Використані джерела:

1. Bhatt, S., Manadhata, P. K., & Zomlot, L. (2014). The operational role of security information and event management systems. IEEE Security & Privacy

2. Vaidyanathan R., Austin T. Getting the best out of your SIEM.

Корж А.Ю.
студентка групи БСДМ-51, ННІКБЗІ ДУІКТ,
Київ, Україна

РОЛЬ USER ACTIVITY MONITORING У МІНІМІЗАЦІЇ РИЗИКІВ ЛЮДСЬКОГО ФАКТОРА В ЦИФРОВІЙ ІНФРАСТРУКТУРІ ОРГАНІЗАЦІЇ

Традиційні підходи до захисту інформаційного периметра, що базувалися на обмеженні зовнішнього доступу, трансформувалися у комплексні системи внутрішнього моніторингу. Подальший розвиток цифрової інфраструктури та зростання складності внутрішніх загроз зумовили необхідність застосування технологій User Activity Monitoring (UAM) як інструменту автоматизації аналізу дій персоналу, що підвищує ефективність виявлення інсайдерських ризиків у сучасних умовах..

Ключові слова: UAM, моніторинг активності, людський фактор, кібербезпека, внутрішні загрози, поведінковий аналіз.

Одним із ключових підходів до забезпечення цілісності інформаційного середовища є моніторинг активності користувачів (UAM), який ґрунтується на зборі та аналізі даних про дії суб'єктів усередині системи. Водночас масштаби корпоративних мереж та велика кількість операцій, що здійснюються персоналом щодня, посилюють проблему «інформаційного шуму», коли надлишок нерелевантних логів сповільнює виявлення реальних інцидентів.

У контексті кібербезпеки автоматизація моніторингу розглядається як чинник, що дозволяє швидко збирати, структурувати та візуалізувати дані про використання робочого часу, доступ до конфіденційних файлів та активність у мережі з меншим обсягом ручного втручання. Технологія UAM дозволяє трансформувати пасивне спостереження у проактивний захист шляхом впровадження механізмів виявлення аномалій [1].

Саме тому, з урахуванням викладеного вище, можна зробити висновок, що UAM як підхід, що базується на безперервному аудиті дій користувачів, має суттєву прикладну цінність для підрозділів кібербезпеки. Це забезпечує можливість підтримувати високий рівень ситуаційної обізнаності щодо внутрішнього стану захищеності системи. Ключовою перевагою тут є можливість ретроспективного аналізу (форензіки), що дозволяє відновити хронологію подій у разі виникнення інциденту.

Методи поведінкового аналізу в межах UAM доцільно розглядати як інструмент підтримки прийняття рішень: для класифікації ризикованих дій, виявлення прихованих каналів витоку даних та пріоритизації сигналів, які потребують негайної уваги аналітика. Представлена логіка узгоджується з підходом, у якому автоматизація за допомогою спеціалізованого ПЗ (наприклад, Teramind) забезпечує швидке виявлення відхилень від «базової лінії» поведінки, тоді як остаточна інтерпретація ризиків та впровадження управлінських рішень залишаються за фахівцем з інформаційної безпеки [2].

Ефективне впровадження систем моніторингу дозволяє не лише мінімізувати прямі збитки від зловмисних дій інсайдерів, а й суттєво знизити ризики, викликані халатністю або необізнаністю персоналу, формуючи культуру кібергігієни всередині організації.

Для забезпечення цілісності цифрової інфраструктури та мінімізації впливу людського фактора доцільно використовувати комплексні платформи, що реалізують повний цикл управління безпекою. На рисунку 1 представлено архітектуру рішення Teramind UAM, яка базується на поєднанні чотирьох фундаментальних етапів: моніторингу (*Monitor*), виявлення (*Detect*), управління (*Manage*) та оперативного реагування (*Respond*).

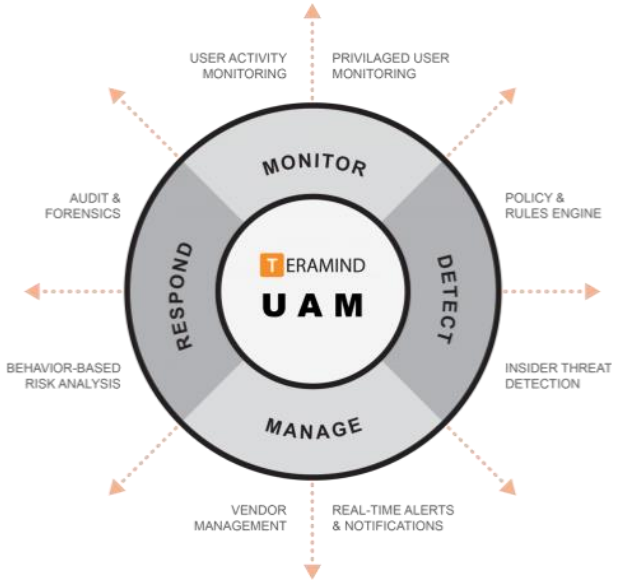

Рис. 1. Функціональна архітектура системи Teramind UAM: цикл безперервного моніторингу та механізми виявлення інсайдерських загроз

Представлена модель демонструє інтеграцію різнорідних інструментів контролю - від моніторингу привілейованих користувачів до інтелектуального аналізу ризиків на основі поведінкових аномалій. Такий підхід дозволяє трансформувати пасивне спостереження у проактивний захист, забезпечуючи підрозділи кібербезпеки необхідним контекстом для швидкого прийняття управлінських рішень

Перелік посилань:

1. What Is User Activity Monitoring (UAM)? Definition & Best Practices. *Network security platform for business | NordLayer*. URL: https://nordlayer.com/learn/threat-management/user-activity-monitoring/ (дата звернення: 14.04.2026).

2. Fullstory. *Intelligent Digital Experience Platform: AI-Powered Behavioral Analytics | Fullstory*. URL: https://www.fullstory.com/blog/what-is-behavior-analytics/ (дата звернення: 15.04.2026).

3. Teramind UAM privacy analysis - Teramind. *Teramind*. URL: https://www.teramind.co/thought-leadership/teramind-uam-privacy-analysis/ (дата звернення: 17.04.2026).

Тихонченко І.О.
студент групи ПД-41, ННІІТ, ДУІКТ,
Київ, Україна

ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ІНФОРМАЦІЇ У МОБІЛЬНІЙ 3D ГРІ ЗІ ЗМІНОЮ ФІЗИЧНИХ ВЛАСТИВОСТЕЙ ОБ'ЄКТА

У роботі розглянуто питання забезпечення інформаційної безпеки у мобільній 3D грі, розробленій із використанням ігрового рушія Unity. Особливу увагу приділено використанню хмарної платформи Firebase для автентифікації користувачів та збереження ігрових даних. Проаналізовано механізми контролю доступу на основі унікального ідентифікатора користувача (UID) та правил безпеки бази даних. Визначено основні загрози та способи їх мінімізації.

Ключові слова: кібербезпека, мобільні ігри, Firebase, автентифікація, захист даних.

У сучасних умовах стрімкого розвитку мобільних технологій комп'ютерні ігри стали важливою складовою цифрового середовища. Значна частина мобільних ігор використовує мережеві сервіси для збереження даних користувачів, синхронізації ігрового прогресу та забезпечення додаткових функціональних можливостей. У зв'язку з цим питання забезпечення інформаційної безпеки набуває важливого значення, оскільки витік або пошкодження даних може призвести до втрати ігрового прогресу, маніпуляції результатами або несанкціонованого доступу до облікових записів користувачів [1].

Дане дослідження присвячене розробці мобільної казуальної 3D гри зі зміною фізичних властивостей об'єкта. Гра створена з використанням ігрового рушія Unity. У процесі проходження рівнів користувач керує об'єктом, змінюючи його фізичні характеристики для подолання перешкод та досягнення фінішу рівня. Для забезпечення збереження ігрового прогресу та авторизації користувачів використовуються хмарні сервіси платформи Firebase.

У системі реалізовано механізм автентифікації користувачів за допомогою сервісу Firebase Authentication. Користувачі можуть входити до системи за допомогою анонімної авторизації (Anonymous Sign-In) або через обліковий запис Google. Після успішної автентифікації кожному користувачеві присвоюється унікальний ідентифікатор (UID), який використовується для доступу до його персональних даних. Такий підхід дозволяє забезпечити розмежування доступу між різними користувачами та запобігти несанкціонованому доступу до інформації інших гравців [2].

Для зберігання ігрових даних використовується база даних Firebase Realtime Database. Дані користувачів організовані у вузлі players, де кожному гравцю відповідає запис з унікальним UID. У базі зберігається інформація про кількість життів гравця (hearts), стан блокування реклами (adBlock), а також інформація про пройдені рівні та отримані бали. Система оцінювання рівнів передбачає нарахування балів за досягнення фінішу, за швидкість проходження рівня та за зібрані бонусні елементи.

Доступ до ігрових даних регулюється правилами безпеки бази даних. Ці правила визначають, які користувачі мають право читати або змінювати певні записи. Приклад реалізації правил доступу наведено на рис. 1.

```
{
  "rules": {
    "players": {
      "$uid": {
        ".read": "auth != null && auth.uid == $uid",
        ".write": "auth != null && auth.uid == $uid"
      }
    }
  }
}
```

Рис. 1. Правила доступу до бази даних у Firebase Realtime Database

У наведеному прикладі використовується перевірка автентифікації користувача та відповідності його унікального ідентифікатора (UID) запису в базі даних. Умова auth != null перевіряє, чи користувач успішно пройшов автентифікацію через сервіс Firebase Authentication. Умова auth.uid == $uid гарантує, що користувач може отримати доступ лише до власних даних у структурі players. Таким чином забезпечується ізоляція інформації між різними гравцями та запобігається можливість несанкціонованого перегляду або зміни чужих ігрових даних.

Однією з потенційних загроз для подібних систем є можливість підміни або маніпуляції ігровими результатами [3]. Такі дії можуть виконуватися шляхом модифікації клієнтського програмного забезпечення або підміни даних, що передаються до бази даних. Для зменшення таких ризиків застосовуються механізми перевірки коректності даних та обмеження доступу до критично важливих параметрів гри. Використання правил доступу Firebase дозволяє значно обмежити можливості сторонніх користувачів змінювати ігровий прогрес інших гравців.

Передача даних між клієнтським додатком та серверною інфраструктурою здійснюється через захищені мережеві протоколи, що забезпечують конфіденційність та цілісність інформації під час її передачі. Використання хмарної платформи дозволяє централізовано керувати політиками доступу до даних, а також оперативно змінювати правила безпеки без необхідності оновлення клієнтського програмного забезпечення.

Використання хмарних сервісів, механізмів автентифікації користувачів та правил доступу до бази даних дозволяє забезпечити захист ігрових даних користувачів та контроль доступу до інформації. Застосування таких механізмів є важливою складовою забезпечення інформаційної безпеки мобільних ігрових застосунків.

Перелік посилань:

1. Sommerville I. *Software Engineering.* — 10th ed. — Boston: Pearson Education, 2016. — Chapter 13, pp. 374. — Режим доступу: https://www.studyhalo.com/media/resources/resources/INF3705/Textbook/Software_Engineering_-_Ian_Sommerville.pdf
2. Firebase Security Rules and Firebase Authentication — Google Developers, 2026. — [Електронний ресурс]. — Режим доступу: https://firebase.google.com/docs/rules/rules-and-auth
3. Bella C. *Cybersecurity in Games.* Cybersecurity and Law, 2025, Nr 2 (14). — pp. 141–147. — Режим доступу: https://www.researchgate.net/publication/399018623_Cybersecurity_in_Games

Ткачов О.С.
Студент групи БСЗМ-61, ННІКБЗІ ДУІКТ,
Київ, Україна

ВИКЛИК СУЧАСНИМ ЗАГРОЗАМ, WI-FI КРИТИЧНИЙ ДЕМАСКУЮЧІЙ ФАКТОР НА ВІЙНІ

У сучасній війні Wi-Fi перетворився на один з головних демаскуючих технічних факторів. На превеликий жаль більшість користувачів сприймають його як безпечний інтерфейс для з'єднання з інтернетом, котрий майже завжди доступний та є у багатьох засобах зв'язку - від смартфона чи пульта керування дроном, до супутникового термінала (Starlink, та інших). Але сьогоденна реальність проста: Wi-Fi мережа, і пристрої з Wi-Fi «світяться» у сотні разів інтенсивніше будь-які супутникові тарілки. Вони дуже легко виявляються та локалізуються засобами радіоелектронної розвідки (РЕР), вони ще й залишають ***Wi-Fi слід***, який можна читати спеціальними системами аналізу даних. Для військових підрозділів і навіть цивільних користувачів у прифронтових районах це означає одне: ***не Wi-Fi забезпечує зв'язок, а Wi-Fi видає вас ворогу***. Саме він стає найбільш зручним маркером для виявлення позицій, визначення маршрутів і збору розвідувальної інформації. Розглянемо чому він становить

небезпеку, які інструменти збирають і використовують ці дані, та які практичні заходи радіо- й кібергігієни допоможуть мінімізувати ризики.

Що таке Wi-Fi слід?

Кожна Wi-Fi мережа та кожен пристрій, оснащений модулем Wi-Fi, постійно генерують набір сигналів, що поширюються далеко за межі приміщення або об'єкта. Саме ці сигнали утворюють так званий Wi-Fi слід (Wi-Fi fingerprint або trail) - своєрідний «цифровий запах», який можна виявляти, фіксувати, відстежувати й аналізувати.

Як проявляє себе Wi-Fi?

Усі Wi-Fi пристрої - від точок доступу й роутерів до смартфонів, смартгодинників, камер відеоспостереження та різноманітних «розумних» гаджетів - постійно передають певні дані у формі службових пакетів. Такі пакети передаються відкрито, тому їх може перехопити будь-який приймач у межах радіуса дії - від звичайного сканера на смартфоні до спеціалізованих засобів РЕР.

• **Beacon (маяк)**: кожна точка доступу через певні проміжки часу передає службові пакети, що містять інформацію про мережу - SSID (назву), BSSID (MAC-адресу точки доступу), стандарт зв'язку, підтримувані швидкості та інші параметри.

Probe-запити: клієнтські пристрої, зокрема смартфони, ноутбуки, планшети чи дрони, надсилають у радіоефір запити для пошуку відомих їм мереж. У таких запитах нерідко міститься перелік «улюблених» SSID, що дає змогу відстежити маршрут пересування або встановити власника пристрою.

• **Інші службові кадри**: кадри асоціації, авторизації та керування також передають метадані, які можуть бути використані для аналізу.

Отже, навіть за відсутності підключення до вашої мережі пристрої не стають **«невидимими»**. І сама мережа, і більшість пристроїв навколо все одно продовжують «проявлятися» в ефірі. Тепер ми розуміємо що перша складова Wi-Fi сліду - це службові пакети даних, які можна фіксувати та зберігати якимось пристроєм. Наприклад - вашим власним смартфоном.

Vendor telemetry та «розумні пристрої»

Wi-Fi слід формується не лише під час підключення до мережі. Сучасні смартфони й IoT-пристрої регулярно здійснюють фонове сканування та передають зібрану інформацію виробнику або стороннім SDK (Software Development Kit). У цих даних зазвичай фіксуються:

BSSID (унікальна MAC-адреса точки доступу),

SSID (символьна назва мережі),

рівень сигналу (RSSI),

час і координати (якщо ввімкнена геолокація),

ідентифікатори пристрою чи додатку.

Таким чином утворюється глобальна «карта Wi-Fi», яку поступово накопичують вендори - зокрема Google, Apple, рекламні платформи та інші. За наявності координат і рівнів сигналу застосування трилатерації для досить точного визначення місцеположення цільового пристрою не становить значної складності.

Різноманітні гаджети постійно оточують нас: смартфони й планшети, бортові системи автомобілів і відеореєстратори, розумні колонки та Smart TV, смартгодинники, контролери розумного дому, «розумні» павербанки, зарядні станції - і це далеко не повний перелік. Більшість таких пристроїв оснащені модулями Wi-Fi або Bluetooth і відповідним програмним забезпеченням, усе це збирає певні дані.

Частина цих пристроїв постійно працює в режимі точки доступу (AP), фактично самостійно поширюючи свій Wi-Fi слід. Однак найважливіше те, що більшість із них збирає та передає телеметрію і дані зі своїх модулів Wi-Fi та Bluetooth/BLE до різних баз даних. Фактично нас оточують мільйони гаджетів і пристроїв, які безперервно виконують роль Wi-Fi та Bluetooth/BLE сканерів.

Background scans і «Airplane mode»

Чимало людей помилково вважають, що достатньо просто вимкнути Wi-Fi та Bluetooth або активувати «Режим польоту», і на цьому збір чи передавання даних повністю припиняються. Насправді ж це не зовсім так.

Окрему загрозу для багатьох сучасних гаджетів становлять фонові сканування, які можуть працювати навіть тоді, коли користувач переконаний, що Wi-Fi вимкнений або пристрій перебуває в режимі польоту.

• **У налаштуваннях смартфонів** часто є функції на кшталт *Wi-Fi scanning* або *Bluetooth scanning для покращення геолокації*. Якщо їх не деактивувати, пристрій може й надалі збирати дані навіть за вимкненого Wi-Fi. Водночас у багатьох гаджетів такі параметри взагалі недоступні для ручного вимкнення.

• **Окремі дослідження та практичні спостереження** вказують на можливість так званого «фейкового airplane mode», коли на екрані модулі виглядають вимкненими, але фактично пристрій продовжує обмін даними. На жаль, з'ясувати, як саме реалізовано «Режим польоту» в конкретній моделі смартфона певного виробника, можна лише шляхом ретельного дослідження.

• **Універсальних публічних доказів**, що всі смартфони завжди поводяться саме так, немає. Водночас уже існують відкриті публікації про виявлені інциденти, тому повністю відкидати цей ризик не можна: він залежить від моделі пристрою, прошивки, налаштувань і наявних SDK.

Це означає, що навіть у режимі, який здається «тихим», пристрій може продовжувати залишати Wi-Fi слід, а також фактично виконувати роль своєрідного PEP-сканера. Для військових це є критично важливим чинником, що вимагає дисципліни та перевірки на рівні кожного конкретного пристрою.

На основі зібраних даних формуються великі бази (Big Data) геолокації Wi-Fi. Публічні приклади - WiGLE.net, Google Location Services, Mozilla MLS (архів), Skyhook. Але існують і комерційні або закриті бази, доступні тільки фахівцям.

Другий рівень - це вендорські телеметрії та сторонні SDK (third-party SDKs). Вендори ОС та пристроїв (Google, Apple, виробники смартфонів) регулярно отримують від гаджетів дані про навколишні мережі: BSSID, SSID, рівень сигналу, timestamp, іноді координати. Сторонні SDK у додатках (рекламні, аналітичні, антифрод-сервіси) теж збирають Wi-Fi дані й відправляють їх у власні бекенди. Користувач часто навіть не здогадується, що додаток «під капотом» передає інформацію про всі точки доступу довкола. Ці масиви даних значно більші й багатші, ніж будь-яка відкрита база. Той, хто має доступ до таких даних, у теорії може:

відстежувати маршрути конкретних пристроїв;

прив'язувати Wi-Fi точки до конкретних осіб чи організацій;

корелювати Wi-Fi записи з іншими ідентифікаторами (GPS, мобільна мережа, IDFA/AAID).

Це і є рівень коли аналітика на основі масивів телеметрії дає можливість будувати детальні картини активності та навіть визначати геолокацію з точністю, достатньою для практичного ураження.

Практичні ризики:

Локалізація позицій і техніки;

Відстеження руху й маршрутів (pattern-of-life, tracing);

Ідентифікація ролей/підрозділів через неймінг і метадані;

Компрометація каналів та систем через WiFi слід мережі чи пристрої;

Найбільша небезпека Wi-Fi полягає в тому, що він створює слід завжди - навіть коли користувач не підключається до мережі. Тому завдання гігієни полягає не у «повному захисті», а в мінімізації сліду. Це поєднання технічних налаштувань і дисципліни користувачів.

Базові правила

Пріоритет дроту - у бойових умовах краще відмовитися від Wi-Fi повністю й використовувати дротові підключення.

Мінімізація часу роботи Wi-Fi - якщо без Wi-Fi ніяк, його радше слід вмикати лише на короткі проміжки часу (time-boxing).

Мінімізація зони покриття. Зниження потужності передавача й екранування (сховані точки під землю, у бліндаж/укриття та т.п.).

Вимкнення непотрібних радіомодулів. Смартфони, ноутбуки, IoT-пристрої повинні мати Wi-Fi і Bluetooth вимкнені поза моментами реального використання.

Моніторинг радіо-гігієни WiFi/Bluetooth/BLE - якщо у підрозділа немає можливості скористатись відповідними засобами РЕР чи моніторингу, то відстеження оточення підрозділу хоча б простими та загальнодоступними програмами WiFi/Bluetooth сканерів для смартфонів – вже дуже ефективний крок.

Рекомендовані технічні заходи:

Сильна криптографія. Варто використовувати WPA3-SAE з обов’язково активованими PMF (802.11w). Якщо така можливість відсутня, слід застосовувати щонайменше WPA2. Використання відкритих мереж є неприпустимим.

Нейтральні SSID. Не рекомендується використовувати назви, які можуть вказувати на приналежність або призначення мережі, наприклад «HQ», «Starlink» чи «Army» номери підрозділів та частин. Доцільно обирати випадкові або нейтральні імена.

Регулярний reset і зміна BSSID. Можливості, способи та інструменти зміни BSSID залежать від конкретного пристрою. Наприклад, для Starlink таким способом може бути Factory reset. Фахівці зазвичай знають, де шукати описи та засоби для виконання таких змін. Регулярна зміна BSSID, а ще краще — його рандомізація, знижує ризик накопичення даних у відповідних базах.

MAC-рандомізація. Її слід вмикати на всіх клієнтських пристроях, водночас пам’ятаючи, що цей захід не є абсолютним, особливо під час асоціації з точкою доступу.

Оптимізація потужності. Використання нижчої потужності передавача дає змогу зменшити дальність поширення та помітність сигналу. Додаткову користь також можуть забезпечити спрямовані антени

Wi-Fi неможливо зробити повністю «невидимим», однак його можна зробити менш помітним, менш тривалим у прояві та безпечнішим. Це досягається завдяки поєднанню технічних налаштувань і належної організаційної дисципліни вашого підрозділу.

Перелік використаних джерел і літератури**:**

- **Victory Drones**. Застосування технологій в умовах війни : курс.
- **Skylinker** : офіційний сайт. URL: https://www.skylinker.io.
- Старосєк А., Бескрестнов С., Варфоломєєв Г. **Розвідка: анонімності не існує**. Київ, 2026.

Харькевич Д.О.
студент групи БСДМ-51, ННІЗІ ДУІКТ,
Київ, Україна

КІБЕРРОЗВІДКА ТА OSINT ЗА ДОПОМОГОЮ ШІ

У тезі розглядається роль штучного інтелекту (ШІ) у сучасній кіберрозвідці та роботі з відкритими даними (OSINT). Наводиться аналіз ключових технологій (наприклад, методів обробки природної мови, комп'ютерного зору, побудови графів зв'язків) та реальних інструментів (Maltego, Paliscope, ChatGPT, Babel X тощо) з практичними прикладами використання. Розглядаються два приклади: використання AI в українському контексті протидії дезінформації і виявлення кіберзагроз, а також збір доказів воєнних злочинів. Акцентовано увагу на етичних і правових обмеженнях (захист персональних даних, вимоги AI Act/EU, ланцюг доказів) і надано рекомендації практикам щодо впровадження OSINT+ШІ (контроль якості даних, підготовка кадрів, міжорганізаційне співробітництво).

Дані з відкритих джерел сьогодні ростуть експоненціально, вийшовши за межі звичних можливостей аналізу. Так, AI вже «стала центральною» для OSINT, дозволяючи за лічені години збирати обсяги інформації [1], на обробку яких раніше потрібно тижні. Наприклад, сучасні моделі кластеризації можуть автоматично групувати схожі повідомлення у соціальних мережах для виявлення нових загроз, суттєво знижуючи навантаження аналітиків. У цьому дослідженні деталізовано принципи використання ШІ в OSINT, огляд доступних інструментів та платформ, приклади застосування в практиці, а також правові й етичні обмеження такого підходу.

Сучасні OSINT-системи аналізують великі потоки даних з різних джерел; ШІ ілюструється технологічними графіками, що нагадують складні цифрові мережі. Інтеграція машинного навчання дає змогу виявляти «ховаються» закономірності: наприклад, на графіках зв'язків підсвічуються раніше непомітні стосунки між підозрюваними. ШІ у розвідці спрощує такі завдання, як перехресна верифікація інформації, виділення наслідків зворотного зв'язку, переклад і класифікація повідомлень. Однак важливо відзначити, що «універсальні» моделі (ChatGPT тощо) призначені для загальних завдань, і їх застосування в офіційних розслідуваннях вимагає додаткових перевірок і підкріплення доказовою базою.

AI-методи в OSINT охоплюють широкий спектр технологій. Серед ключових - обробка природної мови (NLP) [3] для аналізу тексту/медіа, машинне навчання для пошуку патернів і аномалій, та комп'ютерний зір [3] для аналізу зображень, відео та геопросторових даних. Наприклад, інструменти автоматичного розпізнавання осіб або автомобільних номерів дозволяють шукати підозрілих осіб на фотографіях зі зрозумілою для користувача інформацією.

Maltego - графічна платформа для візуалізації OSINT (метадані соціальних мереж, реєстри доменів, Breach API тощо). У 2026 р. Maltego анонсував модуль Monitor із AI-аналізом настроїв соціальних мереж для виявлення загроз у реальному часі [2].

Paliscope Find/Explore (Швеція) - комплекс для офлайн-розслідувань з вбудованим ШІ. Він шукає по численних наборах даних (логи чатів, документи, фото, відео) та використовує розпізнавання облич, голосу, номерів авто тощо. Інші приклади [2]:

Babel X (BabelStreet) - система глибокого аналізу соцмереж і форумів з NLP, що підтримує понад 200 мов (переклад, аналіз тональності) [2].

Recorded Future пропонує ІІ-підсилене зведення кіберзагроз з відкритих джерел (зокрема, інформацію з dark web) через власний граф «Intelligence Graph». [2]

ChatGPT/GPT-4 широко використовують для пошуку зв'язків у тексті, формулювання запитів по-багатомовному або перевірки гіпотез. Також популярні бібліотеки ML (TensorFlow, Hugging Face) для кастомних моделей, а **CrowdTangle, Meltwater** дають приклади SOCMINT (соціальні мережі) з елементами AI (аналіз охоплень, фейків) [2].

На практиці аналітик за допомогою ШІ може швидко сканувати великі набори повідомлень чи документів. Наприклад, на даному фото видно роботу співробітників у темній кімнаті з кількома моніторами, що демонструє умовний офіс OSINT-аналітики. Застосовуючи ML-моделі, вони можуть автоматично виділяти ключові події (наприклад, траєкторії ракет чи «гарячі точки» дезінформації) та будувати графи зв'язків між суб'єктами. Проте аналітики наголошують на необхідності «людино-центричного» контролю: результати ШІ мають перевірятися людиною і при необхідності коригуватися, а невизначеності - явно маркувати.

Виявлення фейкових наративів і доказів воєнних злочинів. В Україні кіберрозвідка масово використовує OSINT з ШІ для боротьби з інформаційними війнами. Наприклад, автоматичний аналіз відео й фото з передової з геолокацією ідентифікує зброю та транспортерів, групуючи їх за регіонами. ШІ-моделі розпізнають осіб (Mugshot, OpenCV) та аналізують контент Telegram-каналів на предмет фейків чи скоординованих атак.

Кіберзагрози та контррозвідка. Інший випадок - виявлення кіберзагроз через соцмережі й чат-боти: модель навчена шукає згадки про розробку злому чи продаж експлойтів. Один європейський центр кібербезпеки навчив нейромережу відсіювати трафік з відкритих джерел і знаходити колективні операції (наприклад, ботнети або DDoS-атаки), що піднімають попередження безпеки.

Застосування ШІ в OSINT підпадає під жорсткі вимоги щодо прав людини і захисту даних. По-перше, потрібно дотримуватися законності збору інформації. Наприклад, відповідно до GDPR та українських норм [4], збір персональних даних можливий лише на законних підставах і з мінімізацією обсягу інформації. Аналітики мають регулярно анонімізувати дані, якщо ідентифікація суб'єкта не потрібна. По-друге, AI-моделі самі по собі «чорні скриньки» - їх рішення слід документувати. Ланцюг збирання та обробки даних (chain of custody) має бути прозорим, щоб при потребі в суді можна було пояснити, звідки взяті докази.

Нещодавно ЄС ухвалив **AI Act** [4] (Регламент 2024/1689), що забороняє певні практики: наприклад, "безцільне скрапінгування" мереж чи відеоспостереження для створення баз розпізнавання облич. Також обмежено застосування біометрії у публічних місцях. У практиці це означає, що OSINT-системи не можуть легально збирати і розпізнавати людей на відео без конкретної правової підстави. Класифікація «високого ризику» означає, що системи з автоматизованою оцінкою доказів (наприклад, аналіз телефонних дзвінків) повинні відповідати жорстким стандартам, включно з перевіркою точності і відсутністю дискримінації. Нарешті, етика передбачає «людину в циклі»: аналізатор повинен оцінювати контекст і зміст даних AI, з огляду на можливі упередження (наприклад, негатив у відношенні певної групи).

- **Людино-центричність:** незважаючи на високий потенціал ШІ, рішення повинні проходити через перевірку фахівця. Використовуйте моделі для швидкого сортування даних, але самостійно оцінюйте їхні результати. Маркуйте невизначеності і уникайте «порожньої впевненості» алгоритмів.
- **Стандарти та прозорість:** впроваджуйте внутрішні політики щодо використання AI-OSINT, що включають протокол аудиту й логування всіх кроків аналізу. Наприклад, документуйте запити до API чи команди ChatGPT, фіксуйте метадані (час, модель, версія). Це критично для ланцюга доказів у суді.
- **Підготовка кадрів:** інвестуйте в навчання аналітиків - як у технічні навички (наприклад, написання ефективних промптів ChatGPT, робота з Python/ML-бібліотеками), так і в розуміння обмежень моделей. Спільно з юристами навчайте команду правовим аспектам (GDPR, AI Act) та етичному мисленню (розпізнавання упереджень).
- **Міжорганізаційна співпраця:** оскільки OSINT та AI стрімко розвиваються, важливо обмінюватися досвідом на конференціях і спільних платформах (хакатони, спільні бази знань). Ураховуйте, що метадані соціальних мереж чи телефонів можуть перетинати

кордони, тому співпраця з іноземними партнерами (наприклад, у межах ЄС) допомагає вчасно виявляти транснаціональні загрози.

Отже, ШІ кардинально змінює OSINT, перетворюючи ручну роботу аналітиків на гібрид з автоматичними алгоритмами. Його використання дає змогу швидше збирати і зіставляти розрізнені дані, але водночас породжує нові ризики - від юридичних колізій до спотворення результатів. Стратегічно важливо розвивати **структуру для довіри** до AI: включно з контролем якості даних, регулюванням алгоритмів та навчанням людей. Пріоритетним є баланс між ефективністю і відповідальністю.

Перелік посилань:

1. OSINT Framework // Open Source Intelligence Tools and Resources. [Електронний ресурс]. Режим доступу: https://osintframework.com
2. Maltego Technologies. Maltego Platform Overview // Maltego. [Електронний ресурс]. Режим доступу: https://www.maltego.com
3. Jurafsky D., Martin J. Speech and Language Processing. 3rd ed. // Stanford University. [Електронний ресурс]. Режим доступу: https://web.stanford.edu/~jurafsky/slp3/
4. Regulation (EU) 2024/1689 (Artificial Intelligence Act) // Official Journal of the European Union. [Електронний ресурс]. Режим доступу: https://eur-lex.europa.eu

Хорольський Костянтин Андрійович
студент групи БСДМ-52, ННІКБЗІ ДУІКТ, Київ, Україна

ФІШИНГ 2.0: ЯК ЗМІНИЛИСЬ АТАКИ З ПОЯВОЮ ШТУЧНОГО ІНТЕЛЕКТУ

Фішинг залишається найпоширенішим вектором кібератак у світі, проте з появою генеративного штучного інтелекту (ШІ) його характер зазнав якісної трансформації. За даними SlashNext, обсяг фішингових повідомлень зріс на 151 % від моменту запуску ChatGPT у листопаді 2022 року. Дослідження Гарвардського університету засвідчило, що AI-згенеровані фішингові листи досягають 54 % показника переходів — на рівні з листами, створеними досвідченими фахівцями з соціальної інженерії. Ці зміни зумовлюють необхідність перегляду традиційних підходів до захисту. У тезах здійснено огляд ключових напрямів впливу ШІ на фішингові атаки та запропоновано актуальні контрзаходи.

Ключові слова: фішинг, штучний інтелект, великі мовні моделі, дипфейк, соціальна інженерія, кібербезпека, голосовий фішинг, OSINT.

Генеративний ШІ як інструмент масштабування фішингу. Поява великих мовних моделей (LLM) фундаментально змінила економіку фішингових атак. Дослідники IBM продемонстрували, що ШІ здатний побудувати повноцінну фішингову кампанію за 5 хвилин і 5 промптів — завдання, на яке досвідчений спеціаліст витрачає близько 16 годин. За оцінками Hazell (2023), генерація 1 000 персоналізованих спір-фішингових листів через LLM коштує лише близько 10 доларів США, що знижує витрати атакуючих до 95 %.

Паралельно на підпільних форумах з'явились «чорноринкові» LLM без етичних обмежень. WormGPT (червень 2023 р.) — модель на базі GPT-J, натренована на шкідливому коді та шаблонах фішингу — генерує бездоганні BEC-листи (Business Email Compromise) за підпискою від 60 доларів. FraudGPT (липень 2023 р.) пропонує створення фішингових сторінок, шкідливого коду та SMS-повідомлень за 200 доларів на місяць. Наприкінці 2024 р. з'явився GhostGPT — Telegram-бот за 50 доларів на тиждень, який генерує поліморфні фішингові листи та підроблені сторінки входу.

Ключовою перевагою AI-згенерованого фішингу є відсутність граматичних і стилістичних помилок, які традиційно слугували індикатором атаки. За даними VIPRE Security, у 2024 р. 40 % BEC-атак уже були згенеровані штучним інтелектом. LLM усувають мовний бар'єр: зловмисник може писати рідною мовою, а модель створює бездоганний текст цільовою мовою із урахуванням корпоративного стилю жертви.

Дипфейки та голосовий фішинг — новий рівень імітації. Технології синтезу голосу та відео відкрили принципово новий вектор атак. Звіт CrowdStrike 2025 зафіксував зростання голосового фішингу (вішингу) на 442 % у другій половині 2024 року [2]. Сучасні інструменти клонування голосу (зокрема ElevenLabs) потребують лише 3 секунди аудіозразка для створення переконливої копії.

Найрезонанснішим інцидентом став випадок з Arup (Гонконг, січень 2024 р.): фінансовий працівник інженерної компанії отримав фішинговий лист нібито від CFO щодо «конфіденційної транзакції». Його запросили на відеоконференцію, де всі учасники були дипфейками керівників компанії, згенерованими з публічно доступних відео. Працівник здійснив 15 переказів на загальну суму 25,6 млн доларів. У липні 2024 р. аналогічна атака на Ferrari була відбита завдяки верифікаційному запитанню, на яке дипфейк не зміг відповісти. В Україні банк ПУМБ наприкінці 2024 — на початку 2025 р. зазнав атак із використанням AI-дипфейків, спрямованих на системи верифікації клієнтів[14].

Автоматизація розвідки та персоналізація атак LLM автоматизують весь ланцюг атаки: від збору OSINT (LinkedIn, соцмережі, прес-релізи) до генерації унікального повідомлення для кожної жертви. Дослідження Heiding et al. (2024) на 101 учаснику показало, що AI-автоматизована OSINT-розвідка створила точні профілі 88 % цілей, а показник переходів за фішинговими посиланнями становив 54 % — на 350 % більше, ніж у контрольній групі з типовим фішингом (12 %).

CERT-UA у звіті за І півріччя 2025 р. підтвердив використання ШІ російськими хакерськими групами для генерації фішингових повідомлень та шкідливого коду (зокрема PowerShell-скрипти у шкідливому ПЗ WRECKSTEEL групи UAC-0219). Загалом фішинг становив 27 % від 3 018 кіберінцидентів, зареєстрованих в Україні за цей період [1]. За оцінками експертів галузі, 9 з 10 кібератак проти України вже використовують елементи ШІ.

Традиційні поради «шукайте граматичні помилки» більше не працюють. Ефективний захист від AI-фішингу потребує комплексного підходу: впровадження AI-систем детекції фішингу (Claude 3.5 Sonnet досягає 97,25 % точності розпізнавання підозрілих листів); обов'язкове використання багатофакторної автентифікації (MFA) та апаратних ключів; запровадження протоколів верифікації фінансових операцій через зворотний дзвінок на відомий номер — саме це врятувало Ferrari; регулярні тренінги з кіберобізнаності з урахуванням AI-загроз. За даними IBM, організації з AI-автоматизованим захистом економлять у середньому 2,2 млн доларів на кожному інциденті витоку даних.

Генеративний ШІ трансформував фішинг зі «спаму з помилками» у високоточну, масштабовану та економічно ефективну зброю соціальної інженерії. Три ключові зміни — зниження вартості атак у десятки разів, зростання якості до рівня експертів та поява дипфейків як нового вектора — створюють безпрецедентні виклики для кібербезпеки. Гонка озброєнь між AI-атакою та AI-захистом уже почалась, і організації, що не інтегрують ШІ у свої системи безпеки, опиняться у програшній позиції. Для України, яка перебуває під постійним тиском кібератак з боку РФ, це питання набуває стратегічного значення.

Перелік посилань:

1. SlashNext. The State of Phishing 2024: Mid-Year Report. SlashNext, 2024. URL: https://slashnext.com/state-of-phishing-2024/

2. CrowdStrike. 2025 Global Threat Report. CrowdStrike, February 2025. URL: https://www.crowdstrike.com/global-threat-report/

3. CERT-UA / Держспецзв'язку. Аналітичний звіт: Російські кібероперації H1 2025. SSSCIP, 2025. URL: https://cip.gov.ua/

4. ПУМБ зазнав кібератак у 2024–2025 роках з використанням діпфейків на базі ШІ. Інтерфакс-Україна, 12 листопада 2025.

Tsarova Sofiia Valeriivna
student of group BSDM-51, ESI IS, SUICT,
Kyiv, Ukraine

ENHANCING EDR THREAT DETECTION CAPABILITIES VIA MACHINE LEARNING

Endpoint security is one of the crucial domains of cybersecurity that protects both companies and individuals from such risks as data breaches or work disruption by technically detecting, analyzing and mitigating threats. Tradional antiviruses, relying on analyzing known digital malware fingerprints, are not sufficient for protecting from modern threat landscape. As new Endpoint Detection and Response (EDR) systems are developed, machine learning is being integrated to address these gaps. This paper explores the application of supervised and unsupervised ML learning techniques, with a specific focus on the Isolation Forest algorithm for efficient endpoint anomalies identification.

Key words: EDR, machine learning, Isolation Forest.

Legacy endpoint security solutions were built around traditional signature threat detection, based on predefined rulesets and datasets of known indicators of compromise. While this approach is still present in modern EDR systems and is relevant for detecting well-documented malware, it becomes increasingly ineffective against zero-days attacks and widely employed techniques, such as fileless malware and living-off-the-land tools exploitation.

To address these vulnerabilities in defense, machine learning (ML) is now heavily integrated within EDR solutions. ML introduces new capabilities to endpoint detection & response systems that allow to detect, analyze and respond to previously unseen threats and sophisticated techniques:

- utilizing collected endpoints data (processes activity, network connections, memory usage, etc.) machine learning models are trained to create normal behavior baselines and detect anomalies more accurately;
- ability of continuous learning from new data let AI-based detection solutions to adapt to evolving attacks techniques;
- machine learning transforms EDR into a proactive system via analyzing historical data and correlating anomalies between distributed endpoints, that allows to forecast and block potential threats before the actual exploiting.

To integrate machine learning with EDR, both supervised and unsupervised learning techniques are used. The supervised learning utilizes labeled datasets of benign and malicious activity patterns to train modern classification. Yet, this approach is limited to recognizing known attack signatures and behavioral consequences. [1]

In order to detect zero-days attacks, the widely recognized unsupervised learning algorithm is Isolation Forest (IF). Specifically designed to detect outliners, it’s predicated on fundamental concept of anomalies deviating significantly that makes them easier to identify.

The IF algorithm idea is randomly choosing features of data in the dataset and its random split value. In case of EDR anomaly detection there can be such features as percentage of CPU or memory usage, destination ports and IP addresses, or parent-child process relationships. The splitting is repeated recursively on each data subset until individual data point are isolated.

The results of this splitting are considered as isolation trees which helps to define anomaly score. It is calculated as the path length of a single data point across all trees in the forest. The shorter is the average path length of certain data point, the more likely it’s defined as an anomaly due to more feature values that noticeably differ from the averages in the dataset. [2]

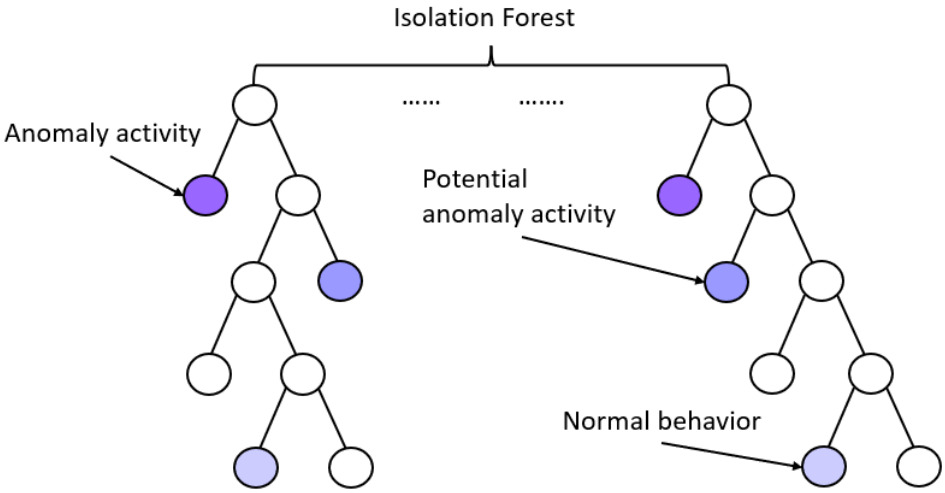

Fig.1 Isolation Forest construction for detecting anomalies

The advantage of Isolation Forest algorithm is its high speed of computing large-scale datasets and its focus on recognizing outliners instead of profiling normal data.

Combination of supervised and unsupervised machine learning models, integrated into EDR solutions, provide fast, accurate and adaptive detection and mitigation of cyberthreats, comparing to traditional signature approach.

However, several challenges have to be taken into consideration to achieve high-performance of ML-based endpoint detection & response systems:

- the diversity of operating systems, hardware configurations and software behaviors can reduce the accuracy of EDR in case of changes in the network. To overcome this, continuous model retraining and time to create normal activity baselines are required;
- while automation accelerates EDR systems capability of detection and response, human expertise remains crucial for contextual judgement and final decision making. Training cybersecurity specialists to interpret ML outputs, manage false positives, and tune detection parameters is essential for providing endpoint security in rapidly evolving threat landscape. [1]

References:

1. Real-Time Endpoint Detection and Response (EDR) with Machine Learning Integration / M. Roberts, J. Turner, E. Williams et al. 2024. URL: https://www.researchgate.net/publication/396621122_Real-Time_Endpoint_Detection_and_Response_EDR_with_Machine_Learning_Integration

2. GeeksforGeeks. Anomaly detection using Isolation Forest - GeeksforGeeks. GeeksforGeeks. URL: https://www.geeksforgeeks.org/machine-learning/anomaly-detection-using-isolation-forest/

Чайківський Віталій Володимирович
студент групи БСДМ-51, ННІКБЗІ ДУІКТ,
Київ, Україна

МАШИННЕ НАВЧАННЯ ЯК ІНСТРУМЕНТ ПІДТРИМКИ КІБЕРРОЗВІДКИ НА ОСНОВІ ВІДКРИТИХ ДЖЕРЕЛ ІНФОРМАЦІЇ

Традиційні підходи до розвідки, які базувалися на обмежених джерелах та ручному аналізі, трансформувалися в розвідку на основі відкритих джерел – OSINT, яка використовує публічно доступні дані. Подальший розвиток кіберпростору та зростання обсягів інформації зумовили необхідність застосування машинного навчання як інструменту автоматизації, аналізу та виявлення прихованих закономірностей, що підвищує ефективність кіберрозвідки в сучасних умовах.

Ключові слова: OSINT, відкриті джерела, кіберрозвідка, машинне навчання, інформація

Одним із ключових підходів до збору інформації та підвищення обізнаності про загрози є кіберрозвідка на основі відкритих джерел (або ж OSINT), яка ґрунтується на публічно або комерційно доступній інформації, спрямованій на закриття визначених інформаційних потреб. Водночас масштаби інформаційного середовища й поява великої кількості джерел інформації посилюють проблему інформаційного перевантаження, коли надлишок нерелевантних даних сповільнює аналітичний цикл і знижує ефективність ручної обробки.

Методи машинного навчання дедалі активніше застосовуються в кібербезпеці як засіб адаптивної, даних-орієнтованої обробки великих потоків даних (зокрема для класифікації, виявлення аномалій, аналізу трафіку та інших захисних завдань). Вищезгадане створює підстави розглядати машинне навчання (machine learning (далі – ML)) не як повноцінну заміну аналітика, а як інструмент його підтримки. Зокрема в частині автоматизації його рутинних операцій, фільтрації шуму та формування пріоритетів для подальшої інтерпретації аналітиком.

У сучасному розумінні OSINT розглядається як продукт, похідний від публічно або комерційно доступної інформації, що використовується для закриття конкретних розвідувальних пріоритетів, вимог або прогалин, а також як «перший» рівень індикаторів розвитку подій у кризових ситуаціях завдяки високій швидкості появи відкритих сигналів [1]. З позиції організації державного рівня OSINT вважається важливим елементом для підтримки ситуаційної обізнаності та формування контексту для рішень, за умови наявності та дотримання процедур конфіденційності даних та громадянських свобод відповідно до чинного законодавства.

Практична цінність OSINT полягає в тому, що організації зазвичай спираються на комбінацію внутрішніх і зовнішніх джерел, а також на інструменти/сенсори/сховища, які забезпечують накопичення та обмін релевантною інформацією про загрози; критичною стає здатність швидко перетворювати дані на придатні для дії індикатори та контекст.

Ключове методологічне обмеження OSINT – масштаб і неоднорідність інформації. Відкриті дані часто містять повтори, шум та неповні відомості, тому підхід спирається виключно на ручну обробку, погано масштабується. Фокус на «інформації розвідувальної цінності» потребує впровадження технічних рішень, які допомагають аналітикам зосередити зусилля на найбільш значущих фрагментах даних.

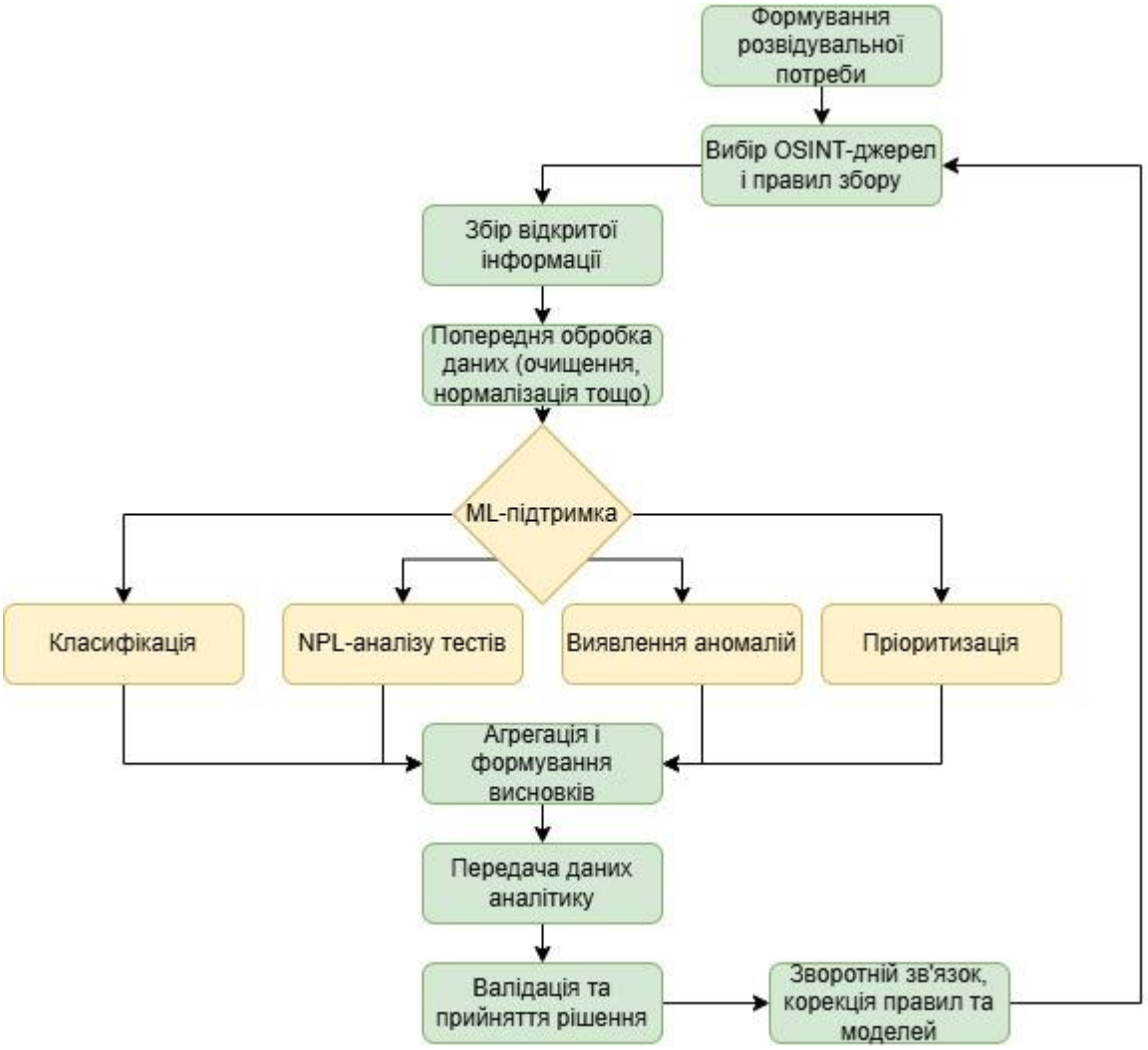

Рис. 1. Використання ML при проведення OSINT-досліджень

Саме тому з огляду на високі обсяги й швидкість надходження відкритих даних, доцільним є впровадження ML як інструменту підтримки аналітичного циклу. У кібербезпеці ML демонструє релевантність у завданнях класифікації, виявлення аномалій, пріоритезація. На рисунку вище схематично продемонстровані ці аспекти при проведення OSINT-дослідження.

Класифікація. У межах OSINT-дослідження класифікація може застосовуватися для багатьох аспектів, зокрема: тематичного групування повідомлень (вразливості/інциденти/кампанії), категоризації об’єктів спостереження (сутності, домени, ресурси, акаунти, артефакти), попередньої оцінки рівня ризику (низький/середній/високий) або типу загрози. Практична цінність полягає в автоматизованому способі відокремлення «сигналів» від «шуму» та прискоренні первинного сортування за критеріями, узгодженими з цілями кіберрозвідки.

NLP (обробка природної мови). Значна частина відкритої інформації має текстову природу, а отже, застосування NLP-методів доцільне для витягання сутностей і фактів, тематичного моделювання, класифікації повідомлень, а також узагальнення тексту для швидкого ознайомлення.

Виявлення аномалій. Для кіберрозвідки виявлення аномалій може виступати механізмом раннього попередження: виявляти нетипові сплески, зміну тематики у потоках повідомлень, узгоджені інформаційні кампанії або появу нових шаблонів поведінки в доступних телеметричних даних. У кібербезпеці загалом підходи на основі виявлення

аномалій розглядаються як важливі, оскільки здатні виявляти відхилення від нормальної поведінки без повної залежності від відомих сигнатур і статичних правил [2].

Пріоритизація. В OSINT пріоритизація є критичною через явище інформаційного перевантаження та обмежені ресурси аналітиків. Практично це може реалізовуватися як ранжування об'єктів за релевантністю до поточних розвідувальних потреб, потенційним впливом, рівнем довіри до джерела або очікуваною оперативною цінністю. У межах організаційних практик обміну інформацією про загрози автоматизація розглядається як чинник, що дозволяє швидко поширювати, перетворювати, збагачувати й аналізувати дані з меншим обсягом ручного втручання.

Саме тому з урахуванням викладеного вище, можна зробити висновок, що OSINT як підхід, що базується на публічно або комерційно доступній інформації, має суттєву прикладну цінність для кіберрозвідки завдяки швидкості появи відкритих сигналів і можливості підтримувати ситуаційну обізнаність. Водночас ключовою проблемою є масштаб і неоднорідність даних, що формує об'єктивний запит на технологічні засоби фільтрації та аналітичної підтримки.

Методи машинного навчання доцільно розглядати як інструмент підтримки кіберрозвідки: для класифікації та первинного сортування потоків OSINT, застосування NLP для роботи з текстовими джерелами, виявлення аномалій та пріоритизації сигналів, які потребують уваги аналітика. Представлена логіка узгоджується з підходом, у якому автоматизація забезпечує швидке перетворення, збагачення й аналіз інформації про загрози, тоді як остаточна інтерпретація та управлінські рішення залишаються за аналітиком.

Перелік посилань:

1. Defense Intelligence Agency. Open Source Intelligence. URL: https://www.dia.mil/About/Open-Source-Intelligence/ (дата звернення: 07.04.2026).

2. Alshammari A., Alsubaie N., Alshahrani S., Alqahtani A. Machine Learning for Open-Source Intelligence Applications. Electronics. 2025. Vol. 14, No. 23. URL: https://www.mdpi.com/2079-9292/14/23/4563 дата звернення: 07.04.2026).

Шабала Євгенія Євгенівна
доцент кафедри кібербезпеки та комп'ютерної інженерії
Київський національний університет будівництва і архітектури
Київ, Україна
Корнійчук Борис Валерійович
доцент кафедри професійної освіти
Київський національний університет будівництва і архітектури
Київ, Україна

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ АНАЛІЗУ ВІДКРИТИХ ДАНИХ У КІБЕРРОЗВІДЦІ

Зростання відкритих даних із соцмереж, онлайн-медіа та державних реєстрів формує великі обсяги неструктурованої інформації, яка відображає поведінку людей, суспільні настрої та соціальні процеси. OSINT і кіберрозвідка використовують ці дані для виявлення закономірностей, прогнозування загроз і підтримки рішень. Штучний інтелект автоматизує аналіз текстів, зображень і поведінкових моделей, допомагаючи швидко фільтрувати інформацію, виявляти аномалії та оцінювати ризики в цифровому середовищі.

Ключові слова: OSINT, кіберрозвідка, штучний інтелект, соціальні мережі, прогнозування загроз

У зв'язку із швидким розвитком цифрових технологій, поширення інтернету та мобільних пристроїв спостерігається зростання обсягів відкритих даних. Щодня соціальні мережі, онлайн-медіа та державні реєстри генерують великі масиви інформації, які активно використовуються в економіці, науці та управлінні.

Одним із головних джерел відкритих даних є соціальні мережі. У світі вже понад 5,6 млрд користувачів соцмереж,і ця кількість постійно зростає [1,]. Мільярди користувачів щодня створюють контент у вигляді текстів, фотографій, відео та коментарів. Ці дані відображають поведінку людей, їхні інтереси, думки та соціальні зв'язки. Внаслідок цього формується великий масив неструктурованої інформації, який може бути використаний для маркетингових досліджень, аналізу громадської думки та прогнозування різних трендів.

Другим важливим джерелом є онлайн-новини та цифрові медіа. Інформація в сучасному світі поширюється миттєво, а новинні ресурси працюють у режимі реального часу. Це створює безперервний потік даних, які можна аналізувати для виявлення інформаційних тенденцій, політичних процесів і соціальних змін.

Третім елементом є державні відкриті дані. Багато урядів, зокрема й в Україні, впроваджують політику відкритості, надаючи доступ до реєстрів, статистичних даних, бюджетної інформації та інших наборів даних. Це підвищує прозорість влади, розвиток електронного урядування та створення нових цифрових сервісів. Варто зауважити, що 90% усієї розвідувальної інформації можна отримати саме з відкритих джерел.[2, с.275]

В інформаційному суспільстві дані стали стратегічним ресурсом, а їх аналіз - ключовим інструментом прийняття рішень. У цьому контексті особливого значення набувають OSINT та кіберрозвідка, які дозволяють отримувати, обробляти й інтерпретувати інформацію з відкритих і цифрових джерел.

І OSINT, і кіберрозвідка базуються на класичному розвідувальному циклі:

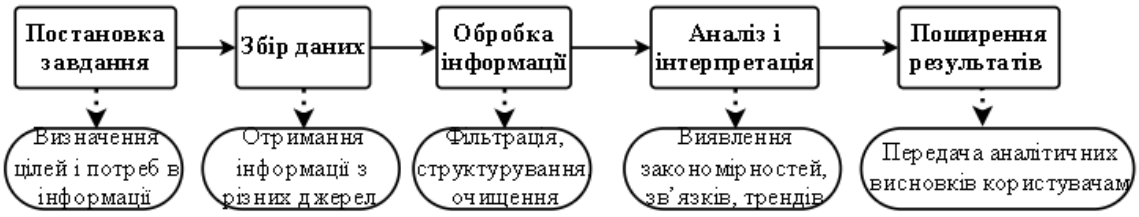

Рис.1. Цикл OSINT і кіберрозвідки

Кіберрозвідка працює з величезними потоками інформації: мережевим трафіком, логами, відкритими джерелами, даркнетом. Штучний інтелект дозволяє швидко обробляти дані, автоматично фільтрувати шум, виділяти релевантну інформацію.

Це суттєво скорочує час від збору даних до отримання аналітичного результату. ШІ активно використовується для виявлення аномалій і потенційних атак.Завдяки ШІ проводиться аналіз поведінки користувачів, виявлення нетипової активності в мережі та розпізнавання нових, раніше невідомих загроз. Також алгоритми машинного навчання можуть знаходити закономірності, які складно помітити людині.

Одна з найважливіших функцій ШІ в кіберрозвідці – це прогнозування загороз. На основі історичних даних системи можна передбачати можливі атаки, визначити найбільш вразливі точки інфраструктури та оцінити ризики. Штучний інтелект значно підсилює можливості OSINT тим, що надає можливість автоматично моніторити соцмережі і новини, аналізувати тексти, виявляти дезінформацію, розпізнавати зображення і відео. ШІ може виявляти приховані зв'язки між подіями, людьми та організаціями.

Для швидкого реагування на нові типи атак ШІ використовується для класифікації malware, аналізу поведінки програм, автоматичного виявлення підозрілих файлів. Також системи із підсиленням ШІ можуть формувати аналітичні звіти, ранжувати загрози за рівнем ризику, допомагати аналітикам у прийнятті рішень.

Основними ШІ інструментами в кіберброзвідці/OSINT є Maltego для побудови зв'язків між людьми, доменами, IP-адресами, SpiderFoot, який автоматизує збір даних із сотень джерел, використовує алгоритми для виявлення загроз і аномалій, ChatGPT для аналізу текстів, Google Lens для розпізнавання об'єктів і тексту на зображеннях, пошуку схожих зображень, Recorded Future, який використовує AI для прогнозування кіберзагроз, аналізує даркнет, форуми, технічні дані, Darktrace для виявлення аномалії в мережі, Shodan для пошуку пристроїв, підключених до інтернету, EOS Land Viewer / USGS Earth Explorer / NASA Earthdata – сервіси для отримання их даних та використовувати їх для наукових, дослідіницьких, агрономічних та інших цілей [3, с.149].

Зі сучасним використанням платформ соціальних мереж споживачі створюють та поширюють більше інформації, ніж будь-коли раніше, деяка з якої є оманливою та не має відношення до реальності [4, с.175]. Прикладом використання ШІ в кіберрозвідці є те, що у соціальних мережах масово поширюється новина про фейкова «надзвичайну подію», яка може викликати паніку.

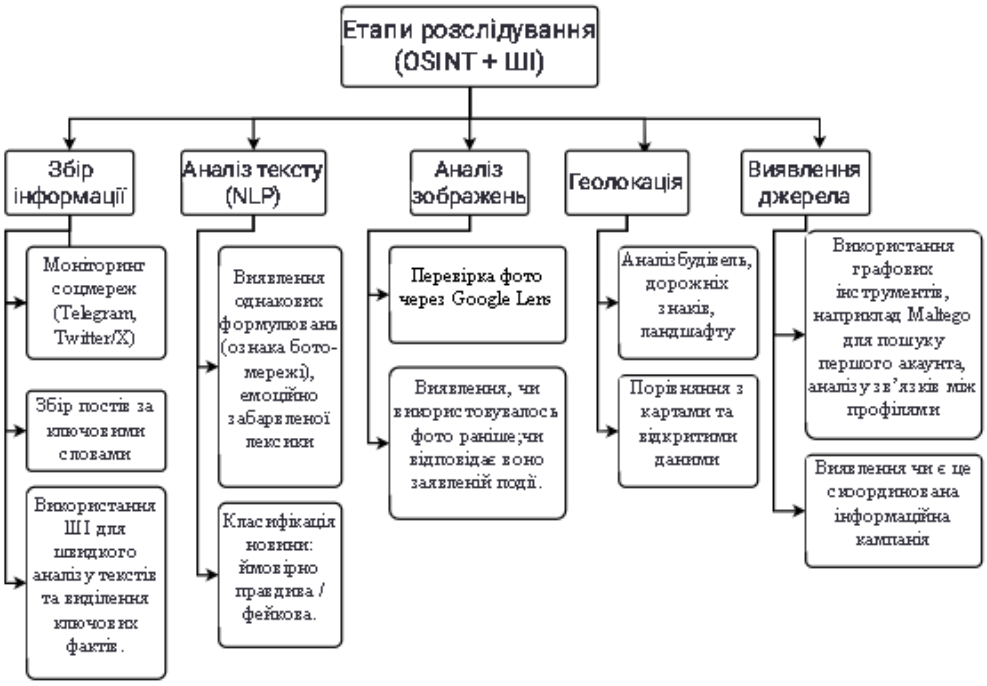

Рис. 2. Етапи розслідування в OSINT із використанням ШІ

На Рис.2 представлена послідовність етапів розслідування в OSINT із використанням штучного інтелекту, що реалізується як багаторівневий процес. Початковим етапом є збір інформації з відкритих джерел, соціальних мереж із застосуванням автоматизованого моніторингу та алгоритмів ШІ для швидкого аналізу текстів і виділення релевантних фактів. Далі здійснюється обробка текстових даних за допомогою методів NLP, що включає виявлення ознак бот-активності, емоційного забарвлення та класифікацію контенту за рівнем достовірності. Наступні етапи відповідають за аналіз зображень (верифікація через інструменти розпізнавання та зворотного пошуку), геолокаційний аналіз (ідентифікація місцевості за візуальними ознаками з подальшим зіставленням із картографічними даними) та встановлення джерела інформації шляхом мережевого аналізу зв'язків між акаунтами. У сукупності ці етапи забезпечують комплексну верифікацію даних, виявлення інформаційних операцій і формування обґрунтованих аналітичних висновків.

Перелік використаної літератури

1. Соцмережі 2026: важлива статистика для комунікацій НУО https://www.prostir.ua/?kb=sotsmerezhi-2026-vazhlyva-statystyka-dlya-komunikatsij-nuo&utm_source=chatgpt.com
2. Думчиков, М.О. Відкрита розвідка (OSINT) у протидії інформаційним війнам: аналітичні інструменти та правові межі / М.О. Думчиков // Аналітично-порівняльне правознавство. – 2025. – Т. 2, № 6. – С. 273-277 – DOI: 10.24144/2788-6018.2025.06.2.43
3. Івкова, В.С. Дослідження існуючих засобів та підходів до проведення OSINT в контексті інформаційної безпеки особи та держави / В.С. Івкова, І.Р. Опірський // Комп'ютерні системи та мережі. – 2025. – Т. 7, № 1. – С. 143–159. – DOI: 10.23939/csn2025.01.143.
4. Назаркевич, М., Коротич, Д., Чолкан, Р. Розроблення методів для виявлення фейкових новин на основі аналізу теорії графів / М. Назаркевич, Д. Коротич, Р. Чолкан // Кібербезпека: освіта, наука, техніка. – 2025. – Т. 3, № 31. – С. 172–187. – DOI: https://doi.org/10.28925/2663-4023.2025.31.997. – Режим доступу: https://csecurity.kubg.edu.ua/index.php/journal/article/view/1110/935

Шкурченко О. А.
Аспірант групи АІКБ-21, ННІКБЗІ, ДУІКТ,
Київ, Україна

ГРУПОВІ КІБЕРЗАГРОЗИ: СУТНІСТЬ, КЛАСИФІКАЦІЯ ТА ОСОБЛИВОСТІ ПРОТИДІЇ

Досліджено сутність та класифікацію групових кіберзагроз у кіберпросторі, що характеризуються координацією дій між множиною суб'єктів, розподілом функцій та спільним використанням інфраструктури. Виокремлено три основні форми групових кіберзагроз: розширені постійні загрози (APT), ботнети та скоординовані кампанії змішаного типу. Проаналізовано системні особливості таких загроз – часову синхронізацію, адаптивність тактик та використання прихованих каналів комунікації, – які ускладнюють їх виявлення традиційними засобами кіберзахисту. Обґрунтовано перспективність застосування методів штучного інтелекту, зокрема алгоритмів кластеризації та виявлення аномалій, для ідентифікації ознак координованої зловмисної активності.

Ключові слова: групові кіберзагрози, розширені постійні загрози (APT), ботнети, координовані кібератаки, штучний інтелект, машинне навчання, виявлення аномалій, MITRE ATT&CK, тактики техніки та процедури (TTP).

Цифрова трансформація суспільства супроводжується ескалацією кіберзагроз, які за своєю природою дедалі частіше набувають колективного, скоординованого характеру. Як зазначено у настанові Національного інституту стандартів і технологій США (National Institute of Standards and Technology, NIST) SP 800-150, спектр суб'єктів загроз охоплює діапазон від окремих автономних зловмисників до добре забезпечених ресурсами груп, що діють координовано в межах злочинного підприємства або від імені держави [1]. На відміну від ізольованих інцидентів, групові кіберзагрози реалізуються організованими структурами з розподіленими ролями, спільною інфраструктурою та узгодженими тактиками, техніками і процедурами (Tactics, Techniques, and Procedures, TTP). Розуміння природи таких загроз є передумовою для побудови ефективних механізмів протидії.

Під груповою кіберзагрозою доцільно розуміти скоординовану зловмисну діяльність у кіберпросторі, що здійснюється множиною суб'єктів – людей, автоматизованих агентів або їх комбінацією, які діють узгоджено для порушення конфіденційності, цілісності або доступності інформаційних ресурсів. Ключовою ознакою, що відрізняє групову кіберзагрозу від сукупності незалежних атак, є наявність координації: часової синхронізації дій, розподілу функцій між учасниками, спільного використання інфраструктури та каналів комунікації.

Аналіз сучасного ландшафту загроз дозволяє виокремити три основні форми групових кіберзагроз.

Перша форма – розширені постійні загрози (Advanced Persistent Threat, APT). За визначенням Sharma et al., APT являють собою цілеспрямовані кібернетичні засоби, мотивовані конкретними намірами, які характеризуються складним дизайном вектора атаки та використанням високоприхованих технік ухилення від виявлення [2]. APT-групи здійснюють тривалі багатоетапні кампанії, що включають початкове проникнення (переважно через цільовий фішинг або експлуатацію вразливостей нульового дня), закріплення у скомпрометованій мережі, бокове переміщення (lateral movement) між вузлами та ексфільтрацію даних. База знань MITRE ATT&CK систематизує понад 140 груп загрозливих суб'єктів, кожна з яких характеризується унікальним набором TTP [3]. Ця систематизація підтверджує, що APT-групи не є випадковими одиничними зловмисниками, а представляють стійкі організаційні структури зі специфічними поведінковими профілями.

Друга форма – ботнети (botnets). Ботнет являє собою мережу скомпрометованих пристроїв під централізованим або розподіленим управлінням через командно-контрольну інфраструктуру (Command and Control, C2). Координація у ботнеті проявляється у синхронних

діях тисяч вузлів – від розподілених атак відмови в обслуговуванні (Distributed Denial of Service, DDoS) до масової розсилки спаму та крадіжки облікових даних. Як демонструє огляд Negera et al., зростання кількості пристроїв Інтернету речей (Internet of Things, IoT) – прогнозовано понад 20 мільярдів підключених пристроїв – суттєво розширює поверхню атаки та створює нові можливості для формування масштабних ботнет-мереж [4].

Третя форма – скоординовані кампанії змішаного типу, де різні підрозділи одного угруповання або кілька взаємопов'язаних угруповань одночасно застосовують різні вектори: фішинг, компрометацію ланцюга постачання (supply chain attack), шкідливе програмне забезпечення (malware), а в деяких випадках – фізичний доступ до інфраструктури [5]. Багатовекторність таких кампаній створює синергетичний ефект, за якого кожен окремий вектор підсилює результативність інших.

Групові кіберзагрози мають низку системних особливостей, що ускладнюють їх виявлення та нейтралізацію. Координація у часі та просторі дозволяє зловмисникам розподіляти навантаження між учасниками, маскуючи загальну картину атаки. Адаптивність групових загроз проявляється у здатності змінювати TTP у відповідь на виявлення, що робить сигнатурні методи захисту малоефективними. Використання прихованих каналів комунікації та шифрування ускладнює ідентифікацію зв'язків між учасниками атаки. Дослідження методів ідентифікації APT-угруповань, представлене Rani et al., систематизує артефакти, що використовуються для встановлення зв'язку між інцидентами та конкретними угрупованнями, поділяючи їх на доказові (evidentiary) та поведінкові (behavioral), що підкреслює складність ідентифікації координованих дій [6].

Традиційні засоби кіберзахисту – антивірусне програмне забезпечення, міжмережеві екрани, системи виявлення вторгнень (Intrusion Detection Systems, IDS) – орієнтовані переважно на відомі загрози та працюють на основі наперед визначених шаблонів. Такі системи не враховують кореляції між діями різних джерел атаки та не здатні виявляти повільні багатоетапні кампанії.

Перспективним напрямом протидії груповим кіберзагрозам є застосування методів штучного інтелекту (Artificial Intelligence, AI). Огляд Buczak та Guven систематизує методи інтелектуального аналізу даних та машинного навчання для виявлення вторгнень, демонструючи ефективність алгоритмів класифікації та кластеризації для ідентифікації аномальних патернів мережевого трафіку [7]. Методи глибинного навчання (Deep Learning, DL), як показує огляд Berman et al., забезпечують додаткові можливості для аналізу складних структур даних, включаючи послідовності подій та багатовимірні часові ряди, що є характерними для координованих атак [8]. Зокрема, методи неконтрольованого навчання – кластеризація (K-Means, DBSCAN) та виявлення аномалій (Isolation Forest, Autoencoders) – є ефективними для ідентифікації нових форм координованих атак, оскільки не потребують попередньо розмічених даних.

Формалізація поняття групової кіберзагрози, класифікація її форм та аналіз системних особливостей створюють теоретичне підґрунтя для розробки спеціалізованих методів і моделей протидії, орієнтованих на виявлення ознак координації, а не лише на ідентифікацію окремих шкідливих дій. Подальші дослідження мають бути спрямовані на розробку адаптивних моделей, здатних до самонавчання в реальному часі, інтеграцію контекстної інформації з множини джерел, а також забезпечення інтерпретованості рішень інтелектуальних систем кіберзахисту.

Список літератури:

1. Johnson, C., Badger, M., Waltermire, D., Snyder, J., & Skorupka, C. (2016). Guide to Cyber Threat Information Sharing (NIST Special Publication 800-150). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-150
2. Sharma, A., Gupta, B. B., Singh, A. K., & Saraswat, V. K. (2023). Advanced Persistent Threats (APT): evolution, anatomy, attribution and countermeasures. Journal of Ambient

Intelligence and Humanized Computing, 14(7), 9355–9381. https://doi.org/10.1007/s12652-023-04603-y

3. MITRE Corporation. (n.d.). ATT&CK Groups. https://attack.mitre.org/groups/
4. Negera, W. G., Schwenker, F., Debelee, T. G., Melaku, H. M., & Ayano, Y. M. (2022). Review of Botnet Attack Detection in SDN-Enabled IoT Using Machine Learning. Sensors, 22(24), 9837. https://doi.org/10.3390/s22249837
5. Шульга, В. П., Іванченко, Є. В., Берестяна, Т. В., & Шкурченко, О. А. (2025). Методи та моделі протидії груповим кіберзагрозам на основі штучного інтелекту. Кібербезпека: освіта, наука, техніка, 2(30), 593–606. https://doi.org/10.28925/2663-4023.2025.30.998
6. Rani, N., Mishra, B., Bajaj, P., & Shukla, S. K. (2024). A Comprehensive Survey of Advanced Persistent Threat Attribution: Taxonomy, Methods, Challenges and Open Research Problems. arXiv preprint arXiv:2409.11415. https://arxiv.org/abs/2409.11415
7. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502
8. Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for cyber security. Information, 10(4), 122. https://doi.org/10.3390/info10040122

Ядлось Д.О.
студент групи ПД-41, ІПЗ ДУІКТ, Київ,
Україна.
Науковий керівник:
Треньов Микита Георгійович,
асистент кафедри ІПЗ ДУІКТ, Київ,
Україна.

ЗАХИСТ ІГРОВИХ ДАНИХ У 3D-ВІДЕОГРІ «SITE-108»

У тезі розглянуто підхід до захисту локальних ігрових даних у проєкті SITE-108 - 3D-відеогрі на Unity з елементами хорору та RPG. Обґрунтовано розподіл даних на некритичні та критичні, доцільність обмеженого використання PlayerPrefs і застосування прикладного шифрування для файлів збереження. Запропонований підхід спрямований на збереження цілісності ігрового прогресу, зменшення ризику підміни збережень та підвищення надійності роботи гри.

Ключові слова: Unity, 3D-відеогра, захист даних, ігрові збереження, PlayerPrefs, AES, цілісність даних.

У сучасних цифрових іграх безпека стосується не лише мережевих сервісів, платежів або автентифікації користувачів, а й збереження локальних ігрових даних. Навіть у проєктах без багатокористувацького режиму файл збереження визначає логіку проходження, доступ до контенту, стан інвентарю та виконання сюжетних подій. Дослідження у сфері game development показують, що питання безпеки в ігрових проєктах часто вирішуються вже після появи проблем, а не на етапі проєктування, оскільки цьому заважають обмеження часу, бюджету та відсутність єдиного підходу до впровадження захисних механізмів [1].

У межах кваліфікаційної роботи “Розробка 3D-відеогри на Unity з елементами хорору та RPG” розробляється проєкт “SITE-108”. Для такої гри важливо зберігати цілісність ігрового прогресу, оскільки втрата або підміна даних впливає на послідовність подій, баланс ресурсів і загальний користувацький досвід. Метою даної тези є обґрунтування підходу до захисту локальних даних у проєкті SITE-108 з урахуванням можливостей Unity та практичних обмежень розробки.

У роботі доцільно розмежувати дані за рівнем критичності. До некритичних належать параметри, зміна яких не порушує сценарну цілісність гри: рівень гучності, роздільна здатність, мовні налаштування, чутливість керування, параметри інтерфейсу. До критичних належать збереження проходження: активна глава, контрольні точки, вміст інвентарю, стан взаємодіючих об'єктів, сюжетні прапорці, відкриті зони та ресурси персонажа. Саме ці дані потребують захисту від випадкового пошкодження або навмисного редагування.

Офіційна документація Unity відверто вказує, що PlayerPrefs зберігається локально без шифрування і не повинен використовуватися для чутливих даних [3]. З огляду на це у проєкті SITE-108 PlayerPrefs доцільно застосовувати лише для некритичних користувацьких параметрів. Таке рішення дозволяє використати стандартний інструмент Unity там, де він справді доречний, і водночас не покладати на нього функції, для яких він не призначений.

Критичні дані збереження доцільно формувати в окрему структуру та серіалізувати у файл. Рекомендації OWASP щодо захисту даних у стані спокою наголошують на двох принципах, важливих для цього підходу: необхідно мінімізувати обсяг чутливої інформації, що зберігається, і за потреби застосовувати сильні криптографічні алгоритми, насамперед AES, а також передбачати окрему перевірку автентичності або цілісності даних [2].

З урахуванням цих рекомендацій у SITE-108 доцільна така схема: налаштування, які не впливають на проходження, зберігаються стандартними засобами Unity, а дані прогресу -

окремо у файлі, де перед записом серіалізований вміст шифрується алгоритмом AES. Для цього ж пакета формується засіб перевірки цілісності, що дозволяє виявити стороннє втручання, часткове пошкодження або некоректне відновлення файлу. Якщо перевірка під час завантаження не проходить успішно, гра не повинна використовувати такі дані безумовно: коректнішим рішенням є відхилення пошкодженого збереження, пропозиція резервної копії або створення нового слоту збереження.

Запропонований підхід є виправданим для проєкту, оскільки поєднує науково обґрунтовані принципи захисту з реально доступними інструментами Unity та C#. Для сюжетної гри це особливо важливо, бо несанкціонована зміна сюжетних прапорців, стану інвентарю або ресурсів персонажа може порушити задуману драматургію та логіку тригерів. Отже, у тезі обґрунтовано підхід, за якого стандартні засоби Unity застосовуються вибірково: PlayerPrefs використовується для некритичних параметрів, а дані прогресу виносяться в окрему систему збереження з прикладним шифруванням і перевіркою цілісності. Практичне значення такого рішення полягає в підвищенні надійності локального зберігання даних та зменшенні ризику підміни ігрових збережень [1-3].

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Klostermeyer P., Amft S., Höltervennhoff S., Krause A., Busch N., Fahl S. Skipping the Security Side Quests: A Qualitative Study on Security Practices and Challenges in Game Development // Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (CCS '24). 2024. P. 2651-2665. DOI: 10.1145/3658644.3690190.

2. OWASP Foundation. Cryptographic Storage Cheat Sheet. URL: https://cheatsheetseries.owasp.org/cheatsheets/Cryptographic_Storage_Cheat_Sheet.html

3. Unity Technologies. PlayerPrefs. Scripting API. URL: https://docs.unity3d.com/6000.3/Documentation/ScriptReference/PlayerPrefs.html

Бойко А.О.
Старший викладач кафедри СТКБ, ННІКБЗІ,
ДУІКТ, Київ, Україна

ВИЯВЛЕННЯ ВЕБ-АТАК У МЕРЕЖЕВОМУ ТРАФІКУ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

У роботі розглянуто проблему виявлення веб-атак у мережевому трафіку з використанням методів машинного навчання. Актуальність дослідження зумовлена зростанням кількості атак на веб-застосунки, а також складністю їх виявлення в умовах значного переважання легітимного трафіку над шкідливим. Як експериментальну основу використано датасет CSE-CIC-IDS2018, що містить реалістичні сценарії мережевих атак, зокрема Brute Force, XSS та SQL Injection. Розглянуто результати бінарної та мультикласової класифікації, а також доцільність використання двоступеневої каскадної архітектури, у якій

перший рівень визначає факт атаки, а другий – її тип. Показано, що каскадний підхід дозволяє підвищити якість розпізнавання рідкісних веб-атак.

Ключові слова: веб-атаки, машинне навчання, виявлення вторгнень, CSE-CIC-IDS2018, дисбаланс класів, XGBoost, SQL Injection, XSS.

Веб-застосунки є одним із найбільш поширених компонентів сучасної інформаційної інфраструктури. Через них користувачі отримують доступ до корпоративних сервісів, баз даних, особистих кабінетів, платіжних систем, адміністративних панелей та інших критично важливих ресурсів.

Проблема виявлення атак на веб-застосунки ускладнюється тим, що вони можуть маскуватися під звичайну взаємодію користувача з веб-сервісом. Наприклад, Brute Force-атака може виглядати як серія типових HTTP-запитів до форми входу, XSS може передаватися як текстовий параметр у запиті, а SQL Injection часто реалізується через модифікацію значень у полях введення або параметрах URL. Якщо система аналізує не повний вміст HTTP-запиту, а лише узагальнені характеристики мережевого потоку, задача стає ще складнішою.

Для дослідження використано датасет CSE-CIC-IDS2018. Офіційний опис набору вказує, що він містить сім сценаріїв атак, зокрема Brute-force, Botnet, DoS, DDoS, Web attacks та інші, а також включає мережевий трафік і приблизно 80 ознак, отриманих за допомогою CICFlowMeter-V3 [1]. У межах експерименту аналізуються дані за 22 та 23 лютого 2018 року, які охоплюють три типи веб-атак: Brute Force на веб-застосунки, XSS та SQL Injection. Вихідний набір сформовано шляхом об’єднання двох CSV-файлів, кожен з яких містить 1 048 575 записів і 81 ознаку. Після очищення даних фінальний набір становив 2 079 952 записи, з яких лише 928 належали до атакового трафіку. Частка атак у вибірці становила приблизно 0,044%, тобто співвідношення між атакою та нормальним трафіком дорівнювало приблизно 1:2241.

Саме дисбаланс класів є центральною проблемою такого дослідження. У більшості реальних корпоративних мереж нормальний трафік значно переважає над шкідливим, тому задача виявлення атак фактично перетворюється на пошук рідкісних атак у великому масиві легітимної активності. Класичні моделі машинного навчання можуть досягати дуже високого значення accuracy, майже завжди передбачаючи клас Benign, але при цьому пропускати значну частину атак. Тому для оцінювання якості виявлення веб-атак недостатньо використовувати лише загальну точність. Набагато важливішими є precision, recall, F1-score, а для мультикласової задачі – F1-macro, оскільки ця метрика однаково враховує якість класифікації як частих, так і рідкісних класів.

Сильний дисбаланс веб-атак у CSE-CIC-IDS2018 також підтверджується в наукових роботах. Зокрема, у дослідженні Zuech, Hancock та Khoshgoftaar зазначено, що весь набір CSE-CIC-IDS2018 містить лише 87 прикладів класу SQL Injection, що ускладнює коректне навчання та оцінювання моделей для цього типу атаки [2]. Автори також підкреслюють, що для веб-атак у цьому наборі даних доцільно враховувати проблему дисбалансу та застосовувати спеціальні підходи до формування навчальної вибірки [3]. Це узгоджується з результатами аналізу наданих матеріалів, де SQL Injection є найрідкіснішим класом: у тестовій вибірці було лише 18 прикладів цього типу атаки.

На першому етапі було розглянуто задачу бінарної класифікації: Benign проти Attack. Метою цього рівня є не визначення конкретного типу атаки, а фіксація самого факту аномальної або потенційно шкідливої активності. Random Forest продемонстрував найкращий баланс між precision і recall: F1-score становив 0,900, а precision – 0,994. XGBoost baseline також показав високий результат із F1-score 0,873.

Окремо було досліджено вплив стратегій балансування класів. Для XGBoost перевірявся параметр scale_pos_weight, який збільшує вагу позитивного класу під час навчання. Зі збільшенням ваги атакового класу recall зростав, однак precision суттєво знижувався. Це означає, що модель починала знаходити більше атак, але водночас генерувала більше хибних спрацювань. У практичних умовах це є важливим компромісом: надмірна

кількість false positive може перевантажити аналітиків SOC, тоді як false negative може означати пропущену атаку. Також розглянуто RandomOverSampler і SMOTE для LightGBM. Вони підвищували recall, але знижували precision, що підтверджує необхідність обирати метод балансування відповідно до конкретних вимог системи захисту.

Для мультикласової класифікації задача ускладнюється тим що, модель повинна не лише відокремити атаку від нормального трафіку, а й правильно визначити її тип. У цьому випадку розглядалися класи Benign, Brute Force-Web, Brute Force-XSS та SQL Injection. Найбільш проблемним виявився клас SQL Injection через дуже малу кількість прикладів і схожість окремих статистичних ознак із Brute Force-Web. У прямій мультикласовій класифікації XGBoost значення F1 для SQL Injection становило 0,69, тоді як у каскадній схемі цей показник зріс до 0,80.

Основною ідеєю каскадної архітектури є поділ загальної задачі на два послідовні рівні. На першому рівні модель визначає, чи є потік нормальним або атакуючим. Якщо потік класифікується як нормальний, рішення фіналізується. Якщо потік визначено як атаку, він передається на другий рівень, де окрема модель класифікує тип атаки: Brute Force-Web, Brute Force-XSS або SQL Injection. Такий підхід зменшує вплив домінуючого класу Benign на класифікатор другого рівня. Результати показали, що двоступенева схема XGBoost binary → XGBoost multiclass досягла F1-macro = 0,896, що на 4,1% вище за прямий XGBoost і на 2,6% вище за Random Forest у мультикласовій постановці. Найбільший приріст якості спостерігався для SQL Injection: F1 збільшився на 0,110. Також було перевірено стабільність результатів при п’яти різних значеннях random_state. Середнє значення F1-macro для каскадного XGBoost становило 0,894 ± 0,002, тоді як для Random Forest – 0,871 ± 0,002, а для прямого XGBoost – 0,861 ± 0,001.

Важливим елементом практичного застосування моделей машинного навчання в кібербезпеці є інтерпретованість. Для систем виявлення атак недостатньо лише отримати передбачення моделі; потрібно також розуміти, які ознаки вплинули на це рішення. Це особливо важливо для аналітиків SOC, які мають перевіряти спрацювання, відокремлювати реальні інциденти від хибних тривог і пояснювати причину класифікації трафіку як шкідливого. Для цього може застосовуватися SHAP – підхід до інтерпретації моделей, який призначає кожній ознаці внесок у конкретне передбачення. У роботі Lundberg та Lee SHAP описано як уніфікований підхід до пояснення прогнозів складних моделей, зокрема ансамблевих алгоритмів [4].

У контексті виявлення веб-атак інтерпретованість дозволяє перевірити, чи модель справді спирається на змістовні характеристики мережевого трафіку, а не на випадкові або технічні артефакти набору даних. У матеріалах дослідження зазначено, що для бінарного класифікатора важливими були ознаки, пов’язані зі швидкістю потоку, кількістю пакетів за секунду, середньою довжиною пакетів, TCP-прапорцями та міжпакетними інтервалами. Для другого рівня каскаду інформативними стали ознаки, пов’язані з обсягом переданих даних у прямому та зворотному напрямках, довжиною заголовків і початковими параметрами TCP-вікна. Це логічно, оскільки різні типи веб-атак формують різні поведінкові профілі на рівні мережевого потоку.

Отже, результати дослідження підтверджують, що задача виявлення веб-атак у flow-based IDS-даних не може розглядатися як проста задача класифікації з високою accuracy. Вона потребує врахування дисбалансу класів, правильного вибору метрик, аналізу рідкісних класів і перевірки стабільності моделей. Особливо важливо не обмежуватися загальними результатами для всього набору даних, а окремо аналізувати якість розпізнавання Brute Force-Web, Brute Force-XSS та SQL Injection. Саме ці класи мають найбільше практичне значення для захисту веб-застосунків, але водночас є найскладнішими для надійного виявлення через малу кількість прикладів і близькість статистичних характеристик.

Перелік посилань:

1. CSE-CIC-IDS2018 on AWS [Electronic resource] / Canadian Institute for Cybersecurity, University of New Brunswick. – Electronic data. – Access mode: https://www.unb.ca/cic/datasets/ids-2018.html

2. Zuech R., Hancock J., Khoshgoftaar T. M. Detecting web attacks using random undersampling and ensemble learners [Electronic resource] // Journal of Big Data. – 2021. – Vol. 8. – Article 75. – Access mode: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00460-8

3. Leevy J. L., Khoshgoftaar T. M. A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 Big Data [Electronic resource] // Journal of Big Data. – 2020. – Vol. 7. – Article 104. – Access mode: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-020-00382-x

4. Lundberg S. M., Lee S.-I. A Unified Approach to Interpreting Model Predictions [Electronic resource] // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – Access mode: https://arxiv.org/abs/1705.07874

Бєланов Владислав Костянтинович
студент групи БСД-43, ННІКБЗІ, ДУІКТ
Київ, Україна

ОСОБЛИВОСТІ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ У ВЕБ-СИСТЕМАХ ІЗ ВИКОРИСТАННЯМ SOCIAL LOGIN

Реєстрація та автентифікація користувачів є критично важливими складовими функціонування веб-систем, оскільки саме на цьому етапі визначається рівень довіри до користувача та формується основа для подальшого контролю доступу. Традиційні підходи, що базуються на використанні логіна та пароля, мають низку обмежень, зокрема пов'язаних із повторним використанням паролів, їх недостатньою складністю та ризиком компрометації облікових даних. У зв'язку з цим поширення набули механізми автентифікації через зовнішніх провайдерів, відомі як Social Login.

Ключові слова: веб-безпека, автентифікація, Social Login, OAuth 2.0, OpenID Connect, токени доступу, redirect URI, облікові записи, кіберзагрози.

Застосування Social Login змінює підхід до забезпечення безпеки у веб-системах, оскільки частина процесів автентифікації делегується стороннім сервісам. Це дозволяє зменшити навантаження на систему в частині зберігання облікових даних, однак водночас створює нові ризики, пов'язані з інтеграцією із зовнішніми сервісами, передачею токенів та обробкою результатів автентифікації.

Однією з ключових особливостей такого підходу є використання протоколів OAuth 2.0 та OpenID Connect, які забезпечують обмін даними між веб-додатком і провайдером ідентифікації. Незважаючи на стандартизацію цих протоколів, їх практична реалізація часто містить помилки, що можуть призводити до вразливостей. Зокрема, некоректна перевірка redirect URI створює передумови для атак типу open redirect, які дозволяють перехоплювати токени доступу або перенаправляти користувачів на шкідливі ресурси.

Окрему увагу слід приділити обробці токенів, що використовуються під час автентифікації. Access Token та ID Token фактично виконують роль облікових даних і тому повинні захищатися відповідним чином. Їх зберігання у незахищених середовищах або передача без використання захищених каналів може призвести до несанкціонованого доступу. У зв'язку з цим необхідно застосовувати захищені механізми передачі даних, обмежувати час життя токенів та контролювати їх використання.

Ще одним важливим аспектом є захист від атак, пов'язаних із підміною або повторним використанням запитів автентифікації. Для цього у протоколах OAuth 2.0 та OpenID Connect використовуються параметри state та nonce, які дозволяють перевіряти цілісність сесії та запобігати несанкціонованому використанню відповідей від провайдера. Їх відсутність або некоректна реалізація створює додаткові ризики для веб-системи.

Особливу складність становить також управління обліковими записами користувачів при використанні Social Login. У випадках, коли один користувач може входити через різні провайдери, виникає необхідність коректного зв'язування акаунтів. Автоматичне об'єднання без додаткової перевірки може призвести до компрометації, якщо зловмисник використовує альтернативний спосіб автентифікації з тим самим ідентифікатором.

Крім того, важливим є обмеження обсягу даних, що отримуються від зовнішніх сервісів. Надмірний збір інформації не лише ускладнює обробку даних, а й підвищує ризики їх витоку. У зв'язку з цим доцільно дотримуватися принципу мінімізації даних і використовувати лише ті атрибути, які є необхідними для функціонування веб-системи.

Таким чином, використання Social Login у веб-системах має як переваги, так і обмеження з точки зору безпеки. Делегування автентифікації дозволяє зменшити частину ризиків, пов'язаних із керуванням обліковими даними, однак потребує додаткового контролю процесів взаємодії із зовнішніми сервісами. Ефективне забезпечення безпеки можливе лише

за умови комплексного підходу, що включає коректну реалізацію протоколів, захист токенів, контроль перенаправлень та управління обліковими записами користувачів.

Перелік посилань:

1. Authentication - OWASP Cheat Sheet Series. *Introduction - OWASP Cheat Sheet Series*. URL: https://cheatsheetseries.owasp.org/cheatsheets/Authentication_Cheat_Sheet.html
2. RFC 6749: The OAuth 2.0 Authorization Framework. *IETF Datatracker*. URL: https://datatracker.ietf.org/doc/html/rfc6749
3. NIST Special Publication 800-63B. URL: https://pages.nist.gov/800-63-3/sp800-63b.html

Шулімова Д.Д.
асистент кафедри СТКБ, ННІКБЗІ ДУІКТ,
Київ, Україна

ЗАСТОСУВАННЯ ДЕРЕВОПОДІБНИХ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ АТАК У КОРПОРАТИВНИХ МЕРЕЖАХ НА ОСНОВІ FLOW-ОЗНАК

Виявлення шкідливої активності в корпоративних мережах є однією з базових задач кіберзахисту, оскільки мережевий трафік має складну структуру, значний обсяг і неоднорідний характер. У таких умовах ускладнюється відокремлення нормальної поведінки від атак, а також інтерпретація результатів роботи систем моніторингу. Практичне значення має не лише здатність моделі виявляти атаки, а й характер допущених помилок: велика кількість хибних спрацювань створює додаткове навантаження на аналіз інцидентів, тоді як пропуски атак безпосередньо впливають на рівень ризику.

Ключові слова: виявлення атак, мережевий трафік, flow-ознаки, машинне навчання, дерева рішень, Random Forest, класифікація, шкідлива активність, DDoS, botnet, web-атаки, матриця помилок.

Одним із підходів до розв’язання цієї задачі є застосування методів машинного навчання, які дозволяють будувати класифікаційні моделі на основі статистичних характеристик мережевого трафіку. При цьому доцільним є використання агрегованого подання даних у вигляді мережевих потоків, для яких обчислюється набір узагальнених показників. Таке представлення, відоме як flow-ознаки, включає характеристики тривалості з’єднання, обсягів переданих даних, кількості пакетів, а також статистику часових інтервалів між ними. Перехід до такого формату дозволяє зменшити розмірність даних і зберегти поведінкові особливості мережевої взаємодії.

Важливою особливістю задачі є те, що різні типи атак по-різному відображаються у просторі ознак. Для атак типу DdoS характерні виражені зміни інтенсивності трафіку, що спрощує їх виявлення. Botnet-активність часто має повторюваний характер і проявляється у регулярних шаблонах взаємодії. У випадку web-атак відмінності від нормального трафіку є менш очевидними, оскільки такі атаки можуть мати схожі статистичні характеристики з легітимною активністю. Це ускладнює побудову моделей, які одночасно забезпечують високу чутливість і низький рівень хибних спрацювань.

Для роботи в таких умовах доцільним є використання деревоподібних методів машинного навчання. Такі моделі здійснюють послідовне розділення простору ознак за допомогою простих правил, що дозволяє поступово відокремлювати шкідливу активність від нормальної. Важливою перевагою є відсутність припущень щодо розподілу даних, що робить ці методи придатними для аналізу неоднорідних вибірок. Крім того, дерева рішень добре працюють із різними типами ознак і не потребують складної попередньої обробки даних.

Окремою перевагою є інтерпретованість результатів. Кожне рішення моделі можна представити у вигляді набору логічних правил, що дозволяє встановити, які саме характеристики трафіку вплинули на класифікацію. Це спрощує аналіз інцидентів і підвищує довіру до результатів роботи моделі. Разом із тим, одиничні дерева рішень можуть бути чутливими до особливостей навчальної вибірки, що впливає на стабільність їх роботи.

Для зменшення цього ефекту застосовуються ансамблеві методи, зокрема Random Forest, який формує множину дерев рішень і об’єднує їх результати. Такий підхід дозволяє зменшити вплив випадкових факторів і забезпечити більш стабільну поведінку моделі, особливо у випадках, коли межа між класами є нечіткою.

Аналіз із використанням набору даних CSE-CIC-IDS2018 показує, що ефективність моделей залежить від типу атаки. Для сценаріїв із чітко вираженими відмінностями між класами обидва підходи забезпечують майже повне відокремлення шкідливого трафіку від нормального. У випадку web-атак спостерігається інша ситуація. Дерева рішень демонструють

здатність виявляти більшість атак, однак генерують значну кількість хибних спрацювань. Ансамблеві методи, навпаки, забезпечують високу точність спрацювань, але допускають більшу кількість пропусків.

Таким чином, результати свідчать про необхідність урахування компромісу між хибними спрацюваннями та пропусками атак при виборі методу детектування. У практичному застосуванні це означає, що налаштування моделі повинно відповідати вимогам до роботи системи моніторингу та допустимому рівню ризику.

Отже, використання деревоподібних методів машинного навчання у поєднанні з аналізом flow-ознак дозволяє ефективно вирішувати задачу виявлення шкідливої активності у корпоративних мережах. Подальші дослідження доцільно спрямувати на покращення відокремлюваності складних типів атак, оптимізацію ознак та адаптацію моделей до змін у структурі мережевого трафіку.

Перелік посилань:

1. A Comparative Study of Two-Stage Intrusion Detection Using Modern Machine Learning Approaches on the CSE-CIC-IDS2018 Dataset. *MDPI*. URL: https://www.mdpi.com/2673-9585/5/1/6?utm_source=chatgpt.com

2. RandomForestClassifier. *scikit-learn*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

3. High Order Residual Distribution Conservative Finite Difference HWENO Scheme for Steady State Problems. *arXiv.org*. URL: https://arxiv.org/abs/2104.06731

Гришко В.Л.
студентка групи БСЗМ-61 СТКБ,
ННІКБЗІ, ДУІКТ
Київ, Україна

АНАЛІЗ ЗАГРОЗ БЕЗПРОВОДНИХ МЕРЕЖ ТА ЗАСОБИ ЗАХИСТУ

У роботі розглянуто актуальні загрози безпеці бездротових мереж Wi-Fi, зумовлені особливостями їх функціонування у відкритому радіоефірі. Проаналізовано основні типи атак, зокрема перехоплення трафіку, атаки типу «людина посередині», підроблені точки доступу (Evil Twin) та деавтентифікаційні атаки. Проаналізовано сучасні методи і засоби забезпечення безпеки, включаючи використання протоколів WPA2/WPA3, механізмів автентифікації IEEE 802.1X, захисту керуючих кадрів та систем моніторингу бездротового середовища. Обґрунтовано необхідність комплексного підходу до захисту Wi-Fi мереж, який поєднує технічні та організаційні заходи.

Ключові слова: бездротові мережі, Wi-Fi, мережеві атаки, Man-in-the-Middle, Evil Twin, deauthentication атаки, захист мереж.

Бездротові мережі Wi-Fi стали невід'ємною складовою сучасної інформаційної інфраструктури, забезпечуючи мобільність користувачів і гнучкість розгортання мережевих сервісів. Їх широке застосування у корпоративному середовищі, освітніх установах та побутових умовах супроводжується зростанням кількості пристроїв, підключених до мережі, що, у свою чергу, розширює поверхню атаки.

Актуальність проблеми підтверджується сучасними статистичними даними, які демонструють як масштаб використання бездротових мереж, так і зростання кількості їх вразливостей. Зокрема, за результатами галузевих досліджень кількість вразливостей у бездротових технологіях зросла з одиничних випадків на початку 2010-х років до понад 900 у 2025 році, що свідчить про суттєве ускладнення бездротового середовища та підвищення інтересу зловмисників до атак на Wi-Fi інфраструктуру [1].

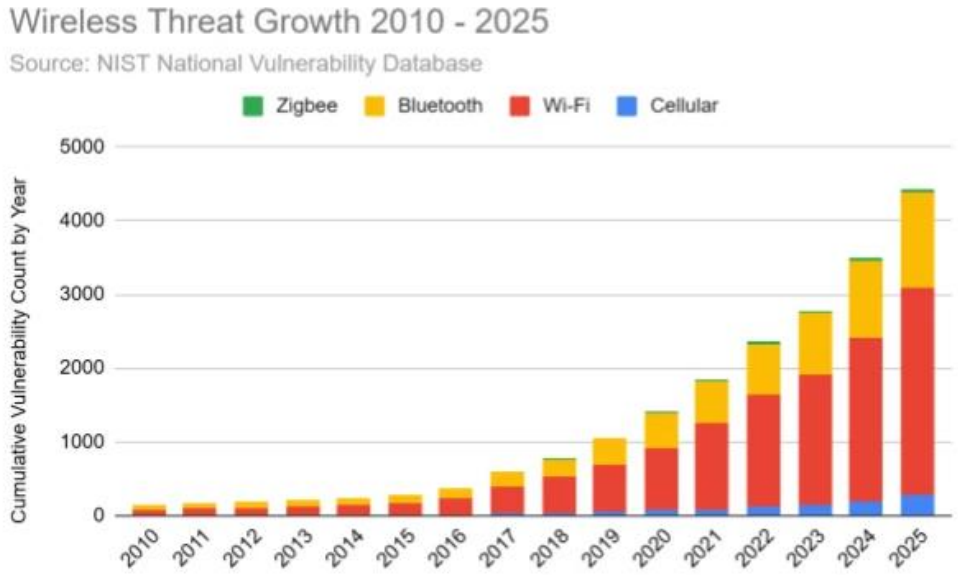

Рис.1. Зростання загроз бездротових мереж [1]

Додатково емпіричні дослідження реальних бездротових мереж показують, що значна частина з них залишається недостатньо захищеною. Наприклад, у вибірках великих міських мереж встановлено, що близько чверті точок доступу працюють без належного шифрування,

тоді як частка впровадження сучасного стандарту WPA3 залишається відносно низькою [2]. Це створює умови для реалізації атак, пов’язаних із перехопленням трафіку та несанкціонованим доступом.

Ключовою особливістю функціонування Wi-Fi мереж є передача даних у спільному середовищі, доступному для будь-якого пристрою в межах радіусу дії точки доступу. Це створює передумови для реалізації атак, спрямованих на порушення конфіденційності, цілісності та доступності інформації. Однією з найбільш поширених загроз є перехоплення мережевого трафіку. У випадках використання слабких або застарілих протоколів шифрування, таких як WEP, зловмисник може отримати доступ до переданих даних шляхом пасивного прослуховування каналу. Навіть за використання більш сучасних протоколів, відкриті мережі без автентифікації залишаються вразливими до збору метаданих і аналізу трафіку.

Суттєву загрозу становлять атаки типу «людина посередині» (Man-in-the-Middle), при яких зловмисник перехоплює та модифікує трафік між клієнтом і точкою доступу. Така атака може реалізовуватися шляхом підміни мережевих параметрів або створення умов, за яких клієнт підключається до контрольованого вузла. Одним із практичних варіантів реалізації подібної атаки є створення підробленої точки доступу, відомої як Evil Twin. У цьому випадку зловмисник розгортає точку доступу з ідентичними параметрами легітимної мережі, змушуючи користувачів підключатися до неї, після чого отримує можливість перехоплення облікових даних та аналізу трафіку.

Окрему категорію становлять атаки на доступність мережі, зокрема деавтентифікаційні атаки. Вони базуються на особливостях протоколів стандарту IEEE 802.11, які передбачають можливість надсилання службових кадрів деавтентифікації без належного криптографічного захисту. Це дозволяє зловмиснику примусово відключати клієнтів від мережі, створюючи відмову в обслуговуванні або формуючи умови для подальших атак [3].

Наведені загрози свідчать про необхідність комплексного підходу до забезпечення безпеки бездротових мереж.

Базовим рівнем захисту є використання сучасних протоколів шифрування, визначених стандартами сімейства IEEE 802.11. Застарілий протокол WEP характеризується критичними криптографічними вразливостями, зокрема використанням слабкого алгоритму RC4 та коротких векторів ініціалізації, що дозволяє відновити ключ шифрування за відносно короткий час. У зв’язку з цим його використання є неприйнятним у сучасних умовах. Протокол WPA2, який базується на алгоритмі AES у режимі CCMP, забезпечує конфіденційність і цілісність переданих даних за умови використання надійних облікових даних. Водночас впровадження WPA3 дозволяє підвищити рівень захисту завдяки використанню протоколу SAE (Simultaneous Authentication of Equals), який забезпечує стійкість до офлайн-атак перебору паролів та унеможливлює відновлення ключа навіть у разі компрометації попередніх сеансів.

У корпоративних мережах доцільним є застосування централізованої автентифікації на основі стандарту IEEE 802.1X, що дозволяє реалізувати контроль доступу на рівні користувачів і пристроїв. Такий підхід забезпечує індивідуальну ідентифікацію клієнтів і зменшує ризик компрометації облікових даних. Додатково важливим є впровадження систем моніторингу бездротового середовища, які дозволяють виявляти несанкціоновані точки доступу та аномальну активність.

Значну роль у підвищенні рівня безпеки відіграє сегментація мережі. Розподіл користувачів на ізольовані сегменти, а також використання гостьових мереж дозволяють

обмежити можливість поширення атаки в межах інфраструктури. Такий підхід є ефективним засобом мінімізації наслідків компрометації окремих вузлів.

Окрім технічних заходів, важливим елементом є організаційна складова забезпечення безпеки. Значна частина інцидентів пов'язана з діями користувачів, які підключаються до незахищених мереж або ігнорують базові правила безпеки. Це підкреслює необхідність підвищення рівня обізнаності користувачів та впровадження політик безпечного використання бездротових мереж [4].

Таким чином, бездротові мережі Wi-Fi залишаються вразливими до широкого спектра загроз, обумовлених як особливостями технології, так і людським фактором. Зростання кількості вразливостей та недостатній рівень впровадження сучасних механізмів захисту підсилюють актуальність проблеми. Ефективне забезпечення безпеки можливе лише за умови комплексного підходу, який поєднує технічні засоби захисту, контроль доступу, моніторинг мережевої активності та організаційні заходи.

Перелік посилань:

1. Wireless security vulnerabilities report 2026. Help Net Security. URL: https://www.helpnetsecurity.com/2026/03/12/report-wireless-security-vulnerabilities-2026/.

2. In Numeris Veritas: An empirical measurement of Wi-Fi integration in industry. ResearchGate. URL: https://www.researchgate.net/publication/395726215.

3. Detection of deauthentication attacks in Wi-Fi networks. Nature Scientific Reports. URL: https://www.nature.com/articles/s41598-025-18947-2.pdf.

4. Top wireless-enabled threats in 2025. Bastille Networks. URL: https://bastille.net/wp-content/uploads/Top-Wireless-Enabled-Threats-in-2025-1.pdf.

Павлюк П.О.
Студент групи БСД-42, ННІКБЗІ, ДУІКТ
Київ, Україна

СИСТЕМА ІДЕНТИФІКАЦІЇ КОРИСТУВАЧІВ ДО РЕСУРСІВ ПРИВАТНОЇ ОРГАНІЗАЦІЇ

У роботі розглянуто підходи до побудови системи ідентифікації користувачів у приватній організації як складової забезпечення контрольованого доступу до інформаційних ресурсів. Проаналізовано основні загрози, пов'язані з компрометацією облікових записів, та обґрунтовано необхідність використання багатофакторної автентифікації. Розглянуто інтеграцію систем ідентифікації з корпоративною інфраструктурою, зокрема застосування централізованих служб каталогів і технологій єдиного входу. Описано підходи до контролю доступу, аналізу поведінкових характеристик користувачів та виявлення аномалій. Окрему увагу приділено захисту облікових даних і управлінню життєвим циклом облікових записів. Визначено ключові вимоги до ефективної системи ідентифікації користувачів.

Ключові слова: ідентифікація користувачів, автентифікація, контроль доступу, кібербезпека, багатофакторна автентифікація, інформаційні ресурси.

Забезпечення контрольованого доступу до інформаційних ресурсів є базовою складовою системи кібербезпеки будь-якої приватної організації. В умовах цифровізації бізнес-процесів значна частина критичних даних зберігається у корпоративних інформаційних системах, що робить їх привабливою ціллю для зловмисників. Основним вектором компрометації залишається отримання несанкціонованого доступу до облікових записів

користувачів, що обумовлює необхідність побудови ефективної системи ідентифікації та автентифікації.

Система ідентифікації користувачів включає сукупність механізмів, що забезпечують встановлення особи користувача, перевірку її автентичності та надання відповідних прав доступу. Вона є складовою більш широкої системи управління доступом і повинна функціонувати у взаємозв'язку з іншими компонентами безпеки, такими як системи моніторингу, журнали подій та засоби виявлення вторгнень [1].

Класичний підхід до автентифікації базується на використанні одного фактору, зазвичай пароля. Однак практика показує, що паролі є вразливими до широкого спектра атак, зокрема перебору, фішингу, витоків баз даних та соціальної інженерії. У зв'язку з цим сучасні системи переходять до використання багатофакторної автентифікації, яка передбачає поєднання кількох незалежних факторів: знання, володіння та біометричних характеристик. Такий підхід значно підвищує стійкість системи до компрометації.

Важливим напрямом розвитку є інтеграція систем ідентифікації з корпоративною інфраструктурою. Використання централізованих служб каталогів дозволяє реалізувати єдину точку управління обліковими записами та політиками доступу. Технології єдиного входу забезпечують користувачам доступ до декількох ресурсів після одноразової автентифікації, що підвищує зручність використання системи без зниження рівня безпеки. При цьому широко застосовуються стандартизовані протоколи автентифікації, які забезпечують взаємодію між різними компонентами інформаційної системи [2].

Окрему роль відіграють моделі контролю доступу. Рольова модель дозволяє групувати користувачів за функціональними обов'язками та призначати їм відповідні права. Атрибутна модель, у свою чергу, забезпечує більш гнучкий підхід, враховуючи контекст доступу, такі як час, місце, тип пристрою або рівень довіри до користувача. Поєднання цих підходів дозволяє реалізувати адаптивну політику безпеки.

Сучасні системи також враховують поведінкові характеристики користувачів. Аналіз таких параметрів, як частота входів, географічне розташування, типові дії в системі, дозволяє виявляти аномалії, які можуть свідчити про компрометацію облікового запису. У таких випадках система може ініціювати додаткову перевірку або обмежити доступ до ресурсів [3].

Значну увагу необхідно приділяти захисту облікових даних. Зберігання паролів у відкритому вигляді є неприпустимим, тому використовуються криптографічні алгоритми хешування з додаванням солі. Передача даних між клієнтом і сервером повинна здійснюватися захищеними каналами зв'язку, що запобігає перехопленню облікових даних. Додатково впроваджуються механізми обмеження кількості спроб входу та автоматичного блокування облікового запису у разі підозрілої активності.

Важливим аспектом є також забезпечення життєвого циклу облікових записів. Це включає створення, модифікацію та видалення облікових записів відповідно до змін у статусі користувача. Неналежне управління цим процесом може призвести до існування неактивних або забутих облікових записів, які можуть бути використані зловмисниками [4].

Таким чином, система ідентифікації користувачів у приватній організації повинна розглядатися як багаторівнева структура, що поєднує технічні, організаційні та аналітичні механізми. Її ефективність визначається не лише використанням сучасних технологій, але й правильним налаштуванням політик безпеки та постійним контролем їх дотримання.

Розробка системи ідентифікації користувачів для приватної організації потребує комплексного підходу, який враховує сучасні кіберзагрози та особливості корпоративної інфраструктури. Використання багатофакторної автентифікації, централізованих систем управління доступом та механізмів аналізу поведінки користувачів дозволяє суттєво знизити ризики несанкціонованого доступу. Запропоновані підходи можуть бути використані для підвищення рівня інформаційної безпеки організацій різного масштабу.

Перелік посилань:

1. Stallings W. Cryptography and network security: principles and practice. 7th ed. Harlow: Pearson Education Limited, 2017. 768 p.

2. Bishop M. Computer security: art and science. 2nd ed. Boston: Addison-Wesley, 2018. 1136 p.

3. Digital identity guidelines: authentication and lifecycle management (SP 800-63B). Gaithersburg: National Institute of Standards and Technology, 2020. 140 p.

4. Ferraiolo D., Kuhn D., Chandramouli R. Role-based access control. 2nd ed. Norwood: Artech House, 2007. 300 p.

Якименко Ю.М.
викладач кафедри УКБЗІ, ННІКБЗІ ДУІКТ,
Київ,Україна

АНАЛІЗ АВТОМАТИЗОВАНОГО ПІДХОДУ ДО РОЗСЛІДУВАННЯ ІНЦИДЕНТІВ В DLP-СИСТЕМІ

Підходи до автоматизації процесів розслідування інцидентів інформаційної безпеки залишаються актуальним в сучасних умовах підвищення кіберзагроз для нормального функціонування інформаційних систем. Процес керування інцидентами є ключовим постачальником інформації, як для модернізації СУІБ, так і для оцінки ефективності використовуваних заходів захисту інформаційних ресурсів організації. В сучасних DLP-Системах розвиток іде саме в напрямку автоматизації цього процесу, тому є актуальнішим і на сьогодні. Пропонуються напрями покращення процесів управління інцидентами інформаційної безпеки завдяки розгляду підходів до розслідування інцидентів.

Ключові слова: автоматизація, безпека, інцидент, інформаційна система, управління.

З ростом числа інформаційних систем і вдосконалюванням інформаційних технологій росте й число інцидентів інформаційної безпеки (ІБ), під якими розуміється одне або кілька небажаних подій безпеки, які впливають на інформаційну безпеку активів систем і можуть привести до негативних наслідків. Такими наслідками можуть бути, наприклад, порушення конфіденційності, цілісності й доступності інформаційних ресурсів, переривання бізнес-процесів і ін. Будь-яка використовувана в організації система захисту інформації, або підготовлена політика ІБ, не може виключати події, які є загрозою для оброблюваної інформації. При експлуатації систем управління інформаційною безпекою (СУІБ) процес керування інцидентами є ключовим постачальником інформації, як для модернізації СУІБ, так і для оцінки ефективності використовуваних заходів захисту інформаційних ресурсів. [1,2]

Організувати даний процес без використання засобів автоматизації представляється важко розв'язуваним завданням, особливо в інформаційних системах організацій. Розробка способу підвищення оперативності аналізу і оцінки інцидентів ІБ можлива тільки за допомогою їх автоматизації та забезпечення адекватності ухвалення рішення по обробці інцидентів ІБ за рахунок інформаційного забезпечення - процедур їх виявлення, аналізу і оцінки на основі бази даних про події і інциденти ІБ. [3]

Розвиток сучасних DLP-Систем іде саме в напрямку автоматизації цього процесу, тому є актуальнішим в процесах управління ІБ. Такі системи вміють збирати, зберігати і категорувати події ІБ, виявляти критичні інциденти, повідомляти про них, надавати зручний інструмент для збору додаткової інформації і в першу чергу – забезпечення оперативного проведення розслідування інцидентів та багато іншого в сфері безпеки. [4]

Але показовим є те, що єдиної методики розслідування інцидентів не існує, організаціями використовуються тільки загальні рекомендації.

Якщо опиратися на вимоги таких нормативних документів, як ДСТУ ISO/IEC 27002, ДСТУ ISO/IEC 27035, NIST SP 800-61 щодо управлінських процесів, то вони пропонуються в основному тільки у вигляді «кращих практик» з побудови СУІБ. [5,6] При роботі із системою

DLP в напрямах покращення процесів управління інцидентами ІБ, особливо в розслідуванні інцидентів такими практиками можна запропонувати наступні процеси:

1. Виявлення і реєстрація подій ІБ.
2. Категорування подій, збір додаткової інформації й виявлення інцидентів ІБ.
3. Оперативне реагування на інцидент (запобігання або усунення наслідків інциденту).
4. Розслідування) інциденту.
5. Реагування на інцидент.
6. Аналіз причин інциденту і отриманих результатів.
7. Підготовка рекомендацій з підвищення загального рівня ІБ (при необхідності).

Збільшення на сьогодні потоку подій ІБ вимагають уваги від компаній-вендорів, щоб сучасні DLP мали розвинені засоби автоматизації керування життєвим циклом інциденту, елементи кейс-менеджменту, інструменти пошуку і аналізу інформації на достатньому рівні для проведення якісного та достатньо глибинного розслідування. Для автоматизації процесу аналізу й оцінки інцидентів ІБ краще розробити автономний програмний засіб**.** Процес для розслідування повинен містить у собі засоби виявлення, вилучення, збереження, документальне оформлення і розшифрування комп'ютерних носіїв - для проведення аналізу доказів і основних причин події та інциденту ІБ.

Перелік використаної літератури

1. Якименко Ю.М., Легомінова С.В., Щавінський Ю.В., Рабчун Д.І. Управління інцидентами інформаційної безпеки. Сучасні методи і засоби: навчальний посібник. Київ: Державний університет телекомунікацій, 2023. 241с.

2. Легомінова С.В., Якименко Ю.М., Мужанова Т. М., Капелюшна Т.В. Вплив управління інцидентами на функціонування системи управління інформаційною безпекою організації. Телекомунікаційні та інформаційні технології. 2025. № 1 (86). С.75-81.

3. Корченко О.Г., Гнатюк С.О., Казмірчук С.В. та ін. Аудит та управління інцидентами інформаційної безпеки: навч.посіб.– Київ : Центр навч.-наук. та наук.-пр. видань НАСБ України, 2014. –190с. URL: https://er.nau.edu.ua/bitstream/NAU/38027/1/Audit%26Incident_15042014.pdf.

4. Якименко Ю.М Підвищення ролі DLP - систем у розслідуванні інцидентів (кіберінцидентів) інформаційної безпеки. Всеукраїнська науково-практична конференція «Цифрова трансформація кібербезпеки» від 27 квітня 2023 року. - Київ: ДУТ, 2023.

Савченко Вадим Володимирович,
студент групи БСДМ-51, ННІКБЗІ ДУІКТ, Київ, Україна
Стожок Максим Романович,
студент групи БСДМ-51, ННІКБЗІ ДУІКТ, Київ, Україна

ШТУЧНИЙ ІНТЕЛЕКТ У РОЗРОБЦІ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ЯК ФАКТОР ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ТА ЯКОСТІ

Розвиток технологій штучного інтелекту суттєво змінює підходи до розробки програмного забезпечення. Використання алгоритмів машинного навчання, обробки

природної мови та автоматизації процесів дозволяє підвищити швидкість створення продуктів, зменшити кількість помилок та оптимізувати витрати ресурсів. ШІ стає не допоміжним інструментом, а ключовим елементом сучасних процесів розробки [1] .

Ключові слова: штучний інтелект, розробка ПЗ, автоматизація, машинне навчання, DevOps.

Сучасна розробка програмного забезпечення характеризується високою складністю та швидкими темпами змін. Команди стикаються з необхідністю швидко створювати якісні продукти в умовах обмежених ресурсів. Штучний інтелект дозволяє вирішувати ці задачі за рахунок автоматизації рутинних процесів, аналізу великих обсягів даних та підтримки прийняття рішень. Інтеграція ШІ у життєвий цикл розробки програмного забезпечення змінює підхід до проєктування, тестування та супроводу систем.

Автоматизація процесів розробки за допомогою ШІ

Одним із ключових напрямів використання ШІ є автоматизація написання коду. Інструменти на основі машинного навчання здатні генерувати фрагменти коду, пропонувати виправлення та оптимізації. Це дозволяє розробникам скоротити час на виконання рутинних задач до 30–40 відсотків. Крім того, автоматизовані системи аналізу коду виявляють помилки ще на ранніх етапах, що знижує витрати на їх виправлення [2].

ШІ також активно використовується у тестуванні програмного забезпечення. Інтелектуальні системи можуть автоматично створювати тест-кейси, визначати критичні сценарії та аналізувати результати тестування. Це підвищує покриття тестами та зменшує ймовірність пропуску критичних помилок.

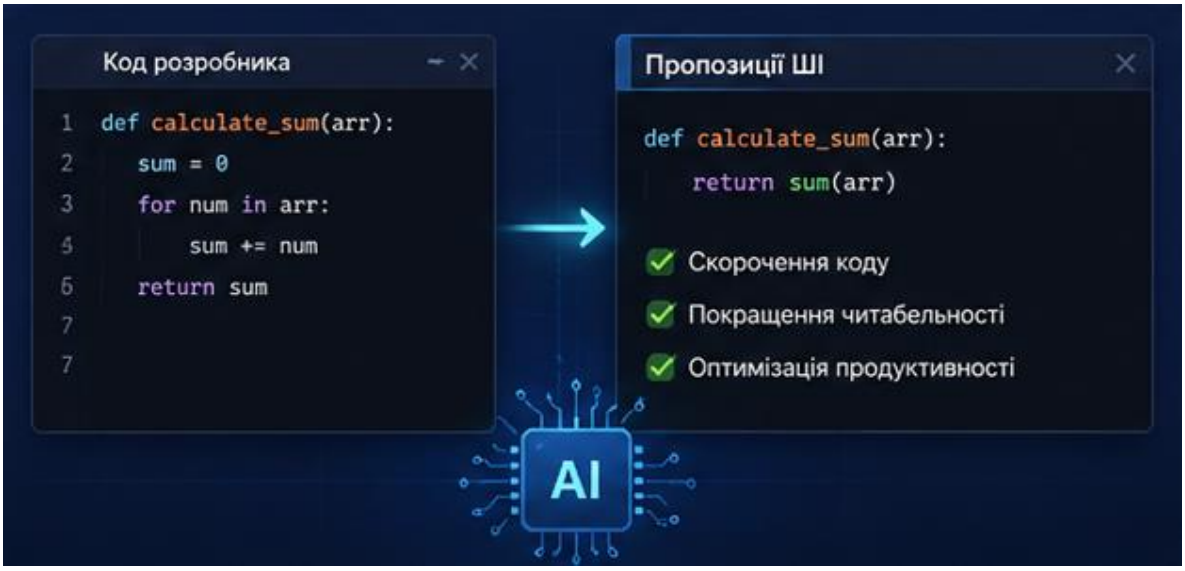

Рис. 1. Автоматизація процесів розробки за допомогою ШІ

Підтримка прийняття рішень у розробці

Штучний інтелект допомагає аналізувати великі обсяги даних, які виникають у процесі розробки. Це включає журнали помилок, метрики продуктивності та історію змін коду. На основі цих даних системи ШІ можуть прогнозувати ризики, визначати слабкі місця та пропонувати оптимальні рішення .

Наприклад, аналіз історичних даних дозволяє оцінити ймовірність виникнення дефектів у певних модулях. Це дає змогу зосередити ресурси на найбільш критичних частинах системи. Такий підхід підвищує ефективність управління проєктами та зменшує кількість критичних збоїв.

Інтеграція ШІ в DevOps та CI/CD

У сучасних практиках DevOps штучний інтелект відіграє важливу роль у автоматизації процесів безперервної інтеграції та доставки. Інтелектуальні системи можуть аналізувати зміни в коді, визначати потенційні проблеми та автоматично запускати необхідні перевірки.

ШІ також використовується для моніторингу систем у реальному часі. Алгоритми здатні виявляти аномалії в роботі програмного забезпечення та оперативно реагувати на них. Це дозволяє зменшити час простою систем і підвищити їх надійність [4].

Рис. 2. Архітектура **DevOps** з використанням ШІ

Вплив ШІ на якість програмного забезпечення

Використання штучного інтелекту безпосередньо впливає на якість програмних продуктів. Автоматичний аналіз коду дозволяє дотримуватися стандартів розробки та уникати типових помилок. Інтелектуальні системи тестування забезпечують більш повне покриття та виявлення складних дефектів.

Дослідження показують, що використання ШІ у тестуванні може зменшити кількість дефектів у продакшені до 50 відсотків. Це значно підвищує довіру користувачів до продукту та знижує витрати на його підтримку [1].

Виклики та обмеження використання ШІ

Попри значні переваги, впровадження ШІ у розробку ПЗ має певні виклики. Одним із них є залежність від якості даних, на яких навчаються моделі. Неправильні або неповні дані можуть призвести до некоректних результатів [3].

Також важливим аспектом є необхідність адаптації команд до нових інструментів. Використання ШІ вимагає нових навичок та підходів до роботи. Крім того, існують питання безпеки та конфіденційності даних, які потребують особливої уваги.

Перспективи розвитку

Штучний інтелект продовжує активно розвиватися та відкриває нові можливості для розробки програмного забезпечення. Очікується подальше вдосконалення інструментів генерації коду, автоматичного тестування та аналізу систем.

У майбутньому ШІ може стати повноцінним учасником процесу розробки, здатним самостійно створювати складні програмні рішення. Це змінить роль розробника, який зосередиться на постановці задач та контролі результатів.

Інтеграція штучного інтелекту в розробку програмного забезпечення вже сьогодні забезпечує суттєві переваги. Вона підвищує ефективність процесів, покращує якість продуктів та дозволяє швидше реагувати на зміни. Подальший розвиток цих технологій визначатиме майбутнє індустрії програмного забезпечення.

Перелік посилань:

1. McKinsey & Company. The economic potential of generative AI: The next productivity frontier. URL: https://www.mckinsey.com (дата звернення: 20.04.2026).
2. GitHub. The State of AI in Software Development. URL: https://github.blog (дата звернення: 20.04.2026).
3. IEEE. Artificial Intelligence in Software Engineering. URL: https://ieeexplore.ieee.org (дата звернення: 20.04.2026).
4. IBM. AI for Software Development. URL: https://www.ibm.com (дата звернення: 20.04.2026).

Кузнецов П.О.
студент групи ПД-41, ННІІТ, ДУІКТ,
Київ, Україна

ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ІНФОРМАЦІЇ У WEB-ЗАСТОСУНКУ "ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ РЕКОМЕНДАЦІЇ КІНОФІЛЬМІВ ЗА ВПОДОБАННЯМИ КОРИСТУВАЧІВ"

У сучасному цифровому середовищі web-застосунки є невід'ємною частиною повсякденного життя користувачів. Вони забезпечують доступ до різноманітних сервісів, зокрема медіаконтенту, але водночас стають об'єктом потенційних кіберзагроз. Тому забезпечення конфіденційності, цілісності та доступності інформації є важливим завданням при розробці web-застосунків [1].

У даній роботі розглядається web-застосунок для пошуку фільмів та формування персонального списку перегляду. Система дозволяє користувачам здійснювати пошук фільмів за назвою, фільтрувати їх за жанрами, роками та рейтингом, а також зберігати обрані фільми у власному списку.

Незважаючи на відсутність складної реєстрації, застосунок оперує користувацькими даними зокрема налаштуваннями та списком фільмів, що потребує належного рівня захисту.

Однією з основних загроз для web-застосунку є атаки типу XSS, які дозволяють зловмисникам впроваджувати шкідливий код у сторінки. Для запобігання цьому всі дані, що вводяться користувачем наприклад, пошукові запити, проходять перевірку та обробку перед відображенням. Також застосовується обмеження на вставку небезпечного HTML-коду.

Іншою поширеною загрозою є несанкціонований доступ до даних. У застосунку дані користувача список перегляду, фільтри зберігаються в локальному сховищі браузера. Для підвищення безпеки використовується перевірка даних перед їх записом та зчитуванням, що дозволяє уникнути пошкодження або підміни інформації.

Застосунок взаємодіє із зовнішнім API TMDB, що також створює певні ризики. Для їх мінімізації використовується захищене з'єднання HTTPS, яке забезпечує шифрування переданих даних і запобігає їх перехопленню [2]. Крім того, API-ключ використовується лише для запитів і не надає доступу до конфіденційної інформації користувачів.

Для забезпечення цілісності даних застосовується контроль коректності отриманої інформації перевірка наявності постера або дати релізу фільму перед відображенням. Це дозволяє уникнути помилок у роботі інтерфейсу та підвищує стабільність системи.

Також важливим аспектом є захист інтерфейсу користувача. Реалізовано обмеження повторного додавання фільмів у список перегляду, що запобігає дублюванню даних. Кнопки взаємодії змінюють свій стан змінюють колір після додавання, що дозволяє користувачу контролювати свої дії та уникати помилок.

Отже, у процесі розробки web-застосунку для пошуку фільмів було враховано основні принципи інформаційної безпеки. Реалізовані механізми захисту дозволяють знизити ризики несанкціонованого доступу, витоку або пошкодження даних. Подальший розвиток системи може включати впровадження авторизації користувачів, розширених механізмів шифрування та серверного зберігання даних.

Перелік посилань:

1. Бондаренко О., Ушкаленко І. (2017). Безпека web-додатків: актуальні проблеми та їх аналіз. Формування ринкової економіки в Україні, Вип. 38, 28-36.
2. Боскін О.О., Корніловська Н.В., Поліщук В.М., Сарафаннікова Н.В. (2023). Безпека веб-додатків та хакерські атаки. Вісник Херсонського національного технічного університету, №3(86), 83–92. https://doi.org/10.35546/kntu2078-4481.2023.3.11
3. OWASP Foundation. OWASP Top Ten Web Application Security Risks. - Режим доступу: https://owasp.org

Душник Володимир Володимирович
Студент групи БСДМ-52, ННІКБЗІ ДУІКТ,
Київ, Україна

КІБЕРРОЗВІДКА ТА OSINT ЗА ДОПОМОГОЮ ШІ

Кіберрозвідка та OSINT із використанням штучного інтелекту є сучасним напрямом інформацій-ної аналітики, що поєднує збір даних із відкритих джерел із можливостями автоматизованого аналізу. Завдяки технологіям машинного навчання та обробки природної мови стає можливим швидке опра-цювання великих обсягів інформації, виявлення прихованих закономірностей, аналіз текстового й ме-діаконтенту, а також ідентифікація потенційних загроз. Такий підхід суттєво підвищує ефективність моніторингу інформаційного простору та відіграє важливу роль у забезпеченні кібербезпеки в умовах зростання цифрових ризиків.

Ключові слова: OSINT, ШІ, Кібберрозвідка.

У сучасному цифровому середовищі обсяги інформації зростають над-звичайно швидкими темпами, що створює нові виклики для її аналізу та ви-користання. Кіберрозвідка та розвідка з відкритих джерел (OSINT) стають ва-жливими інструментами для отримання актуальних даних про події, явища та потенційні загрози. Інтеграція штучного інтелекту в ці процеси відкриває но-ві можливості для автоматизації збору інформації, глибокого аналізу даних і виявлення прихованих закономірностей, що значно підвищує ефективність до-сліджень у сфері безпеки та інформаційних технологій.

Що таке AI OSINT?

Штучний інтелект у сфері OSINT можна описати як сучасний підхід до збору розвідувальної інформації, який підсилюється можливостями AI. Такі те-хнології можна використовувати на різних етапах розслідувань і в різних ситу-аціях, зокрема для:

Виявлення людей, місць, організацій та інших важливих об’єктів у вели-ких масивах даних. Пошуку аномалій або підозрілих елементів у зображеннях чи текстах. Наприклад, AI здатен визначити, чи було зображення створене шту-чно, а також покращувати якість фото, збільшувати їх або допомагати розпізна-ти нечіткі деталі.

Безперервного моніторингу джерел інформації з автоматичними спові-щеннями в реальному часі. Аналізу текстів за допомогою технологій обробки природної мови (NLP), що дозволяє визначати тональність і зміст без ручної роботи. NLP - це галузь машинного навчання, яка навчає системи розуміти й

обробляти людську мову, і використовується, наприклад, у перекладачах, чат-ботах та аналітиці текстів.

Як взаємодіють OSINT та AI

OSINT і штучний інтелект добре доповнюють один одного. Оскільки OSINT передбачає роботу з великими об-сягами публічної інформації, це може бути складно і трудомістко. Інструменти AI, такі як веб-краулери та скрапери, значно спрощують цей процес, дозволя-ючи дослідникам зосередитися на аналізі даних або роботі з іншими видами розвідки, наприклад HUMINT.

Коли варто застосовувати AI в OSINT

Використання AI є особливо ефективним у таких аспектах дослідження:

- автоматизація повторюваних завдань — обробка великих масивів да-них значно швидша, ніж вручну.
- збір інформації — AI може автоматично отримувати дані з новин, со-ціальних мереж, архівів та інших відкритих джерел.
- прогнозування — аналіз закономірностей допомагає робити припуще-нння щодо майбутніх подій.
- аналіз тексту (NLP) — виділення ключових ідей та слів із великих об-сягів інформації.
- виявлення патернів — AI здатен знаходити приховані тенденції, на-приклад у поведінці користувачів соцмереж.
- аналіз медіа — обробка фото та відео для отримання додаткової інфор-мації, включаючи розпізнавання облич чи інтерпретацію мови.

Що може визначати AI

У межах OSINT штучний інтелект допомагає:виявляти дипфейки (аналіз обличчя, звуку, структури зображення),визначати геолокацію за об’єктами та середовищем на фото і відео, перетворювати аудіо в текст і аналізувати зміст, ключові слова та інтонацію, оцінювати натовп (кількість, щільність, характери-стики), знаходити схожі зображення для встановлення джерела та відстеження поширення.

Перелік посилань:

1. “OSINT AI: Як Оптимізувати Своє Розслідування у 2025 році?. URL:https://molfar.com/blog/osint-ai-yak-optymizuvaty-svoe-rozsliduvannya-u-2024-roci

2. ”Штучний інтелект в OSINT: Як покращити розслідування у 2024 році” URL:https://infolight.ua/2024/12/01/shtuchnyj-intelekt-v-osint-yak-pokrashhyty-rozsliduvannya-u-2024-rotsi/

Закаблук Е.Є.
студент, групи КНД-43, ННІІТ, ДУІКТ,
Київ, Україна

ОРГАНІЗАЦІЯ ЗАХИЩЕНОГО ПРОСТОРУ ДЛЯ ЗБЕРІГАННЯ ТА ОБМІНУ КОРПОРАТИВНИМИ ДАНИМИ НА БАЗІ ПРОГРАМНИХ РІШЕНЬ З ВІДКРИТИМ КОДОМ

У роботі досліджено практичну реалізацію засобів контролю доступу, захищеного обміну файлами та колаборації у Nextcloud Hub як корпоративної платформи зберігання даних з відкритим вихідним кодом. Розглянуто механізми розмежування прав між підрозділами через групи LDAP/Active Directory, захист великих файлів засобами публічних посилань з одноразовою автентифікацією, криптографічні вразливості наскрізного шифрування та відповідні контрзаходи. Визначено переваги self-hosted розгортання порівняно з публічними хмарними сервісами в контексті фізичного контролю над корпоративними даними.

Ключові Слова: відкритий вихідний код, корпоративні дані, хмарні сховища.

Постановка задачі. Підприємство щодня формує документи. Розміщення цих файлів на публічних хмарних сервісах означає передачу фізичного контролю над корпоративними даними стороннім серверам. Підписна модель хмарних сервісів створює постійне фінансове навантаження, яке суттєво зростає зі збільшенням кількості співробітників. Окремою проблемою залишається розмежування доступу між підрозділами, бухгалтерія не повинна бачити матеріали відділу продажів, тоді як керівник має доступ до всього. Nextcloud вирішує ці задачі в єдиному власному середовищі.

Мета дослідження. Дослідити засоби контролю доступу, захищеного обміну файлами та колаборації у Nextcloud Hub; визначити механізми розмежування прав між підрозділами та оцінити їхню достатність для захисту корпоративних даних.

Результати дослідження. Nextcloud реалізує ієрархічну файлову структуру: адміністратор створює папки верхнього рівня відповідно до організаційної схеми підприємства, наприклад: /Фінанси, /Маркетинг, /Юридичний, /Спільне. Ключовий інструмент розмежування – групи та права доступу на рівні папки. Кожній папці призначається від одного до чотирьох рівнів дозволів: лише перегляд, завантаження файлів, редагування, повне управління. Групи формуються через панель адміністратора та можуть синхронізуватися з корпоративним LDAP/Active Directory через вбудований модуль User Auth, що суттєво спрощує підготовку нових співробітників у компаніях з існуючою AD-інфраструктурою [1, с. 30].

Для передачі важких файлів – відеороликів, поліграфічних макетів у форматі TIFF/AI, архівів вихідних матеріалів – Nextcloud пропонує механізм публічних посилань з розширеними параметрами безпеки. Адміністратор або власник папки може: встановити термін дії посилання; задати пароль доступу; дозволити лише завантаження без перегляду файлів; активувати водяний знак для зображень. Починаючи з Nextcloud Hub 7, посилання підтримують аутентифікацію через одноразовий код, надісланий на email отримувача, що усуває ризик несанкціонованого перетворення публічного посилання на відкрите. Порівняльне тестування Seafile і Nextcloud показало, що Nextcloud демонструє стабільніші показники завантаження великих файлів при збереженні наскрізного шифрування [2, с. 85].

Дані на сервері зашифровуються за допомогою AES-256. Наскрізне шифрування (E2EE) реалізоване на рівні папки через механізм публічних ключів. Тут важливо враховувати, що оригінальна реалізація E2EE у Nextcloud до версії 3.12 містила вразливість, пов’язану з відсутністю автентичності шифрування публічним ключем RSA-OAEP: зловмисний сервер міг підміняти зашифровані ключі файлів [3, с. 829–830]. Вразливість була частково усунена в оновленнях, втім для чутливих даних рекомендується використовувати SSE разом із VPN-тунелем, а E2EE застосовувати лише для ізольованих папок з документами найвищої конфіденційності. Дослідження Aryan P. демонструє, як ізоляція мережевого трафіку засобами QEMU/KVM суттєво знижує цей ризик на рівні інфраструктури [4, с. 4].

Інтеграція Nextcloud з OnlyOffice Document Server або LibreOffice перетворює сховище на повноцінний офісний простір: кілька користувачів одночасно редагують .docx, .xlsx, .pptx без завантаження файлів на локальні диски. Режим рецензування та коментування повністю сумісний з форматом Microsoft Office, що дозволяє передавати документи зовнішнім партнерам без конвертації. Self-hosted розгортання дає підприємству повний фізичний

контроль над даними при суттєво меншій вартості порівняно з публічними хмарними сервісами.

Висновки. Nextcloud забезпечує повний цикл роботи з корпоративними даними: від зберігання і розмежування доступу між підрозділами до захищеної передачі великих файлів зовнішнім контрагентам і спільного редагування документів. Рольова модель через групи LDAP/AD вирішує задачу організаційного розмежування без потреби в окремих корпоративних DRM-рішеннях. Публічні посилання з терміном дії та одноразовою автентифікацією є практичною альтернативою незахищеному пересиланню матеріалів поштою. Відомі криптографічні вразливості E2EE-компонента мають конкретні технічні контрзаходи і не знецінюють платформу загалом, однак вимагають усвідомленої конфігурації. Self-hosted розгортання дає підприємству повний фізичний контроль над даними, що є ключовою перевагою порівняно з будь-яким публічним хмарним сервісом.

Список використаних джерел

1. Сіренко. О.Є. Моделі та засоби зберігання секретних даних у хмарі : кваліфікаційна робота магістра. НУРЕ, 2021. URL: https://openarchive.nure.ua/bitstreams/1c21fc90-4e1e-4290-90e4-112b08c31f1d/download

2. Dutcher G., Azianyo K., Dissanayake T. P. B. M., Mailewa A. B. Secure Cloud Storage Solution with “Seafile” & “NextCloud”: A Resilient Efficiency Assessment. Advances in Technology. 2024. Vol. 3, no. 2. P. 75–96.
URL: https://journals.sjp.ac.lk/index.php/ait/article/view/7341

3. Albrecht M. R., Backendal M., Coppola D., Paterson K. G. Share with Care: Breaking E2EE in Nextcloud. In: IEEE Euro S&P 2024. P. 828–840. DOI: 10.1109/EuroSP60621.2024.00051. URL: https://eprint.iacr.org/2024/546

4. Aryan P., Shetty S. D. Designing a Secure, Scalable, and Cost-Effective Cloud Storage Solution: A Novel Approach to Data Management using NextCloud, TrueNAS, and QEMU/KVM. arXiv preprint. 2024. arXiv:2412.05091. URL: https://arxiv.org/abs/2412.05091

Аняляк Дем'ян Володимирович
студент 4 курсу, групи КНД-43
Державного університету
інформаційно-комунікаційних технологій

Кібербезпека корпоративних інформаційних систем

Чому безпека корпоративних вебдодатків — це не просто тренд, а необхідність Сьогодні, коли бізнес-процеси практично повністю «переїхали» в онлайн, питання захисту внутрішніх систем стало питанням виживання. Під час роботи над архітектурою своєї системи обліку заявок я прийшов до висновку, що служба техпідтримки — це одна з найпривабливіших цілей для зловмисників. Чому? Бо саме тут акумулюються дані про всі «баги» та слабкі місця інфраструктури компанії. Якщо хакер отримає доступ до такої бази, він фактично отримає мапу вразливостей усієї мережі. Ключові слова: кібербезпека, технічна підтримка, захист даних, цифрова інфраструктура.

Під час проєктування я сфокусувався на трьох моментах, які зазвичай стають «точкою входу» для атак:

По-перше, це слабкі або скомпрометовані паролі — класика, яка досі працює.

По-друге, це ін'єкції (SQLi та XSS), коли через звичайне поле вводу в заявці зловмисник намагається «пропихнути» свій код.

І по-третє, це проблема надлишкових прав, коли звичайний юзер може випадково (або спеціально) залізти в налаштування адмінки.

Як я реалізував захист у Flask, вибір Python та Flask був свідомим. Я хотів побудувати систему, де безпека працює «під капотом», не заважаючи користувачу. В результаті вдалося впровадити кілька важливих рішень. Наприклад, я повністю відмовився від зберігання паролів у відкритому вигляді — тепер використовуються криптографічні хеші бібліотеки Werkzeug. Навіть якщо база даних потрапить у чужі руки, паролі залишаться зашифрованими.

Крім того, використання SQLAlchemy дозволило мені не писати SQL-запити вручну, що автоматично закрило питання з ін'єкціями. А щоб захистити користувачів від підробки запитів, я інтегрував Flask-WTF з генерацією CSRF-токенів. Це такий собі «цифровий підпис» для кожної форми, який підтверджує, що дію виконує саме власник акаунту.

Контроль доступу та аудит, щоб уникнути хаосу з правами, я реалізував модель RBAC. Тепер у кожного своя роль: клієнти просто створюють тікети, виконавці їх обробляють, а адмін бачить повну картину. При цьому кожен крок у системі логується. Це важливо для розслідування інцидентів: завжди можна подивитися, хто, коли і яку дію вчинив. Такий підхід повністю вписується в рекомендації NIST та OWASP.

Список використаних джерел

1. Закон України. *«Про основні засади забезпечення кібербезпеки України».* **https://zakon.rada.gov.ua/laws/show/2163-19**
2. NIST. *Cybersecurity Framework 2.0.* **https://www.nist.gov/cyberframework**
3. Grinberg, M. *Flask Web Development: Developing Web Applications with Python.* — O'Reilly Media, 2018.
4. Microsoft Learn. *Role-Based Access Control (RBAC) Best Practices.* **https://learn.microsoft.com/en-us/azure/role-based-access-control/overview**