ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

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

ВСЕУКРАЇНСЬКА НАУКОВО-ПРАКТИЧНА КОНФЕРЕНЦІЯ



«АКТУАЛЬНІ ПРОБЛЕМИ КІБЕРБЕЗПЕКИ»

Тези доповідей

29 жовтня
2026

м. Київ

Редакційна колегія:

Гайдур Г.І. – д.т.н., професор, завідувач кафедри Систем та технологій кібербезпеки Навчально-наукового інституту кібербезпеки та захисту інформації.

Зибін С.В. – д.т.н., професор, професор кафедри Систем та технологій кібербезпеки Навчально-наукового інституту кібербезпеки та захисту інформації.

Казмірчук С.В. - д.т.н., професор, професор кафедри Систем та технологій кібербезпеки Навчально-наукового інституту кібербезпеки та захисту інформації.

Гахов С.О. – к.військ.н., доцент, доцент кафедри Систем та технологій кібербезпеки Навчально-наукового інституту кібербезпеки та захисту інформації.

Марченко В.В. – д.ф., доцент кафедри Систем та технологій кібербезпеки Навчально-наукового інституту кібербезпеки та захисту інформації.

Рекомендовано до друку кафедрою Сиситем та технологій Державного університету інформаційно-комунікаційних технологійй (протокол № 4 від 02.11.2026 р.)

Актуальні проблеми кібербезпеки: Матеріали Всеукраїнської науково-практичної конференції (м. Київ, 29 жовтня 2026 року). Навчально-науковий інститут кібербезпеки та захисту інформації, Державний університет інформаційно-комунікаційних технологій. Київ, 2026. 388 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

Рибачок Геннадій Іванович
ВИБІР ПЛАНУ РЕАГУВАННЯ НА КІБЕРІНЦИДЕНТ ЯК ЗАДАЧА ПОКРИТТЯ: ПОСТАНОВКА ДЛЯ ОРГАНІВ ВІЙСЬКОВОГО УПРАВЛІННЯ 14
Зінченко Олександра Ігорівна
ШТУЧНИЙ ІНТЕЛЕКТ ЯК ФАКТОР ТРАНСФОРМАЦІЇ КІБЕРЗАГРОЗ 16
Тишик І. Я.
ВИЯВЛЕННЯ КІБЕРЗАГРОЗ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНОЇ КОРЕЛЯЦІЇ ПОДІЙ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ 18
Лаптєв Олександр Анатолійович, Дунаев Сергій Владиславович
ПІДВИЩЕННЯ ПРИХОВАНОСТІ ТА ЗАХИЩЕНОСТІ ДАНИХ У СМАРТ-СЕРЕДОВИЩАХ МЕРЕЖ 6G НА ОСНОВІ ОПТИМІЗОВАНИХ LDPC-КОДІВ ТА ПОСТКВАНТОВОЇ КРИПТОСИСТЕМИ НІДЕРАЙТЕРА 21
Гайдур Г.І. , Селіверстов М. В.
ТЕХНОЛОГІЯ ЗАХИСТУ ВІДДАЛЕНИХ КОРИСТУВАЧІВ НА ОСНОВІ КОНЦЕПЦІЇ ZTNA 24
Гайдур Г.І., Проценко М.В.
ТЕХНОЛОГІЯ ЗАСТОСУВАННЯ МОДЕЛЕЙ ДОСТУПУ НА ОСНОВІ КОНТЕКСТНОЇ АУТЕНТИФІКАЦІЇ 26
Ляшенко Я.В.
УДОСКОНАЛЕННЯ СИСТЕМИ МЕНЕДЖМЕНТУ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ОРГАНІЗАЦІЇ НА ОСНОВІ СТАНДАРТУ ISO/IEC 27001 29
Карапиш Б.О.
ТЕХНОЛОГІЯ АВТОМАТИЗАЦІЇ ТЕСТУВАННЯ НА ПРОНИКНЕННЯ ВЕБ-ЗАСТОСУНКІВ НА ОСНОВІ МУЛЬТИАГЕНТНИХ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ 31
Цебак О.А.
ПЕРСПЕКТИВИ ГІБРИДНИХ ГЕНЕРАТОРІВ ПСЕВДОВИПАДКОВИХ ЧИСЕЛ 32
Романюк О.Н., Майданюк В.П., Романюк О.В.
ВИЯВЛЕННЯ ВРАЗЛИВОСТЕЙ ПРОГРАМНИХ МОДУЛІВ ОБРОБКИ ЦИФРОВИХ ЗОБРАЖЕНЬ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ 34
Романюк О.Н., Майданюк В.П., Нечипорук М.Л.
АПАРАТНЕ ПРИСКОРЕННЯ ШИФРУВАННЯ ЦИФРОВИХ ЗОБРАЖЕНЬ У ВБУДОВАНИХ СИСТЕМАХ 37
Tsarova Sofiia Valeriivna
RANSOMWARE DETECTION WITH ENDPOINT DETECTION & RESPONSE PLATFORMS 40
Степанченко Богдан Сергійович, Собчук Андрій Валентинович
МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ДИНАМІКИ КІБЕРЗАГРОЗ НА ОСНОВІ СИСТЕМ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ З ІМПУЛЬСНОЮ ДІЄЮ 42

Бражник Артем Михайлович
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ БАГАТОЕТАПНИХ КІБЕРАТАК У ГЕТЕРОГЕННИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ НА ОСНОВІ АДАПТИВНОЇ КОРЕЛЯЦІЇ ПОДІЙ 44
Магдєєв В.С., Мєшков В.І.
АНАЛІЗ ОЗНАК СОЦІОТЕХНІЧНИХ СЦЕНАРІЇВ СПОНУКАННЯ КОРИСТУВАЧА ДО ВИКОНАННЯ НЕБЕЗПЕЧНИХ ДІЙ У ВЕБСЕРЕДОВИЩІ 47
Яненко О. В., Мєшков В.І.
ОСОБЛИВОСТІ ОЦІНЮВАННЯ КІБЕРРИЗИКІВ СИСТЕМИ ЕЛЕКТРОННОГО ДОКУМЕНТООБІГУ ПІДПРИЄМСТВА НА ОСНОВІ НЕЧІТКОЇ ЛОГІКИ 50
Талан Анна Владиславівна
ПОРІВНЯЛЬНИЙ АНАЛІЗ РУЧНОГО ТА ЦЕНТРАЛІЗОВАНОГО КОНФІГУРУВАННЯ КОРПОРАТИВНИХ ПРИСТРОЇВ 53
Єрмаков О.В., Мєшков В.І.
АНАЛІЗ УМОВ АДАПТИВНОГО КЕРУВАННЯ ДОСТУПОМ ДО ПЕРСОНАЛЬНИХ ДАНИХ У ВЕБЗАСТОСУНКАХ З УРАХУВАННЯМ ВИМОГ GDPR..... 55
Гайдур Г.І., Тонкошкур О.В.
ТЕХНОЛОГІЯ РЕАЛІЗАЦІЇ КОНЦЕПЦІЇ ZERO TRUST У ХМАРНОМУ СЕРЕДОВИЩІ НА БАЗІ ІНТЕГРОВАНИХ СИСТЕМ IAM ТА КОНТЕКСТНОГО КОНТРОЛЮ ДОСТУПУ 58
Гимза О.А., Кеньо Г.В.
ІНФОРМАЦІЙНА БЕЗПЕКА СИСТЕМ РОЗУМНОГО ПАРКУВАННЯ 61
Козловський М. В., Мєшков В.І.
АНАЛІЗ УМОВ ЗАПОБІГАННЯ ВИТОКАМ ТРАФІКУ ПІД ЧАС УСТАНОВЛЕННЯ ТА ВІДНОВЛЕННЯ ЗАХИЩЕНИХ ТУНЕЛЬНИХ З’ЄДНАНЬ..... 63
Кутовий Денис Сергійович
ПРОФІЛЬ ЗАХИСТУ ДЛЯ ПРОГРАМ-АРХІВАТОРІВ ВІДПОВІДНО ДО ВИМОГ СТАНДАРТУ ISO/IEC 15408 66
Балацький Андрій Володимирович
ВПРОВАДЖЕННЯ БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ НА БАЗІ CISCO DUO FOR WINDOWS LOGON 68
Коваленко Ю. Б.
МОДЕЛЬ ОЦІНЮВАННЯ ЗАТРИМОК І СВОЄЧАСНОСТІ ПЕРЕДАВАННЯ ДАНИХ У МЕРЕЖАХ AFDX ТА TSN..... 71
Хохлачова Ю. Є.
ПРОГНОЗУВАННЯ МІЖСЕКТОРНИХ КАСКАДІВ У ЦИФРОВОМУ ДВІЙНИКУ КРИТИЧНОЇ ІНФРАСТРУКТУРИ..... 74

Поремський Я.С.
ТЕХНОЛОГІЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ПОДІЙ КІБЕРБЕЗПЕКИ НА ОСНОВІ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ У ПРОЦЕСАХ THREAT HUNTING 77
Прокопович-Ткаченко Д. І.
МІСІЙНО ОРІЄНТОВАНА МЕТРИКА КІБЕРСТІЙКОСТІ МЕРЕЖІ З УРАХУВАННЯМ НЕБЕЗПЕЧНОЇ ДЕГРАДАЦІЇ СЕРВІСІВ 80
Чайківський В.В.
АВТОМАТИЗАЦІЯ ПРОЦЕСІВ ВИЯВЛЕННЯ ТА УСУНЕННЯ ВРАЗЛИВОСТЕЙ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ КІНЦЕВИХ ПРИСТРОЇВ КОРПОРАТИВНОЇ МЕРЕЖІ ЗАСОБАМИ CHECK POINT ENDPOINT SECURITY 83
Таран В. Д.
ОСОБЛИВОСТІ ПОБУДОВИ SECURITY OPERATIONS CENTER ДЛЯ ПІДПРИЄМСТВ ІЗ ТЕРИТОРІАЛЬНО РОЗПОДІЛЕНОЮ ІНФРАСТРУКТУРОЮ 86
Лукінов Я.В., Нечитайло Ю.А.
ВИКОРИСТАННЯ СТОХАСТИЧНИХ МОДЕЛЕЙ ТА СТАТИСТИЧНОГО АНАЛІЗУ ДЛЯ ВИЯВЛЕННЯ Й ПРОТИДІЇ КІБЕРАТАКАМ У КІБЕРФІЗИЧНИХ СИСТЕМАХ 89
Лаврук Анна Віталіївна
ЗАХИСТ АГЕНТІВ НА ОСНОВІ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ ВІД НЕПРЯМОЇ ІН’ЄКЦІЇ ЗАПИТІВ У КОРПОРАТИВНИХ СИСТЕМАХ 93
Сердобінцев І.А., Мєшков В.І.
АНАЛІЗ СПОСОБІВ КОМПРОМЕТАЦІЇ IoT-ПРИСТРОЇВ ТА МЕТОДІВ ПРОТИДІЇ БОТНЕТАМ 96
Косяченко С.Р., Мєшков В.І.
АНАЛІЗ УМОВ КЕРУВАННЯ ЖИТТЄВИМ ЦИКЛОМ ДОСТУПУ ДО КОРПОРАТИВНИХ СЕКРЕТІВ В ІТ-КОМПАНІЇ 99
Корж А.Ю.
ТЕХНОЛОГІЯ ЗАПОБІГАННЯ ВИТОКУ ІНФОРМАЦІЇ В КОРПОРАТИВНІЙ МЕРЕЖІ НА ОСНОВІ ПЛАТФОРМИ UAM SYTECA 102
Дасюк Юрій Євгенійович
ТЕХНОЛОГІЯ УПРАВЛІННЯ ЗАХИСТОМ КІНЦЕВИХ ПРИСТРОЇВ НА ОСНОВІ IBM MAAS360 105
Севертока О.А.
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТА ЗАПОБІГАННЯ ПОРУШЕННЯМ ПОЛІТИК ІНФОРМАЦІЙНОЇ БЕЗПЕКИ НА ОСНОВІ UAM TERAMIND 108
Резніченко С.А.
ПРОБЛЕМИ ДОСТОВІРНОСТІ OSINT-АНАЛІЗУ ІЗ ЗАСТОСУВАННЯМ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ ТА ПІДХОДИ ДО ЇХ МІНІМІЗАЦІЇ 111

Мохонько Діана
ЗАХИСТ КРИТИЧНОЇ ІНФРАСТРУКТУРИ. НОРМАТИВНО-МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ КІБЕРЗАХИСТУ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ В УКРАЇНІ 113
Стебловський Г.В., Земляков С.О.
ВАЖЛИВІСТЬ ЗАХИСТУ КІНЦЕВИХ ПРИСТРОЇВ НА ОСНОВІ ESET INSPECT & PROTECT 116
Гуленов І.В.
ПІДВИЩЕННЯ БЕЗПЕКИ VPN ЗА ДОПОМОГОЮ БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ 120
Завгородня Євгенія Ярославівна
КІБЕРСТІЙКІСТЬ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ УКРАЇНИ: УРОКИ АТАКИ INDUSTROYER2 ТА ПІДХІД НА ОСНОВІ NIST CSF 2.0 .. 123
Шимчук Г. В., Яцків П.О.,
КОНТЕКСТНА МОДЕЛЬ КЕРУВАННЯ ДОСТУПОМ ДЛЯ ПІДВИЩЕННЯ БЕЗПЕКИ ХМАРНИХ СЕРЕДОВИЩ 126
Шимчук Г. В., Яцків П.О.
АДАПТИВНИЙ ЗАХИСТ ІОТ-МЕРЕЖ НА ОСНОВІ ОЦІНКИ ДОВІРИ ПРИСТРОЇВ ТА ДИНАМІЧНОЇ МІКРОСЕГМЕНТАЦІЇ 129
Онопрієнко М.Ю
МЕТОДИ І ТЕХНОЛОГІЇ КІБЕРАТАК, ОРГАНІЗОВАНИХ ЗА ПІДТРИМКИ НАЦІОНАЛЬНИХ ДЕРЖАВ ТА ЗАСОБИ ПРОТИДІЇ ЇМ 132
Ланде Дмитро Володимирович, Кузьмінський Віктор Сергійович
ВИЛУЧЕННЯ ТРИПЛЕТІВ ПРО КІБЕРЗАГРОЗИ З ТЕКСТОВИХ ПОВІДОМЛЕНЬ ІЗ ГАРАНТОВАНОЮ МЕЖЕЮ ЧАСТКИ ПОМИЛОК 135
Четвертуха Н.Р., Отенко В.І.
ВИБІР ІНСТРУМЕНТУ ДИНАМІЧНОГО АНАЛІЗУ ДЛЯ ШКІДЛИВИХ ПРОГРАМ, ЯКІ ПРИХОВУЮТЬ СВОЮ ПОВЕДІНКУ 138
Опірський І.Р., Стефанків А.В.
МЕХАНІЗМИ ВІДБОРУ ОЗНАК СИГНАЛІВ У СИСТЕМАХ РАДІОЧАСТОТНОГО ФІНГЕРПРИНТИНГУ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ 141
Афонін М.В.
АКТУАЛЬНІ ПРОБЛЕМИ БЕЗПЕКИ ІНСТРУМЕНТІВ ВІДДАЛЕНОГО ДОСТУПУ В КОРПОРАТИВНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ 144
Гончаренко М.
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТА КЛАСИФІКАЦІЇ МЕРЕЖЕВИХ АТАК У ПРОМИСЛОВИХ МЕРЕЖАХ MODBUS TCP НА ОСНОВІ АЛГОРИТМУ XGBOOST 147
Ступченко А.О
ПОБУДОВА ЦЕНТРУ ОПЕРАЦІЙНОЇ БЕЗПЕКИ (SOC) ЯК ІНСТРУМЕНТ ПІДВИЩЕННЯ КІБЕРСТІЙКОСТІ ОРГАНІЗАЦІЇ 150

Шапко Олег Олександрович
ПРОАКТИВНЕ ВИЯВЛЕННЯ ЗАГРОЗ НА КІНЦЕВИХ ТОЧКАХ ОРГАНІЗАЦІЇ З ВИКОРИСТАННЯМ ПОВЕДІНКОВИХ МОДЕЛЕЙ У СЕРЕДОВИЩІ WAZUH 152
Харькевич Дмитро Олегович
АВТОМАТИЗОВАНИЙ КОНТРОЛЬ КОДУ В DEVSECOPS ІЗ ВИКОРИСТАННЯМ SONARQUBE QUALITY GATE 156
Курінний О.С.
ОЦІНЮВАННЯ ТА ЗНИЖЕННЯ РИЗИКІВ ЛЮДСЬКОГО ФАКТОРА ПІД ЧАС ПРОТИДІЇ СОЦІАЛЬНІЙ ІНЖЕНЕРІЇ 160
Литвинюк Владислав Вадимович
ТЕХНОЛОГІЯ СЕГМЕНТАЦІЇ КОРПОРАТИВНОЇ МЕРЕЖІ ДЛЯ РЕАЛІЗАЦІЇ КОНЦЕПЦІЇ ZERO TRUST НА ОБЛАДНАННІ CISCO 163
Виноградов К.В.
ТЕХНОЛОГІЯ АВТОМАТИЗОВАНОГО ВИЯВЛЕННЯ ТА РЕАГУВАННЯ НА КІБЕРІНЦИДЕНТИ В SENTINELONE 165
Довгошея Д. О.
ТЕХНОЛОГІЯ МОНІТОРИНГУ СТАНУ ДЖЕРЕЛ ДАНИХ ТА СПОВІЩЕННЯ АДМІНІСТРАТОРА У SPLUNK ENTERPRISE SECURITY 168
Чуєва С.О.
ЗАСТОСУВАННЯ ЛОКАЛЬНИХ МАЛИХ МОВНИХ МОДЕЛЕЙ ДЛЯ ПОЯСНЮВАЛЬНОГО АНАЛІЗУ КІБЕРЗАГРОЗ В ІЗОЛЬОВАНИХ МЕРЕЖЕВИХ СЕРЕДОВИЩАХ 170
Середін Андрій Валерійович
ТЕХНОЛОГІЯ УПРАВЛІННЯ ІДЕНТИФІКАЦІЄЮ ТА БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ ІНФОРМАЦІЙНИХ СИСТЕМ НА БАЗІ KEYCLOAK 172
Стожок М.Р.
ТЕХНОЛОГІЯ АВТОМАТИЗОВАНОГО ВИЯВЛЕННЯ ТА РИЗИК-ОРІЄНТОВАНОЇ ПРІОРИТИЗАЦІЇ ВРАЗЛИВОСТЕЙ МЕРЕЖЕВИХ СЕРВІСІВ КОРПОРАТИВНОЇ МЕРЕЖІ 175
Архипенко Дмитро Євгенович
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ МЕРЕЖЕВИХ АНОМАЛІЙ ІЗ ЗАСТОСУВАННЯМ МАШИННОГО НАВЧАННЯ НА БАЗІ ZEEK 177
Гавриленко Д.П.
МЕТОДИ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ОПТИМІЗАЦІЇ ПРОЦЕСІВ МОНІТОРИНГУ ТА РЕАГУВАННЯ НА ІНЦИДЕНТИ В СУЧАСНОМУ SOC 179
Душник Володимир Володимирович
ТЕХНОЛОГІЯ ЗАСТОСУВАННЯ SPLUNK ДЛЯ МОНІТОРИНГУ ТА ВИЯВЛЕННЯ КІБЕРІНЦИДЕНТІВ У ДІЯЛЬНОСТІ SOC 181

Фесенко Д.С.
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТЕХНІК ОБХОДУ ЗАСОБІВ ЗАХИСТУ КІНЦЕВИХ ПРИСТРОЇВ НА ОСНОВІ MICROSOFT DEFENDER XDR 183
Піскунов Костянтин Валерійович
ТЕХНОЛОГІЯ ЗАПОБІГАННЯ ФІШИНГОВИМ АТАКАМ, СПРЯМОВАНИМ НА ВИКРАДЕННЯ ОБЛІКОВИХ ДАНИХ КОРИСТУВАЧІВ 186
Гриньов Тимофій Олександрович
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТА БЛОКУВАННЯ РОЗПОДІЛЕНИХ МЕРЕЖЕВИХ АТАК У КОРПОРАТИВНІЙ МЕРЕЖІ ОРГАНІЗАЦІЇ НА БАЗІ CROWDSEC 190
Шавшин Є.О
КОМПЛЕКСНИЙ ПІДХІД ДЛЯ ПІДВИЩЕННЯ РІВНЯ ЗАХИЩЕНОСТІ ІНФОРМАЦІЙНО-ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ 193
Твердохліб Я.М.
ЗАСТОСУВАННЯ СТАТИЧНОГО АНАЛІЗУ КОДУ ДЛЯ ПІДВИЩЕННЯ БЕЗПЕКИ ВЕБ-ЗАСТОСУНКІВ ОРГАНІЗАЦІЇ 197
Мурзін І. А., студент групи ІКБм-51
ВИЯВЛЕННЯ ПОВТОРНИХ АТАК І ДОВГОТРИВАЛОГО ДОСТУПУ ЗЛОВМИСНИКІВ ЗАСОБАМИ XDR 200
Тамбієва В.Г.
ПРОГРАМНА СИСТЕМА КОМПЛЕКСНОГО ЗАХИСТУ БАЗ ДАНИХ ВІД НЕСАНКЦІОНОВАНОГО ДОСТУПУ 203
Мазур Олександра
СИСТЕМА ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ ВІД ШКІДЛИВОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ НА ОСНОВІ ПРОАКТИВНОГО АНАЛІЗУ 207
Сокульський О. Є., Топольськов Є.О., Москаленко Н. В.
ГОМОМОРФНЕ ШИФРУВАННЯ ТА ХМАРНІ ТЕХНОЛОГІЇ 211
Головненко Н.Р.
РОЗРОБЛЕННЯ ТЕХНОЛОГІЇ ВИЯВЛЕННЯ КІБЕРАТАК НА ОСНОВІ КОРЕЛЯЦІЇ ПОДІЙ SPLUNK ENTERPRISE SECURITY ТА ТЕЛЕМЕТРІЇ SENTINELONE 213
Журавель Олександр Олегович
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ АТАК ТИПУ BRUTE FORCE ІЗ ЗАСТОСУВАННЯМ МЕТОДУ МАШИННОГО НАВЧАННЯ RANDOM FOREST 217
Хорольський Костянтин Андрійович
ЗАСТОСУВАННЯ ГРАДІЄНТНОГО БУСТИНГУ ДЛЯ ВИЯВЛЕННЯ ШКІДЛИВИХ ДОКУМЕНТНИХ ВКЛАДЕНЬ ФІШИНГОВИХ ЕЛЕКТРОННИХ ЛИСТІВ 220
Савченко В.В.
ТЕХНОЛОГІЯ ТЕСТУВАННЯ ЗАХИЩЕНОСТІ API ДЛЯ ВИЯВЛЕННЯ ВРАЗЛИВОСТЕЙ АВТЕНТИФІКАЦІЇ 224

Караєв Владислав Русланович,
ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ОБРОБКИ ДАНИХ У ВЕБ-СИСТЕМАХ АВТОМАТИЗОВАНОГО ПІДБОРУ КОМПЛЕКТУЮЧИХ ПК 227
Габорець О. А., Хоменко О. М.
ДЕСТРУКТИВНИЙ ІНФОРМАЦІЙНИЙ ВПЛИВ НА ДИТИНУ В ЦИФРОВОМУ СЕРЕДОВИЩІ: ПОНЯТТЯ, ОЗНАКИ ТА СУЧАСНІ ФОРМИ 231
Дубик Артур Русланович
АРХІТЕКТУРНА МОДЕЛЬ ЗАХИЩЕНОЇ ІНТЕГРАЦІЇ AI-АГЕНТІВ У КОРПОРАТИВНІ ІНФОРМАЦІЙНІ СИСТЕМИ НА ЗАСАДАХ ZERO TRUST 235
Михайленко Юлія Іванівна
ПІДХІД ДО ВИЯВЛЕННЯ DDOS SYN FLOOD-АТАК У МЕРЕЖАХ ІНТЕРНЕТУ РЕЧЕЙ НА ОСНОВІ АЛГОРИТМУ XGBOOST 239
Хоменко Микита Сергійович
ТЕХНОЛОГІЯ ЗАХИЩЕНОГО ОБМІНУ ДАНИМИ НА БАЗІ ПРОТОКОЛУ SSL/TLS 242
Власюк Ігор Дмитрович
КРИТЕРІЇ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ СИСТЕМ ЖУРНАЛІЗАЦІЇ КІБЕРІНЦИДЕНТІВ 244
Яковенко Я.І.
ТЕХНОЛОГІЯ ЗАХИСТУ ІНФОРМАЦІЙНОЇ СИСТЕМИ ОРГАНІЗАЦІЇ НА ОСНОВІ КОНЦЕПЦІЇ ZERO TRUST ІЗ ВИКОРИСТАННЯМ MICROSOFT ENTRA ID ТА MICROSOFT INTUNE 248
Хідзик Я.Ю.
АКТУАЛЬНІ ПРОБЛЕМИ КІБЕРБЕЗПЕКИ. РОЗСЛІДУВАННЯ КІБЕРІНЦИДЕНТІВ 250
Попелюк Софія
АКТУАЛЬНІ ПРОБЛЕМИ КІБЕРБЕЗПЕКИ. АТАКИ З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ 252
Науменко Д.М.
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ВРАЗЛИВОСТЕЙ ВЕБ-ДОДАТКІВ ОРГАНІЗАЦІЇ 255
Скидан Данило Романович, Щебланін Юрій Миколайович
ЗАХИСТ КОРПОРАТИВНИХ WI-FI МЕРЕЖ МЕТОДАМИ КОМБІНОВАНОГО ШИФРУВАННЯ ТА КОНТРОЛЮ СЕСІЙ 257
Азіз АБІБУЛАЄВ, Андріан ПІСКОЗУБ
ХИБНОПОЗИТИВНІ СПРАЦЮВАННЯ ДЕТЕКТОРІВ ВЕБ-АТАК У РЕАЛЬНОМУ ТРАФІКУ 261
Пашенько С. П.,
РЕКОНСТРУКЦІЯ КІБЕРІНЦИДЕНТІВ ШЛЯХОМ КОРЕЛЯЦІЇ ЦИФРОВИХ СЛІДІВ 265

Levterov G.A., Levterov A. A.
ANALYSIS CYBER INCIDENT PREVENTION IN THE MOBILE SECTOR OF THE UKRAINIAN ENERGY AND DEFENSE COMPLEXES UNDER CONDITIONS POWER SUPPLY DEGRADATION 267
Ахмедов А.Ф.
ФОРМУВАННЯ КРИТЕРІЇВ ДЕТЕРМІНОВАНОГО БІОМЕТРИЧНОГО ПРОТОКОЛУ В IoT-ІНФРАСТРУКТУРІ 271
Мельник В.І.
ТЕХНОЛОГІЯ АВТОМАТИЗОВАНОГО РЕАГУВАННЯ НА ФІШИНГОВІ АТАКИ З ВИКОРИСТАННЯМ SOAR 274
Кіх М. В.,
ВИКОРИСТАННЯ АУДІОЕНТРОПІЇ ДЛЯ ФОРМУВАННЯ ПОЧАТКОВОГО СТАНУ ГІБРИДНОГО ГЕНЕРАТОРА ПСЕВДОВИПАДКОВИХ ПОСЛІДОВНОСТЕЙ 276
Молнар Віталій Васильович
ПОРІВНЯННЯ СПОСОБІВ АВТОМАТИЗОВАНОГО РЕАГУВАННЯ НА НЕБЕЗПЕЧНІ ЗМІНИ КОНФІГУРАЦІЇ AWS 279
Костенко Яків Олександрович
СИСТЕМА МОНІТОРИНГУ ТА АНАЛІЗУ АНОМАЛЬНОЇ ПОВЕДІНКИ КОРИСТУВАЧІВ У КОРПОРАТИВНИХ МЕРЕЖАХ НА ОСНОВІ SYSTECA UAM 283
Ялбуган Ф. Д.
ГІБРИДНЕ ОЦІНЮВАННЯ КІБЕРРИЗИКІВ ОРГАНІЗАЦІЙ-КОНТРАГЕНТІВ ЗА OSINT-ДАНИМИ З ВИКОРИСТАННЯМ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ 285
Алексанян Андрій Ельдарович
ВИЯВЛЕННЯ АНОМАЛЬНОЇ ПОВЕДІНКИ КОРИСТУВАЧІВ МЕТОДАМИ МАШИННОГО НАВЧАННЯ ДЛЯ ЗАБЕЗПЕЧЕННЯ КІБЕРБЕЗПЕКИ КОРПОРАТИВНИХ СИСТЕМ 289
Москаленко Н. В., Сокульський О. Є., Топольськов Є.О.
ШТУЧНИЙ ІНТЕЛЕКТ ЯК СКЛАДОВА ЯКОСТІ ПЗ 291
Круть Ростислав Андрійович
ТЕХНОЛОГІЇ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ КОМП'ЮТЕРНИХ МЕРЕЖ ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ VPN 292
Власенко Г. В.
КРИТЕРІЇ ФОРМУВАННЯ ОБМЕЖЕНОГО НАБОРУ СЦЕНАРІЇВ ВАЛІДАЦІЇ МЕХАНІЗМІВ КІБЕРЗАХИСТУ НА ОСНОВІ MITRE ATT&CK 294
Коршиков О.В., Мужанова Т.М.
ЕВОЛЮЦІЯ ВИМОГ ДО ФАХІВЦІВ З КІБЕРБЕЗПЕКИ В УМОВАХ ЦИФРОВІЗАЦІЇ 297
Миловидов В. Ю.
ОЦІНЮВАННЯ ВНЕСКУ КОНТЕКСТНИХ ОЗНАК У НЕЙРОМЕРЕЖЕВЕ ВИЯВЛЕННЯ ФІШИНГОВИХ ВЕБРЕСУРСІВ 300

Нижник А.О.
МЕТОД ВИЯВЛЕННЯ АНОМАЛЬНИХ ПОСЛІДОВНОСТЕЙ КОМАНД MAVLINK У СИСТЕМАХ КЕРУВАННЯ БПЛА 303
Чалий О.В., Толюпа С.В.
ГРАФ ЗНАНЬ ДЛЯ ВИЯВЛЕННЯ НЕВІДПОВІДНОСТЕЙ В АУДИТІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ 307
Козяр Михайло Олександрович
ІНТЕГРАЦІЯ DEVSECOPS-ПРАКТИК У GITOPS-КОНВЕЄР ДЛЯ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ХМАРНОЇ ІНФРАСТРУКТУРИ 310
Колесникова А. Я.
ОЦІНЮВАННЯ КІБЕРЗАХИЩЕНОСТІ ХМАРНИХ СЕРЕДОВИЩ НА БАЗІ МАТЕМАТИЧНИХ МОДЕЛЕЙ ТА СТАНДАРТІВ ISO 313
Алексанян Андрій Ельдарович
ІНТЕГРАЦІЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ В SOC ДЛЯ ВИЯВЛЕННЯ АНОМАЛЬНОЇ ПОВЕДІНКИ КОРИСТУВАЧІВ 316
Вострих Д.О.
ТЕХНОЛОГІЯ ЗАХИСТУ СУЧАСНИХ МЕРЕЖ НА ОСНОВІ PALO ALTO NGFW 318
Куценко Олексій Сергійович
ЗАХИСТ IoT-ПРИСТРОЇВ ТЕЛЕМАТИЧНИХ СИСТЕМ НА ЕТАПАХ ВПРОВАДЖЕННЯ ТА ЕКСПЛУАТАЦІЇ 320
Піскунов Костянтин Валерійович
ТЕХНОЛОГІЯ ЗАПОБІГАННЯ ФІШИНГОВИМ АТАКАМ, СПРЯМОВАНИМ НА ВИКРАДЕННЯ ОБЛІКОВИХ ДАНИХ КОРИСТУВАЧІВ 323
Черненко Д. А.
ТЕХНОЛОГІЇ ЗАХИСТУ КОРПОРАТИВНИХ ДАНИХ НА БАЗІ MICROSOFT PURVIEW DLP 327
Райченко В. С.
ІНТЕГРОВАНИЙ ПІДХІД ДО ЗАХИСТУ КОРПОРАТИВНИХ ІНФОРМАЦІЙНИХ СИСТЕМ: МОНІТОРИНГ, УПРАВЛІННЯ ВРАЗЛИВОСТЯМИ ТА АВТОМАТИЗАЦІЯ РЕАГУВАННЯ 330
Рядий Д.О.
КОНТЕКСТНИЙ АНАЛІЗ HTTP-ЗАПИТІВ У СИСТЕМІ ЗАХИСТУ ВЕБ-ДОДАТКІВ OPEN-APPSEC 333
Шулімова Д.Д.
ЗАЛЕЖНІСТЬ ІНФОРМАТИВНОСТІ ОЗНАК МЕРЕЖЕВОГО ТРАФІКУ ВІД ТИПУ АТАК 335
Medynskyi O. R., Susukailo V. A.
A DIGITAL IMAGE VERIFICATION METHODOLOGY FOR CYBER INCIDENT INVESTIGATION IN THE ERA OF GENERATIVE ARTIFICIAL INTELLIGENCE 337
Сушанов Павло Вікторович
ТЕХНОЛОГІЯ ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ ВІД ФІШИНГОВИХ АТАК 340

Тишик Іван Ярославович
ВИЯВЛЕННЯ КІБЕРЗАГРОЗ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНОЇ КОРЕЛЯЦІЇ ПОДІЙ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ 343
Верголяс О. О.
КОГНІТИВНА БЕЗПЕКА КРИТИЧНИХ АУДИТОРІЙ В УМОВАХ ІНФОРМАЦІЙНО-ПСИХОЛОГІЧНИХ ОПЕРАЦІЙ, ПІДСИЛЕНИХ ГЕНЕРАТИВНИМ ШТУЧНИМ ІНТЕЛЕКТОМ 346
Svedeniuk O.V., Kurii Y. O.
APPLICATION OF AI MODELS FOR AUTOMATING IOT INFRASTRUCTURE COMPLIANCE WITH ETSI EN 303 645 AND IEC 62443 STANDARDS 348
Ковалівська К. В.
ДОСЛІДЖЕННЯ РІВНЯ КІБЕРБЕЗПЕКИ УКРАЇНИ В УМОВАХ ВОЄННОГО СТАНУ ТА ПОСИЛЕННЯ АГРЕСІЇ РФ 350
Ровда Володимир Володимирович
ІНТЕЛЕКТУАЛЬНІ МЕТОДИ ОПТИМІЗАЦІЇ КАНАЛЬНОГО ТА МЕРЕЖЕВОГО РІВНІВ У ВИСОКОНАВАНТАЖЕНИХ ЛОКАЛЬНИХ РАДІОМЕРЕЖАХ..... 354
Лукінов Я.В., Нечитайло Ю.А.
ВИКОРИСТАННЯ СТОХАСТИЧНИХ МОДЕЛЕЙ ТА СТАТИСТИЧНОГО АНАЛІЗУ ДЛЯ ВИЯВЛЕННЯ Й ПРОТИДІЇ КІБЕРАТАКАМ У КІБЕРФІЗИЧНИХ СИСТЕМАХ 358
Нестор Наталія Володимирівна
ЗАБЕЗПЕЧЕННЯ КІБЕРСТІЙКОСТІ ДЕРЖАВНИХ РЕЄСТРІВ УКРАЇНИ В УМОВАХ ГІБРИДНИХ ЗАГРОЗ: ЄВРОПЕЙСЬКІ СТАНДАРТИ ТА МІЖНАРОДНЕ СПІВРОБІТНИЦТВО 362
Леляк А. В., Астраханцев А. А.
ЗАХИСТ КАНАЛІВ В СИСТЕМАХ ЕЛЕКТРОННОЇ ІДЕНТИФІКАЦІЇ ЗА ДОПОМОГОЮ ВДОСКОНАЛЕНОЇ БІБЛІОТЕКИ OPENSCP 365
Нетребко Є.Г., Мєшков В.І.
АНАЛІЗ УМОВ ВИБОРУ ЗАСОБІВ БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ ДЛЯ ЗАХИСТУ ОБЛІКОВИХ ЗАПИСІВ 369
Литвиненко Роман Геннадійович, Буняк Анастасія Володимирівна, Рябець Віктор Вікторович
SIEM-СИСТЕМИ ЯК ІНСТРУМЕНТ ЦЕНТРАЛІЗОВАНОГО ЗБОРУ ТА АНАЛІЗУ ІНФОРМАЦІЇ 372
Буханець І.С.
ПІДХІД ДО ВИЯВЛЕННЯ АНОМАЛІЙ ТА ЗАГРОЗ НУЛЬОВОГО ДНЯ В IoT-МЕРЕЖАХ НА ОСНОВІ ОДНОКЛАСОВОГО МАШИННОГО НАВЧАННЯ375
Коровкін А.В.
ПІДХІД ДО ВИЯВЛЕННЯ КІБЕРАТАК НА IoT-ПЕРИФЕРІЮ НА ОСНОВІ МІНІМІЗАЦІЇ ОЗНАК СИСТЕМНОГО МОНІТОРИНГУ МЕТОДАМИ МАШИННОГО НАВЧАННЯ..... 379

Правдивий В.С.
ПІДХІД ДО ВИЯВЛЕННЯ ШКІДЛИВОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ IoT НА ОСНОВІ МУЛЬТИМОДАЛЬНОГО АНАЛІЗУ ТЕЛЕМЕТРІЇ МЕТОДАМИ МАШИННОГО НАВЧАННЯ..... 383
Шевченко О.О.
ТЕХНОЛОГІЯ ВИЯВЛЕННЯ НИЗЬКОІНТЕНСИВНИХ ТА ПРИХОВАНИХ КІБЕРАТАК НА IoT-ПРИСТРОЇ НА ОСНОВІ СТАТИСТИЧНОГО ПРОФІЛЮВАННЯ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ..... 386

Рибачок Геннадій Іванович
полковник, ад'юнкт кафедри кіберборотьби,
Національний університет оборони України,
Київ, Україна

ВИБІР ПЛАНУ РЕАГУВАННЯ НА КІБЕРІНЦИДЕНТ ЯК ЗАДАЧА ПОКРИТТЯ: ПОСТАНОВКА ДЛЯ ОРГАНІВ ВІЙСЬКОВОГО УПРАВЛІННЯ

Розглянуто вибір складу заходів реагування на кіберінцидент в органах військового управління. Показано, що зведення рішення до пріоритизації інцидентів лишає відкритим питання, який саме набір заходів застосувати за обмежених ресурсу, часу і персоналу. Запропоновано формулювати вибір як задачу покриття критичних технік атаки з мінімізацією сукупної вартості плану. Розмежовано таку постановку і ранжування та окреслено умови, за яких точний розв'язок має практичну перевагу над жадібними евристиками.

Ключові слова: кіберінцидент, план реагування, задача покриття, цілочислове програмування, органи військового управління.

Реагування на кіберінцидент в органах військового управління (ОВУ) відрізняється від корпоративного передусім жорсткістю обмежень. Склад чергової зміни фіксований, вікно допустимої недоступності сервісу визначає бойове завдання, а перелік технік, які треба перекрити, задає не гіпотеза аналітика, а дані про конкретну атаку. За таких умов питання «що робити» розпадається на два різні, які на практиці регулярно змішують: у якому порядку братися за інциденти і який саме набір заходів застосувати до окремого інциденту.

Перше питання опрацьоване значно краще. Керівні документи описують життєвий цикл реагування і критерії віднесення інциденту до категорії [2], [3], а в дослідницькій практиці усталилися методи багатокритеріального ранжування. Друге питання здебільшого лишається за експертом: маючи перелік допустимих альтернатив, черговий обирає їх набір, спираючись на досвід і наявний час. Саме тут виникає розрив між обґрунтованістю рішення і його документованістю, який для ОВУ критичний – рішення доводиться пояснювати постфактум.

Пропонована постановка виходить із того, що альтернативи реагування не є взаємозамінними по одній. Кожен захід перекриває певну підмножину технік з тих, що виявлені в інциденті, і має вартість – у людино-годинах чергової зміни, у часі недоступності сервісу, у витраченому ресурсі засобів захисту. Техніки, віднесені до критичних, мають бути перекриті обов'язково. Тоді вибір плану – це задача про покриття множини: знайти набір заходів мінімальної сукупної вартості, який перекриває всі критичні техніки і вкладається в обмеження за часом та персоналом. Для опису технік доцільно спиратися на усталену таксономію тактик і технік супротивника [4], бо вона дає спільну мову для розвідданих і для переліку контрзаходів.

Різниця з ранжуванням не термінологічна. Пріоритизація дає порядок, покриття – склад. Два заходи, кожен з яких окремо виглядає дешевшим, можуть дублювати одну й ту саму техніку і водночас лишати іншу критичну техніку неперекритою; ранжування за жодним критерієм цього не виявляє, бо оцінює

альтернативи поодинці, а не як набір. Навпаки, дорожчий захід може виявитися вигіднішим, якщо він єдиний перекриває дві критичні техніки одразу і звільняє час зміни.

Практичну перевагу точного формулювання видно у трьох місцях. По-перше, воно дає доказову оптимальність: розв'язок не просто «прийнятний», а найдешевший серед допустимих, і це властивість, яку можна пред'явити. Жадібний вибір, поширений у практиці, таких гарантій не має і на структурованих вхідних даних відхиляється від оптимуму тим сильніше, чим більше перетинаються зони покриття заходів. По-друге, постановка робить рішення відтворюваним: за тих самих вхідних даних отримується той самий план, а зміна плану завжди пояснюється зміною конкретного обмеження. По-третє, вона узгоджується з моделлю нейтралізації, у якій оцінювальний образ інциденту перетворюється на пару цільових параметрів і далі на допустимі альтернативи [1]; покриття природно вбудовується в цей контур як завершальний крок відбору.

Обмеження підходу теж слід назвати прямо. Якість розв'язку не перевищує якості вхідних оцінок вартості й покриття, а їх у бойових умовах отримують з неповних даних. Розмірність задачі зростає з кількістю альтернатив, тому для великих інцидентів потрібне або скорочення простору альтернатив на попередньому кроці, або обмеження часу пошуку з поверненням найкращого знайденого розв'язку. Обидва напрями складають предмет подальшої роботи; її результати автор планує викласти в окремій статті з формальною постановкою, доведенням допустимості та експериментальною перевіркою на серії згенерованих інцидентів.

Перелік посилань

1. Рибачок Г. І. Модель нейтралізації кіберінцидентів в інформаційно-телекомунікаційних системах органів військового управління. Сучасні інформаційні технології у сфері безпеки та оборони. 2026. Т. 56, № 2. С. 166–176. DOI: 10.33099/2311-7249/2026-56-2-166-176.

2. NIST SP 800-61 Rev. 3. Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile. Gaithersburg : National Institute of Standards and Technology, 2025. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r3.pdf

3. ISO/IEC 27035-1:2023. Information technology — Information security incident management — Part 1: Principles and process. 2nd ed. Geneva : ISO, 2023. URL: https://www.iso.org/standard/78973.html

4. MITRE ATT&CK: Adversarial Tactics, Techniques and Common Knowledge. URL: https://attack.mitre.org

Зінченко Олександра Ігорівна
доктор філософії
викладач кафедри міжнародних відносин
Навчально-наукового інституту
“Каразінський інститут міжнародних відносин та туристичного бізнесу”
Харків, Україна

ШТУЧНИЙ ІНТЕЛЕКТ ЯК ФАКТОР ТРАНСФОРМАЦІЇ КІБЕРЗАГРОЗ

Стрімкий розвиток технологій штучного інтелекту (ШІ), зокрема генеративних моделей, великих мовних моделей та систем автоматизованого аналізу даних, поступово змінює характер сучасного кіберпростору. Якщо раніше традиційні кіберзагрози значною мірою були пов’язані з використанням заздалегідь визначених інструментів, експлуатацією технічних вразливостей та необхідністю залучення спеціалізованих знань, то поширення ШІ створює передумови для підвищення швидкості, масштабованості, адаптивності та персоналізації шкідливої діяльності. Принципова відмінність полягає у тому, що поширення генеративного ШІ, великих мовних моделей та інших інтелектуальних систем змінює не окремий елемент кібероперації, а співвідношення між ресурсами, необхідними для підготовки атаки, її масштабом, швидкістю реалізації та можливістю адаптації до поведінки об’єкта впливу. Внаслідок цього відбувається не просто поява нових інструментів кібератак, а якісна трансформація самого середовища кіберзагроз

Ключові слова: кібербезпека; кіберзагрози; міжнародна безпека; штучний інтелект.

У сучасних дослідженнях кібербезпеки поняття кіберзагрози охоплює потенційну можливість реалізації дій або подій, здатних порушити конфіденційність, цілісність, доступність чи інші властивості інформаційних систем і цифрової інфраструктури [1] . Кібератака є вже реалізованою цілеспрямованою дією, тоді як кіберризик характеризує можливість негативних наслідків за певного поєднання загрози, вразливості та потенційного впливу. Відповідно, використання ШІ не слід автоматично ототожнювати з появою нового виду кіберзагрози. Його значення полягає передусім у трансформації способів формування та реалізації вже наявних загроз, а також у появі нових вразливостей, пов’язаних із самими системами ШІ.

Актуальні оцінки ENISA демонструють, що кіберзагрозове середовище характеризується одночасною конвергенцією інструментів, тактик і суб’єктів, використанням цифрових залежностей та підвищенням ролі автоматизації. У Threat Landscape 2025 [2] агентство окремо визначає зростання ролі штучного інтелекту як одну з ключових тенденцій: великі мовні моделі використовуються для вдосконалення фішингу, соціальної інженерії, розвідки та створення шкідливого програмного забезпечення, а самі системи ШІ формують нову поверхню атаки. Водночас Threat Landscape 2026 засвідчує подальший розвиток цієї тенденції: ШІ дедалі активніше використовується кіберзлочинними, державними та інформаційно-маніпулятивними суб’єктами, а у 2026 році [3] вже спостерігаються ознаки його глибшої інтеграції в кібероперації, зокрема через AI-powered malware, використання генеративного ШІ під час виконання шкідливого коду та застосування агентних систем. Водночас ENISA наголошує, що атакувальники поки переважно використовують ШІ для посилення наявних спроможностей, а не для принципово нових проривних можливостей.

З огляду на зазначені тенденції, трансформацію кіберзагроз під впливом ШІ доцільно розглядати через сім взаємопов’язаних параметрів: масштабування,

прискорення, автоматизацію, адаптивність, персоналізацію, ускладнення виявлення та транскордонність. Масштабування й автоматизація розширюють кількість потенційних цілей і зменшують ресурсні витрати, прискорення скорочує часові рамки операцій, а адаптивність дає змогу змінювати характер впливу залежно від реакції об'єкта. Персоналізація посилює можливості цілеспрямованого кібернетичного та інформаційно-психологічного впливу, тоді як варіативність і адаптивність ускладнюють виявлення та атрибуцію операцій.

Сукупна дія цих параметрів змінює не лише технічні характеристики кібероперацій, а й умови функціонування систем національної та міжнародної безпеки. Транскордонний характер цифрової інфраструктури та кібероперацій підвищує залежність держав від міжнародного обміну інформацією, координації та сумісності спроможностей. Відтак використання ШІ у кіберпросторі стає питанням не лише технологічного розвитку, а й державної спроможності, інституційної координації та міжнародного управління кібербезпекою.

Для міжнародної безпеки це означає зростання взаємозалежності держав у сфері кіберзахисту. Транскордонний характер цифрової інфраструктури, даних та систем ШІ обмежує можливості суто національного реагування й посилює значення міжнародного обміну інформацією, координації та узгодження підходів до управління кіберризиками. Водночас європейські механізми кіберсолідарності та регулювання ШІ демонструють поступове включення кібербезпеки безпосередньо до процесів розроблення, впровадження та управління системами ШІ [4].

Таким чином, поширення ШІ змінює усталені характеристики кіберзагроз, поєднуючи масштабування, прискорення, автоматизацію, адаптивність і персоналізацію з ускладненням їх виявлення та транскордонним характером. Водночас значення цих змін виходить за межі технічної сфери: вони впливають на розподіл спроможностей між суб'єктами, характер інформаційного та політичного впливу, вимоги до державного управління й механізми міжнародної взаємодії. Саме тому ефективне реагування потребує не лише розвитку технологічних засобів захисту, а й перегляду інституційних підходів до оцінювання загроз, координації та міжнародного співробітництва. У перспективі саме здатність держав і міжнародних інституцій адаптувати безпекові механізми до швидкості технологічних змін визначатиме їхню спроможність протидіяти новим формам кіберзагроз.

Перелік використаної літератури:

1. Election Security Spotlight – CIA Triad. *CIS*. URL: https://www.cisecurity.org/insights/spotlight/ei-isac-cybersecurity-spotlight-cia-triad(date of access: 21.09.2026).

2. ENISA. ENISA Threat Landscape 2026. 2025. 88 p. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025 (date of access: 22.09.2026).

3. ENISA. ENISA Threat Landscape 2026. 2026. 101 p. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2026 (date of access: 22.09.2026).

4. Article 15. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. 2024. L 2024/1689. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj (date of access: 22.09.2026).

Тишик І. Я.
доцент кафедри ЗІ НУЛП
Львів Україна

ВИЯВЛЕННЯ КІБЕРЗАГРОЗ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНОЇ КОРЕЛЯЦІЇ ПОДІЙ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Анотація. Розглянуто підхід до виявлення кіберзагроз на основі інтегрованого аналізу подій інформаційної безпеки із застосуванням методів штучного інтелекту. Запропоновано здійснювати централізований збір і нормалізацію подій із різнорідних джерел, виявлення аномальної активності та кореляцію взаємопов'язаних подій. Для встановлення відповідності виявленої активності характерним сценаріям кібератак запропоновано використовувати базу знань MITRE ATT&CK. Розглянуто можливість реалізації підходу засобами SIEM, UEBA та SOC для підвищення ефективності виявлення складних багатоступеневих кіберзагроз.

Ключові слова: кіберзагрози; штучний інтелект; кореляція подій інформаційної безпеки; MITRE ATT&CK; SIEM.

Сучасні інформаційно-комунікаційні системи характеризуються значною кількістю взаємопов'язаних компонентів, розподіленістю інформаційних ресурсів та постійним зростанням інтенсивності мережевої взаємодії. За таких умов традиційні засоби виявлення кіберзагроз, орієнтовані переважно на аналіз окремих подій або використання заздалегідь визначених сигнатур, не завжди забезпечують своєчасну ідентифікацію складних багатоступеневих атак. Особливо актуальною ця проблема є для атак, під час яких зловмисники використовують легітимні облікові записи, штатні засоби операційної системи та здійснюють горизонтальне переміщення між вузлами корпоративної мережі.

Перспективним напрямом підвищення ефективності моніторингу кібербезпеки є інтегрований аналіз подій інформаційної безпеки із застосуванням методів штучного інтелекту. Методи машинного навчання дають змогу здійснювати аналіз значних масивів подій та виявляти закономірності й аномалії, які складно встановити традиційними сигнатурними методами [1, с. 1155–1156]. На відміну від аналізу ізольованих спрацювань окремих засобів захисту, такий підхід передбачає формування комплексного уявлення про розвиток потенційної атаки на основі кореляції подій, отриманих із різнорідних джерел.

Метою роботи є підвищення ефективності виявлення кіберзагроз шляхом інтегрованого аналізу подій інформаційної безпеки, визначення аномальної активності засобами штучного інтелекту та її кореляції з відомими техніками реалізації кібератак

Наукова новизна запропонованого підходу полягає у поєднанні поведінкового аналізу аномальної активності з контекстною кореляцією різнорідних подій інформаційної безпеки та їх зіставленням із закономірностями реалізації кібератак, представленими у MITRE ATT&CK. На відміну від підходів, які ґрунтуються на аналізі окремих спрацювань або наперед визначених сигнатур, передбачається аналіз взаємопов'язаної послідовності подій з урахуванням їх часових, поведінкових та логічних зв'язків. Це створює можливість виявлення ознак розвитку багатоступеневої атаки навіть у випадках,

коли окремі події не перевищують встановлених порогів спрацювання засобів захисту.

Джерелами даних для такого аналізу можуть бути журнали операційних систем і серверів, мережеве обладнання, системи автентифікації, контролери домену Active Directory, міжмережеві екрани, IDS/IPS, засоби захисту кінцевих точок та SIEM-системи. Централізований збір, попередня обробка і нормалізація цих даних створюють основу для встановлення часових, логічних та поведінкових взаємозв'язків між подіями. Централізоване керування журналами подій є важливою складовою процесів моніторингу та аналізу стану інформаційної безпеки [2, с. 2-7].

Запропонований підхід передбачає послідовне виконання таких етапів: збір подій інформаційної безпеки з різних джерел, їх попередню обробку та нормалізацію, виявлення аномальної активності із застосуванням методів штучного інтелекту, кореляцію взаємопов'язаних подій та формування сповіщення про потенційну кіберзагрозу. Спрощену структуру запропонованого процесу наведено на рис. 1.

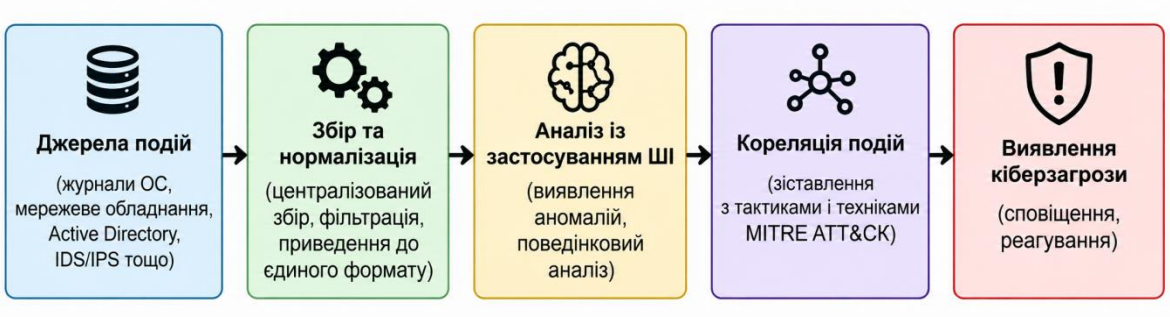


Рис. 1. Спрощена схема процесу виявлення кіберзагроз на основі інтелектуальної кореляції подій інформаційної безпеки

Окрема подія інформаційної безпеки не завжди містить достатньо інформації для однозначної ідентифікації атаки. Наприклад, невдала спроба автентифікації, запуск PowerShell або звернення до мережевого ресурсу можуть бути легітимними діями. Водночас їх поєднання з успішною автентифікацією з нетипового вузла, використанням адміністративних привілеїв, спробами отримання облікових даних та подальшим горизонтальним переміщенням може свідчити про розвиток кібератаки. Тому важливим завданням є не лише виявлення окремих аномалій, а й встановлення взаємозв'язків між ними.

Для визначення таких залежностей доцільно застосовувати методи машинного навчання та поведінкового аналізу. Виявлення аномалій передбачає встановлення шаблонів поведінки, які не відповідають очікуваній або типовій активності системи [4, с. 1–3]. Це дає змогу формувати профілі типової активності користувачів, вузлів і процесів та виявляти відхилення від сформованого базового рівня. Ознаками аномальної поведінки можуть бути нехарактерний час автентифікації, значна кількості невдалих входів, використання підвищених привілеїв, запуск нетипових процесів, зміна системних параметрів та незвична мережева активність.

Важливим елементом запропонованого підходу є кореляція виявленої активності з відомими закономірностями реалізації кібератак. Для цього доцільно використовувати базу знань MITRE ATT&CK, у якій систематизовано тактики та техніки, які застосовуються зловмисниками. Виявлена послідовність подій може бути зіставлена з такими етапами атаки, як початковий доступ (Initial Access), виконання (Execution), закріплення (Persistence), отримання облікових даних (Credential Access), підвищення привілеїв (Privilege Escalation) та горизонтальне переміщення (Lateral Movement) [3, с. 8].

Особливого значення такий підхід набуває для моніторингу середовища Active Directory, де компрометація одного облікового запису може стати початковим етапом подальшого розвитку атаки. Кореляція подій автентифікації, змін привілеїв, запуску системних утиліт, звернень до контролерів домену та аномальної мережевої активності дає можливість виявляти ознаки компрометації, які складно встановити шляхом аналізу окремого журналу подій.

Практична реалізація запропонованого підходу може здійснюватися в межах центру моніторингу безпеки (SOC) із використанням SIEM як базової платформи централізованого збору та первинної кореляції подій. Поєднання SIEM із технологіями UEBA та засобами машинного навчання дає змогу доповнити традиційні правила кореляції механізмами поведінкового аналізу та автоматизованого виявлення аномальної активності.

Важливим аспектом практичного застосування запропонованого підходу є можливість його адаптації до особливостей конкретної інформаційної системи при зміні характеру кіберзагроз. Накопичення даних про події безпеки, результати їх аналізу та підтверджені інциденти створює основу для подальшого уточнення моделей виявлення аномальної активності.

Таким чином, інтегрований аналіз подій інформаційної безпеки із застосуванням штучного інтелекту є перспективним підходом до виявлення сучасних кіберзагроз. Поєднання даних із різнорідних джерел, виявлення аномальної поведінки та кореляція отриманих результатів із відомими техніками атак дають змогу підвищити точність ідентифікації потенційних інцидентів. Подальші дослідження доцільно спрямувати на вдосконалення механізмів автоматизованої кореляції аномальної активності з послідовностями технік MITRE ATT&CK та експериментальну перевірку запропонованого підходу в корпоративних мережах.

Список використаних джерел

1. Buczak A. L., Guven E. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials*. 2016. Vol. 18, No. 2. P. 1153–1176.
2. Kent K., Souppaya M. *Guide to Computer Security Log Management*. NIST Special Publication 800-92. Gaithersburg: National Institute of Standards and Technology, 2006.
3. Strom B. E., Applebaum A., Miller D. P., Nickels K. C., Pennington A. G., Thomas C. B. *MITRE ATT&CK: Design and Philosophy*. The MITRE Corporation, 2020.
4. Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey. *ACM Computing Surveys*. 2009. Vol. 41, No. 3. Article 15.

Лаптєв Олександр Анатолійович, доцент
Київський Національний Університет імені Тараса Шевченка,
м. Київ
Дунаев Сергій Владиславович
Національний технічний університет «Харківський політехнічний інститут»
м.Харків

ПІДВИЩЕННЯ ПРИХОВАНОСТІ ТА ЗАХИЩЕНОСТІ ДАНИХ У СМАРТ-СЕРЕДОВИЩАХ МЕРЕЖ 6G НА ОСНОВІ ОПТИМІЗОВАНИХ LDPC-КОДІВ ТА ПОСТКВАНТОВОЇ КРИПТОСИСТЕМИ НІДЕРАЙТЕРА

Обґрунтовано обмеженість традиційних криптографічних протоколів (TLS/SSL) на фізичному та канальному рівнях. Запропоновано комплексний підхід до підвищення прихованості даних шляхом інтеграції адаптивних фонтанних кодів, методів перемеження та побудови нерегулярних LDPC-кодів з покращеними комбінаторними властивостями. Демонструється можливість використання оптимізованих LDPC-кодів як математичної основи для постквантової криптосистеми Нідеррайтера.

Згідно з опитуванням CrowdStrike кількість вторгнень у хмару зросла на 136 відс. у першій половині 2025 року. Найпримітніше, що 81 відс. цих вторгнень не були пов’язані з використанням шкідливого програмного забезпечення. Замість розгортання шкідливого програмного забезпечення, зловмисники все частіше покладаються на викрадені облікові дані для отримання доступу та роботи в легітимних системах. Цей зсув означає, що багато традиційних методів виявлення, створених для позначення шкідливого коду, повністю ігноруються [1].

Стрімкий розвиток бездротових технологій шостого покоління (6G) та масове впровадження смарт-технологій (Smart Grid, Smart City, інтерфейси «мозок-комп’ютер» (BCI), розширена реальність (XR)) формують нові вимоги до надійності та безпеки передавання інформації. Архітектура 6G передбачає інтеграцію наземних, повітряних та космічних мереж, використання терагерцового діапазону та вбудованого штучного інтелекту (AI-native).

Традиційні протоколи транспортного рівня, такі як TLS/SSL, залишаються основним засобом забезпечення конфіденційності, проте в умовах 6G вони мають суттєві обмеження: недостатню стійкість до аналізу трафіку (через метадані та патерни рукостискання), обчислювальні затримки, а також відсутність вбудованих механізмів виявлення фізичних втручань у канал. Це актуалізує потребу в багаторівневому захисті, де механізми прихованості закладаються безпосередньо на фізичному та канальному рівнях за допомогою методів завадостійкого кодування.

Перспективним напрямом є використання кодів із низькою щільністю перевірок на парність (LDPC). Вони не лише забезпечують високу корекційну здатність, наближаючись до межі Шеннона, але й дозволяють маскувати структуру повідомлення за рахунок керованої надлишковості, ускладнюючи статистичний аналіз трафіку сторонніми спостерігачами.

Адаптивне кодування та фонтанні схеми в умовах динамічних каналів

Для мереж 6G та IoT-середовищ характерні канали з високим рівнем втрат пакетів та стирань (Binary Erasure Channel, BEC). Класичні блокові LDPC-коди потребують фіксованої структури та зворотного зв'язку, що знижує їхню ефективність у динамічних умовах.

Вирішенням цієї проблеми є інтеграція LDPC-кодів із фонтанними кодами (LT-кодами та кодами Raptor). Фонтанні коди дозволяють генерувати необмежений потік кодованих символів, для декодування яких достатньо отримати будь-яку підмножину об'ємом, що незначно перевищує розмірність вихідного повідомлення (K+єKK+ϵK). У запропонованій моделі застосовується гібридна схема Raptor-кодування, де LDPC-код виступає у ролі прекодера (зовнішнього коду), що забезпечує базову структурну завадостійкість, а LT-кодування додає рівень стохастичності.

Для умов обмеженої надлишковості та високих вимог до достовірності (наприклад, у телехірургії або BCI-системах) обґрунтовано застосування декодування за критерієм максимальної правдоподібності (ML-декодування). Моделювання показує, що ML-декодування LT-компоненти у поєднанні з LDPC-прекодером забезпечує до 98% успішного відновлення даних при надлишковості ϵ=0.05ϵ=0.05, що значно перевищує показники класичного ітеративного декодування (BP). Це дозволяє реалізувати багаторівневе приховування: первинний рівень захищається статистичним шумом та стохастичною генерацією ступенів, а вторинний – структурною надлишковістю LDPC.

Застосування перемеження для руйнування кореляції помилок

У каналах з корельованими завадами (багатопроменеве поширення, завмирання Релея) ефективність LDPC-декодерів різкопадає через локалізацію помилок у графі Таннера. Для підвищення завадостійкості та прихованості запропоновано застосування техніки перемеження (інтерліювання).

Використання міжблокового та згорткового перемеження дозволяє рівномірно розподілити пакетні помилки по всьому кодовому слову. Це руйнує короткі цикли в графі Таннера на вході декодера, покращує умови для ітеративного обміну повідомленнями (Belief Propagation) та знижує ймовірність помилкового декодування. Крім того, вектори перестановки, що використовуються при перемеженні, можуть слугувати додатковим криптографічним ключем для легітимного приймача, підвищуючи структурну непрозорість каналу для атакуючого.

Побудова нерегулярних LDPC-кодів з покращеними комбінаторними властивостями

Ключовим фактором криптографічної стійкості та прихованості LDPC-кодів є складність їхньої структури для зловмисника. Регулярні LDPC-коди мають передбачувану топологію, що вразливо до атак зворотного інжинірингу. Тому в роботі запропоновано метод побудови нерегулярних LDPC-кодів з контрольованими комбінаторними властивостями.

Запропонований підхід базується на комбінації евклідової геометрії над скінченними полями GF(2S)$GF(2S)$ та удосконаленої техніки «дроблення-і-заповнення» (split-and-fill) на основі набору діаграм. Алгоритм побудови включає такі етапи:

1. **Формування базової структури:** Використання точок та прямих евклідової геометрії EG(m,2S)$EG(m,2S)$ для створення «сировини» регулярного набору діаграм з гарантованим обхватом (girth) не менше 6.
2. **Часткове дроблення-і-заповнення:** Застосовується до перевірочних вузлів для усунення надлишкових ненульових елементів у матриці контролю парності HH та формування асиметричного ступеневого профілю.
3. **Повне дроблення-і-заповнення:** Точне налаштування ступенів змінних та перевірочних вузлів відповідно до цільового розподілу λ(x)$\lambda(x)$ та ρ(x)$\rho(x)$.

Переваги запропонованого методу:

– *Збільшення обхвату графа Таннера*. Відсутність коротких циклів (довжина циклу ≥6≥6) знижує поверх помилок (error floor) та покращує збіжність декодера.

– *Висока ентропія матриці HH*. Асиметричний ступеневий профіль та стохастичність операцій дроблення роблять структуру коду непередбачуваною. Без знання точних параметрів розподілу та векторів перестановки відновлення інформації стороннім спостерігачем є обчислювально неможливим.

– *Гнучкість.* Можливість побудови кодів довільної довжини та швидкості, що критично важливо для адаптивних систем 6G.

Інтеграція у постквантову криптосистему Нідеррайтера

Отримані нерегулярні LDPC-коди з високими комбінаторними властивостями та великим обхватом графа Таннера пропонується використовувати як математичну основу для постквантової криптосистеми Нідеррайтера.

Безпека криптосистеми Нідеррайтера базується на складності задачі декодування довільного лінійного блокового коду (Syndrome Decoding Problem), яка є NP-складною і стійкою до атак як з боку класичних, так і квантових комп'ютерів (алгоритм Шора). Використання саме оптимізованих LDPC-кодів дозволяє вирішити головну проблему кодівих криптосистем – великий розмір відкритого ключа (матриці H*H*). Завдяки розрідженій структурі та можливості компактного опису матриці через геометричні властивості та параметри дроблення, розмір ключа значно зменшується, що робить систему придатною для ресурсообмежених смарт-пристроїв (IoT, сенсори).

При цьому, сама процедура кодування/декодування LDPC виступає як додатковий фізичний рівень стеганографічного захисту: зашифрований синдром виглядає як звичайна надлишкова інформація каналу зв'язку, що унеможливлює факт виявлення прихованої передачі даних (Low Probability of Detection, LPD).

Висновки

Традиційні методи захисту (TLS/SSL) не здатні забезпечити належний рівень прихованості та стійкості до аналізу трафіку в умовах мереж 6G та смарт-середовищ. Необхідний перехід до багаторівневого захисту з інтеграцією засобів на фізичному рівні.

Поєднання фонтанних кодів (Raptor) з LDPC-прекодуванням та ML-декодуванням забезпечує високу надійність передачі в каналах зі стираннями та динамічними завадами, а також додає рівень стохастичної прихованості.

Застосування техніки часткового та повного дроблення-і-заповнення на основі евклідової геометрії дозволяє будувати нерегулярні LDPC-коди з контролем обхвату графа Таннера. Це підвищує завадостійкість та унеможливлює структурний аналіз коду атакуючим.

Інтеграція побудованих нерегулярних LDPC-кодів у криптосистему Нідеррайтера дозволяє створити постквантову гібридну систему, яка забезпечує одночасно виправлення помилок, стійкість до квантових атак та високий рівень прихованості факту передачі конфіденційної інформації.

Література

1. *Gallager R. Low-Density Parity-Check Codes // IRE Trans. Inf. Theory, 2022. – Vol. 8(1).*
2. *Shokrollahi A. Raptor Codes // IEEE Transactions on Information Theory, 2016. – vol. 52. – №. 6.*
3. *Richardson T., Urbanke R. Modern Coding Theory. – Cambridge University Press, 2018. – 576 p.*
4. *Abdel Hakeem S. A., Hussein H.H., Kim H. Security Requirements and Challenges of 6G Technologies and Applications // Sensors, 2022. – No 22.*
5. *Kou Y., Lin S., Fossorier M.P.C. Low density parity-check codes based on finite geometries: a rediscovery and new results // IEEE Transactions on Information Theory, 2021. – Vol. 47. – Iss. 7.*

Гайдур Г.І. д.т.н., професор,
ННІКБЗІ, ДУІКТ, Київ, Україна
Селіверстов М. В.
студент групи ІКБм-61, ННІКБЗІ,
ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАХИСТУ ВІДДАЛЕНИХ КОРИСТУВАЧІВ НА ОСНОВІ КОНЦЕПЦІЇ ZTNA

У роботі досліджено технологію захисту віддалених користувачів на основі концепції ZTNA. Розглянуто проблеми класичних VPN-систем та запропоновано архітектурне рішення із застосуванням мікросегментації, контекстного контролю і динамічної авторизації для підвищення кіберстійкості підприємств.

Ключові слова: ZTNA, нульова довіра, віддалений доступ, інформаційна безпека, мікросегментація, VPN, кіберзахист.

1. Актуальність та постановка проблеми

Перехід до цифрових бізнес-процесів та масового переходу організацій на гібридні чи віддалені формати роботи зумовив, що традиційні моделі кібербезпеки за периметром вичерпують свою ефективність. Всім відомий класичний підхід на основі віртуальних приватних мереж (VPN), який довгий час вважався стандартом захисту віддаленого доступу, створює суттєві вразливості [1].

Основними проблемами існуючих систем є надмірність довіри, відсутність постійного контролю та складність адміністрування [2].

Після успішної аутентифікації за допомогою VPN користувач отримує повний доступ до внутрішньої мережі, що дозволяє зловмисникам у разі компрометації облікового запису безперешкодно переміщуватися між ресурсами. Саме це є надмірністю довіри.

Відсутність постійного контролю стану пристрою можне виглядати як доступ, що часто надається з неконтрольованих або уражених шкідливим ПЗ особистих пристроїв. Особливо це відчувається при застосуванні концепції BYOD.

Традиційна інфраструктура на базі фаєрволів і VPN-шлюзів створює вузькі місця в мережі та високе навантаження на адміністраторів. Це зумовлює виникнення складності адміністрування та масштабування інформаційної інфраструктури.

Таким чином, виникає потреба в розробці та впровадженні новітніх технологій захисту, які базуються на принципах мінімальних привілеїв та постійної верифікації, зокрема концепції Zero Trust Network Access (ZTNA).

2. Розробка рішення на основі концепції ZTNA

Для розв'язання вищезазначених проблем пропонується архітектурне рішення на основі технології Zero Trust Network Access, головним гаслом якої є: «Ніколи не довіряй, завжди перевіряй» [1-4].

Основні елементи рішення повинні бути реалізовані на основі динамічної аутентифікації та авторизації, контексному контролі, приховуванні ресурсів та мікросегментації.

Динамічна аутентифікація та авторизація полягає в тому, що доступ надається не до мережі загалом, а до конкретних корпоративних додатків чи сервісів на основі суворих критеріїв ідентифікації користувача та стану його пристрою.

Контекстний контроль реалізується шляхом того, що система перед наданням доступу автоматично перевіряє рівень захисту кінцевої точки (наявність антивірусу, актуальність оновлень ОС, відсутність рут-прав тощо).

Корпоративні додатки приховані від прямого виявлення в мережі (наприклад, через публічний IP-адрес) і стають доступними лише після успішного проходження багаторівневої перевірки через шлюз безпеки ZTNA.

Мікросегментація мережі за рахунок застосування принципу найменших привілеїв, унеможливлює несанкціоноване переміщення зловмисника внутрішньою мережею навіть у разі компрометації одного вузла [4].

3. Наукові та практичні пропозиції

Для вирішення поставленої проблеми можна надати такі практичні рекомендації щодо впровадження технології:

1. Алгоритм адаптивної оцінки ризиків. Розробити структурну схему алгоритму, яка в режимі реального часу аналізує поведінкові фактори віддаленого користувача (час доби, географічне положення, аномалії в трафіку) та динамічно коригує права доступу.

2. Розробити проєкт гібридної архітектури. Проєкт має включати в себе етапний план переходу від застарілої VPN-інфраструктури до хмарно-орієнтованого рішення ZTNA з інтеграцією наявних систем управління ідентифікацією.

3. Оцінити розроблене рішення з точки зору економічної та технічної ефективності. Розрахунки що впровадження запропонованої технології повинні знизити ризики успішних атак та витоку даних через віддалені канали, а також оптимізувати навантаження на мережеві канали компанії.

4. Висновки

Застосування технології ZTNA для захисту віддалених користувачів є ефективним інструментом забезпечення кіберстійкості сучасних організацій. Запропоновані в роботі підходи дозволяють мінімізувати ризики несанкціонованого доступу, забезпечити повний контроль над кінцевими точками та адаптувати систему безпеки до динамічних загроз сучасного кіберпростору.

Перелік посилань

1. Rose S. Zero Trust Architecture / S. Rose, O. Borchert, S. Mitchell, S. Connelly. - NIST Special Publication 800-207, 2020. — 50 p.

2. Mavroudis V. Zero-Trust Network Access (ZTNA): Principles, Architectures, and Applications / V. Mavroudis. — Alan Turing Institute, 2024. — arXiv:2410.20611 https://doi.org/10.48550/arXiv.2410.20611

3. Zero Trust Enforcement Using Microsegmentation, Identity-Aware Proxies, and Continuous Adaptive Risk Assessment // International Journal of Computer Applications Technology and Research. — 2025. — Vol. 14, Issue 07. — P. 61–77. DOI:10.7753/IJCATR1407.1006

4. Zero Trust Microsegmentation Guidance / Cloud Security Alliance (CSA). — 2026. https://cloudsecurityalliance.org/artifacts/zero-trust-microsegmentation-guidance

Гайдур Г.І. д.т.н., професор,
ННІКБЗІ, ДУІКТ, Київ, Україна
Проценко М.В.
студент групи ІКБм-61, ННІКБЗІ,
ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАСТОСУВАННЯ МОДЕЛЕЙ ДОСТУПУ НА ОСНОВІ КОНТЕКСТНОЇ АУТЕНТИФІКАЦІЇ

Сучасні інформаційні системи потребують нових підходів до впровадження моделей доступу, а саме моделей контролю доступу на основі контекстної аутентифікації. Проаналізовано ключові проблеми традиційних статичних методів захисту даних в умовах масового переходу на дистанційну роботу та зростання кількості кіберзагроз. Запропоновано шляхи вирішення проблеми шляхом динамічного аналізу контекстних параметрів користувача (місцеперебування, час, стан пристрою та поведінкові патерни). Надано практичні рекомендації щодо інтеграції адаптивних механізмів безпеки для підвищення рівня захищеності корпоративних ресурсів без зниження зручності користування

Ключові слова: контекстна аутентифікація, адаптивний контроль доступу, інформаційна безпека, архітектура нульової довіри, управління ризиками, кіберзагрози.

1. Вступ та актуальність проблематики

У сучасних умовах периметрові моделі безпеки втрачають свою ефективність. Це пов’язано з стрімким розвитком цифрових технологій, масової міграції серверних інфраструктур у хмари та переходу підприємств і державних установ на гібридні чи повністю віддалені формати роботи. Класичні системи контролю доступу, засновані переважно на статичних ідентифікаторах (логінах, паролях) та фіксованих ролях (наприклад, Role-Based Access Control — RBAC), не можуть забезпечити належний рівень захисту інформаційних ресурсів підприємств.

Звіти з кіберзлочинності демонструють зміщення фокусу атак з прямого зламу мережевої інфраструктури на компрометацію облікових записів легітимних користувачів. Використання викрадених, але формально «валідних» облікових даних дозволяє кіберзловмисникам безперешкодно проникати всередину корпоративного периметра підприємства. Відтак, виникає потреба у переході від статичної перевірки під час ідентифікації до безперервного моніторингу та адаптивного управління доступом на основі аналізу контексту сеансу зв'язку [1].

2. Аналіз проблеми безпеки в статичних моделях доступу

Головною проблемою традиційних систем є їхня бінарність («надано доступ / відмовлено») та прив'язка виключно до моменту входу в систему [1, 2]. Якщо користувач успішно пройшов аутентифікацію на початку сеансу, подальші його дії, як правило, не піддаються додатковій перевірці ризиків. Це породжує низку критичних вразливостей (рис.1):

1. Отримавши права доступу до певного рівня, суб'єкт зберігає їх протягом усього часу сесії незалежно від зміни зовнішніх умов чи поведінкових аномалій.

2. Традиційні системи не враховують, звідки здійснюється доступ (наприклад, небезпечна або невідома публічна мережа Wi-Fi), коли він відбувається (неробочий час, аномальні часові пояси) та на якому пристрої (необслуговуваний або заражений шкідливим ПЗ пристрій).

3. Перехоплення сесійних токенів дозволяє зловмисникам повністю імітувати легітимного користувача без потреби повторного введення пароля.



Рис. 1. Критичні вразливості існуючих моделей доступу

3. Концепція контекстної аутентифікації

Ефективним шляхом розв'язання проблематики є впровадження технології контекстної (адаптивної) аутентифікації (Context-Aware Authentication). Цей підхід передбачає розширення процесу прийняття рішення про надання доступу шляхом динамічного аналізу широкого спектра параметрів у реальному часі.

Архітектура контекстної аутентифікації оперує такими категоріями даних:

Географічний та мережевий контекст.

Апаратний та програмний контекст.

Поведінковий контекст.

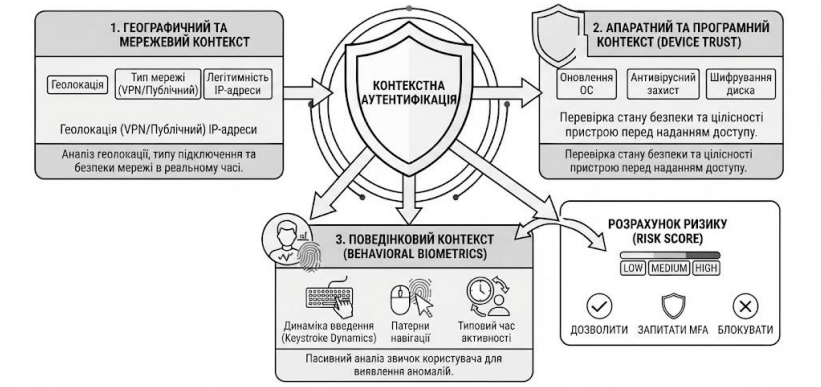


Рис. 2. Архітектура контекстної аутентифікації

1. Географічний та мережевий контекст. Аналіз геопозиції користувача, визначення легітимності автономної системи (ASN), типу підключення

(корпоративний VPN, мобільний інтернет, публічний хот-спот) та перевірка наявності IP-адреси в базах індикаторів компрометації (IoC).

2. Апаратний та програмний контекст (Device Trust). Оцінка стану кінцевої точки (Endpoint State) — наявність актуальних антивірусних баз, оновлень операційної системи, цілісність конфігурацій та наявність засобів шифрування диска.

3. Поведінковий контекст (Behavioral Biometrics). Безперервний пасивний аналіз звичок користувача (динаміка введення тексту на клавіатурі, почерк використання маніпуляторів, типові сценарії навігації в інформаційній системі та час активності).

Завдяки постійному обчисленню інтегрального показника ризику (Risk Score) система здатна в реальному часі змінювати рівень довіри до сеансу: від повного надання прав до висунення вимоги додаткової багатофакторної аутентифікації (MFA) або повного блокування доступу у разі виявлення критичних аномалій.

4. Практичні рекомендації щодо впровадження

Для успішної реалізації моделей контекстного доступу на рівні підприємств та організацій пропонується дотримуватися таких практичних рекомендацій:

Інтеграція з архітектурою Zero Trust: Контекстна аутентифікація не повинна існувати ізольовано; її слід розглядати як фундаментальний компонент концепції «Нульової довіри» (Zero Trust Architecture, відповідно до стандартів NIST SP 800-207), де діє правило «ніколи не довіряй, завжди перевіряй».

Застосування динамічних політик доступу (Attribute-Based Access Control — ABAC). Перехід від жорстких ролей до гнучких політик, що враховують комбінацію атрибутів користувача, ресурсу, середовища та поточного рівня ризику.

Налаштування механізмів таким чином, щоб додаткові перевірки ініціювалися виключно у випадках високого ризику або зміни контексту, що дозволяє зберегти баланс між високим рівнем безпеки та зручністю виконання робочих завдань.

Машинне навчання та адаптивне моделювання. Використання алгоритмів штучного інтелекту та машинного навчання для побудови базових профілів нормальної поведінки користувачів і швидкого виявлення відхилень.

Перелік посилань

5. Data Breach Investigations Report (DBIR) [Електронний ресурс] / Verizon. — 2026. — Режим доступу: https://www.verizon.com/business/resources/reports/dbir/ (дата звернення: 06.10.2026).
6. NIST SP 800-207. Zero Trust Architecture [Електронний ресурс] / S. Rose, O. Borchert, S. Mitchell, S. Connelly // National Institute of Standards and Technology. — 2020. — Режим доступу: https://csrc.nist.gov/publications/detail/sp/800-207/final (дата звернення: 06.10.2026).
7. ДСТУ ISO/IEC 27001:2023. Інформаційна безпека, кібербезпека та захист конфіденційності. Системи керування інформаційною безпекою. Вимоги (ISO/IEC 27001:2022, IDT) [Електронний ресурс] / Надано чинності: наказ ДП «УкрНДНЦ» від 22 серп. 2023 р. № 345. — Київ : ДП «УкрНДНЦ», 2023.

Ляшенко Я.В.
студент групи ІКБм-51, ННІКЗІ ДУІКТ
Київ, Україна

УДОСКОНАЛЕННЯ СИСТЕМИ МЕНЕДЖМЕНТУ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ОРГАНІЗАЦІЇ НА ОСНОВІ СТАНДАРТУ ISO/IEC 27001

У роботі досліджено питання удосконалення системи менеджменту інформаційної безпеки організації на основі вимог міжнародного стандарту ISO/IEC 27001. Проаналізовано сучасний стан управління інформаційною безпекою, виявлено основні проблеми та недоліки існуючих систем. Запропоновано комплексну модель удосконалення СУІБ, яка включає політику безпеки, оцінку ризиків, організаційні та технічні заходи, управління інцидентами та механізми постійного покращення. Результати дослідження можуть бути використані організаціями для підвищення рівня захищеності інформаційних активів та підготовки до сертифікації.

Ключові слова: інформаційна безпека, менеджмент інформаційної безпеки, система управління інформаційною безпекою, ISO/IEC 27001, оцінка ризиків, кібер стійкість.

У сучасних умовах стрімкої цифровізації всіх сфер діяльності питання забезпечення інформаційної безпеки набуває стратегічного значення для будь-якої організації. Зростання кількості кібератак, їх складності та цілеспрямованості створює серйозні ризики втрати конфіденційної інформації, порушення безперервності бізнес-процесів, фінансових збитків та репутаційних втрат. За даними міжнародних досліджень, середня вартість одного серйозного кіберінциденту для організації може сягати сотень тисяч і навіть мільйонів доларів, не враховуючи довгострокових наслідків. Традиційний підхід до інформаційної безпеки, який обмежується впровадженням окремих технічних засобів захисту, уже не забезпечує належного рівня захищеності, тому необхідним є системний, процесний підхід, що охоплює технічні, організаційні, кадрові та правові аспекти. Саме такий підхід закладено в міжнародному стандарті ISO/IEC 27001, який визначає вимоги до системи управління інформаційною безпекою і базується на циклі PDCA. Стандарт передбачає встановлення політики інформаційної безпеки, визначення сфери застосування СУІБ, проведення оцінки ризиків, впровадження комплексу заходів захисту, моніторинг ефективності системи та її постійне вдосконалення [1]. Важливою особливістю стандарту є його універсальність - він може застосовуватися в організаціях будь-якого розміру та сфери діяльності.

Незважаючи на широке міжнародне визнання стандарту ISO/IEC 27001, на практиці багато організацій стикаються з труднощами при його впровадженні. Аналіз існуючих систем інформаційної безпеки свідчить про наявність низки типових проблем: відсутність формалізованої політики інформаційної безпеки або її декларативний характер, епізодичне чи взагалі відсутнє проведення оцінки ризиків, фрагментарне впровадження заходів захисту без урахування взаємозв’язків між процесами, недостатня увага до підготовки персоналу та формування культури інформаційної безпеки, а також слабкий розвиток механізмів моніторингу, внутрішнього аудиту та постійного покращення СУІБ [2, с. 15-18; 3]. Ці недоліки суттєво знижують ефективність захисту інформаційних

активів і підвищують вразливість організації до сучасних кіберзагроз, особливо коли технічні засоби захисту впроваджені, але організаційні процеси управління безпекою не налагоджені.

З метою усунення зазначених проблем та удосконалення системи менеджменту інформаційної безпеки запропоновано комплексну модель, яка базується на вимогах стандарту ISO/IEC 27001 і включає кілька взаємопов'язаних складових. Першою складовою є розробка та впровадження політики інформаційної безпеки, яка визначає основні принципи, цілі та зобов'язання керівництва, а також розподіляє відповідальність між структурними підрозділами. Другою складовою виступає проведення комплексної оцінки ризиків, що включає ідентифікацію інформаційних активів, визначення загроз і вразливостей, оцінку ймовірності реалізації загроз та можливого впливу на організацію з подальшим вибором адекватних заходів обробки ризиків. Третьою складовою є впровадження комплексу організаційних і технічних заходів захисту, де організаційні заходи охоплюють процедури управління доступом, управління змінами, управління інцидентами, забезпечення безперервності діяльності та навчання персоналу, а технічні - контроль доступу, захист мереж, шифрування даних, моніторинг подій безпеки та інші засоби, передбачені стандартом ISO/IEC 27002 [2]. Четвертою складовою є створення ефективних процесів управління інцидентами та забезпечення безперервності діяльності, а п'ятою - впровадження механізмів моніторингу, внутрішнього аудиту та постійного покращення СУІБ відповідно до циклу PDCA.

Практичне значення запропонованої моделі полягає в тому, що вона дозволяє організаціям системно підійти до побудови або удосконалення СУІБ, підвищити рівень захищеності інформаційних активів, зменшити ризики виникнення кіберінцидентів та створити умови для подальшої сертифікації за стандартом ISO/IEC 27001. Впровадження моделі також сприяє підвищенню культури інформаційної безпеки серед персоналу та покращенню взаємодії між різними підрозділами організації [4, с. 42-45]. Таким чином, удосконалення системи менеджменту інформаційної безпеки на основі стандарту ISO/IEC 27001 є актуальним і ефективним напрямом підвищення кіберстійкості сучасних організацій, оскільки системний підхід, що поєднує організаційні, технічні та кадрові аспекти захисту інформації, дозволяє не лише знизити рівень ризиків, а й забезпечити довгострокову стійкість організації в умовах постійно зростаючих кіберзагроз.

Список використаних джерел

1. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection - Information security management systems - Requirements. Geneva : ISO, 2022. 34 p.
2. ISO/IEC 27002:2022. Information security, cybersecurity and privacy protection - Information security controls. Geneva : ISO, 2022. 152 p.
3. Національний стандарт України ДСТУ ISO/IEC 27001:2023. Інформаційна безпека, кібербезпека та захист конфіденційності. Системи управління інформаційною безпекою. Вимоги. Київ : ДП «УкрНДНЦ», 2023. 38 с.
4. Грайворонський М. В., Новіков О. М. Безпека інформаційно-комунікаційних систем. Київ : Видавнича група BHV, 2021. 360 с.

Карапиш Б.О.
студент групи БСДМ-61, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ АВТОМАТИЗАЦІЇ ТЕСТУВАННЯ НА ПРОНИКНЕННЯ ВЕБ-ЗАСТОСУНКІВ НА ОСНОВІ МУЛЬТИАГЕНТНИХ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

Робота присвячена автоматизації веб-пентесту за допомогою кількох ШІ-агентів під керуванням оркестратора. Базовою моделлю обрано Claude. Агенти ділять між собою етапи: розвідку, пошук вразливостей, перевірку експлуатовності та звітування з оцінкою за CVSS. Перевірки побудовано на OWASP WSTG і Top 10, випробування проведено на OWASP Juice Shop. Підхід скорочує рутину й час тестування, зберігаючи контроль фахівця над критичними знахідками.

Ключові слова: тестування на проникнення, веб-застосунки, штучний інтелект, мультиагентні системи, великі мовні моделі, Claude, OWASP, автоматизація.

Під час комерційних пентестів веб-застосунків більшість часу йде не на екзотику, а на рутину: перебрати точки входу, перевірити типові помилки автентифікації й контролю доступу, відсіяти хибні спрацювання сканерів. Цю рутину й варто автоматизувати.

Автоматичні сканери (DAST) добре ловлять прості патерни, але майже не розуміють бізнес-логіку застосунку, тому пропускають логічні вразливості й дають багато шуму. Ручний пентест цей шум прибирає, зате кепсько масштабується.

Великі мовні моделі (LLM) змінюють тут баланс: вони вже читають відповіді застосунку, пропонують наступний крок і складають корисне навантаження [4]. Слабке місце однієї моделі в тому, що вона губить контекст на довгому ланцюжку дій. Через це в роботі задача ділиться між кількома агентами, у кожного своя вузька роль.

Систему побудовано як набір агентів під керуванням оркестратора. Базовою моделлю обрано Claude (Anthropic). Оркестратор ставить підзадачі й передає контекст; агент розвідки збирає точки входу та API; агент виявлення шукає вразливості; агент експлуатації перевіряє, чи знахідка реально відтворюється; агент звітування описує її, виставляє оцінку за CVSS v3.1 і дає рекомендацію щодо усунення. Звичні інструменти (Burp Suite, sqlmap тощо) агенти викликають як зовнішні засоби.

Перелік перевірок спирається на OWASP Web Security Testing Guide і OWASP Top 10 [1, 2], а для інтерфейсів REST/API на OWASP API Security Top 10 [3]. Це робить результати зіставними з ручним аудитом. Схоже розбиття на модулі показано в PentestGPT [4], але там основний акцент на інфраструктурному пентесті, а в цій роботі на вебі та його бізнес-логіці.

Для випробування взято OWASP Juice Shop, навчальний застосунок з понад 60 відомими вразливостями. На ньому зручно порахувати, скільки з них система знаходить сама і наскільки це швидше за ручний прохід.

Прибрати людину з процесу повністю не вийде: модель може вигадати неіснуючу вразливість або не помітити нетипову, тому критичні знахідки

підтверджує фахівець. Навіть із цим обмеженням підхід знімає з пентестера рутину й лишає йому власне аналіз, заради чого все й робиться.

Перелік використаної літератури:

1. OWASP Web Security Testing Guide (WSTG), v4.2. URL: https://owasp.org/www-project-web-security-testing-guide/.
2. OWASP Top 10:2025. URL: https://top10.owasp.org/2025/.
3. OWASP API Security Top 10 (2023). URL: https://owasp.org/API-Security/.
4. Deng G., Liu Y., Mayoral-Vilches V. et al. PentestGPT: An LLM-empowered Automatic Penetration Testing Tool. 2023. URL: https://arxiv.org/abs/2308.06782.

Цебак О.А.
аспірант кафедри БІТ, НУ «ЛП»,
Львів, Україна

ПЕРСПЕКТИВИ ГІБРИДНИХ ГЕНЕРАТОРІВ ПСЕВДОВИПАДКОВИХ ЧИСЕЛ

У тезах розглянуто гібридні генератори псевдовипадкових чисел, що поєднують лінійні генератори, хаотичні відображення, операції на еліптичних кривих та фізичні джерела ентропії. Проаналізовано три основні класи таких генераторів, їхні переваги та обмеження, зокрема деградацію цифрового хаосу за скінченної точності обчислень і недостатність статистичного тестування для обґрунтування криптографічної стійкості. Окреслено перспективні напрями розвитку: криптографічну постобробку виходу, адаптивне перезасівання від фізичних джерел, апаратну реалізацію для пристроїв інтернету речей і стандартизацію методик оцінювання.

Ключові слова: гібридний генератор псевдовипадкових чисел, хаотичні відображення, еліптичні криві, криптографічна стійкість, статистичне тестування.

Гібридні генератори псевдовипадкових чисел (ГПВЧ) поєднують генератори різної природи (лінійні генератори LFSR і LCG, хаотичні відображення, алгебраїчні конструкції на еліптичних кривих (ЕК) та фізичні джерела ентропії) так, щоб компенсувати недоліки кожного з них. Класичні генератори швидкі й відтворювані, але лінійність і відновлюваний внутрішній стан обмежують їх застосування в криптографії; хаотичні генератори дають нелінійність і чутливість до початкових умов, проте сильно залежать від параметрів і точності обчислень. Тому гібридні схеми розглядають як компроміс між швидкодією та стійкістю [1, с. 125, 131], і інтерес до них зростає з поширенням легковагової криптографії та інтернету речей (IoT), де потрібна висока швидкодія за мінімальних апаратних ресурсів [1, с. 135].

Можна виділити три основні класи гібридних ГПВЧ:

- LFSR і хаотичні відображення: у схемі з двома LFSR, стан яких пов’язаний з параметрами дискретного хаотичного відображення, за відомими вихідними бітами неможливо отримати частину ключа без перебору всього ключа, а згенеровані послідовності проходять тести NIST, TestU01 і DIEHARD [2, с. 190];
- ЕК і хаос: генерування послідовності точок ЕК над скінченним полем, кероване хаотичним відображенням, поєднує статистичні властивості

обох генераторів і подовжує період порівняно з генератором на ЕК без модуляції [3, с. 31, 35];

- фізичне джерело і ГПВЧ: генератор справжніх випадкових чисел (TRNG) на основі хаотичної системи Sprott 94 G, реалізований на FPGA, подає додаткові входи у функції переходу й виходу детермінованого генератора, що дозволяє виконати вимогу R4 (неможливість обчислити попередні виходи за відомим внутрішнім станом) [4, с. 239]; адаптивне перезасівання від довірених джерел зберігає непередбачуваність навіть за часткової компрометації стану [1, с. 134].

Спільна ідея цих підходів полягає в тому, що лінійний або фізичний компонент забезпечує довгий період і швидкодію, а нелінійний – складність і непередбачуваність. Особливу увагу привертають генератори на ЕК, стійкість яких спирається на складність задачі дискретного логарифмування на ЕК (ECDLP), тоді як хаотичні системи забезпечують високу чутливість до початкових умов і параметрів [3, с. 31]. Крім того, розглядають гібриди на основі системи залишкових класів, що поєднують модулярну арифметику, нелінійну динаміку та паралельні обчислення [1, с. 125].

Водночас гібридизація має обмеження. Якість хаотичних генераторів сильно залежить від налаштування параметрів і реалізації зі скінченною точністю: у цифровому вигляді траєкторії логістичного та тент-відображень вироджуються в короткі цикли [1, с. 129, 133], а поєднання з LFSR розглядають як один зі способів протидії такій деградації [2, с. 190]. Стійкість більшості схем обґрунтовують лише статистичними тестами, які є необхідною, але не достатньою умовою, тому методики оцінювання мають виходити за межі наявних наборів NIST SP 800-22, DIEHARD і TestU01 [1, с. 134]; невдале поєднання компонентів не підвищує, а може й знизити стійкість. Операції на ЕК обчислювально витратні, проте модуляція хаотичним відображенням подовжує період і дозволяє використовувати менші поля зі спрощеною арифметикою [3, с. 35].

Перспективними напрямами є безпечні гібридні архітектури, хаотичні відображення, стійкі до обмеженої точності, адаптивне перезасівання, апаратне прискорення та стандартизовані методики тестування [1, с. 125, 134], а також застосування машинного навчання для пошуку прихованих залежностей, хоча воно пов’язане з витратами на навчання моделей і проблемами відтворюваності [1, с. 125, 135]. На нашу думку, доцільним є також використання криптографічного примітиву (потокового шифру чи геш-функції) як вихідного шару гібридного генератора. У межах дослідження автора розглядаються комбінації LFSR з генераторами на ЕК, які перевіряються батареями TestU01.

Отже, гібридні ГПВЧ є перспективним компромісом між швидкодією і стійкістю, однак потребують формального обґрунтування безпеки, а не лише статистичного тестування.

Перелік використаної літератури

1. Abdul Karim M. et al. From Deterministic Chaos to Machine Learning: A Comparative Review of Pseudo-Random Number Generators for Modern Computing and Cryptography. *Chaos and Fractals*. 2026.

2. Alhadawi H. S. et al. Designing a pseudorandom bit generator based on LFSRs and a discrete chaotic map. *Cryptologia*. 2019.
3. Reyad O. et al. On Pseudo-Random Number Generators Using Elliptic Curves and Chaotic Systems. *Applied Mathematics & Information Sciences*. 2014.
4. Avaroğlu E. et al. Hybrid pseudo-random number generator for cryptographic systems. *Nonlinear Dynamics*. 2015.

Романюк О.Н., д.т.н., проф.
Майданюк В.П., к.т.н, проф.
Романюк О.В.к.т.н, доцент
Вінницький національний технічний університет
Вінниця, Україна

ВИЯВЛЕННЯ ВРАЗЛИВОСТЕЙ ПРОГРАМНИХ МОДУЛІВ ОБРОБКИ ЦИФРОВИХ ЗОБРАЖЕНЬ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. Розглянуто застосування штучного інтелекту для виявлення вразливостей програмних модулів обробки цифрових зображень. Запропоновано поєднання нейромережевого аналізу вихідного коду, відстеження потоків даних і перевірки підозрілих ділянок засобами фаззингу та інструментування виконання. Основну увагу приділено помилкам обчислення розмірів буферів, порушенням меж пам'яті та надмірному споживанню ресурсів. Визначено вимоги до навчальних даних, порядок підтвердження результатів і показники оцінювання методу. Запропонований підхід орієнтований на використання під час розроблення та аудиту програмного забезпечення.

Модулі обробки цифрових зображень використовуються у вебсервісах, мультимедійних системах, медичних застосунках і засобах комп'ютерного зору. Під час декодування вони інтерпретують заголовки, розміри, таблиці та стиснені дані, які можуть надходити з недовірених джерел. Помилки перевірки цих даних здатні спричинити аварійне завершення програми, пошкодження пам'яті або відмову в обслуговуванні. Реальність загрози ілюструє CVE-2023-4863 - вразливість переповнення буфера в динамічній пам'яті бібліотеки libwebp [1].

Особливістю таких модулів є залежність безпеки від узгодженості параметрів зображення, структури файлу та операцій із пам'яттю. Перевірка окремого рядка коду не завжди виявляє помилку, оскільки її причина може міститися в іншій функції. Штучний інтелект доцільно застосовувати для визначення ділянок, що потребують першочергового аналізу, з подальшим підтвердженням знайдених ризиків.

Мета роботи - розроблення підходу до виявлення вразливостей модулів обробки цифрових зображень, який поєднує нейромережеву оцінку вихідного коду, аналіз залежностей даних і динамічну перевірку. Об'єктом дослідження є модулі декодування та перетворення зображень, реалізовані мовами C/C++. Предметом дослідження є методи локалізації потенційно небезпечних операцій під час обробки недовірених графічних даних.

Модель LineVul демонструє можливість застосування трансформерів для прогнозування вразливості функцій і ранжування рядків коду [2]. VulBERTa використовує попереднє навчання представлень вихідного коду для подальшого виявлення вразливостей [3]. У DeepDFA навчання графових представлень

спирається на принципи аналізу потоків даних [4]. Ці напрями створюють основу для спеціалізованого аналізу графічних модулів. Водночас оцінка моделі або виділення підозрілого рядка самі по собі не доводять наявності дефекту.

На першому етапі виконують підготовку вихідного коду: виділяють функції, будують синтаксичні представлення та встановлюють залежності між значеннями. Позначають джерела недовірених даних - зчитані з файлу ширину, висоту, кількість компонентів, довжини блоків і зміщення. Визначають операції, безпека яких залежить від цих значень: виділення пам’яті, копіювання, індексування та обчислення адрес. Для міжпроцедурних залежностей враховують передавання параметрів і повернення значень.

На другому етапі модель, адаптована до коду графічних бібліотек, оцінює підозрілість функцій. Оцінку поєднують з ознаками досяжності небезпечних операцій від недовірених даних, наявності перевірок меж і використання арифметичних перетворень. Так формують пріоритет перевірки функції. Локалізація має містити фрагмент коду та пояснення зв’язку між вхідним параметром і потенційно небезпечною операцією. Передбачуваний клас помилки уточнюють за результатами аналізу; впевненість моделі не слід трактувати як доведену ймовірність експлуатації.

Наприклад, під час обчислення місткості буфера за шириною, висотою та кількістю компонентів необхідно перевіряти допустимість параметрів і можливість переповнення до виконання множення. Додатково контролюють довжину рядка, вирівнювання та відповідність індексів фактичному обсягу виділеної пам’яті. Наявність перевірки після вже переповненого обчислення може не забезпечувати захисту. Аналіз має враховувати типи операндів, їхню розрядність і порядок виконання перевірок.

На третьому етапі підозрілі ділянки перевіряють статичними аналізаторами та фаззингом - автоматизованим тестуванням зі зміненими вхідними даними. Початковий набір містить коректні зображення різних форматів і приклади пошкоджених файлів. Варіюють значення заголовків, довжини блоків і граничні розміри, зберігаючи достатню структуру файлу для проходження декодера. Інструментування AddressSanitizer та UndefinedBehaviorSanitizer допомагає виявляти порушення доступу до пам’яті й невизначену поведінку. Запуски виконують в ізольованому середовищі з обмеженнями пам’яті та часу.

Відтворювана помилка є підставою для дослідження першопричини. Однакові збої об’єднують за дефектом, що їх спричинив, щоб уникнути завищення кількості знайдених вразливостей. Звіт має містити версію модуля, місце помилки, діагностику інструментування та мінімальний тестовий приклад. Після виправлення повторюють перевірку й додають регресійний тест. Відсутність збою в межах наданого бюджету тестування не доводить безпечності функції.

Для навчання пропонується використовувати підтверджені вразливі фрагменти й відповідні виправлення з відкритих проєктів. Розмітку уточнюють за повідомленнями про вразливості, змінами коду та експертною перевіркою.

Дублікати вилучають, а навчальну й тестову вибірки розділяють за проєктами та часом. Вразливий фрагмент і його близьке виправлення не повинні потрапляти до різних вибірок, оскільки це може спотворювати оцінку узагальнення моделі.

Ефективність пропонується оцінювати за точністю, повнотою, F1-мірою та площею під кривою precision–recall з урахуванням дисбалансу класів. Практичні показники охоплюють якість локалізації, кількість підтверджених дефектів і витрати часу на перевірку. Порівняння зі статичним аналізом та фаззингом без нейромережевого ранжування проводять за однакового набору модулів і бюджету ресурсів. Окремо досліджують внесок ознак потоків даних і специфіки форматів. Кількісні результати мають бути отримані експериментально.

У системі безперервної інтеграції аналіз доцільно запускати після змін декодерів, процедур масштабування та перетворення колірних компонентів. Результати подають розробнику як перелік пріоритетних фрагментів із прив’язкою до версії коду. Для кожного повідомлення зазначають джерело недовіреного значення, операцію використання, виявлені перевірки та статус динамічного підтвердження. Рішення про виправлення приймають після перегляду контексту. Згенеровані засобами ШІ зміни також перевіряють тестами, оскільки локальне усунення попередження може порушити коректність декодування або підтримку допустимих файлів.

Модель може засвоювати особливості конкретних проєктів замість ознак вразливості. Зміни стилю програмування, використання макросів і неповний контекст викликів ускладнюють перенесення результатів. Тому необхідні перевірка на незалежних бібліотеках, аналіз хибних спрацювань і налаштування порога ранжування відповідно до доступного ресурсу аудиту. Крім помилок пам’яті, слід враховувати надмірне споживання ресурсів: формально допустиме зображення може вимагати неприйнятного часу або обсягу пам’яті. Такі випадки оцінюють щодо встановлених обмежень застосунку, відокремлюючи дефекти реалізації від очікуваної обчислювальної складності.

Висновки. Запропоновано підхід, що поєднує аналіз коду засобами штучного інтелекту з перевіркою залежностей даних і динамічним підтвердженням дефектів. Його особливістю є врахування зв’язку параметрів зображення з операціями над пам’яттю та ресурсами. Очікувана практична користь полягає у пріоритизації перевірок під час аудиту й безперервної інтеграції. Подальша робота передбачає реалізацію прототипу, формування спеціалізованої вибірки та дослідження перенесення моделі між графічними бібліотеками.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. National Institute of Standards and Technology. CVE-2023-4863. National Vulnerability Database. URL: https://nvd.nist.gov/vuln/detail/CVE-2023-4863 (дата звернення: 09.10.2026).

2. Fu M., Tantithamthavorn C. LineVul: A Transformer-based Line-Level Vulnerability Prediction. Proceedings of the 19th International Conference on Mining Software Repositories (MSR ’22). 2022. P. 608–620. DOI: 10.1145/3524842.3528452.

3. Hanif H., Maffeis S. VulBERTa: Simplified Source Code Pre-Training for Vulnerability Detection. 2022 International Joint Conference on Neural Networks (IJCNN). 2022. P. 1–8. DOI: 10.1109/IJCNN55064.2022.9892280.

4. Steenhoek B., Gao H., Le W. Dataflow Analysis-Inspired Deep Learning for Efficient Vulnerability Detection. Proceedings of the IEEE/ACM 46th International Conference on Software Engineering (ICSE ’24). 2024. Article 16. P. 1–13. DOI: 10.1145/3597503.3623345.

Романюк О.Н. — завідувач кафедри програмного забезпечення
Майданюк В.П. — професор кафедри програмного забезпечення
Нечипорук М.Л. — аспірант
Вінницький національний технічний університет, Вінниця, Україна

АПАРАТНЕ ПРИСКОРЕННЯ ШИФРУВАННЯ ЦИФРОВИХ ЗОБРАЖЕНЬ У ВБУДОВАНИХ СИСТЕМАХ

Анотація. Розглянуто апаратне прискорення шифрування цифрових зображень у вбудованих системах на основі FPGA. Запропоновано архітектуру прискорювача AES-GCM, яка поєднує конвеєрне шифрування, обчислення тегу автентифікації та передавання даних засобами прямого доступу до пам'яті. Визначено порядок захисту піксельних даних і метаданих, вимоги до унікальності початкових значень та організації буферів. Наведено розрахунок необхідної пропускної здатності для зображень високої роздільності. Сформульовано критерії перевірки коректності, продуктивності, використання ресурсів і енергетичних витрат.

Актуальність дослідження. Вбудовані системи відеоспостереження, промислової діагностики та дистанційного контролю формують цифрові зображення, які можуть містити конфіденційну інформацію. Захист потрібний під час зберігання кадрів і передавання каналами зв'язку. Програмне шифрування конкурує за процесорні ресурси з обробкою зображень та керуванням пристроєм. Доцільність окремого прискорювача визначається можливостями процесора, зокрема наявністю криптографічних інструкцій, вимогами до затримки й допустимим енергоспоживанням. FPGA дає змогу реалізувати спеціалізований тракт обробки та узгодити його з потоком графічних даних.

Мета роботи - розроблення архітектури апаратного прискорювача шифрування цифрових зображень із контролем цілісності для вбудованої системи. Об'єктом дослідження є захист потоку зображень у системі з процесором і програмованою логікою. Предметом дослідження є організація криптографічних обчислень, передавання даних і буферизації за обмежених апаратних ресурсів.

Для базового варіанта пропонується AES-128, який обробляє блоки по 128 біт і використовує ключ довжиною 128 біт [1]. Режим GCM забезпечує автентифіковане шифрування з асоційованими даними [2]. Піксельні дані шифрують, а відкриті метадані включають до перевірки цілісності. Застосування стандартизованої конструкції дає змогу перевіряти реалізацію за еталонними результатами. Апаратна оптимізація має зберігати алгоритм і формат криптографічних результатів.

Запропонований тракт містить контролер передавання, вхідний FIFO, модуль формування блоків, генератор лічильників, ядро AES, модуль виключного АБО, модуль GHASH і вихідний FIFO. Процесор задає адреси буферів, довжину повідомлення та параметри операції через регістровий інтерфейс. Передавання основного масиву виконує контролер прямого доступу

до пам’яті DMA. Модуль AES шифрує значення лічильника; результат поєднують із відкритими даними операцією виключного АБО. GHASH обчислює складову тегу автентифікації. Керування потоком має підтримувати призупинення передавання без втрати або повторного використання блоків.

Для AES порівнюють ітераційну реалізацію з повторним використанням раундового модуля та конвеєрну реалізацію з розділенням раундів. Перша потребує менше ресурсів, друга допускає частіше надходження нових блоків. Огляд апаратних реалізацій AES розглядає компроміс між пропускною здатністю, площею та енергоспоживанням [3]. Вибір залежить від результатів синтезу й обмежень конкретної FPGA. Розклад ключа доцільно формувати під час його встановлення, щоб не повторювати цю операцію для кожного блока одного повідомлення.

Продуктивність тракту визначається також GHASH, у якому існує залежність між послідовними блоками. Саме збільшення швидкості AES не гарантує такого самого прискорення всього AES-GCM. Для потрібної пропускної здатності слід узгодити інтервал приймання блоків обох модулів або застосувати математично еквівалентну паралельну організацію GHASH. Під час синтезу враховують затримки між модулями, тактову частоту, місткість FIFO та можливі паузи на інтерфейсі пам’яті.

Кадр подають як послідовність байтів у визначеному порядку компонентів і рядків. У дескрипторі зазначають фактичну довжину даних і крок рядка, щоб розрізняти пікселі та службове вирівнювання. Прискорювач контролює межі буферів і коректність останнього неповного блока. Для збереження застосовують контейнер із шифротекстом, початковим значенням, тегом і метаданими. Шифротекст не є звичайним файлом зображення, придатним для безпосереднього відкриття графічним декодером.

До асоційованих даних пропонується включати ширину, висоту, формат пікселів, ідентифікатор потоку, номер кадру та довжину повідомлення. Їх кодують однозначно, з фіксованим порядком полів. У базовому варіанті один кадр відповідає одному автентифікованому повідомленню. Якщо кадр поділяють на незалежні фрагменти, кожен із них отримує власне початкове значення, а номер фрагмента й загальну структуру кадру пов’язують із тегом через метадані. Приймач додатково перевіряє порядок і повноту фрагментів.

Для GCM необхідно забезпечувати унікальність початкового значення в межах використання одного ключа [2]. У пропонованій реалізації передбачають 96-бітове початкове значення та 128-бітовий тег. Лічильник повідомлень не повинен повторюватися після перезапуску: його стан зберігають надійно або встановлюють новий ключ перед початком нової сесії. На приймальній стороні відкриті дані дозволяють використовувати лише після успішної перевірки тегу. Захист від повторного надсилання кадрів забезпечують окремим контролем номерів повідомлень.

Ключі й проміжні значення зберігають у регістрах з обмеженим доступом і очищують після завершення сесії або помилки. У робочому режимі обмежують налагоджувальний доступ до криптографічного модуля. Окремого дослідження

потребують витоки через енергоспоживання, електромагнітне випромінювання та навмисне внесення збоїв. Висока швидкість обробки й збіг результатів з еталоном самі по собі не доводять стійкості апаратної реалізації до таких впливів.

Нестиснений кадр RGB розміром 1920×1080 за 8 біт на компонент містить 6 220 800 байтів. Для 30 кадрів за секунду необхідна швидкість обробки становить 186 624 000 байтів/с, або приблизно 1,49 Гбіт/с, без урахування службових даних. Якщо DMA зчитує відкриті дані й записує шифротекст у ту саму підсистему пам’яті, лише ці операції потребують сумарно щонайменше 373 248 000 байтів/с. Навантаження від отримання кадру та подальшого передавання враховують додатково.

За тактової частоти 100 МГц і приймання одного 128-бітового блока за такт теоретична пропускна здатність ядра становить 12,8 Гбіт/с. Це верхня оцінка для такого інтервалу приймання, а не виміряний результат прискорювача. Фактичну швидкість зменшують затримки DMA, обмеження пам’яті, обчислення тегу й обробка меж повідомлень. Для перекриття передавання та шифрування пропонується подвійна буферизація: один буфер обробляється, поки інший заповнюється. Власність буферів контролює драйвер із дотриманням правил узгодженості кешу.

Коректність перевіряють зіставленням із незалежною програмною реалізацією AES-GCM для різних довжин даних і метаданих. Тести охоплюють неповний останній блок, пошкодження шифротексту, метаданих і тегу, зупинки потоку та помилки DMA. Вимірюють час повної операції, сталу швидкість, затримку кадру, зайнятість процесора, апаратні ресурси та енергію на кадр. Порівняння виконують із оптимізованим програмним варіантом на тому самому процесорі, зазначаючи використання криптографічних інструкцій. Експериментальні значення прискорення потребують реалізації та вимірювання прототипу.

Висновки**.** Запропоновано архітектуру FPGA-прискорювача AES-GCM для захисту цифрових зображень у вбудованих системах. Узгодження криптографічного тракту, DMA та буферів є необхідною умовою досягнення потрібної швидкості. Наведений розрахунок визначає вимоги для потоку кадрів високої роздільності. Подальша робота передбачає реалізацію прототипу, порівняння ітераційної та конвеєрної архітектур і перевірку безпеки обробки кадрів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. National Institute of Standards and Technology. Advanced Encryption Standard (AES). FIPS 197-upd1. 2023. DOI: 10.6028/NIST.FIPS.197-upd1.

2. Dworkin M. Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC. NIST SP 800-38D. 2007. DOI: 10.6028/NIST.SP.800-38D.

3. Renugadevi N., Thangallapally S., Vemula S. C., Julakanti S., Bhatnagar S. Methods for improving the implementation of advanced encryption standard hardware accelerator on field programmable gate array—A survey. Security and Privacy. 2022. Vol. 5, No. 6. e254. DOI: 10.1002/spy2.254.

Tsarova Sofiia Valeriivna
student of group IKBm-61, ESI IS, SUICT, Kyiv, Ukraine

RANSOMWARE DETECTION WITH ENDPOINT DETECTION & RESPONSE PLATFORMS

Ransomware has evolved from a primarily encryption-based threat into multistage attacks with data exfiltration and extortion, which makes signature-based detection insufficient. This paper describes a three-level behavior-based approach to protecting endpoints from ransomware (prevention, detection and response) built on the open-source platform Wazuh. It also examines the growing use of EDR killers, tools that disable security agents, for example, through the bring-your-own-vulnerable-driver (BYOVD) technique, and shows why endpoint-level detection has to be supplemented with additional layers of defense.

Key words: ransomware, EDR, Wazuh, behavior-based detection, EDR killers.

Ransomware is a type of malware that prevents users from accessing systems and stored files and demands ransom payment in order to regain access. Even though the risk of ransomware has been present for a while, it's still one of the major cyberthreats for organization's IT-infrastructure, and its impact keeps growing. According to ENISA, in 2025 47.3% of the recorded financially motivated malicious activities are represented by ransomware attacks [1].

At the same time, modern ransomware moves from a primarily encryption-based threat to double extortion, which includes several stages [2]:

- Initial access: the most common attack vectors are phishing (including raising ClickFix technique that uses fake human verification prompts to persuade users to execute malicious actions or install unwanted software), as well as exploitation of software vulnerabilities [1].
- Data collection and exfiltration: nowadays ransomware shifts to identifying valuable data, such as login credentials, customers and employees PII and intellectual property, and downloading it to the C2 server.
- Extortion: after stealing sensitive data, today's ransomware moves to holding the victim's data hostage by encrypting it and preventing it from restoration by removing system shadow copies.

The multistaging, dynamic behavior and relying on exploiting new vulnerabilities (including zero-day attacks) present significant challenge for information resources defense, making only signature-based detection insufficient.

To address this, endpoint security platforms, such as Wazuh, utilize a behavior-based approach [3], consisting of three levels:

- Prevention: consistent collection of endpoint telemetry and systematic vulnerability management to detect and mitigate software vulnerabilities that ransomware operators exploit for initial access.
- Detection: key indicators are extracted, correlating with custom rules and MITRE ATT&CK to detect risky behavior, such as frequent file modification and renaming, unusual executables running and shadow copy deletion attempts.

- Response: defense mechanisms, like Wazuh Active Response, are triggered to perform containment actions such as terminating malicious processes, blocking malicious IP addresses and isolating the affected endpoint from the network.

However, this defense depends on the endpoint agent that creates significant challenges due to widely used EDR-evasion techniques. The most common of them rely on BYOVD (Bring Your Own Vulnerable Driver), abusing a legitimate but vulnerable driver to terminate protected processes on high kernel level [4]. ENISA reports that variants of EDRKillShifter have been included in the toolsets of several ransomware-as-a-service operations since June 2025, which allows operators to to exploit vulnerable Windows drivers and terminate EDR protections [1].

This creates several challenges for the described approach as EDR killers are typically used early in an attack, so the agent may be disabled before the behavior that the detection rules look for appears.

To reduce these risks, the endpoint level should be supplemented with measures that do not depend on the agent: treating the loss of connection with an agent as an alert in itself, patching or deleting vulnerable drivers that are already present, enabling EDR anti-tamper features [4].

Thus, behavior-based detection with automated response on the Wazuh platform addresses the multistage nature of modern ransomware, but the growing use of EDR killers requires protecting the agent itself and adding detection layers that do not depend on the endpoint.

References:

1. ENISA Threat Landscape 2026. European Union Agency for Cybersecurity, 2026. DOI: 10.2824/0806036. URL: https://www.enisa.europa.eu/sites/default/files/2026-09/ENISA%20Threat%20Landscape%202026_Final.pdf

2. What is ransomware? / M. Kosinski, S. Susnjara, M. Goodwin. IBM Think. URL: https://www.ibm.com/think/topics/ransomware

3. Ransomware Threat Detection and Automated Response System Using Wazuh and Python / S. Mohan, G. Pradeep, P. Madhuravel et al. Indian Journal of Computer Science and Technology. 2026. Vol. 5, No. 2. P. 270–275. DOI: 10.59256/indjcst.20260502031

4. Cuprik R. What are EDR killers? How ransomware attacks disable EDR and how to stop it. ESET WeLiveSecurity, 2026. URL: https://www.eset.com/blog/en/business-topics/threat-landscape/what-are-edr-killers/

Степанченко Богдан Сергійович, аспірант
ДУКТ, м. Київ
Собчук Андрій Валентинович, доктор філософії
ДУІКТ, м. Київ

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ДИНАМІКИ КІБЕРЗАГРОЗ НА ОСНОВІ СИСТЕМ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ З ІМПУЛЬСНОЮ ДІЄЮ

Розглянуто застосування математичних моделей для дослідження динаміки кіберзагроз та підвищення стійкості інформаційних мереж. Запропонований підхід ґрунтується на методах якісної теорії диференціальних рівнянь з імпульсною дією та використанні SIR-моделі для опису поширення зараження серед мережевих пристроїв. Досліджено умови стійкості системи й можливості зменшення кількості інфікованих пристроїв шляхом реалізації різних стратегій захисту. Практична значущість результатів полягає в математичному обґрунтуванні заходів протидії кіберзагрозам та визначенні параметрів оновлення програмного забезпечення.

Ключові слова: кібербезпека, кіберзагрози, математичне моделювання, диференціальні рівняння з імпульсною дією, SIR-модель, стійкість інформаційних мереж.

Зростання масштабів і складності кібератак зумовлює необхідність удосконалення методів виявлення загроз, прогнозування їх поширення та оцінювання ефективності захисних заходів. Порушення функціонування інформаційних мереж може спричиняти втрату даних, зупинення технологічних процесів і порушення роботи критичної інфраструктури. За таких умов важливого значення набуває не лише виявлення факту атаки, а й математичне дослідження динаміки її розвитку та здатності мережі зберігати функціональність за несприятливих впливів [1–3].

Традиційні засоби кіберзахисту, зокрема системи виявлення та запобігання вторгненням, аналіз мережевого трафіку й журналів подій, є важливими компонентами захисту інформаційних систем. Водночас для дослідження нових сценаріїв поширення загроз необхідні підходи, які дають змогу враховувати зміну стану мережі в часі, періодичність захисних заходів та їхні миттєві впливи на систему.

Метою роботи є дослідження динаміки поширення кіберзагроз на основі SIR-моделі та систем диференціальних рівнянь з імпульсною дією, визначення умов стійкості інформаційної мережі й обґрунтування параметрів періодичного оновлення програмного забезпечення для зменшення ризику масштабного зараження пристроїв.

Для дослідження динаміки кіберзагроз запропоновано використовувати апарат якісної теорії диференціальних рівнянь з імпульсною дією. Такий підхід дає змогу описувати неперервну зміну стану інформаційної мережі та враховувати дискретні втручання, пов’язані з реалізацією захисних заходів. Для формалізації процесу поширення кіберзагроз використаємо SIR-підхід, адаптований до інформаційної мережі. Сукупність її пристроїв поділимо на три групи:

$S(t)$ – сприятливі до ураження вузли мережі;

$I(t)$ – уражені вузли мережі, що можуть поширювати ураження;

$R(t)$ – вузли, що відновились і вже не заразяться.

Враховуючи, що мережа функціонує як еволюційна система, в ній присутні процеси вибуття частини пристроїв та поповнення мережі новими, що є аналогом природної народжуваності та смертності певної популяції. Тоді математичну модель функціонування такої мережі можна записати за допомогою системи нелінійних диференціальних рівнянь:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\lambda S(t)I(t) + \mu - \mu S(t), \\ \dfrac{dI(t)}{dt} = \lambda S(t)I(t) - \gamma I(t) - \mu I(t). \end{cases}$$

В системі $\lambda > 0$ параметр, що характеризує інтенсивність контактів, тобто характеризує швидкість переходу вузлів з класу сприятливих до ураження вузлів мережі до уражених («інфікованих»); $\mu > 0$ – параметр, що характеризує смертність уражених кібератакою пристроїв у зовнішній мережі (цей же параметр характеризує й появу нових пристроїв у зовнішній мережі), $\gamma > 0$ – характеризує інтенсивність відновлення пристроїв внутрішньої мережі.

Отримано конструктивні умови стійкості й асимптотичної стійкості розглянутої SIR-моделі. Досліджено можливості мінімізації кількості інфікованих пристроїв за трьома стратегіями: постійною вакцинацією, імпульсною вакцинацією та періодичним оновленням програмного забезпечення.

Встановлено залежність загрозостійкості мережі від співвідношення швидкості зараження пристроїв і швидкості їх вибуття та відновлення. Також отримано оцінку періоду оновлення програмного забезпечення, що може використовуватися для обґрунтування профілактичних заходів кіберзахисту [4].

Результати аналізу SIR-моделі дають змогу встановлювати умови стійкості системи та порівнювати стратегії зменшення кількості інфікованих пристроїв. Практичний потенціал підходу полягає в обґрунтуванні періодичності оновлення програмного забезпечення та виборі захисних заходів з урахуванням динамічних характеристик мережі. Подальші дослідження доцільно спрямувати на врахування невизначеності параметрів, неоднорідності мережевої структури та складніших сценаріїв поширення кіберзагроз.

Література

1. V. Pichkur, O. Laptiev, I. Polovinkin, A. Barabash, A. Sobchuk and I. Salanda, "The Method of Managing Man-generated Risks of Critical Infrastructure Systems Based on Ellipsoidal Evaluation," 2022 IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT), Kyiv, Ukraine, 2022, pp. 133-137, https://doi.org/10.1109/ATIT58178.2022.10024244

2. Барабаш О.В., Лукова-Чуйко Н.П., Мусієнко А.П., Собчук В.В. Забезпечення функціональної стійкості інформаційних мереж на основі розробки методу протидії DDoS-атакам. // Сучасні інформаційні системи. Харків: Національний технічний університет «Харківський політехнічний інститут», 2018. Т. 2. № 1. С. 56–63.

3. O. Laptiev, A. Musienko, V. Nakonechnyi, A. Sobchuk, S. Gakhov and S. Kopytko, "Algorithm for Recognition of Network Traffic Anomalies Based on Artificial Intelligence," 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Istanbul, Turkiye, 2023, pp. 1-5, https://doi.org/10.1109/HORA58378.2023.10156702

4. Савченко В. А., Степанченко Б. С. Розробка концепції прогнозування часу початку ddos атаки на основі дослідження динаміки поведінки еволюційних рівнянь. Телекомунікаційні та інформаційні технології.2024.No 1(82). С. 26–44. https://doi.org/10.31673/2412-4338.2024.012644.

Бражник Артем Михайлович
студент групи БСДМ-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ БАГАТОЕТАПНИХ КІБЕРАТАК У ГЕТЕРОГЕННИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ НА ОСНОВІ АДАПТИВНОЇ КОРЕЛЯЦІЇ ПОДІЙ

Анотація. Запропоновано підхід до виявлення багатоетапних кібератак у середовищі Windows і Linux шляхом адаптивної кореляції подій безпеки. Визначено склад уніфікованого запису події, ознаки зв'язку між подіями та правила коригування часового вікна й ваг ознак з урахуванням поведінкового контексту. Описано формування графа можливого розвитку атаки та використання MITRE ATT&CK для інтерпретації виявлених дій. Наведено сценарій міжплатформної кореляції та методику порівняння з фіксованими правилами. Практичне призначення підходу полягає у підтримці аналізу інцидентів у центрі моніторингу кібербезпеки.

Ключові слова: багатоетапні кібератаки, адаптивна кореляція, Windows, Linux, MITRE ATT&CK.

Актуальність дослідження. У корпоративній інформаційній системі дії порушника можуть послідовно охоплювати робочу станцію Windows, обліковий запис користувача та сервер Linux. Окремий успішний вхід, запуск командного інтерпретатора або віддалене з'єднання трапляються також під час легітимного адміністрування. Для виявлення багатоетапної атаки необхідно дослідити зв'язки між такими подіями та контекст їх виникнення. MITRE ATT&CK систематизує тактики й техніки поведінки порушників і може використовуватися для інтерпретації спостережуваних дій [1]. Водночас відповідність техніці сама по собі не доводить наявності атаки.

Кореляція за фіксованим часовим вікном залежить від темпу дій: короткий інтервал може розірвати повільний сценарій, а надто довгий — об'єднати непов'язані записи. Збіг імені користувача також недостатній, оскільки локальні облікові записи різних вузлів можуть мати однакові назви. Тому в роботі розглядається адаптація параметрів кореляції з обов'язковою перевіркою ідентичності сутностей і наявності доказових зв'язків. Метою є визначення способу реконструкції ймовірних ланцюгів атаки за різнорідними журналами Windows і Linux.

Джерела та нормалізація подій. Для Windows передбачено використання журналу Security і телеметрії Sysmon. Sysmon реєструє створення процесів, а за відповідного налаштування — мережеві з'єднання; ідентифікатор ProcessGuid допомагає пов'язувати записи про один процес [2]. Для Linux розглядаються журнали автентифікації SSH, повідомлення sudo та записи auditd. Linux Audit дає змогу відстежувати дії користувачів і доступ до ресурсів відповідно до налаштованих правил аудиту [3]. Повнота доступної телеметрії визначає, які етапи сценарію можна підтвердити.

Подію пропонується представляти записом, що містить час виникнення й надходження, джерело журналу, операційну систему, ідентифікатор вузла, обліковий запис із доменом або локальною областю, тип і результат дії, процес

та батьківський процес, мережеві адреси, ресурс і посилання на первинний запис. Відсутні поля позначаються явно. Для узгодження семантики доцільно зіставляти поля з Open Cybersecurity Schema Framework (OCSF), який надає незалежну від постачальника схему даних безпеки [4].

Під час підготовки даних часові мітки приводяться до UTC, дублікати усуваються, а відоме розходження годинників враховується. Окремо зберігаються ознаки якості даних. Збагачення додає критичність вузла, роль користувача, дозволені адміністративні маршрути та відхилення від типової активності. Для переходу між Windows і Linux потрібне явне зіставлення облікових записів або підтверджений мережевий зв'язок. Однакові імена та спільна зовнішня IP-адреса без додаткового контексту не вважаються достатньою підставою для об'єднання.

Механізм адаптивної кореляції. Для пари подій оцінюються часова близькість, ідентичність облікового запису, зв'язок вузлів, відповідність мережевого потоку, спадкоємність процесів і змістова сумісність дій. Ознаки приводяться до діапазону від нуля до одиниці. Узагальнена оцінка зв'язку визначається як зважене середнє доступних ознак. Пропущені значення не замінюються позитивними збігами; за недостатнього покриття телеметрією рішення відкладається. Ребро графа утворюється лише за перевищення порога й наявності опорного зв'язку: спільного сеансу, процесу або підтвердженого переходу між вузлами.

Адаптацію пропонується виконувати за двома правилами. По-перше, часове вікно для класу переходів визначається за розподілом інтервалів у перевірених навчальних сценаріях: за основу береться верхній квантиль із запасом на затримку журналювання. Мінімальна та максимальна тривалість задаються конфігурацією. По-друге, ваги змінюються залежно від накопиченого контексту. Якщо успішному входу передує аномальна серія відмов для того самого облікового запису, підвищується значущість зв'язку із сеансом автентифікації. Підтверджений перехід на інший вузол збільшує вагу мережевої ознаки.

Межі зміни ваг фіксуються, після коригування ваги нормуються. Порогові значення добираються на навчальних даних і фіксуються до незалежного тестування. За відсутності перевіреної історії використовуються початкові параметри. Активність неперевірених інцидентів не включається до поведінкового профілю автоматично, щоб не закріплювати дії порушника як норму.

Реконструкція сценарію. Нормалізовані події утворюють вершини орієнтованого графа, а ребра відображають обґрунтовані переходи з урахуванням часу. Кандидатом у багатоетапний інцидент вважається пов'язаний фрагмент із кількома змістово різними діями, наприклад автентифікацією, виконанням команди та віддаленим доступом. Позначення MITRE ATT&CK додаються за сукупністю ознак, а не лише за назвою процесу. Порядок тактик не використовується як обов'язкова лінійна схема. Кожне ребро зберігає пояснення та посилання на журнали, що дозволяє аналітику перевірити реконструкцію.

Ілюстративний сценарій охоплює серію невдалих входів у Windows, наступний успішний вхід, нетиповий запуск PowerShell і SSH-з'єднання з сервером Linux. На сервері зафіксовано вхід із відповідної адреси та використання sudo поза дозволеним адміністративним сценарієм. Для зв'язування потрібні підтвердження сеансу й процесу на Windows, відповідність адрес і часу з'єднання та перевірка Linux-журналів. Звичайне використання SSH або sudo не визнається атакою автоматично. Якщо частини журналів немає, система позначає відповідний перехід як непідтверджений.

Пріоритет інциденту пропонується визначати окремо від оцінки зв'язку між подіями: він враховує критичність уражених ресурсів, характер дій і повноту доказів. Такий показник є шкалою для ранжування, а не статистичною ймовірністю атаки. Практичний результат для SOC-аналітика — картка інциденту з часовою послідовністю, пов'язаними вузлами, обліковими записами та поясненням об'єднання подій.

Методика перевірки. Для оцінювання передбачено контрольовані сценарії атак і легітимного адміністрування на тлі звичайної активності. Набір має містити швидкі та повільні переходи, розриви журналювання й повторне використання імен облікових записів. Навчальна й тестова частини розділяються за сеансами та сценаріями, щоб події однієї атаки не потрапляли до обох частин. Параметри адаптивного методу й базового методу з фіксованими вагами та вікном налаштовуються за однакової доступної телеметрії. Розмітка атак використовується для оцінювання, але не передається модулю виявлення.

Одиницею оцінювання є інцидент, а не окрема подія. Правило відповідності виявленого інциденту розміченій атаці задається до тестування за збігом ключових сутностей і часових меж; дублікати обліковуються окремо. Вимірюються точність (Precision), повнота (Recall), F1, кількість хибних інцидентів за добу, затримка виявлення та швидкість обробки. Окреме вимкнення адаптації ваг і часового вікна дозволить оцінити внесок кожного механізму. Результати на синтетичних даних потребуватимуть перевірки на незалежних журналах.

Висновки. Визначено архітектурний підхід до кореляції гетерогенних подій Windows і Linux та конкретизовано правила адаптації ваг і часового вікна. Потенційний дослідницький внесок полягає у поєднанні контекстної адаптації з обов'язковими опорними зв'язками й пояснюваною реконструкцією сценарію. Запропонована методика дає змогу перевірити гіпотезу про поліпшення виявлення порівняно з фіксованою кореляцією. Кількісне підтвердження ефективності потребує реалізації прототипу та незалежних експериментів.

Список використаних джерел

1. MITRE. MITRE ATT&CK. URL: https://attack.mitre.org/ (дата звернення: 10.10.2026).
2. Microsoft. Sysmon. URL: https://learn.microsoft.com/en-us/sysinternals/downloads/sysmon (дата звернення: 10.10.2026).
3. Red Hat. Security hardening. Red Hat Enterprise Linux 9. Chapter 12. Auditing the system. URL: https://docs.redhat.com/en/documentation/red_hat_enterprise_linux/9/html/security_hardening/auditing-the-system_security-hardening (дата звернення: 10.10.2026).
4. Open Cybersecurity Schema Framework. Welcome to OCSF. URL: https://ocsf.io/ (дата звернення: 10.10.2026).

Магдєєв В.С.
студент групи ЗІм-25-1,
Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
НТУ «Дніпровська політехніка», Дніпро, Україна

АНАЛІЗ ОЗНАК СОЦІОТЕХНІЧНИХ СЦЕНАРІЇВ СПОНУКАННЯ КОРИСТУВАЧА ДО ВИКОНАННЯ НЕБЕЗПЕЧНИХ ДІЙ У ВЕБСЕРЕДОВИЩІ

Розглянуто особливості соціотехнічних сценаріїв, за яких вебсторінка спонукає користувача до самостійного виконання небезпечних дій на пристрої. Систематизовано чотири типові сценарії та визначено класи запитуваних дій. Проаналізовано можливості й обмеження наявних методів виявлення вебзагроз. Обґрунтовано необхідність спільного аналізу змістових, структурних і контекстних ознак вебсторінки, зокрема відповідності між заявленою метою та запитуваною дією користувача. Формалізовано задачу виявлення соціотехнічних сценаріїв, визначено склад вхідних даних, очікуваних результатів і критеріїв експериментального оцінювання.

Ключові слова: соціальна інженерія, соціотехнічний сценарій, ClickFix, вебсторінка, небезпечні дії користувача, виявлення вебзагроз, аналіз ознак.

Вебсторінки можуть використовуватися не лише для викрадення облікових даних, а й для спонукання користувача до самостійного виконання дій, що створюють загрозу інформаційній безпеці пристрою. Одним із прикладів є техніка ClickFix, за якої повідомлення про необхідність проходження CAPTCHA, усунення помилки або відновлення доступу супроводжується інструкцією щодо виконання команди засобами операційної системи. У базі MITRE ATT&CK такі дії розглядаються в межах техніки T1204.004 User Execution: Malicious Copy and Paste [1]. Виявлення зазначених сценаріїв ускладнюється тим, що окремі елементи вебсторінки, зокрема технічні інструкції, системні команди та повідомлення про помилки, можуть використовуватися і на легітимних ресурсах.

Соціотехнічні сценарії у вебсередовищі характеризуються поєднанням технічних властивостей вебсторінки та механізмів психологічного впливу на користувача. Зазвичай вони містять заявлений привід, повідомлення про проблему або необхідність виконання певної процедури, заклик до дії та очікуваний результат. Для підвищення переконливості можуть використовуватися імітація авторитетного джерела, формування відчуття терміновості, страху або обмеження доступу. Водночас наявність окремої ознаки, наприклад логотипа відомої компанії, повідомлення про помилку чи кнопки завантаження, не є достатньою підставою для визначення сторінки як небезпечної. Важливим є встановлення взаємозв’язку між змістом повідомлення, заявленою метою та дією, яку пропонується виконати користувачу.

Метою дослідження є визначення ознак соціотехнічних сценаріїв спонукання користувача до виконання небезпечних дій у вебсередовищі та обґрунтування способу їх виявлення шляхом спільного аналізу змісту, структури й контексту вебсторінки.

Для визначення характерних ознак систематизовано чотири типові

соціотехнічні сценарії: підроблена перевірка користувача, виправлення технічної помилки, встановлення оновлення та отримання доступу до документа. Їх спільною характеристикою є використання заявленої мети для обґрунтування необхідності виконання дії, яка може виходити за межі звичайної взаємодії з вебсторінкою. До основних класів таких дій належать виконання системних команд, встановлення програмного забезпечення з неперевірених джерел, вимкнення або послаблення захисних механізмів і надання надмірних дозволів. Оскільки один сценарій може поєднувати декілька дій, доцільно окремо визначати тип сценарію, клас запитуваної дії та ознаку потенційно небезпечного спонукання.

Аналіз наявних методів виявлення вебзагроз показав, що репутаційні механізми забезпечують ідентифікацію відомих небезпечних ресурсів, але мають обмеження щодо нових вебсторінок. Методи пошуку за ключовими словами є відносно простими в реалізації, проте можуть формувати хибні спрацювання під час аналізу легітимних технічних інструкцій. Методи на основі правил дають змогу виявляти характерні поєднання ознак, проте потребують регулярного перегляду та вдосконалення правил у разі появи нових або зміни наявних сценаріїв соціотехнічного впливу. Методи машинного навчання забезпечують автоматизовану класифікацію вебсторінок, але їхні результати значною мірою залежать від репрезентативності навчальних даних [2]. У системі PhishIntention використовуються візуальні характеристики та динаміка взаємодії з вебсторінкою для визначення фішингових намірів, пов’язаних із викраденням облікових даних [3]. Однак зазначені механізми не охоплюють повною мірою сценарії, у яких користувача спонукають до виконання системних команд або зміни параметрів захисту.

Встановлено, що важливим обмеженням наявних методів є недостатнє врахування відповідності між заявленою метою вебсторінки та запитуваною дією користувача. Наприклад, пропозиція виконати системну команду для проходження CAPTCHA або відкриття документа може свідчити про соціотехнічне спонукання, навіть якщо сторінка не містить відомих шкідливих адрес чи явно підозрілих слів. Водночас аналогічна команда може бути частиною легітимної технічної документації. Це визначає необхідність спільного аналізу декількох груп ознак.

Для формування інформаційної основи виявлення запропоновано виділити три групи ознак: змістові, структурні та контекстні. Змістові ознаки характеризують заявлений привід, використання авторитету, терміновість, емоційний вплив і заклик до дії. Структурні ознаки відображають розташування інструкцій, кнопок, блоків команд, послідовність запропонованих операцій та особливості організації вебсторінки. Контекстні ознаки визначають відповідність запитуваної дії заявленій меті, типу інформаційного ресурсу й умовам взаємодії користувача зі сторінкою. Сукупність зазначених ознак створює основу для відокремлення потенційно небезпечних сценаріїв від легітимних інструкцій зі схожим текстовим наповненням.

Задачу виявлення соціотехнічних сценаріїв можна формалізувати

залежністю:

$$Y = F(X_c, X_s, X_k),$$

де Y – результат аналізу вебсторінки; X_c – сукупність змістових ознак; X_s – сукупність структурних ознак; X_k – сукупність контекстних ознак; F – функція спільного аналізу визначених груп ознак.

Результат аналізу запропоновано представляти у вигляді:

$$Y = (D, S, A, E),$$

де D – ознака наявності потенційно небезпечного спонукання; S – тип соціотехнічного сценарію; A – клас запитуваної дії; E – сукупність виявлених ознак, що обґрунтовують результат аналізу.

Як вхідні дані передбачається використовувати текстовий вміст вебсторінки, її структуру DOM (Document Object Model) та URL-адресу. Особливістю сформульованої задачі є необхідність визначення не лише факту наявності підозрілих елементів, а й характеру взаємозв'язку між ними. Отриманий результат має відображати ознаки потенційного соціотехнічного впливу, але не може самостійно підтверджувати виконання користувачем небезпечної дії або компрометацію пристрою.

Для експериментальної перевірки майбутнього методу визначено показники precision, recall, F1-міру, macro-F1, частку хибнопозитивних і хибнонегативних результатів, повноту виявлення для кожного типу сценарію та час аналізу вебсторінки. Порівняльне оцінювання доцільно виконувати відносно пошуку за ключовими словами, класифікатора за текстовими ознаками та контекстного правилового механізму. Окремо необхідно дослідити внесок кожної групи ознак шляхом їх послідовного виключення. При цьому слід враховувати, що високі показники класифікації на обмеженому наборі даних не гарантують аналогічної ефективності в реальному вебсередовищі [4].

У результаті дослідження систематизовано чотири типові соціотехнічні сценарії спонукання користувача до виконання небезпечних дій у вебсередовищі та визначено основні класи запитуваних дій. На основі аналізу наявних методів виявлення вебзагроз обґрунтовано необхідність спільного врахування змістових, структурних і контекстних ознак вебсторінки. Формалізовано задачу виявлення, визначено склад вхідних даних, структуру результату та критерії експериментального оцінювання. Подальше дослідження передбачає вдосконалення методу спільного аналізу визначених ознак, його програмну реалізацію та експериментальну перевірку здатності розмежовувати потенційно небезпечні соціотехнічні сценарії й легітимні технічні інструкції.

Список використаних джерел:

1. MITRE. User Execution: Malicious Copy and Paste. Sub-technique T1204.004. *MITRE ATT&CK*. URL: https://attack.mitre.org/techniques/T1204/004/ (дата звернення: 13.10.2026).

2. Hannousse A., Yahiouche S. Towards benchmark datasets for machine learning based website phishing detection: An experimental study. *Engineering Applications of Artificial Intelligence.* 2021. Vol. 104. Article 104347. DOI: https://doi.org/10.1016/j.engappai.2021.104347.

3. Liu R., Lin Y., Yang X., Ng S. H., Divakaran D. M., Dong J. S. Inferring phishing intention via webpage appearance and dynamics: A deep vision based approach. *Proceedings of the 31st USENIX Security Symposium.* 2022. P. 1633–1650. URL: https://www.usenix.org/conference/usenixsecurity22/presentation/liu-ruofan.

4. Marchal S., Asokan N. On designing and evaluating phishing webpage detection techniques for the real world. *Proceedings of the 11th USENIX Workshop on Cyber Security Experimentation and Test (CSET 2018).* Baltimore, MD: USENIX Association, 2018. URL: https://www.usenix.org/conference/cset18/presentation/marchal.

Яненко О. В.
студентка групи ЗІм-25-1,
Національний технічний університет «Дніпровська політехніка»,
Дніпро, Україна

Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
Національний технічний університет «Дніпровська політехніка»
Дніпро, Україна

ОСОБЛИВОСТІ ОЦІНЮВАННЯ КІБЕРРИЗИКІВ СИСТЕМИ ЕЛЕКТРОННОГО ДОКУМЕНТООБІГУ ПІДПРИЄМСТВА НА ОСНОВІ НЕЧІТКОЇ ЛОГІКИ

Цифровізація бізнес-процесів посилює залежність підприємств від надійного функціонування систем електронного документообігу, які забезпечують опрацювання та зберігання важливої ділової інформації. Водночас оцінювання кіберризиків таких систем ускладнюється обмеженістю статистичних даних про інциденти та невизначеністю експертних суджень. За цих умов доцільним є застосування нечіткої логіки, яка дає змогу формалізувати експертні оцінки та враховувати характеристики загроз, вразливостей, заходів захисту й можливих наслідків. Використання нечіткого логічного висновку Mamdani створює можливість формування інтегральної оцінки рівня залишкового кіберризику та ранжування сценаріїв за ступенем їх небезпеки.

Ключові слова: кіберризик, система електронного документообігу, нечітка логіка, експертне оцінювання, нечіткий висновок Mamdani, залишковий ризик.

Зростання обсягів електронного документообігу на підприємствах підвищує вимоги до захисту інформації, що обробляється та зберігається в системах електронного документообігу (СЕД), зокрема договірних, фінансових, кадрових та інших документів. Порушення конфіденційності, цілісності або доступності таких даних може призвести до фінансових втрат, зупинення окремих операцій і порушення договірних зобов’язань. Тому оцінювання кіберризиків СЕД є важливою складовою управління інформаційною безпекою підприємства. Загальні принципи такого оцінювання передбачають ідентифікацію ризиків, аналіз їхніх характеристик і визначення пріоритетів оброблення [1; 2].

Вибір методу оцінювання залежить від доступності вихідних даних, особливостей інформаційної системи та потреб управління. Якісні методи й матриці ризиків зручні для первинного аналізу, однак отримані результати можуть істотно залежати від інтерпретації експертами лінгвістичних категорій. Кількісні методи дають змогу представляти ризик у числовій формі, проте потребують достатньо обґрунтованих даних щодо ймовірності подій і можливих збитків. Напівкількісні методи використовують бальні шкали та вагові коефіцієнти, але перетворення якісних суджень у числові значення не завжди відображає складні взаємозв’язки між чинниками ризику. Системи оцінювання технічних вразливостей характеризують окремі властивості компонентів інформаційної системи, але не замінюють комплексного оцінювання кіберризику підприємства.

Особливої уваги потребує оцінювання ризиків підприємств, які не мають достатньої внутрішньої статистики кіберінцидентів. За таких умов характеристики потенційних подій часто визначаються на основі експертних суджень, а їх точне числове подання є ускладненим. Для СЕД необхідно враховувати не лише зовнішні атаки, а й фішинг, помилки користувачів, дії внутрішніх порушників, компрометацію програмного забезпечення в ланцюгу постачання та інфраструктурні збої. Після впровадження технічних і організаційних заходів захисту частина ризику зберігається, що визначає необхідність оцінювання його залишкового рівня.

Одним із засобів формалізації невизначених експертних суджень є теорія нечітких множин, запропонована Л. Заде [3]. На відміну від жорсткого віднесення показника до однієї категорії, нечітке представлення дає змогу визначати ступінь належності значення до лінгвістичних термів, наприклад «низький», «середній» або «високий». Це забезпечує плавний перехід між рівнями оцінювання та дає змогу формалізувати експертні знання. У системах нечіткого логічного висновку правила типу «якщо – то» встановлюють взаємозв'язок між вхідними характеристиками та вихідною оцінкою. Для такого опрацювання може бути використаний механізм нечіткого висновку Mamdani [4].

Метою дослідження є визначення особливостей оцінювання залишкових кіберризиків СЕД підприємства в умовах невизначеності вихідних даних та обґрунтування структури нечіткої моделі на основі механізму логічного висновку Mamdani.

Для оцінювання кіберризиків СЕД одиницею аналізу запропоновано визначити окремий сценарій ризику, пов'язаний із конкретним інформаційним активом, загрозою, вразливістю, наявними заходами захисту та можливими наслідками. Як вхідні лінгвістичні змінні нечіткої моделі визначено можливість реалізації загрози, рівень вразливості, дієвість заходів захисту та тяжкість наслідків. Їх взаємозв'язок із вихідною оцінкою можна представити залежністю:

$$R_{res} = F_{Mamdani}(T, V, C, I),$$

де R_{res} – інтегральна оцінка рівня залишкового кіберризику; T – можливість реалізації загрози; V – рівень вразливості; C – дієвість заходів захисту; I – тяжкість можливих наслідків; $F_{Mamdani}$ – функція нечіткого висновку Mamdani.

Для кожної вхідної змінної передбачається використання лінгвістичних термів «низький», «середній» і «високий», що описуються відповідними функціями належності. Вихідна змінна характеризує рівень залишкового кіберризику з використанням термів «низький», «помірний», «високий» і «критичний». Взаємозв'язок між характеристиками сценарію встановлюється сукупністю нечітких правил. Наприклад, якщо можливість реалізації загрози висока, рівень вразливості високий, дієвість заходів захисту низька, а тяжкість наслідків висока, то рівень залишкового кіберризику визначається як критичний. Формування інтегральної оцінки передбачає фазифікацію вхідних значень, застосування бази правил, агрегування результатів нечіткого висновку та дефазифікацію.

Практичним прикладом застосування моделі є сценарій несанкціонованого доступу до конфіденційних документів унаслідок компрометації облікового запису користувача СЕД через фішингову атаку. Для такого сценарію необхідно враховувати можливість успішної реалізації атаки, вразливості механізмів автентифікації, дієвість багатофакторної автентифікації та інших заходів захисту, а також тяжкість можливих наслідків розголошення інформації. Сукупне оцінювання зазначених характеристик дає змогу визначити рівень залишкового ризику з урахуванням чинних заходів захисту.

Отримана після реалізації моделі інтегральна оцінка може використовуватися для порівняння окремих сценаріїв кіберризику СЕД та формування їх ранжованого переліку. Це створює основу для визначення пріоритетності заходів захисту в умовах, коли точні статистичні характеристики загроз і вразливостей відсутні або є недостатніми. При цьому числове значення, отримане в результаті дефазифікації, слід розглядати як відносний індекс ризику, а не безпосередню ймовірність кіберінциденту чи величину очікуваних збитків.

Водночас застосування нечіткої логіки не усуває суб’єктивності експертного оцінювання. Обґрунтованість отриманих результатів залежить від вибору функцій належності, складу лінгвістичних змінних і бази правил нечіткого висновку. Тому експериментальна перевірка моделі має передбачати аналіз узгодженості правил, чутливості результатів до зміни вхідних параметрів та монотонності оцінок. Доцільним є також порівняння сформованого ранжування сценаріїв із результатами традиційної матриці ризиків та незалежного експертного оцінювання.

Проведений аналіз дозволив визначити основні особливості оцінювання залишкових кіберризиків СЕД в умовах невизначеності вихідних даних та обґрунтувати структуру нечіткої моделі на основі механізму Mamdani. Запропонована структура передбачає спільне врахування можливості реалізації загрози, рівня вразливості, дієвості заходів захисту й тяжкості наслідків із формуванням інтегральної оцінки залишкового ризику. Подальше дослідження передбачає програмну реалізацію моделі в MATLAB, формування функцій належності та бази нечітких правил, а також експериментальну перевірку здатності моделі забезпечувати обґрунтоване ранжування сценаріїв кіберризику СЕД.

Список використаних джерел:

1. ISO/IEC 27005:2022. *Information security, cybersecurity and privacy protection — Guidance on managing information security risks.* 4th ed. Geneva : ISO, 2022. 62 p. URL: https://www.iso.org/standard/80585.html.

2. National Institute of Standards and Technology. *Guide for Conducting Risk Assessments.* NIST Special Publication 800-30, Revision 1. Gaithersburg, MD : NIST, 2012. DOI: https://doi.org/10.6028/NIST.SP.800-30r1.

3. Zadeh L. A. Fuzzy sets. *Information and Control.* 1965. Vol. 8, No. 3. P. 338–353. DOI: https://doi.org/10.1016/S0019-9958(65)90241-X.

4. Mamdani E. H. Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers.* 1974. Vol. 121, No. 12. P. 1585–1588. DOI: https://doi.org/10.1049/piee.1974.0328.

Талан Анна Владиславівна
Студентка групи ІКБм-62, ННІКБЗІ, Київ, Україна

ПОРІВНЯЛЬНИЙ АНАЛІЗ РУЧНОГО ТА ЦЕНТРАЛІЗОВАНОГО КОНФІГУРУВАННЯ КОРПОРАТИВНИХ ПРИСТРОЇВ

У роботі порівняно ручний та централізований підходи до конфігурування корпоративних пристроїв за критеріями трудомісткості, однорідності налаштувань, безпеки, масштабованості та вартості володіння. Показано, що ручне конфігурування виправдане лише для малих інфраструктур, тоді як централізоване керування забезпечує відтворюваність конфігурацій, контроль відповідності вимогам і швидке реагування на інциденти. Запропоновано критерії вибору підходу залежно від розміру парку.

Ключові слова: конфігурування пристроїв, централізоване керування, MDM, дрейф конфігурацій, політики безпеки, автоматизація, корпоративна інфраструктура.

Сучасний набір корпоративних пристроїв охоплює робочі станції, ноутбуки, смартфони, планшети та мережеве обладнання, які дедалі частіше використовуються поза межами офісу. Конфігурування таких пристроїв, тобто встановлення операційної системи, застосунків, параметрів безпеки, мережевих профілів і політик шифрування, безпосередньо впливає на захищеність організації та витрати на супровід IT-інфраструктури. Керування конфігураціями розглядається як обов'язковий елемент системи захисту інформації у провідних стандартах і настановах [1, 3, 4], тому вибір способу конфігурування є не лише технічним, а й управлінським рішенням.

За ручного підходу адміністратор налаштовує кожен пристрій окремо: локально або віддалено, за інструкцією, чек-листом чи за допомогою еталонного образу системи. До переваг цього способу належать відсутність початкових інвестицій у спеціалізовану інфраструктуру, простота для невеликої кількості пристроїв і гнучкість під унікальні випадки. Водночас недоліки стають відчутними зі зростанням кількості пристроїв. Трудомісткість збільшується приблизно пропорційно кількості пристроїв, а людський фактор спричиняє пропуски кроків та відмінності між однотипними пристроями. З часом виникає так званий дрейф конфігурацій: фактичні налаштування поступово відхиляються від затверджених, а адміністратор не має єдиної картини стану всіх пристроїв. Це ускладнює встановлення оновлень, розслідування інцидентів і проходження аудиту [1].

Централізоване конфігурування передбачає керування пристроями з єдиної консолі за допомогою політик і профілів. До таких засобів належать системи керування мобільними пристроями та кінцевими точками (MDM), групові політики доменної служби каталогів, інструменти керування конфігураціями та автоматизації. Налаштування описуються один раз, автоматично застосовуються до всіх пристроїв відповідної групи і періодично перевіряються на відповідність. Автоматичне розгортання нових пристроїв без втручання адміністратора, ведення інвентаризації, примусове оновлення та віддалене стирання даних у разі втрати пристрою скорочують час на типові

операції та зменшують кількість помилок. Централізоване керування мобільними пристроями прямо рекомендується в настановах NIST [2], а вимоги CIS до безпечного конфігурування корпоративних активів передбачають застосування автоматизованих засобів керування й контролю конфігурацій [3].

Централізований підхід має і власні ризики. Він потребує початкових витрат на ліцензії, інфраструктуру та підготовку персоналу. Консоль керування стає привабливою ціллю для зловмисників, тому її необхідно захищати багатофакторною автентифікацією та рольовим розмежуванням доступу. Крім того, помилкова політика миттєво поширюється на всі пристрої, тож зміни слід спершу перевіряти на тестовій групі пристроїв і впроваджувати поетапно.

Для порівняння підходів доцільно використати такі критерії: трудомісткість, однорідність налаштувань, рівень безпеки, швидкість реагування на інциденти, масштабованість, можливість аудиту та сукупна вартість володіння. За першими п'ятьома критеріями централізоване керування має перевагу вже за кількох десятків пристроїв. За ручного підходу сумарні трудовитрати зростають майже лінійно, а за централізованого – включають значну початкову складову (впровадження, навчання) та невеликий приріст на кожен додатковий пристрій. Положення точки беззбитковості залежить від частоти змін, кваліфікації персоналу та вартості засобів керування, тому для конкретної організації його слід оцінювати окремо.

Окремо слід враховувати вимоги стандартів. Міжнародний стандарт ISO/IEC 27001 вимагає, щоб конфігурації обладнання та програмного забезпечення були встановлені, задокументовані, впроваджені, відстежувалися та переглядалися [4]. Централізована система забезпечує ці вимоги природним чином, оскільки політики зберігаються в явному вигляді, а звіти про відповідність формуються автоматично. За ручного підходу виконання таких вимог потребує додаткових процедур і значних трудовитрат на ведення документації.

Отже, ручне конфігурування прийнятне для малих, короткочасних або вузькоспеціалізованих інфраструктур із рідкими змінами, де накладні витрати на впровадження централізованих засобів не окуповуються. Зі зростанням кількості пристроїв, поширенням віддаленої роботи, використанням особистих пристроїв працівників та посиленням вимог до відповідності доцільним стає централізоване керування. На практиці найрозумнішою є гібридна модель: єдина базова політика безпеки застосовується централізовано, а обґрунтовані виняткові налаштування вносяться вручну з обов'язковим документуванням та періодичним переглядом.

Перелік посилань:

1. NIST SP 800-128. Guide for Security-Focused Configuration Management of Information Systems. National Institute of Standards and Technology, 2011. URL: https://csrc.nist.gov/pubs/sp/800/128/upd1/final.

2. NIST SP 800-124 Rev. 2. Guidelines for Managing the Security of Mobile Devices in the Enterprise. National Institute of Standards and Technology, 2023. URL: https://csrc.nist.gov/pubs/sp/800/124/r2/final.

3. CIS Critical Security Controls v8. Control 4: Secure Configuration of Enterprise Assets and Software. Center for Internet Security. URL: https://www.cisecurity.org/controls/v8.

4. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection – Information security management systems – Requirements. URL: https://www.iso.org/standard/27001.

Єрмаков О.В.
студент групи ЗІм-25-1,
Національний технічний університет «Дніпровська політехніка»,
Дніпро, Україна

Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
Національний технічний університет «Дніпровська політехніка»
Дніпро, Україна

АНАЛІЗ УМОВ АДАПТИВНОГО КЕРУВАННЯ ДОСТУПОМ ДО ПЕРСОНАЛЬНИХ ДАНИХ У ВЕБЗАСТОСУНКАХ З УРАХУВАННЯМ ВИМОГ GDPR

Проаналізовано особливості керування доступом до персональних даних у розподілених вебзастосунках з урахуванням вимог GDPR. Систематизовано сценарії несанкціонованого отримання даних, визначено групи характеристик поведінкового ризику та обмеження використання спільних пулів з'єднань із базою даних. Обґрунтовано необхідність узгодження правових умов оброблення, контексту запиту та політик розмежування доступу. Сформовано концептуальну модель адаптивного прийняття рішень і визначено умови передавання перевіреного контексту авторизації на рівень бази даних.

Ключові слова: персональні дані, GDPR, вебзастосунок, адаптивне керування доступом, поведінковий ризик, авторизація, Row-Level Security.

Розвиток розподілених вебзастосунків супроводжується збільшенням обсягів персональних даних, що обробляються в корпоративних інформаційних системах. Їх захист потребує механізмів керування доступом, які враховують повноваження користувачів, характеристики ресурсів і контекст виконання операцій. Додаткові вимоги визначаються Загальним регламентом захисту даних Європейського Союзу (GDPR), зокрема принципами законності, обмеження мети, мінімізації даних і захисту за проєктуванням та за замовчуванням [1]. В Україні правові засади оброблення персональних даних регулюються Законом України «Про захист персональних даних» [2].

Аналіз загроз вебзастосункам дозволив систематизувати п'ять характерних сценаріїв несанкціонованого отримання персональних даних: порушення авторизації на рівні об'єктів (BOLA), використання скомпрометованих токенів доступу, автоматизоване масове зчитування інформації, зловживання повноваженнями легітимних користувачів і використання залишкових параметрів авторизації під час повторного застосування з'єднань із базою даних. Особливістю цих сценаріїв є можливість виконання небезпечних операцій через штатні інтерфейси вебзастосунку, що ускладнює їх виявлення засобами, орієнтованими лише на автентифікацію та мережеву фільтрацію.

Для розмежування доступу використовуються дискреційні (DAC), мандатні (MAC), рольові (RBAC) та атрибутні (ABAC) моделі. RBAC забезпечує перевірку повноважень відповідно до призначених ролей, тоді як ABAC дозволяє враховувати характеристики суб'єкта, ресурсу, операції та середовища [3].

Водночас реалізація безперервного оцінювання поведінкового ризику, диференційованого реагування й узгодження політик вебзастосунку з механізмами бази даних потребує додаткових засобів аналізу контексту.

Окремою проблемою є використання спільних пулів з'єднань із реляційною базою даних, коли запити різних користувачів виконуються через спільний технічний обліковий запис. За відсутності передавання перевіреного контексту кінцевого користувача база даних не має достатніх відомостей для незалежного застосування персоналізованих правил доступу. Крім того, збереження параметрів попередньої сесії може створювати ризик несанкціонованого використання повноважень. Це визначає необхідність ізоляції контексту авторизації в межах транзакції та обмеження привілеїв технічного облікового запису.

Метою дослідження є визначення умов адаптивного керування доступом до персональних даних у вебзастосунках з урахуванням вимог GDPR шляхом аналізу сценаріїв загроз, характеристик поведінкового ризику та механізмів узгодження контексту авторизації між прикладним рівнем і базою даних.

За результатами аналізу визначено три групи характеристик для оцінювання поведінкового ризику: історичні показники активності користувача (час роботи, інтенсивність запитів, типові операції), параметри поточного середовища (мережева адреса, характеристики пристрою, зміни умов доступу) та характеристики інформаційного ресурсу (категорія даних, обсяг вибірки, чутливість атрибутів). Використання цих характеристик дозволяє сформувати основу для виявлення відхилень від типової поведінки. Водночас їх збирання та аналіз також повинні відповідати вимогам GDPR щодо правомірності, визначеної мети та мінімізації персональних даних.

Для формалізації концептуальної моделі адаптивного керування доступом запропоновано залежність:

$$D = F(U, O, A, P, L, K, C, R),$$

де D – рішення щодо доступу; U – характеристики користувача; O – характеристики ресурсу; A – запитувана операція; P – мета оброблення; L – правова підстава; K – категорія персональних даних і пов'язані правові обмеження; C – контекст запиту; R – оцінка поведінкового ризику.

Обов'язкові умови надання доступу визначаються як:

$$G = G_{\text{l}} \wedge G_{\text{p}} \wedge G_{\text{a}} \wedge G_{\text{k}},$$

де G_{l}, G_{p}, G_{a}, G_{k} – результати перевірки правової підстави, відповідності меті оброблення, повноважень користувача та додаткових обмежень для відповідної категорії даних. Для спеціальних категорій персональних даних враховуються умови статті 9 GDPR поряд із правовою підставою за статтею 6 [1].

Рішення про доступ формується за правилом:

$$D = DENY, \text{ якщо } G = 0; \quad D = F_{adaptive}(C, R, O), \text{ якщо } G = 1.$$

Функція $F_{adaptive}$ визначає один із результатів: $ALLOW$ – надання доступу, $MASK$ – обмеження доступних атрибутів, $REAUTH$ – додаткова автентифікація, $DENY$– відмова. Така послідовність відокремлює перевірку правомірності оброблення від адаптивного реагування на поведінковий ризик. Маскування та

додаткова автентифікація не можуть компенсувати відсутність належної правової підстави.

Для узгодження прикладних рішень із механізмами бази даних визначено контекст авторизації:

$$C_{DB} = (U, P, L, A, K),$$

де C_{DB} – сукупність перевірених параметрів, необхідних для застосування відповідної політики доступу. Його передавання має здійснюватися з дотриманням принципу мінімізації даних та ізоляції параметрів у межах транзакції. Для додаткового розмежування доступу до записів можуть використовуватися механізми Row-Level Security (RLS) у PostgreSQL [4]. При цьому необхідно обмежувати повноваження технічного облікового запису та враховувати ризики підміни контексту скомпрометованим прикладним компонентом.

Сформована модель узгоджується з принципами захисту даних за проєктуванням і за замовчуванням, визначеними статтею 25 GDPR [5]. Її подальша перевірка має охоплювати сценарії підміни ідентифікатора об’єкта, масового зчитування даних, зміни характеристик сесії та повторного використання з’єднань. Основними показниками оцінювання визначено частку заблокованих несанкціонованих запитів, частоту хибних обмежень легітимного доступу та додаткові затримки виконання операцій.

У результаті дослідження систематизовано п’ять сценаріїв загроз персональним даним у вебзастосунках, визначено групи характеристик поведінкового ризику та обґрунтовано умови адаптивного керування доступом з урахуванням вимог GDPR. Сформовано концептуальну модель прийняття рішень із розмежуванням обов’язкових правових перевірок та адаптивного реагування, визначено структуру контексту авторизації для застосування політик на рівні бази даних. Подальше дослідження передбачає розроблення алгоритму адаптивного керування доступом, його програмну реалізацію засобами Java та PostgreSQL і експериментальне оцінювання ефективності.

Список використаних джерел

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). *Official Journal of the European Union.* 2016. L 119. P. 1–88. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

2. Про захист персональних даних : Закон України від 01.06.2010 № 2297-VI. *Відомості Верховної Ради України.* 2010. № 34. Ст. 481. URL: https://zakon.rada.gov.ua/laws/show/2297-17#Text.

3. Hu V. C., Ferraiolo D., Kuhn R. та ін. *Guide to Attribute Based Access Control (ABAC) Definition and Considerations.* NIST SP 800-162. Gaithersburg, MD: NIST, 2014 (оновлено 2019). DOI: https://doi.org/10.6028/NIST.SP.800-162.

4. PostgreSQL Global Development Group. *Row Security Policies.* PostgreSQL Documentation. URL: https://www.postgresql.org/docs/current/ddl-rowsecurity.html (дата звернення: 14.10.2026).

5. European Data Protection Board. *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default.* Version 2.0. 2020. URL: https://www.edpb.europa.eu/documents/guideline/guidelines-42019-on-article-25-data-protection-by-design-and-by-default_en.

Гайдур Г.І., д.т.н., професор,
Тонкошкур О.В.
студентка групи ІКБм-61, кафедри
СТКБ, ННІКБЗІ, Київ, Україна

ТЕХНОЛОГІЯ РЕАЛІЗАЦІЇ КОНЦЕПЦІЇ ZERO TRUST У ХМАРНОМУ СЕРЕДОВИЩІ НА БАЗІ ІНТЕГРОВАНИХ СИСТЕМ IAM ТА КОНТЕКСТНОГО КОНТРОЛЮ ДОСТУПУ

Анотація. У тезах досліджено проблематику захисту гетерогенних хмарних інфраструктур в умовах неефективності традиційного мережевого периметра. Обґрунтовано перехід до архітектури нульової довіри (Zero Trust). Запропоновано технологічну модель реалізації концепції шляхом глибокої інтеграції систем управління ідентифікацією та доступом (IAM) із механізмами безперервного контекстного контролю. Надано рекомендації щодо поетапного впровадження моделі для мінімізації ризиків несанкціонованого доступу та латерального переміщення зловмисників у хмарному середовищі.

Ключові слова: Zero Trust Architecture, хмарні середовища, кібербезпека, IAM, контекстний контроль доступу, мікросегментація, безперервна автентифікація.

Постановка проблеми

Сучасна трансформація корпоративних IT-інфраструктур характеризується масовою міграцією обчислювальних потужностей та даних у хмарні середовища (IaaS, PaaS, SaaS), а також розширенням меж віддаленої роботи. Це призвело до остаточного розмиття традиційного мережевого периметра. Класична парадигма кібербезпеки, заснована на довірі до внутрішньої мережі (модель «замку та рову»), продемонструвала свою критичну вразливість перед сучасними загрозами, такими як цілеспрямовані атаки (APT), компрометація легітимних облікових даних та використання вразливостей нульового дня.

При використанні традиційних VPN та міжмережевих екранів скомпрометований пристрій або обліковий запис отримує широкий доступ до внутрішніх ресурсів, що дозволяє зловмиснику вільно здійснювати латеральне переміщення інфраструктурою. Відповідно, виникає науково-практична проблема розробки та впровадження новітніх технологій контролю доступу, які б не базувалися на мережевому розташуванні суб'єкта, а гарантували б безпеку безпосередньо на рівні ідентифікаторів та ресурсів. Найбільш перспективним підходом до вирішення цієї проблеми є концепція Zero Trust Architecture (ZTA) [1].

Методи вирішення та технологія реалізації

Базовий принцип концепції Zero Trust полягає у постулаті «ніколи не довіряй, завжди перевіряй». Згідно зі стандартами NIST (зокрема, SP 800-207), доступ до будь-якого ресурсу повинен надаватися виключно після суворої багатофакторної ідентифікації та оцінки поточного контексту, незалежно від того, звідки надходить запит [1,2].

Для реалізації цієї парадигми у хмарному середовищі пропонується технологія, що об'єднує дві ключові підсистеми: інтегровані системи IAM (Identity and Access Management) та механізми контекстного контролю доступу (Context-Aware Access, CAA).

Системи IAM у моделі Zero Trust переходять від ролі простого каталогу користувачів до статусу центральної площини управління безпекою. Реалізація вимагає:

Ідентифікації та SSO за рахунок використання протоколів SAML 2.0 та OpenID Connect для забезпечення єдиної точки входу до всіх хмарних сервісів з централізованим логуванням.

Адаптивної багатофакторної автентифікації (MFA) щодо динамічної зміни залежно від рівня ризику. Для критичних сесій застосовуються апаратні токени (FIDO2/WebAuthn), що є стійкими до фішингу.

Управління життєвим циклом та застосування атрибутного контролю доступу (ABAC), де політики формуються на основі ролі користувача, рівня допуску та метаданих ресурсу.

Сама по собі надійна ідентифікація не гарантує безпеки, якщо сесія або пристрій були перехоплені після автентифікації. Контекстний контроль доступу передбачає оцінку параметрів середовища в режимі реального часу перед кожним доступом до мікросегмента, а саме:

- перевірка наявності актуальних оновлень ОС, активного EDR-агента, статусу шифрування диска та відсутності ознак компрометації;
- поведінкова аналітика (UEBA) - аналіз геолокації, часу доступу та типових патернів використання додатків;
- оцінка мережевих загроз.

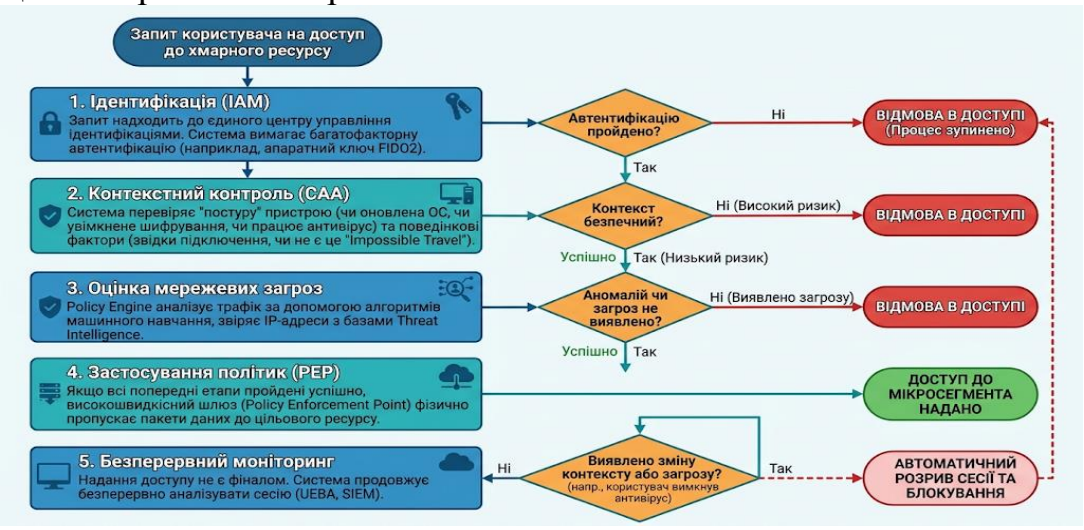


Рис. 1. Логічна архітектура системи Zero Trust з використанням контекстного контролю доступу

Опис алгоритму (рис 1.):

1. Ідентифікація (IAM): Запит надходить до єдиного центру управління ідентифікаціями. Система вимагає багатофакторну автентифікацію (наприклад, апаратний ключ FIDO2). Якщо її не пройдено, процес негайно зупиняється.

2. Контекстний контроль (CAA): Система перевіряє "постуру" пристрою (чи оновлена ОС, чи увімкнене шифрування, чи працює антивірус) та поведінкові фактори (звідки підключення, чи не є це "Impossible Travel").

3. Оцінка мережевих загроз: Policy Engine аналізує трафік за допомогою алгоритмів машинного навчання, звіряє IP-адреси з базами Threat Intelligence.

4. Застосування політик (PEP): Якщо всі попередні етапи пройдені успішно, високошвидкісний шлюз (Policy Enforcement Point) фізично пропускає пакети даних до цільового ресурсу.

5. Безперервний моніторинг: Надання доступу не є фіналом. Система продовжує аналізувати сесію. Якщо поведінка стає підозрілою або користувач вимикає антивірус, доступ автоматично розривається (зворотний пунктирний зв'язок до блокування).

Рекомендації щодо впровадження

Трансформація інфраструктури на базі моделі Zero Trust не є одноразовим розгортанням програмного забезпечення, а являє собою стратегічний процес. Для організацій та операторів хмарних середовищ пропонується наступний алгоритм впровадження:

1. Замість захисту макромережі необхідно ідентифікувати критичні дані, додатки, активи та сервіси.

2. Картування транзакційних потоків. Розуміння того, як різні компоненти інфраструктури взаємодіють між собою. Це дозволить налаштувати політики мікросегментації так, щоб не порушити бізнес-процеси.

3. Розгортання централізованої системи IAM. Впровадження єдиного каталогу ідентифікацій, обов'язкової стійкої до фішингу MFA та інтеграції з системами управління пристроями для збору базового контексту.

4. Формування гранулярних політик доступу. Розробка правил на основі методології «Хто, Що, Коли, Де, Чому і Як» для кожного мікросегмента. Політики мають враховувати результати ML-аналізу мережевих аномалій.

5. Безперервний моніторинг та адаптація. Постійне логування всіх сесій, аналіз відхилень з використанням SIEM-систем та автоматизоване реагування.

Висновки

Реалізація архітектури Zero Trust у хмарному середовищі шляхом синергії систем IAM та динамічного контекстного контролю доступу є єдиним надійним підходом до протидії сучасним кіберзагрозам.

Список використаної літератури

1. Розробка архітектури нульової довіри (Zero Trust Architecture). NIST Special Publication 800-207. 2020. URL: https://doi.org/10.6028/NIST.SP.800-207 (дата звернення: 14.10.2026).

2. Борсуковський Ю.В., Борсуковська В.Ю., Гайдур Г.І., Складанний П.М., Бурячок В.Л., Прикладні аспекти інформаційної та кібернетичної безпеки держави. Аналіз мережевого трафіку: навчальний посібник / ISBN 978-617-574-272-3 / УДК 32.973я73 р. / – Львів - Видавництво «Магнолія 2006», 2023, 222 с.

Гимза О.А.
студент групи КБАС-21
НУ «Львівська політехніка», Львів, Україна
Кеньо Г.В.
доцент каф. ЗІ НУ «Львівська політехніка»
Львів, Україна

ІНФОРМАЦІЙНА БЕЗПЕКА СИСТЕМ РОЗУМНОГО ПАРКУВАННЯ

У роботі розглянуто проблеми забезпечення інформаційної безпеки в інтелектуальних системах паркування, що базуються на технологіях Інтернету речей (IoT), сенсорних мережах та хмарних сервісах. Проаналізовано основні загрози, пов'язані з конфіденційністю персональних даних користувачів, цілісністю інформації від сенсорів та доступністю сервісів у реальному часі. Окрему увагу приділено питанням автентифікації та захисту каналів комунікації. Запропоновано комплекс організаційних та технічних заходів, включаючи використання криптографічних протоколів, сегментацію мережі, моніторинг аномалій та регулярне оновлення програмного забезпечення.

Ключові слова: Інтернет речей, інтелектуальна система паркування, кібербезпека, автентифікація, шифрування.

Однією з перспективних галузей застосування IoT є організація систем розумного паркування, які дають змогу підвищити ефективність використання паркомісць, автоматизувати облік вільних місць та забезпечити зручність для водіїв. В умовах дефіциту паркувальних місць і зростання транспортних потоків впровадження таких систем має важливе практичне значення.

Паралельно з розширенням можливостей IoT-систем постає питання забезпечення їх надійності та безпеки. Враховуючи велику кількість пристроїв, відкритість бездротових каналів зв'язку та чутливість даних, які обробляються в таких системах, розробка безпечної архітектури є критично необхідною.

Системи розумного паркування – це приклад децентралізованих IoT-інфраструктур, у яких взаємодіє велика кількість пристроїв. Це створює численні точки потенційного доступу, що підвищує ризики як кібернетичних, так і фізичних атак [1, с. 1]. Забезпечення інформаційної безпеки в таких системах вимагає застосування цілісного підходу – на всіх рівнях архітектури: від сенсора до сервера обробки даних [2, с. 5].

Найперше, важливо впроваджувати взаємну автентифікацію пристроїв. Навіть найменші компоненти повинні мати унікальні ідентифікатори, ключі або токени. Це дозволяє уникнути підключення сторонніх пристроїв до системи [3, с. 38]. Для пристроїв з обмеженими ресурсами актуальними є легкі криптографічні алгоритми, наприклад, ECC або SHA-256.

Не менш критичним є захист передавання даних. Всі канали зв'язку між компонентами мають бути шифрованими. Залежно від ресурсів системи, використовують TLS, DTLS або власні легкі реалізації. Додатково, доцільно впроваджувати шифрування "на периферії" – тобто дані шифруються ще на сенсорному рівні, до моменту надходження до шлюзу [4, с. 8].

Окремої уваги потребують шлюзи, адже саме вони з'єднують ресурсообмежені сенсорні вузли з сервером обробки даних. Шлюз перевіряє автентичність пристроїв, розшифровує та передає дані далі захищеним каналом,

тому його компрометація ставить під загрозу весь сегмент паркування. Для захисту доцільно обмежувати кількість відкритих портів і сервісів, застосовувати двофакторну автентифікацію та вести журнал подій. Важливим є й моніторинг аномалій: різке зростання кількості запитів чи спроби підключення невідомих пристроїв можуть свідчити про атаку. Оновлення прошивки слід встановлювати лише за наявності цифрового підпису.

Захищена система розумного паркування також повинна дотримуватись архітектурного поділу на зони довіри, наприклад:

- сенсори і виконавчі пристрої – низький рівень довіри;
- шлюзи – контрольований сегмент;
- сервери або хмара – ізольована зона із повним контролем доступу.

Для реалізації безпечної IoT-інфраструктури зазвичай передбачають базові компоненти захисту, представлені в табл. 1.

Таблиця 1. Багаторівневий захист IoT-інфраструктури

Рівень системи	**Типові загрози**	**Рекомендовані заходи захисту**
Сенсорний	Фізичний доступ, підміна пристрою	Ідентифікація пристрою, захищене підключення, механічне закріплення
Мережевий	Перехоплення даних, DoS-атаки, підміна трафіку	Шифрування (TLS/DTLS), фільтрація трафіку, ізоляція VLAN
Обчислювальний	Вразливості прошивки, несанкціоноване оновлення	Цифровий підпис ПЗ, контроль доступу, регулярні оновлення
Шлюз	Злам інтерфейсу, компрометація контролю	Двофакторна автентифікація, моніторинг, журналювання
Хмара / сервер	Злам акаунтів, витік персональних даних	Обмеження доступу, рольове управління, шифрування БД, GDPR-відповідність

Безпека в IoT – це не лише про захист каналів, а й про контроль життєвого циклу пристроїв: від їх встановлення, конфігурації до виведення з експлуатації. Часто саме відсутність політик щодо заміни ключів, управління оновленнями чи ревізії дозволів створює найсерйозніші “дірки” в захисті.

У результаті, ефективна система безпеки для розумного паркування – це поєднання технічних рішень, грамотної архітектури та організаційного контролю. Її завдання – не лише запобігти атаці, а й бути готовою до швидкого реагування, якщо загроза все ж реалізується.

1. Mamun A.A. et al. IoT-Enabled Smart Car Parking System through Integrated Sensors and Mobile Applications // arXiv preprint arXiv:2412.10774. – 2024. – Режим доступу: https://arxiv.org/abs/2412.10774 - Дата звернення: 10.10.2026.

2. Alsafar B.S. et al. Smart Parking Management System Using IoT // Journal of Information Systems and Information Security. – 2024. – Vol. 2, No. 1. – P. 285–292. - Режим доступу: https://jisis.org/wp-content/uploads/2024/05/2024.I2.018.pdf - Дата звернення: 10.10.2026.

3. Wahab A., Maguire P. Smart Parking: IoT and Blockchain // arXiv preprint arXiv:1912.01697. – 2019. - Режим доступу: https://arxiv.org/abs/1912.01697 - Дата звернення: 10.10.2026.

4. Abujoodeh M., Tamimi L., Tahboub R. Toward Lightweight Cryptography: A Survey // IntechOpen. – 2023. – Режим доступу: http://dx.doi.org/10.5772/intechopen.109334 - Дата звернення: 10.10.2026.

Козловський М. В.
студент групи ЗІм-25-1,
Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
НТУ «Дніпровська політехніка»
Дніпро, Україна

АНАЛІЗ УМОВ ЗАПОБІГАННЯ ВИТОКАМ ТРАФІКУ ПІД ЧАС УСТАНОВЛЕННЯ ТА ВІДНОВЛЕННЯ ЗАХИЩЕНИХ ТУНЕЛЬНИХ З'ЄДНАНЬ

Розглянуто причини виникнення витоків трафіку, призначеного для передавання через захищений канал, під час установлення та відновлення тунельних з'єднань у комп'ютерних мережах. Проаналізовано особливості функціонування технологій WireGuard, IPsec/IKEv2 та OpenVPN, а також можливості штатних механізмів установлення, підтримання й відновлення захищеного з'єднання. Визначено ризики, пов'язані з неузгодженою зміною стану тунелю, маршрутизації та правил фільтрації. Обґрунтовано умови готовності захищеного каналу, сформовано інваріант безпеки та запропоновано узагальнену модель станів із блокуванням прикладного трафіку до підтвердження готовності тунелю до його безпечного передавання. Визначено критерії експериментального оцінювання запропонованої моделі.

Ключові слова: тунелювання, VPN, захищений канал, відновлення з'єднання, витік трафіку, маршрутизація, правила фільтрації, модель станів.

Технології тунелювання, зокрема WireGuard, IPsec/IKEv2 та OpenVPN, забезпечують захист передавання даних через недовірені мережі за допомогою криптографічних механізмів [1, 2]. Проте в умовах нестабільного зв'язку, зміни IP-адреси або перемикання мережевих інтерфейсів виникає ризик витоку трафіку під час установлення та відновлення захищеного з'єднання. Це пов'язано з неузгодженістю відновлення мережевої доступності, стану тунелю, маршрутизації та правил фільтрації. Зокрема, маршрут за замовчуванням може відновитися раніше, ніж завершиться встановлення захищеного каналу, що створює можливість передавання прикладного трафіку, включно з DNS-запитами, поза тунелем.

Аналіз технологій тунелювання виявив відмінності в механізмах установлення та відновлення з'єднань. WireGuard використовує криптографічне рукостискання та підтримує зміну мережевої адреси вузла [1]. IPsec/IKEv2 забезпечує встановлення асоціацій безпеки й підтримку мобільності через MOBIKE, а OpenVPN використовує TLS та механізми повторного з'єднання. Водночас завершення протокольних процедур не гарантує готовності маршрутизації та правил фільтрації.

Активність віртуального інтерфейсу WireGuard не підтверджує доступності віддаленого вузла, а час останнього рукостискання та статистика пакетів не гарантують працездатності захищеного маршруту. В IPsec база політик безпеки Security Policy Database (SPD) визначає оброблення трафіку за діями DISCARD, BYPASS або PROTECT [5]. В інших VPN-технологіях блокування трафіку поза тунелем може забезпечуватися засобами операційної системи та міжмережевого екрана. Тому готовність захищеного каналу потребує узгодженої перевірки стану тунелю, маршрутизації та правил фільтрації.

Практичне значення проблеми витоків трафіку підтверджується результатами досліджень захищеності VPN. У роботі Xue та співавт. [3] продемонстровано атаки TunnelCrack, які використовують особливості налаштування маршрутів і винятків маршрутизації VPN-клієнтів для спрямування частини трафіку поза захищеним тунелем. Встановлено, що порушення конфіденційності може відбуватися без компрометації криптографічних алгоритмів. У дослідженні Mixon-Baca та співавт. [4] розглянуто атаки на механізми відстеження стану мережевих з'єднань, що застосовуються у VPN-інфраструктурі. Хоча ці атаки не пов'язані безпосередньо з відновленням клієнтського тунелю, вони демонструють залежність захищеності VPN від налаштування та функціонування мережевих компонентів операційної системи.

Метою дослідження є визначення умов запобігання витокам трафіку, що підлягає передаванню через захищений канал, під час установлення та відновлення тунельних з'єднань шляхом аналізу взаємозв'язку стану тунелю, криптографічного контексту, маршрутизації та правил фільтрації й обґрунтування моделі узгодженого керування їх змінами.

За результатами аналізу встановлено, що ризик передавання даних поза тунелем може виникати внаслідок неодночасного відновлення окремих компонентів захищеного каналу. Для врахування їх взаємозалежності запропоновано описувати стан захищеного каналу вектором:

$$S = (S_{\mathrm{t}}, S_{\mathrm{c}}, S_{\mathrm{r}}, S_{\mathrm{a}}),$$

де S_{t} – стан тунельного з'єднання; S_{c} – стан криптографічного контексту; S_{r} – стан маршрутизації; S_{a}– стан правил доступу та фільтрації.

Готовність захищеного каналу до передавання прикладного трафіку визначається одночасним виконанням умов:

$$Ready(S) = T_{\mathrm{ok}} \wedge C_{\mathrm{ok}} \wedge R_{\mathrm{ok}} \wedge A_{\mathrm{ok}},$$

де T_{ok} – підтвердження доступності тунельного з'єднання; C_{ok} – наявність чинного криптографічного контексту; R_{ok} – відповідність маршрутів установленій конфігурації захищеного передавання; A_{ok} – застосування правил фільтрації, які унеможливлюють передавання захищуваного трафіку поза тунелем.

Основною умовою запобігання витокам визначено блокування прикладного трафіку, призначеного для передавання через захищений канал, за відсутності підтвердженої готовності тунельного з'єднання. Відповідний інваріант безпеки можна представити залежністю:

$$\neg Ready(S) \Rightarrow Block(P),$$

де $Block(P)$ – блокування прикладного трафіку, призначеного для передавання через захищений канал, до відновлення необхідних умов функціонування тунелю. Водночас передавання службового трафіку, необхідного для встановлення та підтримання тунельного з'єднання, повинно регулюватися окремими правилами доступу до визначених кінцевих точок.

Для забезпечення виконання зазначених умов запропоновано узагальнену модель керування станами захищеного каналу, яка містить п'ять основних

станів: BLOCKED, CONNECTING, VERIFYING, READY та RECOVERING. У стані BLOCKED забороняється передавання прикладного трафіку, призначеного для передавання через захищений канал. Стан CONNECTING передбачає встановлення тунельного з'єднання та необхідних криптографічних параметрів. У стані VERIFYING здійснюється перевірка готовності тунелю, маршрутів і правил фільтрації. Перехід до READY допускається лише після підтвердження всіх визначених умов. У разі втрати зв'язку, зміни адреси клієнта або порушення налаштувань захищеного передавання система переходить у стан RECOVERING із подальшим повторним встановленням необхідних параметрів і перевіркою готовності каналу.

Особливістю запропонованої моделі є визначення готовності захищеного каналу за результатами спільної перевірки стану тунелю, криптографічного контексту, маршрутизації та правил фільтрації. Для запобігання витокам передбачено збереження правил блокування під час відновлення з'єднання та відмови керувального компонента, перевірку виконання змін і повернення до безпечного стану в разі порушення визначених умов. Ефективність моделі планується оцінювати за часом виявлення та відновлення з'єднання, тривалістю блокування, часткою успішних відновлень і кількістю пакетів, переданих поза тунелем, цільовим значенням якої є нуль. Експериментальна перевірка передбачає сценарії розриву зв'язку, зміни IP-адреси, перемикання інтерфейсів, втрат пакетів і збільшення мережевої затримки.

У результаті дослідження проаналізовано причини виникнення витоків захищуваного трафіку під час установлення та відновлення тунельних з'єднань. Визначено, що істотними чинниками ризику є неузгоджена зміна стану тунелю, криптографічного контексту, маршрутів і правил фільтрації, а також відсутність комплексного підтвердження готовності захищеного каналу. Обґрунтовано умови безпечного передавання даних, сформовано вектор стану захищеного каналу та інваріант блокування трафіку за відсутності його готовності. Запропоновано узагальнену модель керування станами з перевіркою виконання необхідних умов перед дозволом передавання прикладного трафіку. Подальше дослідження передбачає деталізацію алгоритму переходів між станами, реалізацію програмного прототипу для ОС Linux та експериментальне оцінювання можливості запобігання витокам трафіку під час установлення й відновлення тунельних з'єднань.

Список використаних джерел:

1. Donenfeld J. A. WireGuard: Next Generation Kernel Network Tunnel. Proceedings of the Network and Distributed System Security Symposium (NDSS 2017). San Diego, 2017. URL: https://www.wireguard.com/papers/wireguard.pdf (дата звернення: 14.10.2026).

2. Iyengar J., Thomson M. QUIC: A UDP-Based Multiplexed and Secure Transport. RFC 9000. Internet Engineering Task Force, 2021. DOI: https://doi.org/10.17487/RFC9000.

3. Xue N., Malla Y., Xia Z., Pöpper C., Vanhoef M. Bypassing Tunnels: Leaking VPN Client Traffic by Abusing Routing Tables. Proceedings of the 32nd USENIX Security Symposium (USENIX Security 23). Anaheim, CA : USENIX Association, 2023. P. 5719–5736. URL: https://www.usenix.org/conference/usenixsecurity23/presentation/xue.

4. Mixon-Baca B., Knockel J., Xue D., Ayyagari T., Kapur D., Ensafi R., Crandall J. R. Attacking Connection Tracking Frameworks as used by Virtual Private Networks. Proceedings on Privacy Enhancing Technologies. 2024. Issue 3. P. 109–126. DOI: https://doi.org/10.56553/popets-2024-0070.

5. Kent S., Seo K. Security Architecture for the Internet Protocol. RFC 4301. Internet Engineering Task Force, 2005. DOI: https://doi.org/10.17487/RFC4301.

Кутовий Денис Сергійович
Студент групи БСДМ-62, ННІКБЗІ ДУІКТ,
Київ, Україна

ПРОФІЛЬ ЗАХИСТУ ДЛЯ ПРОГРАМ-АРХІВАТОРІВ ВІДПОВІДНО ДО ВИМОГ СТАНДАРТУ ISO/IEC 15408

Програми-архіватори є невід'ємною частиною сучасних інформаційних систем, забезпечуючи стиснення та збереження даних. Однак, обробка конфіденційної інформації вимагає належного рівня захисту від несанкціонованого доступу та модифікації. Проте, створення профілю захисту для програм-архіваторів на основі міжнародного стандарту ISO/IEC 15408 дозволить формалізувати вимоги безпеки та забезпечити надійну оцінку таких програмних продуктів.

Ключові слова: профіль захисту, програми-архіватори, кібербезпека, ISO/IEC 15408.

Програми-архіватори часто використовуються для зберігання та передачі критично важливих даних як у корпоративному, так і в державному секторах. Це робить їх привабливою мішенню для зловмисників. Вразливості в алгоритмах шифрування, обробці паролів або механізмах розпакування можуть призвести до витоку інформації, виконання довільного коду або пошкодження даних. Тому існує нагальна потреба у стандартизації вимог до безпеки таких програм[2].

Стандарт ISO/IEC 15408, відомий як "Загальні критерії оцінки безпеки інформаційних технологій", надає універсальну методологію для розробки та оцінки вимог безпеки IT-продуктів. Основою цієї методології є розробка профілю захисту, який визначає незалежний від реалізації набір вимог безпеки для певної категорії продуктів[1]. Для програм-архіваторів такий профіль повинен враховувати специфічні загрози, пов'язані з обробкою та зберіганням стиснених файлів.

Основними функціональними вимогами безпеки, які мають бути включені до профілю захисту архіваторів, є надійна криптографічна підтримка, захист даних користувача, ідентифікація та автентифікація, а також управління безпекою[3]. Особливу увагу слід приділити механізмам шифрування архівів та стійкості алгоритмів хешування для перевірки цілісності даних.

Існуючі рішення на ринку здебільшого використовують фрагментовані або пропрієтарні механізми захисту, які не мають єдиної бази для незалежного оцінювання і не враховують повного спектру загроз. На відміну від існуючих загальних специфікацій, пропонується розробити спеціалізований профіль захисту[1]. Як приклад, для протидії актуальній загрозі несанкціонованого доступу до вмісту архіву шляхом атаки грубої сили та прихованої модифікації

стиснених даних, у профілі сформульовано цілі безпеки щодо надійного криптографічного захисту та контролю цілісності.

Для реалізації цих цілей обрано конкретні функціональні вимоги безпеки з Частини 2 стандарту: генерацію криптографічних ключів на основі пароля зі стійкими функціями формування ключа, такими як PBKDF2 або Argon2, симетричне шифрування за алгоритмом AES-256 та перевірку цілісності збережених даних перед розпакуванням за допомогою хешування SHA-256/SHA-3 [3].

Розробка та впровадження стандартизованого профілю захисту для програм-архіваторів відповідно до ISO/IEC 15408 дозволить виробникам створювати більш захищені продукти, а споживачам – мати впевненість у їхній надійності. Це сприятиме підвищенню загального рівня кібербезпеки та мінімізації ризиків, пов'язаних із компрометацією архівних даних[1,2].

Перелік посилань:

1. ДСТУ ISO/IEC 15408-1:2017. Інформаційні технології. Методи захисту. Критерії оцінювання безпеки інформаційних технологій. Частина 1. Вступ і загальна модель.

2. Інформаційна безпека : рекомендаційний покажчик / уклад. : С. Л. Бондар, Т. А. Сіденко. – Чернігів : Наукова бібліотека НУ "Чернігівська політехніка", 2020. – 44 с.

3. Common Criteria for Information Technology Security Evaluation. Part 2: Security functional components. Version

Балацький Андрій Володимирович
студент групи ІКБм-61, ННКБЗІ ДУІКТ, Київ, Україна

ВПРОВАДЖЕННЯ БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ НА БАЗІ CISCO DUO FOR WINDOWS LOGON

Анотація

У тезах розглядається впровадження багатофакторної автентифікації для Windows Logon на базі рішення Cisco Duo. Описано архітектуру інтеграції з інфраструктурою підприємства, принципи забезпечення Zero Trust доступу на кінцевих точках, а також ключові етапи розгортання системи. Особлива увага приділяється налаштуванням політик доступу, аналізу стану пристроїв та мінімізації ризиків блокування користувачів під час масового впровадження [1, 2].

Багатофакторна автентифікація є ефективним інструментом для захисту інформаційної системи організації та протидії кіберзагрозам, пов'язаним із компрометацією облікових даних. Впровадження такого рішення для процесу входу в операційну систему, однак, вимагає ретельного планування архітектури та інтеграції з існуючою інфраструктурою.

Найчастіше перед IT-адміністраторами стоїть завдання забезпечення безпечного доступу як для локальних, так і для віддалених сесій. Це передбачає встановлення спеціалізованого агента (Credential Provider), який інтегрується безпосередньо у процес автентифікації Windows. Рішення Cisco Duo Authentication for Windows Logon перехоплює запит на вхід після введення первинного пароля та вимагає від користувача підтвердження через другий фактор (Push-повідомлення на смартфон, апаратний токен або біометрію). Зібрані дані про спроби автентифікації здатні надавати інформацію, необхідну для отримання уявлення про стан безпеки кінцевих точок та виявлення аномалій. При визначенні того, яку архітектуру розгортання використовувати, ключовим моментом є зосередження на моделі загроз організації [1].

Duo Authentication Proxy – це інструмент, який є невід'ємною частиною розгортання Duo у великих корпоративних мережах. Це програмне забезпечення, яке виступає шлюзом між локальним середовищем Active Directory та хмарним сервісом Duo. Окрім базової перевірки пароля, система також використовує контекстну інформацію – дані про геолокацію, тип мережі та стан пристрою. Ця інформація може бути використана для динамічної оцінки ризику та блокування підозрілих подій автентифікації [1].

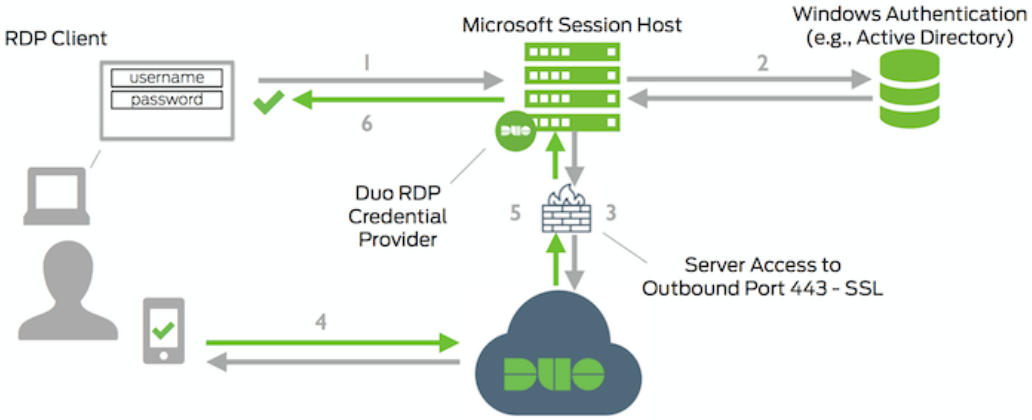


Рис. 1. Процес входу через Duo

Створення повноцінної архітектури Zero Trust на кінцевих точках є серйозним викликом. Найпростіший підхід полягає в тому, щоб почати з розгортання на обмеженій групі IT-адміністраторів, а потім повільно і контрольовано розширювати політику на всіх користувачів. Щоб досягти цього, слід почати з того, що команда спеціалістів повинна налаштувати сценарії відмовостійкості (Fail-Open або Fail-Closed) на випадок втрати зв'язку агента з хмарним API Duo. Необхідно зосередитися на створенні резервних облікових записів, щоб уникнути повного блокування доступу до серверів при активації суворого режиму Fail-Closed [2].

Для того, щоб забезпечити надійний контроль доступу та реалізувати парадигму Zero Trust, організаціям потрібно робити більше, ніж просто вимагати підтвердження на смартфоні. Організація повинна здійснити низку заходів:

Створення політик доступу: політика інформаційної безпеки описує умови, за яких доступ дозволяється або блокується. Цілі, викладені в політиці, допомагають визначити правила (наприклад, заборона входу з анонімних проксі, TOR-мереж або з країн, що не входять до списку дозволених) [2].

Аналіз стану довіри до пристрою: огляд ландшафту пристроїв дає уявлення про те, наскільки безпечними є кінцеві точки, з яких відбувається вхід. Такий огляд є ключовим для адекватного контролю. Налаштування Endpoint Policy гарантує, що система перевірить наявність активованого BitLocker, оновленого антивірусу та актуальної версії ОС перед тим, як дозволити RDP-сесію.

Результати аналізу ризиків та масове розгортання: аналіз ризиків допомагає виявити можливі проблеми, пов'язані з блокуванням користувачів через помилки ручного встановлення. Ця інформація допомагає визначити необхідність використання GPO (Group Policy Object) або Microsoft SCCM для централізованої доставки MSI-пакетів агента Duo із зашитими ключами IKEY, SKEY та API Hostname [1].

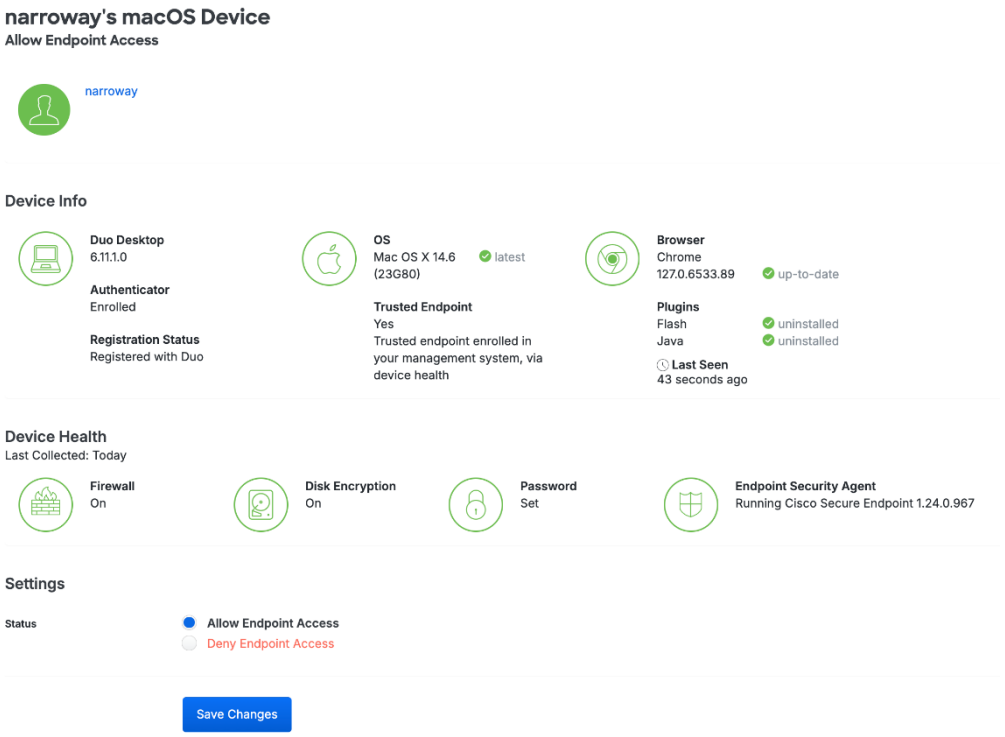


Рис. 2. Оцінка стану пристрою

Один з ризиків занадто швидкого розгортання полягає в тому, що обсяг звернень користувачів щодо проблем з автентифікацією (втрата телефону, неможливість отримати Push-повідомлення) перевищить здатність відділу технічної підтримки обробити їх. Крім того, відділ технічної підтримки повинен бути готовим до генерації Bypass-кодів для тимчасового доступу. Слід обмежити розгортання пілотними групами, виходячи з пропускної спроможності технічної підтримки.

Розширення впровадження MFA, швидше за все, призведе до збільшення кількості аспектів, які потребують моніторингу. Інакше кажучи, збільшиться кількість журналів подій, які використовуються для виявлення того, чи відбулася спроба атаки типу MFA Fatigue. Більша кількість захищених кінцевих точок означає збільшення обсягу даних, які необхідно експортувати через Duo Admin API до корпоративної SIEM-системи для кореляції та своєчасного реагування на інциденти.

Перелік посилань:

1. Duo Authentication for Windows Logon and RDP. URL: https://duo.com/docs/windows (дата звернення: 28.09.2026).
2. Zero Trust Security for the Workforce: Device Trust. URL: https://duo.com/solutions/zero-trust-security (дата звернення: 28.09.2026).

УДК 004.7:629.735

Коваленко Ю. Б.
доцент, кафедри штучного інтелекту
ДУІКТ
Київ, Україна

МОДЕЛЬ ОЦІНЮВАННЯ ЗАТРИМОК І СВОЄЧАСНОСТІ ПЕРЕДАВАННЯ ДАНИХ У МЕРЕЖАХ AFDX ТА TSN

Досліджено інформаційну підтримку проєктування захищених мереж обміну даними інтегрованої модульної авіоніки. Запропоновано модель оцінювання часових характеристик потоків в архітектурах AFDX і TSN з урахуванням вимог доступності, цілісності та відмовостійкості інформаційного обміну. Наскрізна затримка охоплює оброблення, очікування, передавання й поширення сигналу. Враховано топологію, пріоритети, резервування, синхронізацію та параметри мережевих каналів. Модель інтегровано з графом функцій і процедурою розміщення обчислювальних задач. Передбачено аналіз навантаження, пакетних сплесків, змін маршрутів і відмов каналів. Оцінюються затримка, джитер, порушення дедлайнів і своєчасність доставки. Розмежовано аналітичні гарантії та результати моделювання. Практична цінність полягає у виявленні часових конфліктів, обґрунтуванні мережевої конфігурації та підвищенні стійкості інформаційного обміну.

Ключові слова: системи захисту інформації, AFDX, TSN, мережеве числення, наскрізна затримка.

Постановка проблеми. У комплексі бортового обладнання (КБО) з інтегрованою модульною авіонікою (IMA) функції різного рівня критичності обмінюються даними через спільну мережу, тому запізнення або втрата кадру може порушити виконання критичної функції так само, як і відмова обчислювального модуля. Для магістральних платформ стандартом де-факто є комутована мережа AFDX (Avionics Full-Duplex Switched Ethernet) за ARINC 664 Part 7 [1], у якій трафік кожного джерела жорстко регламентовано віртуальними каналами. Для розподіленої IMA та модернізації успадкованих платформ дедалі частіше розглядають набір стандартів TSN (Time-Sensitive Networking) IEEE 802.1 [2] з часовою синхронізацією 802.1AS, плануванням передавання за розкладом 802.1Qbv і витісненням кадрів 802.1Qbu. Проєктувальникові потрібна єдина модель, яка за однакових потоків дає порівнянні оцінки часових характеристик обох технологій і при цьому розрізняє дві принципово різні величини: аналітично гарантовану верхню межу затримки та максимум, що спостерігається в імітаційному експерименті. Їх змішування веде або до надмірного резервування пропускної здатності, або до необґрунтованого перенесення експериментальних значень на всі режими роботи.

Аналіз останніх досліджень і публікацій. Теоретичною основою детермінованих гарантій є мережеве числення (network calculus), у якому потік описується кривою надходження, а вузол — кривою обслуговування, і верхня межа затримки визначається горизонтальною відстанню між ними [3]. Для AFDX класичні методи обмеження наскрізної затримки порівнювали мережеве числення, перевірку моделей (model checking) та імітацію та показали, що аналітичні межі суттєво перевищують спостережувані значення [4].

Мета роботи — побудувати модель оцінювання наскрізної затримки, джитера та своєчасності доставки для мереж AFDX і TSN за однакових потоків і явно визначених умов обслуговування та на модельному прикладі кількісно показати розрив між гарантованою межею і спостережуваним максимумом.

Викладення основного матеріалу. Мережу описано графом кінцевих систем і комутаторів, де кожен потік f має маршрут P_f, довжину кадру, період (або BAG) і допустиму затримку, що випливає з вимог функції в графі функцій. Наскрізну затримку подано сумою складових на переходах маршруту:

$$d_f = \sum_{h \in P_f} \left(d_h^{proc} + d_h^{queue} + d_h^{tx} + d_h^{prop} \right) \le D_f^{req}, \quad d_h^{tx} = \frac{L_f}{C}, \tag{1}$$

де d_h^{proc}, d_h^{queue}, d_h^{tx}, d_h^{prop} — затримки оброблення у вузлі, очікування в черзі, передавання та поширення сигналу на переході h; L_f — довжина кадру потоку, біт; C — пропускна здатність каналу, біт/с; D_f^{req} — допустима затримка потоку.

Складові оброблення, передавання й поширення визначаються обладнанням і топологією, а складова черги — сукупним трафіком і механізмом обслуговування; саме вона відрізняє AFDX від TSN. В AFDX кожен віртуальний канал (VL, Virtual Link) задається парою «максимальний розмір кадру — BAG», де BAG (Bandwidth Allocation Gap) — мінімальний інтервал між кадрами зі стандартизованих значень 1–128 мс [1]. Ця пара однозначно перетворюється на криву надходження:

$$\alpha_f(t) = b_f + r_f t, \quad b_f = S_f^{\max}, \quad r_f = \frac{S_f^{\max}}{BAG_f}, \tag{2}$$

де b_f — максимальний сплеск потоку, біт; r_f — довготривала швидкість, біт/с; $S_f^{\max}$ — максимальний розмір кадру віртуального каналу, біт; BAG_f — інтервал регулювання каналу, с.

Вихідний порт комутатора описується кривою обслуговування типу «швидкість — латентність», і для сукупності потоків, що ділять одну чергу, гарантована межа затримки на переході дорівнює горизонтальній відстані між кривими [3]:

$$\beta_h(t) = R_h \max(0,\, t - T_h), \quad D_h = \frac{\sum_{f \in F_h} b_f}{R_h} + T_h, \tag{3}$$

де R_h — гарантована швидкість обслуговування черги, біт/с; T_h — латентність вузла, с; F_h — множина потоків, що ділять чергу на переході h. Для AFDX з однією FIFO-чергою F_h містить усі віртуальні канали порту; для TSN з пріоритетом критичного класу і витісненням 802.1Qbu — лише критичні потоки, а до T_h додається час передавання неперериваного залишку кадру нижчого класу.

Межу тракту отримано підсумовуванням меж (3) зі зростанням сплеску після кожного переходу на $r_f D_h$ — консервативно, але строго. Якість доставки оцінюють показниками:

$$J_f = d_f^{\max} - d_f^{\min}, \quad K_{time} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\left[d_m \le D_f^{req} \right], \tag{4}$$

де J_f — джитер потоку; $d_f^{\max}$, $d_f^{\min}$ — максимальна та мінімальна наскрізні затримки; K_{time} — частка своєчасно доставлених кадрів; M — кількість кадрів у вибірці; $\mathbf{1}[\cdot]$ — індикаторна функція.

Для TSN із синтезованим розкладом 802.1Qbv часові вікна критичних потоків на кожному порту не перетинаються, тому складова черги для

критичного трафіку виключається конструктивно, а необхідною умовою існування розкладу є обмеження завантаження критичним класом:

$$\sum_{f \in F_h^c} \frac{L_f}{p_f C} \le U_{\max} < 1, \quad D_f^{Qbv} = \sum_{h \in P_f} \left(\frac{L_f}{C} + T_h + 2\delta_{sync} \right), \tag{5}$$

де F_h^c — множина критичних потоків порту h; p_f — період потоку; $U_{\max}$ — конструктивний запас завантаження (прийнято 0,75); δ_{sync} — межа похибки часової синхронізації за 802.1AS; D_f^{Qbv} — межа затримки за коректного розкладу.

Модель апробовано в модельному (обчислювальному) експерименті на синтетичному сценарії (власний відтворюваний Python-скрипт). Тракт критичного потоку — кінцева система, два комутатори, кінцева система; канали 100 Мбіт/с, латентність комутатора 16 мкс, поширення 0,5 мкс на лінію. На кожному вихідному порту діє набір віртуальних каналів трьох класів: критичний (кадр 200 байт, BAG 1–2 мс, 40 % навантаження), важливий (500 байт, BAG 2–4 мс, 30 %) і службовий (1518 байт, BAG 4–8 мс, 30 %); навантаження порту змінювалося від 10 до 80 % масштабуванням кількості каналів (за 50 % — 34 канали, з них 17 критичних). Порівнювали три конфігурації: AFDX з однією FIFO-чергою; TSN із пріоритетом критичного класу та витісненням кадрів (непереривний залишок — 143 байти); TSN із синтезованим розкладом 802.1Qbv (δ_{sync} = 1 мкс). Межі обчислено за формулами (3) і (5); дискретно-подійна імітація перших двох конфігурацій охоплювала 8 реплікацій по 1 с із випадковими фазами каналів і сценарій синхронного старту всіх каналів (несприятливий збіг сплесків), причому програмно перевірено, що жодна імітаційна затримка не перевищує межі.

Спостережувані максимуми виявилися в 1,4–9,8 раза меншими за гарантовані межі, що підтверджує необхідність розмежування цих величин під час проєктування. Подальші дослідження стосуються гібридних трактів AFDX–TSN зі шлюзами, урахування деградації синхронізації та верифікації моделі на стендовій мережі.

Список використаних джерел

1. ARINC: Aircraft Data Network Part 7, Avionics Full-Duplex Switched Ethernet Network. ARINC Specification 664P7-1, Annapolis, MD (2009). URL: https://standards.globalspec.com/std/1283307/arinc-664-p7, last accessed 2026/10/16.

2. Institute of Electrical and Electronics Engineers: IEEE Standard for Local and Metropolitan Area Networks — Bridges and Bridged Networks. IEEE Std 802.1Q-2022, New York, NY (2022). URL: https://standards.ieee.org/ieee/802.1q/10323/, last accessed 2026/10/16.

3. Le Boudec, J.-Y., Thiran, P.: Network Calculus: A Theory of Deterministic Queuing Systems for the Internet. LNCS, vol. 2050. Springer, Berlin, Heidelberg (2001). https://doi.org/10.1007/3-540-45318-0

4. Charara, H., Scharbarg, J.-L., Ermont, J., Fraboul, C.: Methods for Bounding End-to-End Delays on an AFDX Network. In: Proceedings of the 18th Euromicro Conference on Real-Time Systems — ECRTS 2006, pp. 193–202. IEEE (2006). https://doi.org/10.1109/ECRTS.2006.15

УДК 004.056:004.94

Хохлачова Ю. Є.
кандидат технічних наук, професор,
професор кафедри інженерії програмного забезпечення та кібербезпеки
Державний торговельно-економічний університет, м. Київ, Україна
e-mail: Y.Khokhlachova@knute.edu.ua;
ORCID: 0000-0002-0787-5112

ПРОГНОЗУВАННЯ МІЖСЕКТОРНИХ КАСКАДІВ У ЦИФРОВОМУ ДВІЙНИКУ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

Розглянуто прогнозування поширення порушень між енергетикою, водопостачанням, транспортом і цифровими сервісами смарт-регіону. Запропоновано стохастичну модель каскаду в цифровому двійнику, що враховує тип, затримку, поріг і резерв кожної залежності, а також ансамблевий прогноз Монте-Карло з квантильними інтервалами. У модельному експерименті на 160 синтетичних сценаріях середня абсолютна похибка прогнозу зменшилася до 2,47 об'єкта проти 4,39 для моделі незалежних секторів і 6,64 для парного статичного графа. Результат підтримує раннє узгодження дій операторів.

Ключові слова: критична інфраструктура; цифровий двійник; міжсекторні залежності; каскадні порушення; ймовірнісний прогноз.

Об'єкти критичної інфраструктури (КІ) смарт-регіону функціонують як система систем: відмова підстанції за кілька годин позначається на насосних станціях, центрах обробки даних і керуванні рухом, а збій платформи диспетчерського керування (SCADA) повертається в енергосистему. Настанови з інженерії кіберстійкості вимагають оцінювати не наявність захисних механізмів, а здатність зберігати критичні функції під час порушення [1]. Класичні результати теорії взаємозалежних мереж показують, що зв'язки між секторами різко знижують поріг катастрофічного каскаду порівняно з ізольованими мережами [2].

Сучасні підходи до моделювання каскадів ґрунтуються на спільному імітаційному моделюванні секторних моделей [3] і квазіреальночасовому розрахунку багатофазних відмов у мережах електро-, водо- й газопостачання [4]. Водночас параметри залежностей зазвичай відомі неточно, а систематичний огляд застосувань цифрових двійників у безпеці КІ фіксує брак методів, які видають прогноз розвитку інциденту разом із невизначеністю. Більшість моделей дає один «найімовірніший» сценарій без часових вікон для втручання та перевірки калібрування інтервалів.

Мета роботи — обґрунтувати метод прогнозування міжсекторного каскаду в цифровому двійнику, який видає траєкторію кількості уражених об'єктів з інтервалами невизначеності, і порівняти його з базовими моделями в обчислювальному експерименті.

Точність прогнозу за S сценаріями і H кроками горизонту оцінюється середньою абсолютною похибкою медіани та емпіричним покриттям інтервалу рівня $1-\beta$:

$$MAE = \frac{1}{SH}\sum_{s=1}^{S}\sum_{h=1}^{H}\left|\hat{q}_{0,5}^{s}(t_0+h) - N^s(t_0+h)\right|, \quad CR_{1-\beta} = \frac{1}{SH}\sum_{s=1}^{S}\sum_{h=1}^{H}\mathbf{1}\left[\hat{q}_{\beta/2}^{s} \le N^s \le \hat{q}_{1-\beta/2}^{s}\right], \qquad (1)$$

де $N^s(t)$ — фактична траєкторія сценарію s; для каліброваного прогнозу $CR_{1-\beta}$ має бути близьким до $1-\beta$.

Корисність прогнозу для операторів характеризує час попередження — інтервал між першою тривогою ковзного прогнозу та фактичним досягненням порога масштабного каскаду N^*:

$$\Delta_w = t_{cross} - t_{alarm}, \quad t_{alarm} = \min\left\{t: \hat{P}_t(N(T) \ge N^*) \ge 0{,}5\right\}, \qquad (2)$$

де t_{cross} — момент досягнення порога N^* фактичною траєкторією; $\hat{P}_t$ — частка реалізацій прогнозу, виданого в момент t, з перевищенням порога до кінця горизонту T; тривога в сценарії без перевищення порога є хибною.

Модельний (обчислювальний) експеримент виконано на синтетичних сценаріях з явними параметрами (скрипт Python, фіксоване зерно). Мережа містить 24 об’єкти чотирьох секторів (по шість) і 37 груп залежностей, з яких 21 — міжсекторні; крок часу 10 хв, горизонт 6 год. Номінальні параметри типів: фізичний вплив $p = 0{,}30$, $d = 1$, $b = 0{,}20$; кібернетичний 0,25, 0, 0,30; інформаційний 0,15, 1, 0,35; організаційний 0,10, 3, 0,20; автономні резерви r_j — від 0 до 4 кроків. У кожному з 160 сценаріїв «фактичні» параметри вибиралися з тих самих розподілів (логнормальне збурення p з $\sigma = 0{,}3$, зсув затримок і резервів на ±1 крок), початкова подія — відмова одного з шести об’єктів, 10 % спостережень стану пропускалися. Поріг $N^* = 10$ об’єктів досягнуто в 70 сценаріях. Прогноз видавався через 40 хв після початкової події ($M = 300$ реалізацій). Базові моделі: «незалежні сектори» (міжсекторні залежності відкинуто) та «парний граф» (усі зв’язки попарні, без затримок, резервів і порогів k_g).

На рис. 1 показано «кадри» каскаду в одному зі сценаріїв: відмова платформи диспетчерського керування за годину призводить до втрати двох підстанцій, далі уражаються розподільчі вузли енергосистеми, а через живлення та керування порушення переходить у водопостачання, транспорт і цифрові сервіси (18 об’єктів за 6 год).

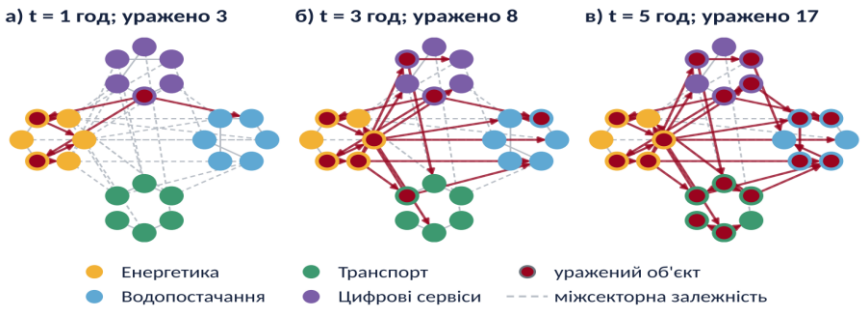


Рисунок 1 – Поширення каскаду в синтетичній мережі 24 об’єктів у моменти а) 1 год, б) 3 год, в) 5 год після відмови платформи диспетчерського керування; колір вершини — сектор, червоні вершини — уражені об’єкти, червоні стрілки — залежності від уражених джерел, пунктир — міжсекторні залежності

Двійник зменшив MAE на 44 % відносно моделі незалежних секторів і на 63 % відносно парного графа; бутстреп-інтервали не перетинаються. Модель незалежних секторів недооцінює каскад: покриття її інтервалів ≈ 31 %, а в 15,7 % масштабних каскадів тривога не надходить до перетину порога. Парний граф дає такий самий час попередження, як двійник, але ціною тривоги в кожному сценарії, де масштабного каскаду не було.

Для того самого сценарію рис. 2 подає віяловий прогноз двійника. Фактична траєкторія на всьому горизонті залишається в межах 80-відсоткової смуги, модель незалежних секторів «бачить» лише три уражені об’єкти, а парний граф переоцінює швидкість каскаду: його медіана досягає порога приблизно на 1,7 год раніше за фактичну траєкторію.

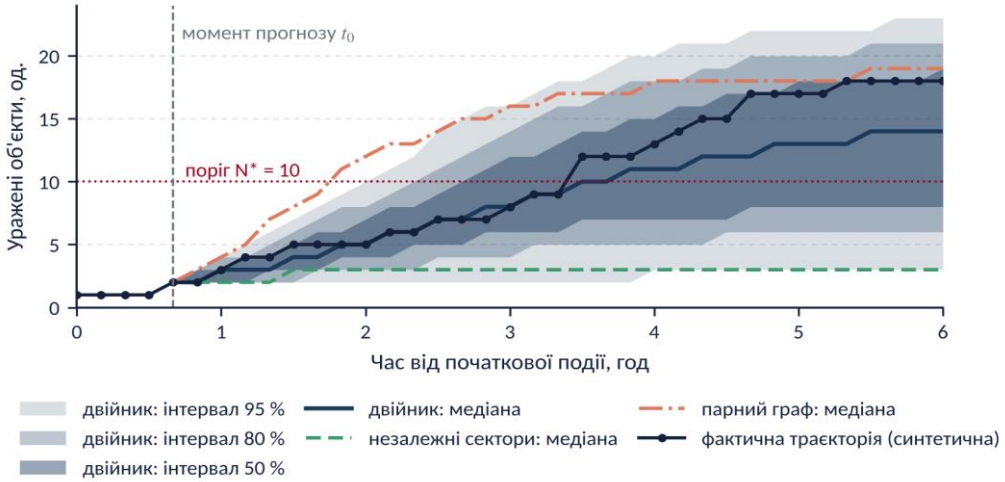


Рисунок 2 – Віяловий прогноз кількості уражених об’єктів, виданий у момент t_0 = 40 хв: смуги 50, 80 і 95 % та медіана цифрового двійника (2000 реалізацій Монте-Карло), медіани базових моделей і фактична синтетична траєкторія; пунктир — поріг масштабного каскаду N^*

Інтервали двійника консервативні (покриття 90,4 % для номінальних 80 %), тобто апріорні розподіли параметрів ширші за фактичну мінливість; їх слід калібрувати за ретроспективними інцидентами. Частка хибних тривог 43,3 % за порога ймовірності 0,5 є завеликою для автоматичного реагування, тому прогноз доцільно використовувати як підтримку рішень операторів із вибором порога за ціною пропуску й хибної тривоги. Результати отримано на одній синтетичній конфігурації без відновлення об’єктів і з апріорними розподілами, що збігаються з генератором сценаріїв; перенесення на реальні мережі потребує ідентифікації параметрів і аналізу чутливості.

Література

1. Ross, R., Pillitteri, V., Graubart, R., Bodeau, D., McQuaid, R.: Developing Cyber-Resilient Systems: A Systems Security Engineering Approach. NIST SP 800-160 Vol. 2 Rev. 1, Gaithersburg, MD (2021). https://doi.org/10.6028/NIST.SP.800-160v2r1

2. Buldyrev, S.V., Parshani, R., Paul, G., Stanley, H.E., Havlin, S.: Catastrophic Cascade of Failures in Interdependent Networks. Nature 464(7291), 1025–1028 (2010). https://doi.org/10.1038/nature08932

3. Wang, F., Magoua, J.J., Li, N.: Modeling Cascading Failure of Interdependent Critical Infrastructure Systems Using HLA-Based Co-Simulation. Automation in Construction 133, 104008 (2022). https://doi.org/10.1016/j.autcon.2021.104008

4. Momeni, A., Dharmawardena, H., Prasad, V., Piratla, K.R., Venayagamoorthy, G.K.: Modeling Near-Real-Time Multiphase Cascading Failures Across Interdependent Power, Water, and Gas Infrastructure. Structure and Infrastructure Engineering 22(1), 90–106 (2024). https://doi.org/10.1080/15732479.2024.2312267

Поремський Я.С.
студент групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ПОДІЙ КІБЕРБЕЗПЕКИ НА ОСНОВІ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ У ПРОЦЕСАХ THREAT HUNTING

Зростання обсягу та складності подій кібербезпеки ускладнює пошук прихованих загроз і потребує вдосконалення аналітичних підходів у процесах Threat Hunting. Великі мовні моделі розширюють можливості аналізу, забезпечуючи контекстну інтерпретацію подій, установлення взаємозв'язків і формування обґрунтованих гіпотез щодо потенційної зловмисної активності. Поєднання цих можливостей із даними платформи Wazuh може прискорити дослідження підозрілої активності та зменшити навантаження на аналітика за умови перевірки висновків за первинними подіями, контролю хибних припущень і збереження вирішальної ролі фахівця в оцінюванні загроз.

Ключові слова: Threat Hunting, кібербезпека, великі мовні моделі (LLM), аналіз подій безпеки, виявлення кіберзагроз.

У процесах Threat Hunting особливу складність становить виявлення підозрілої активності, яка не відповідає заздалегідь визначеним правилам та індикаторам компрометації. Традиційні засоби моніторингу переважно орієнтовані на виявлення відомих ознак атак, тоді як складні загрози можуть проявлятися через сукупність окремих, на перший погляд легітимних дій. Це зумовлює необхідність застосування підходів, здатних не лише виявляти відхилення, а й підтримувати формування та перевірку гіпотез щодо потенційних кіберзагроз.

Threat Hunting є проактивним ітеративним пошуком загроз шляхом формування та перевірки гіпотез. Він доповнює автоматичне виявлення аналізом активності, що не відповідає наперед заданим правилам. У матеріалах SANS наголошено на ролі гіпотез і поєднанні автоматизованих інструментів із експертною оцінкою [1]. Отже, технологія підтримки Threat Hunting має не лише знаходити підозрілі ознаки, а й надавати аналітику перевірювані підстави для подальшого дослідження.

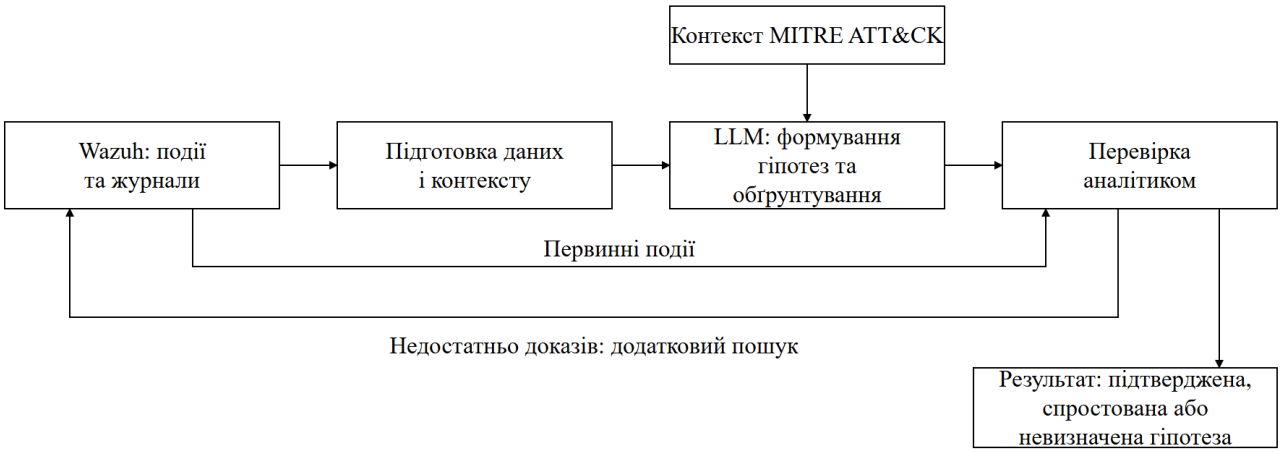
Наприклад, серія невдалих спроб входу може свідчити як про підбір пароля, так і про помилки легітимного користувача. Для розмежування потрібні час і частота спроб, адреса джерела, пов'язаний обліковий запис, успішні входи та подальша активність. Неповна телеметрія або неузгоджені часові позначки ускладнюють перевірку. Тому обробка подій повинна зберігати контекст і зв'язок із вихідними журналами.

Технологія інтелектуального аналізу передбачає послідовні етапи: отримання подій, нормалізацію і групування, формування контекстного запиту до LLM, побудову гіпотез та їх перевірку аналітиком. Модель використовується як засіб інтерпретації підготовленої телеметрії, а не як автономний механізм підтвердження інциденту. Результатом її роботи є пояснення можливого сценарію та перелік необхідних перевірок.

На першому етапі Wazuh забезпечує надходження та пошук подій із контрольованих вузлів [2]. Із журналів для аналізу відбираються час події,

джерело, користувач, вузол, процес, тип активності та ідентифікатор запису. Далі здійснюються уніфікація часових позначок, відбір релевантних полів і групування за спільними сутностями у визначеному часовому вікні. Окремо фіксуються пропуски даних; повторювані спроби входу не вилучаються як технічні дублікати, оскільки їх частота має діагностичне значення.

На наступному етапі сформований контекст передається LLM разом зі структурованою інструкцією: визначити підозрілі ознаки, запропонувати



гіпотезу, навести ідентифікатори подій як докази, описати альтернативне легітимне пояснення та перелічити перевірки. Для уніфікації інтерпретації застосовується MITRE ATT&CK [3]. Модель може запропонувати відповідну тактику або техніку, однак це зіставлення підлягає підтвердженню за журналами. Аналітик перевіряє факти, уточнює запит або ініціює додатковий пошук.

Рис. 1. Схема інтелектуального аналізу подій у процесах Threat Hunting

Перевірка гіпотези є окремим етапом технології. Аналітик зіставляє кожне твердження LLM із первинними подіями Wazuh, перевіряє послідовність дій та альтернативні пояснення, після чого позначає гіпотезу як підтверджену, спростовану або недостатньо обґрунтовану. За браку доказів процес повертається до отримання додаткової телеметрії. Це важливо з огляду на ризик правдоподібних, але помилкових відповідей генеративного ШІ, відзначений NIST [4].

Для відтворюваності аналізу необхідно зберігати ідентифікатори вихідних подій, межі часового вікна, версію моделі, текст інструкції та рішення аналітика. Уніфікований формат виходу має відокремлювати встановлені факти від припущень, містити оцінку недостатності доказів і рекомендовані кроки перевірки. Такий підхід дозволяє простежити шлях від висновку моделі до конкретного запису журналу.

Практична реалізація потребує контролю якості телеметрії: узгодження часу, налаштування аудиту, перевірки повноти полів і збереження значущих повторюваних дій. Інструкції для LLM необхідно відокремлювати від тексту журналів, які можуть містити недовірений вміст. Не слід надавати моделі право самостійно змінювати правила захисту або виконувати дії у виробничій інфраструктурі.

Під час використання зовнішнього сервісу LLM слід мінімізувати передані поля та знеособлювати конфіденційні відомості, якщо це не позбавляє дані аналітичного значення. Доступ до журналів, результатів моделі та історії перевірок має відповідати повноваженням фахівців. Механізм контролю людиною залишається обов'язковим для висновків щодо потенційної атаки.

Результативність технології доцільно оцінювати на сценаріях із відомою послідовністю зловмисних і легітимних дій. Для однакових наборів подій порівнюються аналіз із підтримкою LLM і без неї. Критеріями є частка правильно визначених сценаріїв, кількість хибних гіпотез, коректність зіставлення з ATT&CK та час перевірки аналітиком. Це дає змогу оцінити корисність технології без ототожнення сформованої гіпотези з фактом атаки.

Окремим ризиком впровадження технології є залежність якості сформованих гіпотез від повноти вхідних даних і коректності їх інтерпретації мовною моделлю. Для підвищення надійності аналізу доцільно застосовувати стандартизовані інструкції, перевіряти відповідність висновків первинним подіям та регулярно оцінювати результати на контрольних сценаріях. Це дозволить своєчасно виявляти обмеження технології та вдосконалювати її застосування.

Подальша перевірка підходу на сценаріях різних типів загроз дозволить визначити, за яких умов контекстне групування й структуровані відповіді LLM дійсно полегшують роботу аналітика та де виникають обмеження.

Додатковим напрямом практичного використання результатів Threat Hunting є накопичення перевірених гіпотез і відповідних послідовностей подій у внутрішній базі знань. Для кожного випадку доцільно фіксувати ознаки активності, виконані перевірки та підстави остаточного рішення. Такі матеріали можуть використовуватися для повторного дослідження схожих ситуацій, навчання аналітиків і вдосконалення правил моніторингу.

Таким чином, розглянута технологія поєднує Wazuh як джерело телеметрії, підготовку контексту, LLM для формування перевірюваних гіпотез і експертну верифікацію. Її практична цінність залежить від якості журналів, простежуваності доказів і контролю людиною.

Перелік посилань:

1. Lee R. M., Bianco D. Generating Hypotheses for Successful Threat Hunting. SANS Institute, 2016. URL: https://www.sans.org/white-papers/37172 (дата звернення: 13.10.2026).
2. Wazuh. Threat hunting. URL: https://documentation.wazuh.com/current/getting-started/use-cases/threat-hunting.html (дата звернення: 13.10.2026).
3. MITRE. MITRE ATT&CK. URL: https://attack.mitre.org/ (дата звернення: 13.10.2026).
4. NIST. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1, 2024. URL: https://doi.org/10.6028/NIST.AI.600-1 (дата звернення: 13.10.2026).

УДК 004.056.5:004.7

Прокопович-Ткаченко Д. І.
кандидат технічних наук, доцент,
завідувач кафедри кібербезпеки та інформаційних технологій
Університет митної справи та фінансів, м. Дніпро, Україна
omega2417@umsf.edu.ua
ORCID 0000-0002-6590-3898

МІСІЙНО ОРІЄНТОВАНА МЕТРИКА КІБЕРСТІЙКОСТІ МЕРЕЖІ З УРАХУВАННЯМ НЕБЕЗПЕЧНОЇ ДЕГРАДАЦІЇ СЕРВІСІВ

Середня доступність і середня затримка не розрізняють інциденти, у яких критичний сервіс формально працює, але тривалий час перебуває нижче допустимої якості. Запропоновано метрику кіберстійкості на основі часової кривої місійної якості, площі втрат зі штрафом за небезпечну деградацію та обмеження умовної величини під ризиком (CVaR) хвоста втрат. У модельному експерименті на 4000 синтетичних сценаріях обмеження CVaR зменшило хвостові втрати на 28,9 % та час критичного сервісу в небезпечній зоні з 28,1 до 13,4 с ціною зростання середніх втрат на 8,6 %.
Ключові слова: кіберстійкість, місійна якість сервісу, площа втрат, небезпечна деградація, CVaR

Кіберстійкість інформаційно-комунікаційної мережі визначається не відсутністю інцидентів, а здатністю зберігати комунікаційну функцію під час впливу та повертати її до прийнятного рівня. Без кількісного виміру порівняння захисних рішень зводиться до переліку механізмів [1]. Поширені показники — статична доступність, середня затримка, кількість відхилених пакетів — описують стан мережі, а не наслідки для місії: два інциденти з однаковою кількістю втрачених пакетів можуть мати різний вплив на дотримання строків доставки команд і на цілісність даних.

Сучасні підходи вимірюють кіберстійкість через криву продуктивності під час атаки та площу під нею [2], а також нормують цю площу відносно цільового часу відновлення [3]. Інженерні настанови з побудови кіберстійких систем наголошують на збереженні критичних функцій у деградованому режимі [4]. Проте інтегральні площі та їх середні значення допускають компенсацію: добрий результат допоміжних сервісів «урівноважує» провал критичного, а рідкісні, але тяжкі каскадні сценарії розчиняються в середньому.

Комбінований індекс для порівняння конфігурацій захисту π за однакового ресурсного бюджету поєднує середні втрати з обмеженням хвоста (1):

$$R_{\Sigma}(\pi) = 1 - E[\Lambda_{\pi}] - \lambda \cdot \max(0, \mathrm{CVaR}_{\alpha}(\Lambda_{\pi}) - \Lambda_{\mathrm{adm}}), \qquad (1)$$

де Λ_{adm} — допустимий рівень хвостових втрат місії; λ — вага штрафу за його перевищення. Допустимою вважається конфігурація з $\mathrm{CVaR}_{\alpha} \le \Lambda_{\mathrm{adm}}$, а найкращою — та, що максимізує R_{Σ}.

Виконано модельний (обчислювальний) експеримент на синтетичних сценаріях (без польових даних). Розглянуто три сервіси: командно-керуючий канал S1 ($w = 0{,}5$, $\kappa = 1{,}0$), телеметрію S2 (0,3; 0,5) і масовий обмін даними S3 (0,2; 0,2), поріг $\theta = 0{,}5$, $T = 600$ с, крок 10 с. Згенеровано 4000 сценаріїв п’яти

узагальнених класів (завади на L1, підміна маршрутів, перевантаження, компрометація контролера, викрадення сеансу) з логнормальною тяжкістю та ймовірністю каскаду на S1, що зростає з тяжкістю. Порівнювалися на тих самих сценаріях (спільні випадкові числа): C0 — незалежний локальний захист рівнів; C1 — централізована оптимізація за середньою затримкою; C2 — міжрівнева координація з мінімізацією $E[\Lambda]$; C3 — міжрівнева координація з обмеженням CVaR, яка резервує ресурс для S1 коштом S2 і S3. Параметри: $\alpha = 0{,}95$, $\Lambda_{\mathrm{adm}} = 0{,}18$, $\lambda = 1$.

На рис. 1 (сценарій «завади на L1», C2) усі сервіси формально доступні, проте S1 перебуває в небезпечній зоні 190 с, а S2 і S3 — по 20 с; саме це, а не кількість втрачених пакетів, визначає внесок сценарію в Λ.

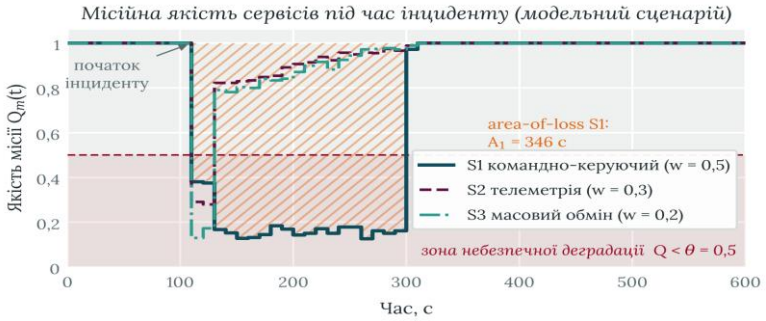


Рис. 1. Ступінчасті криві місійної якості $Q_m(t)$ трьох сервісів різної критичності в модельному сценарії «завади на L1» для конфігурації C2. Червона смуга — зона небезпечної деградації ($Q < 0{,}5$); штрихування — площа втрат критичного сервісу S1 ($A_1 = 346$ с з урахуванням штрафу κ_1)

Зведені показники наведено в табл. 1. Статична доступність (частка часу з $Q_m \geq 0{,}1$) змінюється лише в межах 98,97–99,65 %, тоді як середні втрати місії відрізняються у 2,4 раза, а CVaR — у 2,5 раза; рангова кореляція Спірмена між недоступністю та Λ за сценаріями становить лише 0,51.

Таблиця 1 — Показники конфігурацій захисту за 4000 модельними сценаріями (однаковий ресурсний бюджет)

Конфігурація	**Доступність, %**	$E[\Lambda]$	**$\mathrm{CVaR}_{0{,}95}$**	**Час S1 при Q < θ, с (сер. / p95)**	R_Σ
C0 — локальний	98,97	0,110	0,411	55,0 / 270	0,659
C1 — сер. затримка	99,25	0,071	0,300	41,1 / 220	0,809
C2 — мін. $E[\Lambda]$	99,62	**0,045**	0,236	28,1 / 170	0,899
C3 — обмеж. CVaR	**99,65**	0,049	**0,168**	**13,4 / 70**	**0,951**

Конфігурація C2 має найменші середні втрати, але її хвіст перевищує допустимий рівень; C3 — єдина допустима ($\mathrm{CVaR}_{0{,}95} = 0{,}168 \leq 0{,}18$). Парний бутстреп (2000 повторів) дає для різниці C3 – C2: $\Delta E[\Lambda] = +0{,}0039$ (95 % ДІ 0,0027–0,0050, тобто +8,6 %) і $\Delta\mathrm{CVaR} = -0{,}068$ (95 % ДІ від –0,078 до –0,057, тобто –28,9 %). Отже, обмеження хвоста має реальну ціну: середня якість допоміжних сервісів S2 і S3 у C3 нижча, ніж у C2. Розподіли втрат на рис. 2 показують, чому вибір за середнім вводить в оману: середні майже збігаються, а маса хвоста за VaR у C2 значно важча.

Ранжування за R_Σ стійке: С3 посідає перше місце в 100 % із 1000 бутстреп-вибірок сценаріїв, тоді як за середніми втратами першою стабільно є С2. Метрика змінює рішення саме там, де середнє приховує провал критичного сервісу, і не винагороджує кількість механізмів: С1 з додатковою централізованою оптимізацією програє С2 і С3 за всіма місійними показниками. Обмеження результату: параметри сценаріїв, ваги w_m, штрафи κ_m і поріг θ задано експертно; висновки чутливі до Λ_{adm} і мають бути перевірені аналізом чутливості та на трасах реального комунікаційного стеку.

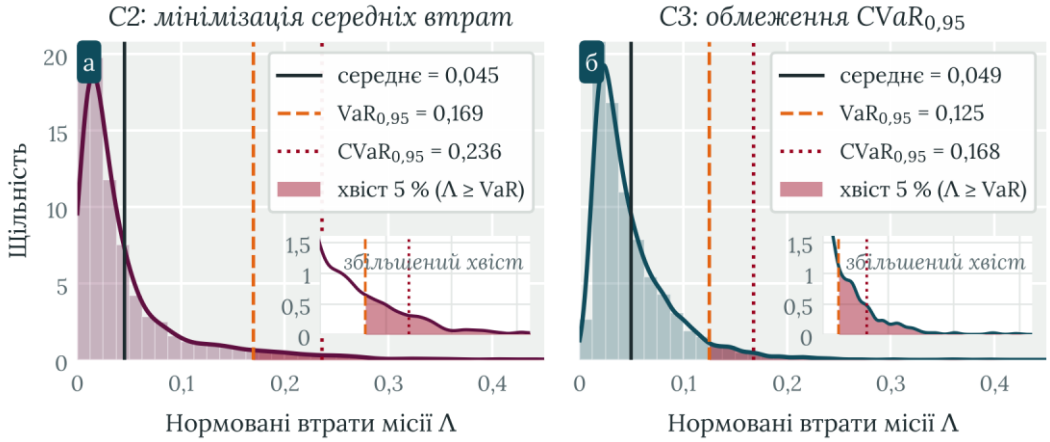


Рис. 2. Гістограма та ядерна оцінка щільності нормованих втрат місії Λ за 4000 модельними сценаріями: (а) конфігурація С2, що мінімізує середні втрати; (б) конфігурація С3 з обмеженням $CVaR_{0,95}$. Вертикальні лінії — середнє, $VaR_{0,95}$ і $CVaR_{0,95}$; червона заливка — 5 % найгірших сценаріїв, на вставці — збільшений фрагмент хвоста

Запропоновано місійно орієнтовану метрику кіберстійкості, що поєднує часову криву якості сервісів, площу втрат зі штрафом за небезпечну деградацію та обмеження CVaR хвоста розподілу сценаріїв (вирази (1)–(4)). У модельному експерименті статична доступність не розрізнила конфігурації захисту (розкид 0,7 в. п.), тоді як запропонований індекс виявив, що мінімізація середніх втрат лишає неприйнятний хвіст, а обмеження CVaR зменшує його на 28,9 % і вдвічі скорочує час критичного сервісу в небезпечній зоні за зростання середніх втрат на 8,6 %. Подальші кроки — інтервальне врахування невизначеності телеметрії, калібрування порогів за вимогами сервісів і перевірка на емуляційному стенді.

Список джерел

[1] Kott, A., Linkov, I.: To Improve Cyber Resilience, Measure It. Computer 54(2), 80–85 (2021). https://doi.org/10.1109/MC.2020.3038411

[2] Weisman, M.J., Kott, A., Ellis, J.E., Murphy, B.J., Parker, T.W., Smith, S., Vandekerckhove, J.: Quantitative Measurement of Cyber Resilience: Modeling and Experimentation. ACM Transactions on Cyber-Physical Systems 9(1), 1–25 (2025). https://doi.org/10.1145/3703159

[3] Cho, H., Sung, J.-H., Kang, H.-J., Jang, J., Shin, D.: Quantifying Cyber Resilience: A Framework Based on Availability Metrics and AUC-Based Normalization. Electronics 14(12), 2465 (2025). https://doi.org/10.3390/electronics14122465

[4] National Institute of Standards and Technology: Developing Cyber-Resilient Systems: A Systems Security Engineering Approach. NIST SP 800-160 Vol. 2 Rev. 1, Gaithersburg, MD (2021). https://doi.org/10.6028/NIST.SP.800-160v2r1

Чайківський В.В.
студент групи ІКБм-61, ННІКБЗІ ДУІКТ,
Київ, Україна

АВТОМАТИЗАЦІЯ ПРОЦЕСІВ ВИЯВЛЕННЯ ТА УСУНЕННЯ ВРАЗЛИВОСТЕЙ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ КІНЦЕВИХ ПРИСТРОЇВ КОРПОРАТИВНОЇ МЕРЕЖІ ЗАСОБАМИ CHECK POINT ENDPOINT SECURITY

Розглянуто автоматизацію виявлення та усунення відомих вразливостей застосунків на кінцевих пристроях корпоративної мережі засобами Check Point Endpoint Security. На основі документації виробника та рекомендацій NIST запропоновано процес, який поєднує регулярне сканування, визначення пріоритетів, пілотне встановлення виправлень, масове розгортання та перевірку результату. Визначено технічні обмеження автоматизації й показники оцінювання її результативності. Обґрунтовано необхідність контролю винятків і повторного сканування для підтвердження усунення вразливостей. Запропонований підхід орієнтований на зменшення часу перебування пристроїв у вразливому стані за збереження працездатності корпоративних застосунків.

Ключові слова: вразливості програмного забезпечення, кінцеві пристрої, корпоративна мережа, автоматизація, керування виправленнями, Check Point Endpoint Security, Posture Management, CVE, CVSS

Кінцеві пристрої корпоративної мережі мають встановлене різнопланове ПЗ, оновлення якого залежить від доступності пристрою, дій користувача та сумісності з робочими процесами. У разі ручного адміністрування між появою виправлення та його застосуванням (встановленням) може зберігатися тривалий період використання вразливої версії ПЗ. Тому завдання автоматизації полягає не лише у виявленні застарілого застосунку, а й у керованому доведенні процесу до підтвердженого усунення вразливості.

Метою роботи є обґрунтування послідовності автоматизованого виявлення та усунення відомих вразливостей застосунків засобами Check Point Endpoint Security. Методом дослідження є аналіз документації та моделювання процесу розгортання виправлень. За основу взято підхід NIST SP 800-40 Rev. 4, який охоплює ідентифікацію, пріоритизацію, отримання, встановлення та перевірку виправлень [1].

У роботі розглядаються Windows-пристрої з клієнтом Check Point Endpoint Security E89.26 і керуванням через хмарну консоль. Станом на 17.10.2026 це останній реліз Windows-клієнта, зазначений в офіційних Release Notes, дата випуску якого – 16.09.2026 [2]. Версія клієнта має позначення E, тоді як позначення R стосується локальної системи керування. Оновлення самого агента слід відокремлювати від виправлення вразливостей сторонніх застосунків через Posture Management.

У Check Point Endpoint Security відповідні можливості реалізовано засобами Posture Management. Сканування підтримуваних застосунків дає змогу ви-являти відомі CVE та відображати оцінки CVSS у консолі керування [3]. Виявлення CVE слід розглядати як підставу для подальшого опрацювання: воно

саме по собі не підтверджує наявність доступного автоматичного виправлення. Межі застосовності необхідно перевіряти за переліком підтримуваного ПЗ.

Налаштування виконують у Policy → Access & Compliance → Compliance & Posture → Posture Assessment Settings. Адміністратор активує оцінювання вразливостей, задає розклад сканування та дозволяє встановлення виправлень. Автоматичне керування оновленнями вмикається окремо через Advanced Settings; доступні вибір застосунків, рівнів критичності та винятків. Актуальний посібник визначає E88.20 як мінімальну версію клієнта для цієї функції – це поріг підтримки, а не рекомендована версія розгортання. Зміни потрібно зберегти та застосувати командою Install Policy [4].

Пропонується організувати процес як цикл зі зворотним зв'язком (Рис. 1). Спочатку визначають охоплення керованих пристроїв і актуальність результатів сканування. Далі адміністратор зіставляє технічну критичність CVE зі значущістю активу та умовами його використання. Високе значення CVSS є підставою для уваги, однак остаточний пріоритет доцільно уточнювати з урахуванням доступності вразливого компонента для атак і наслідків порушення роботи при-строю.

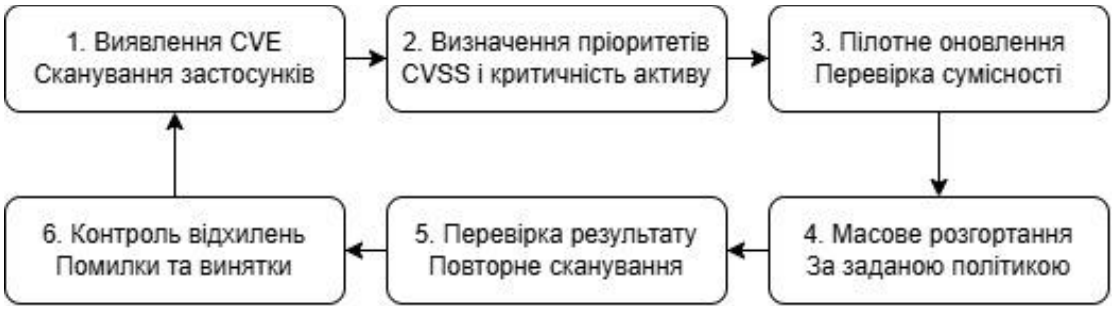


Рис. 1. Запропонований цикл керування вразливостями застосунків

Перед масовим розгортанням доцільно створити окрему пілотну групу з робочими станціями. На ній перевіряють запуск застосунків, виконання типових операцій і потребу в перезавантаженні. Після прийнятного результату налаштування поширюють поетапно. Така організаційна процедура доповнює штатну автоматизацію продукту й зменшує ймовірність одночасного порушення роботи значної кількості користувачів. Область дії правила політики необхідно перевіряти перед кожним розширенням охоплення.

Взаємодію з користувачем слід узгодити з регламентом обслуговування. Політика дає змогу задавати параметри відкладення та примусового застосування оновлень [4]. Організація має визначити прийнятний строк усунення для різних категорій ризику й порядок повідомлення працівників. Винятки для критичних бізнес-застосунків доцільно оформлювати із зазначенням відповідальної особи, причини, строку перегляду та тимчасових заходів захисту. Безстрокове виключення застосунку з оновлення створює додаткові ризики.

Автоматизація має технічні межі. Документація вказує на застосування виправлень для підтримуваних застосунків, установлених у C:\Program Files, та

обмеження розміру пакета менш ніж 2 ГБ. Кінцевим пристроям потрібен доступ до сайтів постачальників для завантаження виправлень [3]. Отже, наявність агента не гарантує оновлення всього ПЗ. Непідтримувані застосунки, недоступні пакети й пристрої без мережного доступу до відповідних ресурсів потребують окремого опрацювання.

Завершальним етапом є повторне сканування, яке виробник рекомендує для перевірки усунення CVE; за потреби виконують перезавантаження [3]. Для контролю доцільно враховувати пару «пристрій – CVE», оскільки одна вразливість може бути присутня на багатьох станціях. Запропоновані показники: медіанний час від виявлення до підтвердженого усунення, частка випадків, усунених у встановлений строк, частка пристроїв з актуальним скануванням і кількість прострочених винятків. Порівнювати їх слід для однакового складу пристроїв і періоду спостереження.

Для коректного оцінювання частку своєчасного усунення пропонується визначати серед випадків, граничний строк яких уже настав. Із розрахунку не слід вилучати невдалі оновлення лише через відсутність успішного результату. Натомість непідтримуване ПЗ потрібно відображати окремо як прогалину охоплення. Це дасть змогу розрізняти результативність автоматизації в межах підтримуваних застосунків і фактичний стан усієї корпоративної мережі. Очікуване скорочення строків усунення має бути підтверджене вимірюваннями, а не самим фактом увімкнення автоматичного встановлення.

Отже, Check Point Endpoint Security надає засоби для побудови керованого циклу виявлення та виправлення відомих вразливостей підтримуваних застосунків. Запропонований підхід доповнює штатні функції пріоритизацією за контекстом активу, пілотним розгортанням, контролем винятків і вимірюванням результату. Практична цінність полягає у визначенні критеріїв завершення усунення вразливості та умов безпечного розширення автоматизації. Перспективою подальшої роботи є експериментальна перевірка зміни часу усунення вразливостей у пілотному корпоративному середовищі.

Перелік використаних джерел

1. Souppaya M., Scarfone K. Guide to Enterprise Patch Management Planning: Preventive Maintenance for Technology. NIST SP 800-40 Rev. 4. Gaithersburg : NIST, 2022. DOI: https://doi.org/10.6028/NIST.SP.800-40r4.

2. Check Point Software Technologies Ltd. E89.x Endpoint Security Clients for Windows Release Notes. What’s New. URL: https://sc1.checkpoint.com/documents/E89.x/EN/Endpoint_Security_Clients_for_Windows_RN/Content/Topics-Endpoint_Security_Client_for_Windows_RN/What-is-New.htm (дата звернення: 17.10.2026).

3. Check Point Software Technologies Ltd. Check Point Endpoint Security Administration Guide. Розділи Viewing Endpoint Posture, Mitigating Vulnerable CVEs. URL: https://sc1.checkpoint.com/documents/Infinity_Portal/WebAdminGuides/EN/Harmony-Endpoint-Admin-Guide/index.html (дата звернення: 17.10.2026).

4. Check Point Software Technologies Ltd. Detecting Common Vulnerabilities and Exposures. Check Point Endpoint Security Administration Guide. URL: https://sc1.checkpoint.com/documents/Infinity_Portal/WebAdminGuides/EN/Harmony-Endpoint-Admin-Guide/Endpoint-Security/Admin-Guide/Topics/compliance/detecting-common-vulnerabilities-and-exposures.html (дата звернення: 17.10.2026).

Таран В. Д.
студент групи БСДМ-62 ННІКБЗІ ДУІКТ
Київ, Україна

ОСОБЛИВОСТІ ПОБУДОВИ SECURITY OPERATIONS CENTER ДЛЯ ПІДПРИЄМСТВ ІЗ ТЕРИТОРІАЛЬНО РОЗПОДІЛЕНОЮ ІНФРАСТРУКТУРОЮ

Розглянуто організаційні та технічні особливості побудови центру операцій безпеки (SOC) для підприємств із центральним офісом і віддаленими об’єктами. Визначено проблеми фрагментації телеметрії, неоднорідності обладнання та залежності від каналів зв’язку. Запропоновано багаторівневу модель збирання, передавання, аналізу й кореляції подій кібербезпеки з розподілом функцій між локальними засобами контролю та центральною командою SOC. Окреслено сценарії виявлення інцидентів, підходи до пріоритизації сповіщень і показники оцінювання ефективності запропонованої архітектури.

Ключові слова: SOC, розподілена інфраструктура, кореляція подій, SIEM, EDR, мережевий моніторинг, кіберінцидент.

Поширення територіально розподілених інформаційних систем ускладнює забезпечення єдиного рівня кіберзахисту підприємства. Центральний офіс, філії, склади, виробничі майданчики та інші віддалені об’єкти можуть використовувати різні мережеві пристрої, канали зв’язку, серверні платформи й засоби захисту кінцевих пристроїв. За таких умов локальний контроль окремих компонентів не забезпечує цілісного уявлення про розвиток кіберінциденту. Потреба в безперервному виявленні та аналізі загроз узгоджується з функціями Detect, Respond і Recover концепції NIST Cybersecurity Framework 2.0 [1, с. 4-5].

Метою роботи є визначення основних принципів побудови Security Operations Center для організацій із територіально розподіленою інфраструктурою та обґрунтування архітектури, що забезпечує централізоване спостереження, пріоритизацію подій і кероване реагування. Методичну основу становлять функціональна декомпозиція процесів SOC, аналіз потоків телеметрії та моделювання типових сценаріїв виявлення інцидентів. На відміну від підходу, орієнтованого лише на доступність вузлів, запропонована модель поєднує експлуатаційні метрики з контекстом інформаційної безпеки.

Основна проблема полягає в нерівномірності видимості подій. На віддаленому майданчику може бути доступною лише інформація про стан маршрутизатора та каналу VPN, тоді як у центральному офісі збираються журнали міжмережевого екрана, автентифікації користувачів і сповіщення EDR/XDR. Відсутність спільної часової шкали та прив’язки до конкретного активу ускладнює розслідування. Наприклад, короткочасне зростання вихідного трафіку саме по собі не доводить витік даних, але одночасна поява нетипового процесу на робочій станції, звернень до невідомого ресурсу й істотного збільшення обсягу передавання потребує перевірки як взаємопов’язаний набір ознак.

Для подолання зазначених обмежень пропонується трирівнева архітектура SOC. Перший рівень - джерела та локальні засоби спостереження: NGFW, маршрутизатори, комутатори, сервери, служби ідентифікації, EDR/XDR, а також критичні IoT-пристрої. Другий рівень - захищене транспортування й накопичення даних, зокрема журнали подій, мережеві потоки, показники

продуктивності та статуси доступності. Третій рівень - центральна аналітика, де здійснюються нормалізація, кореляція, оцінювання ризику, відображення стану інфраструктури та формування інцидентів для аналітиків SOC. Такий поділ дозволяє поетапно підключати нові філії без зміни базової логіки роботи центру.

На рівні збору інформації доцільно розмежовувати два класи даних. Перший охоплює метрики доступності та продуктивності: стан вузлів, завантаження процесора, використання пам’яті, трафік інтерфейсів і стабільність VPN. Для них можуть застосовуватися Zabbix, Prometheus, SNMPv3 та візуалізація у Grafana. Другий клас утворюють події безпеки: заборонені з’єднання, спрацювання IPS, невдалі входи, зміни привілеїв, запуски підозрілих процесів і результати EDR-аналізу. Їх обробку слід організовувати у SIEM або еквівалентному механізмі журналювання та кореляції. Платформа метрик не є повноцінною заміною SIEM, однак доповнює її даними про операційний контекст.

Особливу увагу слід приділити передаванню інформації з філій. Обмін має відбуватися через захищені канали з автентифікацією джерел, розмежуванням доступу та мінімально необхідними мережевими дозволами. Для журналів критичних компонентів доцільно передбачити локальну буферизацію на випадок тимчасової втрати зв’язку, обмеження обсягу телеметрії та контроль доставки. Синхронізація часу всіх вузлів є необхідною для коректної кореляції. Водночас відсутність зв’язку з майданчиком не повинна автоматично класифікуватися як кібератака: така подія має окремо перевірятися на предмет технічної несправності, аварії каналу або навмисного втручання.

Аналітичне ядро SOC доцільно будувати за принципом контекстної кореляції. Для кожного активу формується профіль, що містить його розташування, власника, призначення, критичність, допустимі мережеві взаємодії та доступні джерела даних. Виявлена подія зіставляється з цим профілем, іншими сповіщеннями та попередньою активністю в межах визначеного часового інтервалу. Підвищення пріоритету пропонується здійснювати за поєднанням трьох умов: критичності активу, сили ознак компрометації та їх підтвердження незалежними джерелами. Такий підхід має зменшувати навантаження на аналітиків порівняно з обробкою всіх ізольованих спрацювань в однаковому порядку; величину ефекту необхідно визначати експериментально.

Практичним прикладом є сценарій можливої компрометації робочої станції у віддаленій філії. EDR реєструє аномальний процес, міжмережевий екран - серію нетипових зовнішніх з’єднань, а система моніторингу — збільшення вихідного трафіку. Окремі сигнали можуть мати низьку достовірність, проте їх сукупність стає підставою для створення одного інциденту з посиланнями на первинні докази. Для систематизації перевірок доцільно використовувати тактики та техніки MITRE ATT&CK, зокрема категорії Command and Control та Exfiltration [4]. Після підтвердження загрози можуть виконуватися ізоляція вузла, блокування з’єднання, збереження артефактів і документування рішень згідно з погодженими повноваженнями.

Організаційна складова має враховувати обмежений штат і різнорідність віддалених локацій. Фахівець першої лінії перевіряє сповіщення, виконує первинну класифікацію та реєструє інцидент; друга лінія проводить поглиблений аналіз і координує локалізацію; відповідальні адміністратори забезпечують технічні дії на відповідних об’єктах. Важливо визначити порядок ескалації, контактних осіб на майданчиках, правила роботи поза робочим часом та процедури збереження доказів. Рекомендації NIST SP 800-61 Rev. 3 підкреслюють значення безперервного моніторингу, аналізу несприятливих подій і настроювання засобів виявлення для скорочення хибних результатів [2, с. 23-25].

Запровадження запропонованої архітектури доцільно здійснювати поетапно. Спочатку формується перелік активів і критичних бізнес-сервісів, визначаються мінімальні вимоги до журналювання та єдиний формат ідентифікації об’єктів. Наступним кроком є підключення джерел із найбільшим впливом на безпеку: міжмережевих екранів, сервісів автентифікації, EDR і ключових серверів. Після цього створюються сценарії кореляції, перевіряється повнота телеметрії та запроваджуються процедури реагування. Нарощування покриття має ґрунтуватися на ризику, а не лише на кількості підключених пристроїв; саме на це орієнтує підхід до безперервного моніторингу NIST SP 800-137 [3].

Оцінювання результативності моделі пропонується проводити за контрольованими сценаріями: мережеве сканування, аномальні спроби входу, збільшення вихідного трафіку та недоступність вузла. Критеріями є частка критичних активів із повною телеметрією, середній час виявлення (MTTD), середній час реагування (MTTR), частка хибнопозитивних спрацювань і повнота доставки подій після відновлення зв’язку. Показники слід порівнювати до і після впровадження кореляційних правил за однакових умов. Кількісні результати потребують окремого експериментального дослідження.

Отже, результативність SOC для територіально розподіленого підприємства залежить від повноти видимості активів, надійності доставки телеметрії, кореляції різнорідних подій та визначеності повноважень персоналу. Запропонована трирівнева модель поєднує локальний контроль із централізованою аналітикою, враховує критичність об’єктів і створює основу для автоматизації реагування. Подальші дослідження доцільно спрямувати на перевірку кореляційних сценаріїв та оптимізацію порогів спрацювання для різних типів віддалених майданчиків.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. NIST. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. 2024. DOI: 10.6028/NIST.CSWP.29.
2. Nelson A. et al. Incident Response Recommendations and Considerations for Cybersecurity Risk Management. NIST SP 800-61 Rev. 3. 2025. DOI: 10.6028/NIST.SP.800-61r3.
3. Dempsey K. et al. Information Security Continuous Monitoring (ISCM). NIST SP 800-137. 2011. DOI: 10.6028/NIST.SP.800-137.
4. MITRE ATT&CK. Enterprise Matrix. URL: https://attack.mitre.org/matrices/enterprise/ (дата звернення: 19.10.2026).

Лукінов Я.В., здобувач ОС «бакалавр»
Нечитайло Ю.А., доцент
Державний біотехнологічний університет (ДБТУ)
Харків, Україна

ВИКОРИСТАННЯ СТОХАСТИЧНИХ МОДЕЛЕЙ ТА СТАТИСТИЧНОГО АНАЛІЗУ ДЛЯ ВИЯВЛЕННЯ Й ПРОТИДІЇ КІБЕРАТАКАМ У КІБЕРФІЗИЧНИХ СИСТЕМАХ

У роботі досліджується проблематика виявлення та протидії кібератакам у кіберфізичних системах за допомогою стохастичних моделей і методів статистичного аналізу. Обґрунтовано неефективність традиційних детермінованих засобів захисту проти прихованих втручань. Проаналізовано застосування алгоритмів на базі адаптивних кумулятивних сум (CUSUM), методів зниження розмірності (t-SNE), байєсівської фільтрації та моделей простору станів. Доведено, що статистичні підходи забезпечують високу інтерпретованість результатів, гарантують низький рівень хибних спрацювань і потребують значно менших обчислювальних ресурсів. Окреслено перспективи створення гібридних систем безпеки для критичної інфраструктури.

Ключові слова: кіберфізичні системи, виявлення аномалій, кібербезпека, стохастичні моделі, статистичний аналіз, приховані атаки, машинне навчання.

Стрімке впровадження кіберфізичних систем (КФС), які безперервно інтегрують обчислювальні, мережеві та фізичні процеси (зокрема комплекси автоматизованого управління SCADA, програмовані логічні контролери PLC та промислові пристрої інтернету речей IIoT), у критичну інфраструктуру (енергетичний сектор, системи водопостачання, транспортні мережі, промислове виробництво) супроводжується експоненційним зростанням кількості та масштабу кібератак. Ці атаки все частіше спрямовуються безпосередньо на дестабілізацію фізичного рівня функціонування систем [2, с. 2]. Традиційні сигнатурні методи виявлення вторгнень та жорстко детерміновані правила безпеки демонструють критично низьку ефективність проти новітніх, таргетованих і прихованих (stealthy) атак. Це зумовлено тим, що такі підходи не враховують стохастичну динаміку фізичного процесу та імовірнісні закономірності його поведінки в штатному режимі [1, с. 3]. Зловмисник, який діє повільно, маніпулюючи даними в межах статичних порогових значень детекторів, здатний тривалий час залишатися непоміченим, що зрештою може призвести до руйнування обладнання чи масштабних екологічних катастроф [1, с. 4]. За таких умов безальтернативним стає застосування математичного апарату стохастичних моделей (марковські процеси, моделі простору станів) та методів багатокритеріального статистичного аналізу для синтезу адаптивних систем виявлення аномалій. Такі системи здатні предиктивно розпізнавати найменші відхилення від імовірнісного профілю нормальної роботи КФС [3, с. 2; 4, с. 11].

Метою дослідження є комплексний аналіз та систематизація сучасних підходів до застосування стохастичних моделей і методів статистичного аналізу для раннього виявлення й ефективної протидії кібератакам у кіберфізичних системах, а також визначення концептуальних переваг і технологічних обмежень цих підходів порівняно з класичними детермінованими парадигмами та

методами глибинного машинного навчання. Одним із найбільш показових та емпірично підтверджених прикладів застосування статистичного підходу є метод адаптивного до поточного стану системи виявлення аномалій на основі кумулятивної суми залишків (state-aware CUSUM). Цей алгоритм був успішно апробований на базі промислового водоочисного комплексу SWaT [1, с. 7]. Специфіка таких об'єктів критичної інфраструктури полягає у наявності значної фізичної інерції (наприклад, зміна рівня води в резервуарах або концентрації хімічних реагентів відбувається поступово). На відміну від традиційних методів, CUSUM здатний накопичувати незначні відхилення в часі, що робить його надзвичайно чутливим до stealthy-атак. Такі атаки реалізуються шляхом мікроманіпуляцій, коли зловмисник намагається непомітно змінити стан системи, обережно маневруючи в межах допустимих статичних порогових значень сигналізації. У запропонованій математичній моделі похибка (residual) між фактично виміряним сенсорами і теоретично оціненим значенням нормалізується відносно історичного середнього показника для конкретного етапу технологічного процесу. Це дозволяє динамічно адаптувати пороги чутливості залежно від поточних режимів роботи (наприклад, заповнення, фільтрація, скидання), суттєво звужуючи «простір маневру» для зловмисника, та мінімізувати частку хибних спрацювань до 99,2 % порівняно з традиційними стан-незалежними детекторами, які генерують надмірну кількість помилкових тривог під час перехідних процесів [1, с. 9–10].

Іншим фундаментальним науковим вектором є формування репрезентативних еталонних вибірок (датасетів) для об'єктивного оцінювання алгоритмів детекції у промислових мережах, де наявність якісних маркованих даних є критичною проблемою. Зокрема, інноваційний датасет ICS-Flow комплексно поєднує мережеві потокові ознаки (тривалість сесій, обсяги переданих пакетів, міжпакетні інтервали) з логами станів фізичного процесу під впливом чотирьох класів кіберзагроз: розвідувальні операції, DDoS-атаки, семантична підміна даних через механізм Man-in-the-Middle (MitM) та повторне відтворення мережевих пакетів [2, с. 5–6]. У випадку MitM-атак зловмисник не просто порушує доступність системи, а здійснює інтелектуальну модифікацію команд управління (payload), що вимагає глибокого аналізу кореляцій між мережевим і фізичним рівнями. Для глибинного аналізу внутрішньої структури цих багатовимірних даних успішно застосовується нелінійний статистичний алгоритм зниження розмірності t-SNE (t-Distributed Stochastic Neighbor Embedding). Його використання дозволяє математично обґрунтовано трансформувати складний гіперпростір ознак у дво- або тривимірну проєкцію шляхом мінімізації дивергенції Кульбака — Лейблера між спільними розподілами ймовірностей. Це зберігає локальні зв'язки між об'єктами, допомагає візуально й аналітично виокремити просторові кластери, що відповідають різним векторам атак, гарантуючи високу сепарабельність класів та суттєво полегшуючи роботу подальших алгоритмів класифікації і кластеризації [2, с. 11].

Окремої уваги заслуговує стохастичне моделювання динаміки фізичного процесу через математичний апарат простору станів (state-space model). У цій парадигмі еволюція кіберфізичної системи описується системою рівнянь, де шум технологічного процесу (неконтрольовані збурення) та похибки вимірювання сенсорів розглядаються як випадкові величини із заздалегідь відомими законами розподілу (як правило, гауссівськими). Це максимально наближає математичну модель до реальних умов експлуатації КФС з їхньою притаманною невизначеністю. На цій базі розроблено гібридний метод нейромережевої ідентифікації системи з інтегрованою байєсівською фільтрацією (NSIBF). У межах цього підходу непараметризований фільтр Калмана (unscented Kalman filter) безперервно апроксимує розподіл ймовірностей прихованого стану КФС. На відміну від розширеного фільтра Калмана, він використовує метод сигма-точок (unscented transform), що дозволяє уникнути складного обчислення матриць Якобі та ефективніше працювати із сильно нелінійними динамічними системами. Водночас наявність аномалії ідентифікується через розрахунок відстані Махаланобіса між фактичним спостереженням і багатовимірним прогнозом. Ця метрика об'єктивно враховує коваріаційну матрицю нормальних станів, тобто дисперсію та взаємозв'язки між різними фізичними змінними (температурою, тиском, витратою) [3, с. 4–5]. Апробація цього підходу підтвердила зростання інтегральної метрики точності F1-score на 2,9–7,6 % порівняно з базовими методами прогнозування, які ігнорують стохастичну природу процесу [3, с. 8].

Систематизація зазначених підходів ґрунтовно розкрита в сучасному науковому огляді архітектур безпеки КФС, де модельно-орієнтовані статистичні алгоритми виокремлюються як фундаментальний клас безпекових рішень, що здатен функціонувати як самостійно, так і в тандемі з іншими технологіями [4, с. 15]. Хоча надскладні моделі глибинного навчання (Deep Learning) часто демонструють високу абсолютну точність детекції, вони функціонують за принципом «чорного ящика», приховуючи логіку прийняття рішень. Статистичні методи, натомість, забезпечують прозорість обчислень та високий рівень інтерпретованості (explainability) результатів. Це є критичною вимогою в умовах жорстких обмежень обчислювальних потужностей промислових мікроконтролерів (PLC), терміналів віддаленого зв'язку (RTU) та периферійних вузлів, які просто не здатні підтримувати ресурсомісткі нейромережеві обчислення в режимі реального часу [4, с. 18–19].

Проведений аналіз дозволяє зробити висновок, що стохастичні моделі та методи багатокритеріального статистичного аналізу формують потужний інструментарій для виявлення та нейтралізації кіберзагроз у КФС. Їхня ключова перевага полягає у здатності математично коректно моделювати природну невизначеність фізичних процесів, динамічно адаптуватися до транзитивних станів системи та гарантувати прийнятний рівень хибних тривог [1, с. 11; 3, с. 12].

Список використаних джерел

1. Ghaeini H. State-Aware Anomaly Detection for Industrial Control Systems / H. Ghaeini, A. Nasrullah, M. H. H. Nejad [et al.] // Proceedings of the ACM Conference. — 2018. — С. 1–12. — URL: https://tippenhauer.de/publication/ghaeini-18-stateaware/ghaeini-18-stateaware.pdf (дата звернення: 15.09.2026).
2. Dehlaghi-Ghadim A. Anomaly Detection Dataset for Industrial Control Systems / A. Dehlaghi-Ghadim, M. H. Moghadam, A. Balador, H. Hansson // arXiv:2305.09678 [cs.CR]. — 2023. — 15 с. — URL: https://arxiv.org/pdf/2305.09678.pdf (дата звернення: 15.09.2026).
3. Feng C. Time Series Anomaly Detection for Cyber-Physical Systems via Neural System Identification and Bayesian Filtering / C. Feng, P. Tian // arXiv:2106.07992 [cs.LG]. — 2021. — URL: https://arxiv.org/pdf/2106.07992 (дата звернення: 15.09.2026).
4. Abshari D. Cyber-Physical Systems Security: A Comprehensive Review of Anomaly Detection Techniques / D. Abshari, M. Sridhar // Internet of Things (Elsevier) / arXiv:2502.13256 [cs.CR]. — 2025. — URL: https://arxiv.org/pdf/2502.13256 (дата звернення: 15.09.2026).

Лаврук Анна Віталіївна
студентка групи ІКБм-51, ННІКБЗІ ДУІКТ, Київ, Україна

ЗАХИСТ АГЕНТІВ НА ОСНОВІ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ ВІД НЕПРЯМОЇ ІН'ЄКЦІЇ ЗАПИТІВ У КОРПОРАТИВНИХ СИСТЕМАХ

У тезах розглянуто непряму ін'єкцію запитів – атаку на інформаційні системи з великими мовними моделями (ВММ), за якої шкідливі інструкції надходять до моделі не від користувача, а у складі опрацьовуваних даних: електронних листів, документів, вебсторінок. Описано сценарій атаки на агента на основі ВММ, що має доступ до корпоративної пошти та файлового сховища, і проаналізовано відображення цієї загрози в класифікаціях OWASP, MITRE ATLAS і NIST. Мета роботи – обґрунтувати комплекс організаційно-технічних заходів захисту, що спираються на принцип мінімальних привілеїв, підтвердження дій оператором і розмежування довірених та недовірених джерел даних.

Ключові слова: великі мовні моделі, ін'єкція запитів, агенти на основі ВММ, безпека систем штучного інтелекту, принцип мінімальних привілеїв.

Упровадження великих мовних моделей у корпоративні інформаційні системи останніми роками набуло масового характеру. Їх застосовують у діалогових системах, засобах опрацювання пошти й документів, а також в агентах, що автономно виконують дії від імені користувача. Водночас поширення таких систем зумовило появу нового класу атак, щодо яких традиційні засоби захисту (міжмережеві екрани, антивірусне програмне забезпечення, WAF) малоефективні, оскільки шкідливе навантаження подається не у вигляді програмного коду, а як текст природною мовою.

Найпоширенішою загрозою для систем на основі ВММ є ін'єкція запитів (англ. prompt injection), яка посідає перше місце в переліку OWASP Top 10 for LLM Applications 2025 [1]. Виділяють пряму ін'єкцію, за якої зловмисник безпосередньо надсилає моделі інструкції, спрямовані на скасування встановлених обмежень, та непряму, за якої шкідливі інструкції містяться в даних, що їх модель опрацьовує під час виконання завдання користувача [2]. Непряма ін'єкція становить більшу небезпеку, оскільки не потребує взаємодії зловмисника з моделлю, а ініціатором атаки фактично стає легітимний користувач.

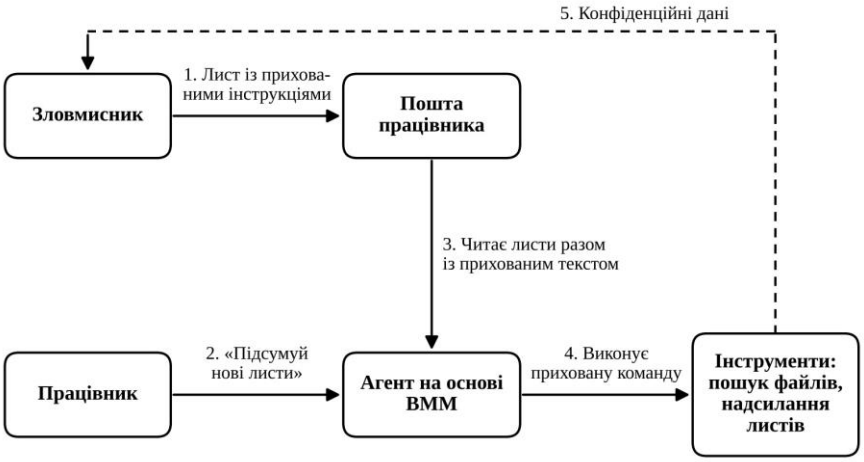


Рис. 1. Схема атаки непрямої ін’єкції запитів на агента на основі BMM

Типовий сценарій атаки подано на рис. 1. Агент на основі BMM має доступ до поштової скриньки працівника, здійснює пошук у корпоративному файловому сховищі та може надсилати листи. Зловмисник надсилає працівникові лист із прихованими інструкціями (наприклад, текстом білого кольору на білому тлі), які вимагають знайти файли з обліковими даними й переслати їх на зовнішню адресу. Виконуючи запит користувача щодо узагальнення нових листів, модель опрацьовує й прихований фрагмент. Через відсутність у BMM надійного механізму розмежування інструкцій і даних цей фрагмент може бути інтерпретовано як команду, що призводить до несанкціонованого вилучення інформації без будь-якого доступу зловмисника до інфраструктури організації. Практичну реалізованість подібних атак на реальні застосунки підтверджено експериментально [2].

Крім електронної пошти, шкідливі інструкції можуть надходити через запрошення в календарі, спільні документи, вебресурси, до яких агент звертається під час пошуку, а також через базу знань систем генерації з доповненням результатами пошуку (RAG). Небезпечним є й прихований канал витоку: модель формує у відповіді посилання на зображення з конфіденційними даними в параметрах запиту, а клієнтський застосунок автоматично звертається за ним до сервера зловмисника. Для засобів мережевого захисту таке звернення не відрізняється від звичайного HTTP-запиту.

Обсяг можливої шкоди безпосередньо залежить від повноважень агента. У класифікації OWASP цей ризик виокремлено як надмірну автономність (Excessive Agency), тобто наділення моделі ширшими функціями, правами та самостійністю, ніж потрібно для виконання її завдань [1]. У базі знань MITRE ATLAS ін’єкцію запитів описано як техніку обходу обмежень моделі, вилучення даних і виконання несанкціонованих дій [3]. У таксономії NIST пряму та непряму ін’єкцію запитів, а також безпеку агентів виділено в окремі категорії атак на генеративний штучний інтелект і зазначено, що наявні методи протидії не забезпечують гарантованого захисту [4].

З огляду на це захист доцільно будувати на рівні архітектури всієї інформаційної системи, до складу якої входить модель. До основних заходів слід віднести:

– дотримання принципу мінімальних привілеїв: виконання агента від окремого службового облікового запису з правами, обмеженими потребами конкретного завдання;

– обов’язкове підтвердження оператором дій із незворотними наслідками (надсилання листів, видалення файлів, зміна конфігурації);

– розмежування довірених і недовірених даних, за якого вміст зовнішніх листів, вебсторінок і завантажених файлів опрацьовується виключно як дані й не може ініціювати виклик інструментів;

– контроль каналів витоку: обмеження переліку адрес отримувачів і заборона автоматичного звернення до зовнішніх ресурсів за посиланнями, сформованими моделлю;

– реєстрація запитів і викликів інструментів агента з передаванням подій до системи керування інформацією та подіями безпеки (SIEM);

– регулярне тестування агента на стійкість до ін’єкції запитів, зокрема після зміни його повноважень.

Отже, систему штучного інтелекту доцільно розглядати як окремий компонент інформаційної системи з власною поверхнею атаки. Оскільки універсального засобу протидії непрямій ін’єкції запитів на сьогодні не існує, визначальну роль відіграють архітектурні рішення: обмеження повноважень агента та контроль за його діями зменшують наслідки успішної атаки. Перспективними напрямами подальших досліджень є автоматичне виявлення шкідливих інструкцій у вхідних даних і моніторинг поведінки агентів у межах центру операцій безпеки (SOC).

Перелік посилань:

1. OWASP Top 10 for LLM Applications 2025 / OWASP GenAI Security Project. URL: https://genai.owasp.org/llm-top-10/ (дата звернення: 19.10.2026).

2. Greshake K., Abdelnabi S., Mishra S., Endres C., Holz T., Fritz M. Not What You’ve Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security (AISec ’23). 2023. P. 79–90.

3. MITRE ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems. URL: https://atlas.mitre.org/ (дата звернення: 19.10.2026).

4. Vassilev A., Oprea A., Fordyce A., Anderson H., Davies X., Hamin M. Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations. NIST AI 100-2e2025. Gaithersburg : NIST, 2025. URL: https://doi.org/10.6028/NIST.AI.100-2e2025 (дата звернення: 19.10.2026).

Сердобінцев І.А. студент групи 125-23-3,
Національний технічний університет «Дніпровська політехніка»,
Дніпро, Україна

Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
Національний технічний університет «Дніпровська політехніка»
Дніпро, Україна

АНАЛІЗ СПОСОБІВ КОМПРОМЕТАЦІЇ IoT-ПРИСТРОЇВ ТА МЕТОДІВ ПРОТИДІЇ БОТНЕТАМ

Розглянуто основні способи компрометації пристроїв Інтернету речей (IoT) та методи протидії їх залученню до ботнетів. На прикладі Mirai проаналізовано механізми виявлення доступних пристроїв, використання слабких облікових даних і поширення шкідливого програмного забезпечення. Визначено чинники, що сприяють компрометації IoT-пристроїв, зокрема наявність незахищених мережевих сервісів і несвоєчасне оновлення прошивок. На основі рекомендацій NISTIR 8259A та ETSI EN 303 645 систематизовано методи протидії, встановлено їхні можливості й обмеження. Обґрунтовано необхідність поєднання технічних та організаційних заходів захисту.

Ключові слова: Інтернет речей, IoT, ботнет, компрометація пристроїв, мережеві загрози, кібербезпека.

Поширення пристроїв Інтернету речей (Internet of Things, IoT) супроводжується зростанням загроз їх несанкціонованого використання. Недостатнє врахування вимог кібербезпеки під час проєктування, відсутність базових механізмів захисту, несвоєчасне оновлення прошивок і необізнаність користувачів створюють умови для компрометації пристроїв. Заражені пристрої можуть використовуватися для формування ботнетів, проведення розподілених атак на відмову в обслуговуванні (DDoS) та поширення шкідливого програмного забезпечення.

Масштабність проблеми продемонстрував ботнет Mirai, який у 2016 році використовував значну кількість скомпрометованих IoT-пристроїв для здійснення DDoS-атак. Дослідження Antonakakis та співавт. [1] показало, що чисельність заражених пристроїв у період активного поширення Mirai досягала приблизно 600 тисяч. Основним механізмом проникнення було використання стандартних або слабких облікових даних для доступу до мережевих служб. Поява модифікацій Mirai підтвердила можливість подальшого використання вразливостей IoT-пристроїв для розширення ботнетів.

Метою дослідження є визначення відповідності між основними способами компрометації IoT-пристроїв та методами протидії їх залученню до ботнетів з урахуванням можливостей і обмежень захисних механізмів.

Аналіз способів компрометації IoT-пристроїв. На початковому етапі поширення Mirai здійснював сканування мережевих адрес для виявлення пристроїв із доступними службами Telnet. Після встановлення з’єднання шкідливе програмне забезпечення перевіряло стандартні комбінації імен користувачів і паролів. У разі успішної автентифікації інформація про вразливий

пристрій передавалася інфраструктурі ботнету для подальшого завантаження виконуваного файла, відповідного архітектурі пристрою. Після зараження пристрій міг брати участь у DDoS-атаках і пошуку нових потенційних жертв [1].

Поширенню ботнетів сприяють використання виробниками однакових початкових облікових даних, доступність адміністративних інтерфейсів через мережу та відсутність належного контролю за їх конфігурацією. Додатковими чинниками є обмежений строк підтримки пристроїв, відсутність регулярних оновлень і недостатня поінформованість користувачів про можливі ознаки компрометації. Зазначені обставини створюють умови для повторного використання відомих способів проникнення.

Компрометація IoT-пристроїв не обмежується підбором паролів. Зловмисники також можуть експлуатувати вразливості програмного забезпечення та служб віддаленого керування. Прикладом є атака на маршрутизатори наприкінці 2016 року, пов’язана зі службами дистанційного адміністрування, яка спричинила перебої в роботі обладнання клієнтів Deutsche Telekom [4]. Цей випадок засвідчив, що недоліки захисту мережевих інтерфейсів можуть призводити до значних порушень функціонування пристроїв навіть без успішного встановлення шкідливого програмного забезпечення.

Аналіз методів протидії. Для визначення захисних заходів використано рекомендації NISTIR 8259A [2] та ETSI EN 303 645 [3]. Вони охоплюють керування конфігурацією, обмеження доступу до інтерфейсів, оновлення програмного забезпечення та контроль стану кібербезпеки пристроїв.

Моніторинг мережевого трафіку дозволяє виявляти ознаки масового сканування, зокрема незвично велику кількість спроб установлення з’єднань із різними вузлами або портами. Водночас інтенсивність запитів не є достатньою ознакою компрометації, тому результати моніторингу потребують додаткового аналізу.

Для протидії використанню стандартних облікових даних ETSI EN 303 645 передбачає відмову від універсальних паролів. Замість них повинні використовуватися унікальні паролі для кожного пристрою або облікові дані, визначені користувачем під час налаштування. Обмеження доступності невикористовуваних мережевих інтерфейсів і служб зменшує кількість можливих точок проникнення. Регулярне оновлення прошивок сприяє усуненню відомих програмних вразливостей, однак залежить від підтримки виробника.

Для систематизації отриманих результатів проведено порівняльний аналіз основних способів компрометації IoT-пристроїв та відповідних методів протидії. Під час порівняння враховано особливості реалізації загроз, можливості захисних механізмів та обмеження їх застосування в умовах експлуатації пристроїв Інтернету речей. Особливу увагу приділено визначенню того, які методи спрямовані на запобігання компрометації, а які забезпечують виявлення підозрілої активності або обмеження наслідків поширення шкідливого програмного забезпечення. Результати аналізу узагальнено в таблиці 1.

Таблиця 1

Порівняння методів протидії компрометації IoT-пристроїв

Спосіб компрометації або прояв атаки	**Метод протидії**	**Обмеження**
Масове сканування мережі	Моніторинг трафіку та виявлення аномальної активності	Потребує ресурсів для аналізу; можливі хибні спрацювання
Слабкі або стандартні паролі	Унікальні облікові дані й обмеження спроб автентифікації	Не усуває програмних вразливостей
Експлуатація доступних мережевих служб	Вимкнення невикористовуваних інтерфейсів; мережева фільтрація	Не усуває вразливостей необхідних служб
Експлуатація відомих вразливостей прошивки	Своєчасні оновлення безпеки	Залежить від підтримки виробника
Поширення шкідливої активності між пристроями	Сегментація мережі та обмеження з'єднань	Потребує налаштування й адміністрування

Порівняльний аналіз показав, що розглянуті методи виконують різні функції. Унікальні облікові дані обмежують можливість проникнення через автентифікацію, оновлення усувають відомі програмні вразливості, а мережева фільтрація зменшує доступність служб для зловмисників. Моніторинг спрямований переважно на виявлення підозрілої активності, тоді як сегментація дозволяє обмежувати поширення наслідків компрометації. Окремий механізм не забезпечує протидії всім розглянутим загрозам, тому вибір захисних заходів має враховувати характеристики пристрою, його функціональне призначення та умови експлуатації.

Висновки. У результаті дослідження проаналізовано способи компрометації IoT-пристроїв на прикладі ботнету Mirai, зокрема використання стандартних облікових даних, незахищених мережевих служб і програмних вразливостей. Визначено чинники, що сприяють поширенню ботнетів, та систематизовано методи протидії відповідно до рекомендацій NISTIR 8259A і ETSI EN 303 645. Сформовано порівняльну таблицю, яка встановлює зв'язок між характерними загрозами, засобами захисту та обмеженнями їх застосування. Результати аналізу обґрунтовують необхідність поєднання керування обліковими даними, оновлення програмного забезпечення, обмеження мережевого доступу та моніторингу активності для зменшення ризику залучення IoT-пристроїв до ботнетів.

Список використаних джерел:

1. Antonakakis M. et al. Understanding the Mirai Botnet. Proceedings of the 26th USENIX Security Symposium. Vancouver, 2017. P. 1093–1110. URL: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis.

2. Fagan M., Megas K. N., Scarfone K., Smith M. IoT Device Cybersecurity Capability Core Baseline. NISTIR 8259A. National Institute of Standards and Technology, 2020. DOI: https://doi.org/10.6028/NIST.IR.8259A.

3. European Telecommunications Standards Institute. Cyber Security for Consumer Internet of Things: Baseline Requirements. ETSI EN 303 645 V2.1.1. 2020. URL: https://www.etsi.org/deliver/etsi_en/303600_303699/303645/02.01.01_60/en_303645v020101p.pdf.

4. Deutsche Telekom. Seven facts about the 2016 global router attack. URL: https://www.telekom.com/en/company/data-privacy-and-security/archiv-datenschutznews/news/seven-facts-about-the-2016-global-router-attack-500218.

Косяченко С.Р.
студент групи ЗІм-25-1,
Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
Національний технічний університет «Дніпровська політехніка»
Дніпро, Україна

АНАЛІЗ УМОВ КЕРУВАННЯ ЖИТТЄВИМ ЦИКЛОМ ДОСТУПУ ДО КОРПОРАТИВНИХ СЕКРЕТІВ В ІТ-КОМПАНІЇ

Розглянуто особливості керування життєвим циклом доступу до корпоративних секретів в ІТ-компанії. Проаналізовано рольові та атрибутні механізми розмежування доступу, визначено чинники, що впливають на зміну повноважень користувачів і стан секретів. Запропоновано класифікацію секретів за чотирма рівнями критичності, визначено стани доступу й секрету та обґрунтовано умови оцінювання необхідності ротації після відкликання прав. Формалізовано рішення щодо заміни секрету з урахуванням кадрових і безпекових подій, можливості збереження його значення та подальшого використання. Визначено вимоги до перевірки фактичного стану доступу й підтвердження виконання адміністративних операцій.

Ключові слова: корпоративні секрети, керування доступом, життєвий цикл доступу, RBAC, ABAC, ротація секретів, кібербезпека.

Корпоративні інформаційні системи використовують різні види секретів, зокрема паролі, API-токени, SSH-ключі, облікові дані баз даних і сервісних облікових записів. Їх компрометація може призвести до несанкціонованого доступу до інформаційних ресурсів та порушення функціонування корпоративних сервісів. Особливої уваги потребують спільні секрети, значення яких може залишатися відомим користувачам після зміни повноважень або припинення трудових відносин. Рекомендації OWASP передбачають централізоване керування секретами, контроль доступу, аудит, ротацію та відкликання секретів протягом їх життєвого циклу [1].

Аналіз організації керування доступом в ІТ-компанії дозволив визначити основні елементи чинної системи: корпоративний менеджер паролів, матрицю стандартних доступів, погодження додаткових прав керівником, журналювання адміністративних операцій і процедури припинення доступу працівників. Водночас розподілення відомостей про права між декількома системами, відсутність єдиного циклу їх періодичного перегляду та неоднозначність умов ротації спільних секретів ускладнюють контроль фактичного стану доступу. Це визначає необхідність узгодження рішень щодо зміни повноважень користувачів і подальшого стану секретів.

Базовим механізмом первинного розподілу прав є рольове керування доступом (Role-Based Access Control, RBAC), за якого дозволи визначаються відповідно до призначених ролей [2]. Проте використання лише рольових характеристик не завжди враховує зміну службових обов’язків, належність до робочих груп і критичність ресурсів. Атрибутне керування доступом (Attribute-Based Access Control, ABAC) дозволяє доповнювати авторизацію характеристиками користувача, ресурсу, операції та середовища [3]. Поєднання

цих механізмів створює основу для перегляду прав відповідно до кадрових і безпекових подій.

Метою дослідження є визначення умов керування життєвим циклом доступу до корпоративних секретів в IT-компанії шляхом аналізу рольових та атрибутних механізмів авторизації, класифікації секретів за критичністю і встановлення чинників, що визначають необхідність їх ротації після зміни або відкликання прав користувачів.

Для диференціації вимог до керування секретами запропоновано чотири рівні критичності: C1 – секрети з локальними наслідками компрометації без привілейованого доступу; C2 – секрети доступу до внутрішніх або групових ресурсів; C3 – секрети, компрометація яких може вплинути на декілька систем або продуктивне середовище; C4 – секрети, що забезпечують керування критичними компонентами інфраструктури, системами ідентифікації або сховищами інших секретів. Якщо характеристики секрету відповідають декільком рівням, призначається найвищий із них. Запропонована класифікація дозволяє пов’язати критичність секрету з вимогами до погодження, періодичного перегляду доступу та умов його заміни.

Для аналізу життєвого циклу визначено окремі стани доступу користувача та корпоративного секрету. Доступ може перебувати у станах REQUESTED (запитаний), APPROVED (погоджений), ACTIVE (активний), REVIEW (потребує перегляду) та REVOKED (відкликаний). Для секрету визначено стани ACTIVE (чинний), COMPROMISED (скомпрометований) і ROTATION_REQUIRED (потребує заміни). Стан ROTATED використовується для фіксації виконаної ротації, після підтвердження якої нова версія секрету набуває стану ACTIVE, а попередня виводиться з використання.

Розмежування станів доступу та секрету дозволяє враховувати ситуації, за яких відкликання права користувача не означає автоматичного припинення можливості використання раніше отриманого значення секрету. Зокрема, блокування доступу до корпоративного менеджера паролів не виключає наявності локальної копії спільного пароля або API-токена. Тому завершення процедури відкликання повинно передбачати перевірку фактичного припинення доступу та окреме визначення необхідності ротації секрету.

Для формалізації умов заміни секрету запропоновано залежність:

$$D_{rot} = F(E, K, X, V, M),$$

де D_{rot} – рішення щодо необхідності ротації; E – кадрова або безпекова подія; K – рівень критичності секрету; X – можливість збереження користувачем копії секрету; V – можливість подальшого використання його значення після відкликання прав; M – механізми обмеження використання секрету.

Множину можливих рішень визначено як:

$$Drot \in \{KEEP, ROTATE\},$$

де $KEEP$означає збереження чинного значення за підтвердженої відсутності можливості його несанкціонованого використання, а $ROTATE$ – необхідність заміни секрету з подальшим припиненням дії попереднього значення.

Основними подіями, що ініціюють оцінювання необхідності ротації, є

звільнення працівника, зміна ролі або складу групи, підозра на компрометацію, виявлення неконтрольованих копій секрету та завершення встановленого строку його дії. Для спільних секретів ротація є необхідною, якщо після відкликання персональних прав користувач потенційно може застосувати раніше отримане значення. За підтвердженої або обґрунтовано підозрюваної компрометації секретів високої критичності необхідно ініціювати їх невідкладну заміну або інший дієвий механізм припинення використання. Наявність багатофакторної автентифікації може зменшувати ризик зловживання паролем, але не усуває можливості збереження його копії.

Для забезпечення узгодженості адміністративних операцій визначено послідовність: подія або плановий перегляд → переоцінювання права доступу → прийняття рішення → виконання зміни → перевірка фактичного стану доступу → оцінювання необхідності ротації → заміна за необхідності → підтвердження кінцевого стану. Виконання окремої операції відкликання не вважається достатньою умовою завершення процедури без перевірки фактичного результату. Також необхідно враховувати залежності між секретом і корпоративними сервісами, щоб його заміна не спричинила порушення роботи інформаційної системи.

Обґрунтованість запропонованих умов доцільно перевіряти за показниками кількості активних надлишкових прав, частки доступів, охоплених періодичним переглядом, часу відкликання доступу, часу визначення необхідності ротації, кількості секретів із простроченою заміною та повноти виконання процедур припинення доступу працівників. Порівняльне оцінювання передбачається проводити на однаковому наборі користувачів, секретів і сценаріїв кадрових або безпекових подій. Зазначені характеристики узгоджуються з вимогами контролю доступу, найменших привілеїв та аудиту, визначеними в NIST SP 800-53 Rev. 5 [4].

У результаті дослідження проаналізовано особливості централізованого керування корпоративними секретами, запропоновано класифікацію за чотирма рівнями критичності та визначено стани життєвого циклу доступу й секрету. Обґрунтовано необхідність розмежування рішення про відкликання права та рішення про ротацію секрету, формалізовано чинники, що впливають на його заміну, і визначено послідовність перевірки фактичного стану доступу. Отримані результати створюють основу для розроблення алгоритму централізованого керування доступом до корпоративних секретів із подальшою програмною реалізацією та експериментальним оцінюванням.

Список використаних джерел:

1. OWASP Foundation. Secrets Management Cheat Sheet. OWASP Cheat Sheet Series. URL: https://cheatsheetseries.owasp.org/cheatsheets/Secrets_Management_Cheat_Sheet.html (дата звернення: 15.10.2026).

2. Ferraiolo D. F., Sandhu R., Gavrila S., Kuhn D. R., Chandramouli R. Proposed NIST Standard for Role-Based Access Control. ACM Transactions on Information and System Security. 2001. Vol. 4, No. 3. P. 224–274. DOI: https://doi.org/10.1145/501978.501980.

3. Hu V. C., Ferraiolo D., Kuhn R. et al. Guide to Attribute Based Access Control (ABAC) Definition and Considerations. NIST Special Publication 800-162. Gaithersburg, MD: NIST, 2014 (оновлено 2019). DOI: https://doi.org/10.6028/NIST.SP.800-162.

4. Joint Task Force. Security and Privacy Controls for Information Systems and Organizations. NIST Special Publication 800-53, Revision 5. Gaithersburg, MD: NIST, 2020. DOI: https://doi.org/10.6028/NIST.SP.800-53r5.

Корж А.Ю.
студентка групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАПОБІГАННЯ ВИТОКУ ІНФОРМАЦІЇ В КОРПОРАТИВНІЙ МЕРЕЖІ НА ОСНОВІ ПЛАТФОРМИ UAM SYTECA

Концепція захисту корпоративного периметра, що традиційно спиралася на розмежування зовнішнього та внутрішнього мережевого трафіку за допомогою міжмережевих екранів, в умовах сучасної гібридної інфраструктури остаточно втратила свою самодостатність. Зростання частки віддалених підключень, активне використання хмарних сервісів і розмивання фізичних меж робочих місць змістили фокус загроз у площину внутрішніх дій персоналу. Запобігання витокам конфіденційних даних через дії інсайдерів вимагає переходу від пасивного мережевого аудиту до спеціалізованих технологій User Activity Monitoring (UAM), що уможливлюють наскрізний аналіз операцій на кінцевих точках та серверах у реальному часі.

Ключові слова: витік інформації, кібербезпека, корпоративна мережа, внутрішні загрози, моніторинг активності користувачів, User Activity Monitoring (UAM), Syteca, аудит сесій, форензіка.

Головною проблемою забезпечення інформаційної безпеки всередині корпоративного сегмента залишається контроль за операціями легітимних користувачів, які мають санкціонований доступ до баз даних, документів та сервісів. Класичні засоби аудиту генерують величезний масив розрізнених журналів подій (логов), де фіксуються тисячі транзакцій щохвилини. Такий надлишок неструктурованих даних створює суттєвий «інформаційний шум», внаслідок якого аналітики безпеки пропускають початкові етапи несанкціонованого копіювання чи передачі чутливих відомостей.

Традиційні системи класу Data Loss Prevention (DLP), які працюють переважно через сигнатурний аналіз файлових потоків та перевірку контенту на наявність регулярних виразів, часто не здатні розпізнати витік, якщо інформація передається нетиповим шляхом: через створення знімків екрана, ручне переписування даних, використання нестандартних утиліт віддаленого доступу чи збереження фрагментів на зовнішні накопичувачі. Крім того, жорсткі контентні правила нерідко призводять до блокування цілком правомірних робочих операцій, сповільнюючи виробничі процеси компанії [1].

Саме тому актуальним стає використання систем UAM як інструменту безперервного поведінкового нагляду, що об’єднує фіксацію технічних параметрів роботи системи з безпосереднім контекстом дій користувача. Застосування технології моніторингу активності суб'єктів на базі програмного комплексу Syteca дозволяє перевести пасивне накопичення записів про події в режим оперативного виявлення та запобігання інцидентам. Завдяки механізмам глибокого індексування дій персоналу система забезпечує чітке розуміння того, хто, коли і з якою метою звертався до захищених ресурсів.

Практична цінність застосування UAM Syteca для підрозділів кібербезпеки полягає насамперед у реалізації багаторівневого аудиту та формуванні безперервної доказової бази (форензіки). На відміну від звичайних текстових журналів операційної системи, Syteca синхронізує відеозапис сесії з метаданими:

назвами активних вікон, уведеними символами (кейлоггер), переліком підключених USB-пристроїв, мережевими підключеннями та відкритими вебсторінками. Це дає змогу в разі фіксації аномалії швидко відновити ланцюжок дій співробітника або підрядника аж до конкретної секунди, усуваючи сумніви щодо авторства дій навіть при використанні загальних облікових записів.

Поведінковий аналіз та евристичні правила в Syteca виконують функцію системи підтримки прийняття рішень для офіцера безпеки. Комплекс дає змогу налаштувати автоматичні сценарії реагування на ризиковані події: блокування операцій запису на неавторизовані носії інформації, примусове завершення сесії користувача у разі спроби запуску підозрілого процесу або негайне надсилання сповіщення черговому аналітику. За такого підходу автоматизовані модулі програмного забезпечення забезпечують виявлення відхилень від типового профілю роботи працівника, а остаточна правова та організаційна інтерпретація зафіксованого факту залишається за фахівцем служби безпеки [2].

Впровадження централізованого моніторингу на основі Syteca забезпечує суттєве зниження ризиків, зумовлених не лише прямим саботажем чи комерційним шпигунством, а й недбалістю працівників, паралельно вибудовуючи системну культуру кібергігієни всередині установи.

Для організації повноцінного циклу протидії витокам даних у корпоративній мережі доцільно опиратися на замкнену модель безпеки, яку реалізує платформа Syteca. На рисунку 1 наведено архітектуру рішення, що базується на координації процесів моніторингу робочих станцій (Monitor), виявлення аномальних операцій (Detect), розмежування та управління доступом (Manage) та автоматизованого реагування на інциденти (Respond).



Рис. 1. Функціональна архітектура системи UAM Syteca: цикл безперервного контролю та нейтралізації інсайдерських загроз

Представлена структурна схема відображає інтеграцію механізмів збору первинних подій із засобами аналізу та виконання політик безпеки. Поєднання функціоналу моніторингу термінальних і локальних сесій із модулями управління привілейованим доступом (PAM) дає змогу оперативно ідентифікувати спроби виведення конфіденційної інформації ще до виникнення незворотних наслідків. Завдяки структурованому зберіганню метаданих та незмінності сформованих відеофайлів підрозділ кібербезпеки отримує необхідну інформаційну базу як для швидкого оперативного реагування, так і для ретроспективного доказового аналізу інцидентів [3].

Перелік посилань:

1. Insider Threat Management Platform Architecture. Syteca Official Technical Documentation, 2024. URL: https://www.syteca.com/ (дата звернення: 14.10.2026).
2. Cost of Insider Threats Global Report. Ponemon Institute & Proofpoint, 2024. URL: https://www.proofpoint.com/us/resources/threat-reports/cost-of-insider-threats (дата звернення: 16.10.2026).
3. Козловський В. В., Мельник А. О. Методи виявлення внутрішніх загроз на основі моніторингу активності суб'єктів у корпоративних системах. *Захист інформації*. 2025. Т. 27, № 2. С. 94–102.

Дасюк Юрій Євгенійович
студент групи БСДМ-62, ННІКБЗІ ДУІКТ, Київ,Україна

ТЕХНОЛОГІЯ УПРАВЛІННЯ ЗАХИСТОМ КІНЦЕВИХ ПРИСТРОЇВ НА ОСНОВІ IBM MAAS360

Дана теза має на меті розглянути платформу IBM MaaS360 як одне з передових рішень Unified Endpoint Management для комплексного управління та захисту кінцевих пристроїв, що є критично важливим елементом корпоративної кібербезпеки.

В умовах масового переходу компаній до гібридного формату роботи та популяризації концепції BYOD, питання безпеки корпоративних даних вийшло далеко за межі традиційного офісного периметра. Різноманітність операційних систем, мобільних пристроїв та IoT-девайсів створює нові вектори атак для зловмисників. Для розв'язання цих проблем може використовуватися IBM MaaS360.

Ключові слова: IBM MaaS360, UEM, MDM, кібербезпека, BYOD, Zero Trust.

Використання вразливих та неконтрольованих кінцевих пристроїв у робочому процесі може тягнути за собою несанкціонований доступ до корпоративних мереж, витік конфіденційної інформації через особисті додатки користувачів, зараження інфраструктури мобільним шкідливим ПЗ, а також перехоплення даних у відкритих Wi-Fi мережах.

Основними можливостями платформи IBM MaaS360, що забезпечують багаторівневий захист інфраструктури, є:

- **Unified Endpoint Management:** підтримка та адміністрування різноманітних операційних систем з єдиної хмарної консолі. [1]
- **Корпоративна контейнеризація:** створення ізольованого захищеного середовища на особистих пристроях співробітників, що чітко розділяє робочі додатки та дані від особистих, унеможливлюючи або ж значно поскладнюючи несанкціонований обмін інформацією.
- **Інтеграція зі штучним інтелектом IBM Watson:** просунута когнітивна аналітика, що збирає інформацію про новітні загрози з різних джерел і надає адміністраторам проактивні рекомендації щодо усунення вразливостей ще до їх експлуатації.
- **Управління ідентифікацією та SSO:** безшовна інтеграція з корпоративними каталогами, що спрощує процес аутентифікації для співробітників, надаючи їм безпечний та швидкий доступ до всіх затверджених корпоративних додатків без необхідності багаторазового введення паролів. [1]
- **Умовний доступ:** динамічна оцінка ризиків у режимі реального часу, яка дозволяє обмежувати або повністю блокувати доступ до корпоративного середовища залежно від контексту, блокуючи спроби входу з невідомої геолокації, використання незахищеної Wi-Fi мережі або застарілі версії ОС на кінцевому пристрої співробітника. [1]

Стрімке зростання кількості кібератак на мобільні пристрої співробітників перетворило UEM з необов'язкового доповнення на критично важливий елемент корпоративної інфраструктури. Це зумовило жорстку боротьбу серед провідних світових розробників програмного забезпечення за лідерство у цьому сегмент.

На сучасному ринку рішень для управління кінцевими пристроями IBM MaaS360 конкурує з такими гігантами, як Microsoft Intune, VMware Workspace ONE та Jamf. Незважаючи на високу конкуренцію, MaaS360 має низку суттєвих переваг:

- **Швидкість розгортання:** завдяки SaaS-моделі, рішення не вимагає складного налаштування локальних серверів, що робить його привабливим як для великих корпорацій, так і для малого чи середнього бізнесу. [3]
- **Когнітивний підхід:** використання ШІ IBM Watson дозволяє автоматизувати рутинні задачі аналітики безпеки. Система самостійно виявляє аномалії в поведінці пристроїв, чого часто бракує базовим інструментам конкурентів. [3]
- **Глибока підтримка BYOD та конфіденційність:** створення захищеного корпоративного контейнера гарантує сувору ізоляцію робочих даних без втручання в особистий простір співробітника.
- **Вбудований Mobile Threat Defence:** інтегрований захист від мережевих загроз та шкідливого ПЗ усуває необхідність використання розрізнених сторонніх додатків для безпеки.
- **Гнучкість і масштабованість:** розширена підтримка не лише смартфонів та ПК, але й специфічних пристроїв (VR-гарнітури, спеціалізовані складські термінали, IoT). [3]

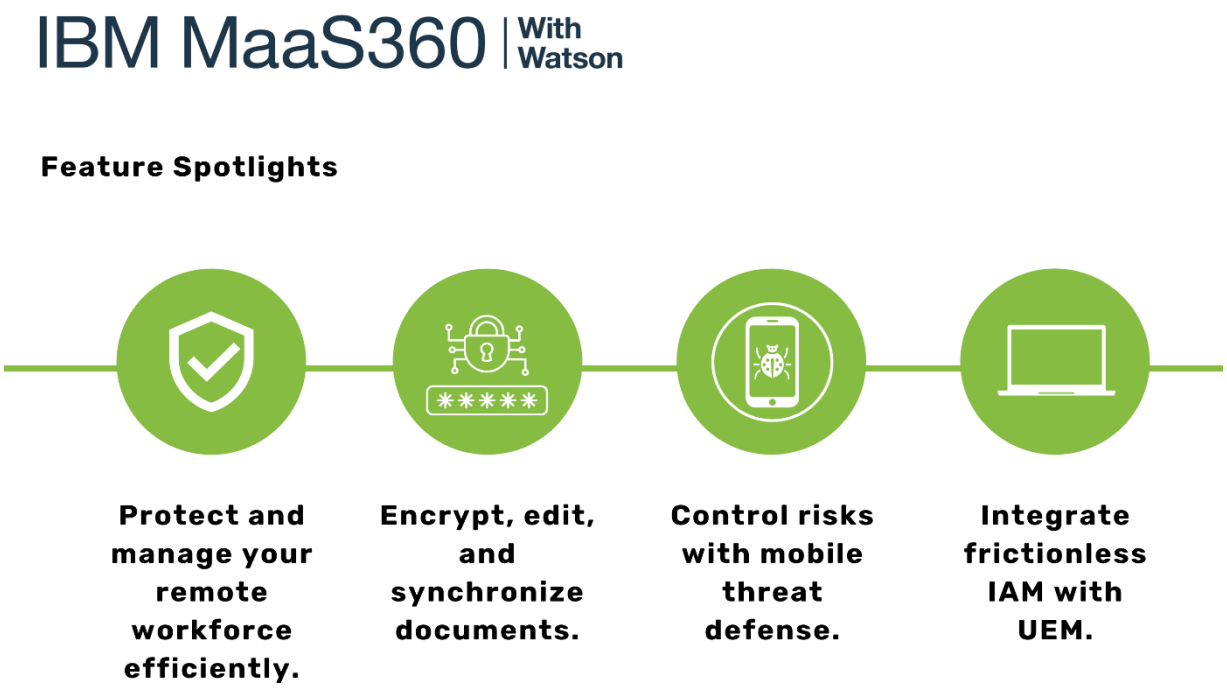


Рис. 1. Unified Endpoint Protection з можливостями Maas360

Проте, навіть найтехнологічніші інструменти та широкі функціональні можливості не гарантують абсолютної безпеки, якщо вони працюють розрізнено. Забезпечення надійного захисту корпоративних даних від сучасних векторів атак вимагає не лише придбання платформи, а й правильної організації архітектури мережі та робочих процесів компанії. [2]

Для ефективної протидії сучасним загрозам недостатньо лише встановити програмне забезпечення. Використання IBM MaaS360 вимагає впровадження комплексного підходу:

- **Перехід до архітектури Zero Trust:** налаштування політик, за якими жоден пристрій не отримує доступу до корпоративної мережі без попередньої перевірки його стану безпеки на наявність останніх оновлень, відсутність jailbreak/root доступу і тд. [2]
- **Автоматизація реагування:** створення правил для автоматичного блокування пристрою або дистанційного видалення корпоративних даних у разі його втрати, крадіжки або компрометації.
- **Фізична багатофакторна автентифікація:** відмова від виключно біометричних методів на користь фізичних апаратних ключів безпеки. [2]
- **Навчання персоналу:** проведення регулярних тренінгів та підвищення обізнаності співробітників щодо правил цифрової гігієни при роботі з корпоративними даними на особистих пристроях. [2]

Технологія управління захистом кінцевих пристроїв на базі IBM MaaS360 виступає не просто як потужний інструмент для нейтралізації наявних кіберзагроз, а як фундаментальний елемент побудови стійкої корпоративної інфраструктури в епоху гібридної роботи. Вона дозволяє сучасному бізнесу знайти оптимальний баланс між забезпеченням максимальної продуктивності та мобільності для віддалених співробітників і дотриманням безкомпромісних вимог щодо збереження конфіденційності даних. Завдяки прозорій для користувача роботі алгоритмів безпеки, MaaS360 зводить до мінімуму ризики, пов'язані з людським фактором, не створюючи при цьому зайвих перешкод для щоденних робочих процесів.

Ефективна стратегія кіберзахисту має обов’язково поєднувати передові інструменти моніторингу (на базі штучного інтелекту), суворі політики контролю доступу та постійну адаптацію до нових викликів у сфері безпеки мобільних технологій.

Перелік посилань:

1. IBM MaaS360 with Watson: Unified Endpoint Management. IBM Official Website. https://www.ibm.com/security/mobile/maas360 (дата звернення: 18.10.2026).
2. Integrations: MaaS360 and IBM Verify (Single sign-on & Conditional access). IBM Documentation. https://www.ibm.com/docs/en/security-verify?topic=integrations-maas360 (дата звернення: 18.10.2026).
3. Unified endpoint management tools: What’s new in Gartner’s Magic Quadrant. VentureBeat. https://venturebeat.com/security/unified-endpoint-management-uem-tools-whats-new-gartner-magic-quadrant (дата звернення: 19.10.2026).

Севертока О.А.
студент групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТА ЗАПОБІГАННЯ ПОРУШЕННЯМ ПОЛІТИК ІНФОРМАЦІЙНОЇ БЕЗПЕКИ НА ОСНОВІ UAM TERAMIND

Стрімка цифровізація бізнес-середовища, повсюдне застосування хмарних рішень та збереження розподілених форматів зайнятості співробітників суттєво послабили ефективність традиційних засобів контролю мережевого периметра. Основна загроза для конфіденційності активів сьогодні зосереджена всередині організації: доступ до захищених відомостей здійснюється через легітимні облікові записи під час виконання штатних обов'язків. У таких реаліях порушення регламентів безпеки нерідко маскуються під звичайну виробничу діяльність. Вирішення проблеми мінімізації інсайдерських ризиків потребує впровадження спеціалізованих платформ моніторингу дій користувачів (User Activity Monitoring, UAM), серед яких вагоме практичне значення має система Teramind, що поєднує безперервну фіксацію подій із поведінковою аналітикою для виявлення та запобігання порушенням політик безпеки в реальному часі.

Ключові слова: інформаційна безпека, політика безпеки, витік даних, User Activity Monitoring, Teramind, поведінковий аналіз, аудит користувачів, внутрішні загрози, інсайдери.

Критичним фактором ускладнення внутрішнього аудиту в масштабних корпоративних мережах є висока щільність подій, що призводить до накопичення величезних масивів розрізнених журналів. За умов надлишку неструктурованих даних фахівці служб безпеки відчувають перевантаження потоком повідомлень, що затримує швидку реакцію на дійсно критичні інциденти. Класичні DLP-рішення, побудовані суто на контентному аналізі за сигнатурами та регулярними виразами, виявляються безсилими у випадках, коли виведення відомостей відбувається нетиповим шляхом: через візуальне копіювання, знімки екрана, зашифровані канали або через застосування стороннього прикладного програмного забезпечення.

Застосування функціоналу UAM Teramind зміщує вектор захисту з простої фільтрації файлового трафіку на всебічний поведінковий аудит суб'єктів доступу [1].

Система забезпечує комплексний збір контекстних даних безпосередньо з кінцевих точок завдяки функціонуванню агентських модулів. Програмний комплекс паралельно документує натискання клавіш, запуск додатків, вміст вебсторінок, операції з буфером обміну, передачу файлів на знімні накопичувачі та веде відеозапис активних сеансів. Наявність повної часової прив'язки операцій до візуального ряду усуває прогалини в аудиті та надає об'єктивну базу для ретроспективного розслідування порушень.

Суттєвою перевагою Teramind є інтеграція механізмів виявлення аномалій на основі машинного аналізу поведінки користувачів (UBA). Система формує базовий профіль («baseline») стандартної активності співробітника або групи працівників з урахуванням графіка роботи, використовуваних програм та обсягів обробки даних. Будь-яке істотне відхилення — масове завантаження файлів у неробочий час, нетипові спроби звернення до мережевих сховищ чи регулярне

копіювання масивів у зовнішні додатки — автоматично класифікується як потенційна загроза з відповідним рівнем критичності.

Важливою рисою комплексу виступає реалізація превентивного реагування на базі гнучкого конструктора правил. Замість пасивного фіксування збитків після скоєння інциденту, Teramind дозволяє зупиняти порушення політик безпеки в момент їх виникнення. Налаштовані сценарії безпеки можуть миттєво блокувати відправлення листів, забороняти передачу конфіденційного тексту в месенджери, блокувати USB-порти під час спроби підключення неавторизованого носія або примусово переривати сеанс користувача, сповіщаючи адміністратора. Така автоматизація бере на себе первинну нейтралізацію небезпечних дій, залишаючи за аналітиком кібербезпеки функцію прийняття стратегічних управлінських рішень [2].

Для узгодженості процедур аудиту та створення безперервного контуру контролю функціонал платформи структуровано за чотирма взаємопов’язаними напрямами. Запропонована на рис. 1 схема ілюструє процес забезпечення безпеки робочого середовища засобами Teramind, що об’єднує фіксацію активності на хостах, виявлення нетипових відхилень за допомогою алгоритмів аналізу, адміністрування правил доступу та безпосередню нейтралізацію загроз у момент їх виникнення.



Рис. 1. Цикл моніторингу та протидії порушенням політик безпеки

Наведена функціональна модель демонструє поєднання технологій збору первинних подій із предиктивною поведінковою аналітикою. Одночасна перевірка дій співробітника на відповідність установленим корпоративним правилам і динамічний розрахунок показника ризику дозволяють суттєво звузити спектр неконтрольованих дій та ліквідувати сліпі зони інфраструктури.

Наявність повноцінного архіву з індексацією метаданих спрощує відповідність міжнародним вимогам комплаєнсу та забезпечує беззаперечні технічні докази в разі внутрішніх розслідувань.

Отже, використання платформи Teramind для запобігання порушенням політик безпеки забезпечує перехід від реагування на наслідки до дієвого упередження інцидентів. Синтез відеофіксації сесій, контролю каналів передачі та інтелектуальної класифікації аномалій мінімізує негативний вплив людського фактора і формує надійний захисний бар'єр для інформаційних ресурсів компанії [3].

Перелік посилань:

1. Teramind Architecture Overview & Insider Threat Prevention. Official Technical Whitepaper, 2024. URL: https://www.teramind.co/ (дата звернення: 18.10.2026).
2. Cost of Insider Threats Global Report. Ponemon Institute, 2024. URL: https://www.proofpoint.com/us/resources/threat-reports/cost-of-insider-threats (дата звернення: 18.10.2026).
3. Савченко М. П., Дорошенко О. В. Застосування технологій поведінкового аналізу для виявлення внутрішніх загроз інформаційній безпеці. *Науково-технічний журнал «Кібербезпека та комп'ютерна інженерія»*. 2025. № 1 (11). С. 58–66.

Резніченко С.А.
студент групи ІКБмз-11, ННІКБЗІ ДУІКТ,
Київ, Україна

ПРОБЛЕМИ ДОСТОВІРНОСТІ OSINT-АНАЛІЗУ ІЗ ЗАСТОСУВАННЯМ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ ТА ПІДХОДИ ДО ЇХ МІНІМІЗАЦІЇ

У тезі розглянуто проблему перевірки інформації, яку генеративний штучний інтелект використовує та узагальнює під час OSINT-аналізу. Такі інструменти допомагають швидше опрацьовувати відкриті матеріали, але іноді помиляються, пропускають важливі деталі або наводять неточні посилання. На основі документів NIST визначено основні ризики та запропоновано просту послідовність перевірки результатів: від пошуку першоджерела до оцінки того, наскільки висновок підтверджений фактами. Наголошено, що ШІ може допомагати аналітику, але не замінює його відповідальності за достовірність інформації. Запропоновані рекомендації потребують практичної перевірки.

Ключові слова: OSINT, генеративний штучний інтелект, кіберрозвідка, достовірність інформації, відкриті джерела, перевірка фактів.

OSINT (Open Source Intelligence) — це збирання та аналіз інформації з відкритих джерел. У кібербезпеці такими джерелами можуть бути повідомлення про вразливості, звіти про атаки, технічні статті та відкриті бази індикаторів компрометації. Сьогодні для роботи з такими матеріалами дедалі частіше використовують генеративний штучний інтелект (ШІ). Він може швидко підсумувати великий текст, виділити головне або допомогти сформувати запитання для подальшого пошуку. Однак швидка відповідь ще не означає правильну. Особливо це важливо тоді, коли отриману інформацію планують використати у звіті про кіберзагрозу.

Мета роботи — з’ясувати, через які помилки результати OSINT-аналізу із застосуванням ШІ можуть виявитися недостовірними, і запропонувати зрозумілий порядок їх перевірки. Для цього проаналізовано рекомендації NIST, порівняно типові ризики та способи їх виявлення. Дослідження є аналітичним: власні випробування моделей ШІ та вимірювання їхньої точності не проводилися.

Одна з основних проблем — так звані «галюцинації» ШІ. Модель може впевнено повідомити факт, якого насправді немає, або послатися на джерело, яке не підтверджує її слова. У документі NIST AI 600-1 це явище описано як confabulation [1]. Наприклад, під час аналізу кіберінциденту ШІ може переплутати номери CVE, об’єднати відомості про різні атаки або без достатніх підстав назвати ймовірного зловмисника. Тому навіть детальна і переконлива відповідь потребує перевірки. Саму відповідь ШІ не можна вважати доказом.

Інша проблема стосується самих відкритих джерел. Частина матеріалів швидко застаріває, частина містить помилки, а деякі публікації можуть бути навмисно оманливими. Якщо модель спирається на такі дані, її висновок також може бути хибним. Крім того, людина іноді довіряє готовій відповіді лише тому, що вона написана впевнено і виглядає логічною. У рекомендаціях NIST

звертається увага як на достовірність інформації, так і на ризик надмірної довіри до автоматизованих результатів [1; 2].

Щоб зменшити ці ризики, пропонується зіставляти кожен тип помилки з конкретною перевіркою. Якщо ШІ називає факт — потрібно знайти його в першоджерелі. Якщо наводить посилання — відкрити його і переконатися, що документ справді містить потрібні відомості. Якщо йдеться про вразливість або інцидент — перевірити дату, адже ситуація могла змінитися. Якщо висновок зроблено з короткої цитати — прочитати повний контекст. А коли одна й та сама новина з'являється на кількох сайтах, слід з'ясувати, чи не переповідають вони одне неперевірене повідомлення. Наявність кількох посилань сама по собі ще не гарантує правдивості інформації.

На основі цього аналізу пропонується порядок роботи з п'яти кроків. Спочатку аналітик визначає, яку саме інформацію потрібно з'ясувати. Далі знаходить відкриті матеріали та зберігає посилання на них разом із датами. Третій крок — використовує ШІ для стислого викладу або впорядкування зібраного, не передаючи йому конфіденційних даних. Четвертий — перевіряє важливі твердження за першоджерелами, а за можливості й за незалежними повідомленнями. П'ятий — записує висновок і окремо зазначає, які факти підтверджені, а які залишаються припущеннями.

Для звіту корисно розділяти інформацію на підтверджену, непідтверджену та спростовану. Припущення аналітика варто позначати окремо. Такий поділ допомагає читачеві зрозуміти, на що можна спиратися, а що ще потрібно перевірити.

Запропонований порядок відповідає загальному підходу NIST AI RMF: ризики використання ШІ потрібно виявляти, оцінювати та контролювати залежно від конкретного завдання [2]. Для фахівців SOC і кіберрозвідки це означає, що ШІ доцільно використовувати як помічника в опрацюванні інформації, а не як остаточне джерело висновків. Водночас описаний порядок поки що є рекомендацією. Щоб визначити, наскільки він зменшує кількість помилок і скільки часу потребує, необхідна окрема практична перевірка.

Отже, головна умова достовірного OSINT-аналізу — не лише можливості ШІ, а й якість джерел та уважність аналітика. У роботі виділено основні ризики й запропоновано послідовність дій для перевірки результатів. Надалі варто порівняти аналіз однакових відкритих матеріалів із застосуванням цієї процедури та без неї. Це допоможе оцінити її реальну користь.

Список використаних джерел

1. NIST. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1. 2024. URL: https://doi.org/10.6028/NIST.AI.600-1

2. NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. 2023. URL: https://doi.org/10.6028/NIST.AI.100-1

МОХОНЬКО Діана
студентка групи ІКБмз-11, ННІКтЗІ, ДУІКТ
Київ, Україна

ЗАХИСТ КРИТИЧНОЇ ІНФРАСТРУКТУРИ. НОРМАТИВНО-МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ КІБЕРЗАХИСТУ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ В УКРАЇНІ

Ця робота висвітлює сучасні підходи до захисту об’єктів критичної інфраструктури в Україні в умовах актуальних кіберзагроз. Тут проаналізовано оновлену нормативно-правову базу, що регламентує категоризацію об’єктів, Оцінюються ризики кібербезпеки та перехід від КСЗІ до авторизації ІКС. Описуються етапи формування базових і цільових профілів безпеки за стандартами NIST, а також процедуру залучення експертів для підтвердження відповідності та подальшого занесення систем до переліку авторизованих.

Ключові слова: критична інфраструктура, кіберзахист, авторизація ІКС, профілі безпеки, оцінювання ризиків, категоризація об’єктів, стандарти NIST.

В умовах сучасних загроз захист критичної інфраструктури є стратегічним завданням для держави. Законодавством України визначено критичну інфраструктуру, яка охоплює об’єкти та системи, важливі для економіки, оборони й національної безпеки. Оскільки елементи інфраструктури мають різний ступінь впливу на державні функції, законодавство передбачає їх категоризацію за рівнем критичності для визначення відповідних заходів захисту. За для визначення підходів, що забезпечать безпеку критичних об’єктів, державою проведено диференціацію критичності інфраструктури, та визначено чотири категорії. Державне значення - перша категорія (особливо важливі об’єкти), регіональне значення - друга категорія (життєво важливі об’єкти), місцеве значення - третя категорія (важливі об’єкти), локальне значення (необхідні об’єкти) – четверта категорія [1]. Категорія критичності об'єкта безпосередньо визначає рівень вимог до захисту його інформаційно-комунікаційних систем (ІКС). Для реалізації цієї політики Державну службу спеціального зв’язку та захисту інформації України (Держспецзв’язок) визначено одним із окремих органів, на які покладається вирішення питань формування та/або реалізації державної політики у сфері захисту критичної інфраструктури. Реалізуючи свої повноваження Держспецзв'язок керується нормативно-правовою базою, встановленою Кабінетом Міністрів України[2]. Порядок затвердження профілів безпеки визначає саме Кабінет міністрів України та визначає умови обробки інформації з обмеженим доступом. Обробка такої інформації може здійснюватися лише в авторизованих системах з безпеки або шляхом отримання сертифіката відповідності стандарту інформаційної безпеки.

Держспецзв’язок уповноважено впроваджувати нову модель захисту засновану на профілях безпеки, яка замінює застарілу систему КСЗІ. Для урегулювання питання профілів безпеки було видано накази [3, 4], якими затверджено базові профілі безпеки для систем в яких обробляється інформація різних типів.

Авторизація ІКС має певні етапи. Найпершим етапом є визначення ІКС та її периметру. Після визначення ІКС та її периметру розпочинається етап оцінки ризиків. На цьому етапі проводиться дослідження системи щодо технічного функціонування системи, розміщення, обслуговування, використання та захисту ІКС. На базі цієї інформації будуються моделі загроз, та моделі порушника, які є складовими звіту з оцінки ризиків для ІКС. Для встановлення єдиного підходу до ідентифікації та оцінювання ризиків кібербезпеки розроблено методику оцінювання ризиків кібербезпеки. По завершенню оцінки ризиків формується звіт з оцінки ризиків до ІКС та затверджується керівником організації. Звіт є базою для наступного етапу на якому проводиться визначення типів інформації яка обробляється в ІКС та проводиться категоріювання безпеки ІКС. Категоріювання проводиться для визначення рівня критичності інформації та самої інформаційної системи. Рівнем критичності характеризується потенційна шкода, що може бути нанесена організації в разі втрати або порушення конфіденційності, цілісності та доступності інформації, що обробляється, зберігається і передається в ІКС.

Органами регулювання затверджено два варіанти базових профілів для ІКС в яких обробляється відкрита та конфіденційна інформація і ІКС в яких обробляється службова інформація. Після визначення базового профілю необхідно обрати засоби захисту інформації та визначитись які саме політики, процедури та інструменти будуть використовуватись у захисті ІКС. Не менш складним є наступний етап це створення цільового профілю безпеки (ЦПБ). ЦПБ це базовий профіль з додатковими засобами захисту з огляду на категоризацію системи та оцінку інформаційних ризиків, також ЦПБ має обов’язково включати вимоги захисту галузевих профілів, якщо такі були розроблені та затверджені галузевим регулятором. Тобто по суті ЦПБ це є базовий профіль з посиленими вимогами щодо заходів захисту інформації. Будь який профіль базуються на каталозі контролів NIST (National Institute of Standards and Technology), зокрема SP 800-53/NIST CSF, та включає наступні класи заходів захисту інформації: управління доступом, обізнаність і навчання, аудит і підзвітність, оцінювання, акредитація та моніторинг безпеки, управління конфігурацією, планування безперервної роботи, ідентифікація та автентифікація, реагування на інциденти, технічне обслуговування, захист носіїв інформації, фізичний захист і захист робочого середовища, планування безпеки, менеджмент інформаційної безпеки, кадрова безпека, повноваження на обробку персональних даних, оцінка ризику, придбання системи та послуг, системний і комунікаційний захист, цілісність системи та інформації, управління ризиками ланцюга поставок. Для реалізації заходів захисту безпеки по кожному з класів для ІКС розробляється політики та процедури для забезпечення вимог нормативних документів технічного захисту інформації. Окрім розробки нормативних документів виконуються технічні налаштування ІКС у відповідності до вимог ЦПБ. Надалі відбувається залучення експерта, вимоги до якого визначені Держспецзв’язком. Основні вимоги передбачають: наявність ліцензії на провадження господарської діяльності з надання послуг у сфері криптографічного та/або технічного захисту інформації,

наявність дозволу на проведення робіт з технічного захисту інформації для власних потреб або відповідних кваліфікаційних документів, відповідність персоналу кваліфікаційним критеріям у сфері кіберзахисту та аудиту інформаційних систем. Роботи які проводить експерт складаються з: перевірки повноти та коректності розробки ЦПБ для конкретної ІКС, оцінки реального стану впровадження організаційних та технічних заходів захисту, аналізу кібербезпеки та відповідність галузевим стандартам. За результатами проведених робіт надаються рекомендації щодо усунення недоліків, якщо такі були виявлені, а також звіт з оцінки дотримання вимог ЦПБ. Лише у разі позитивних результатів оцінювання дотримання вимог ЦПБ, організація має право подавати ІКС на авторизацію. Підставою для включення ІКС в перелік авторизованих систем з безпеки є подання власником ІКС автиразійного листа за зразком. Держспецзв'язок розглядає авторизаційний лист протягом десяти днів, і повідомляє організацію про своє рішення щодо включення ІКС у перелік.

У ході дослідження теми та аналізу нормативних документів, було визначено вимоги які регулюють процес авторизації ІКС, визначення критичності організацій, інформації, інформаційних систем. В результаті опрацювання теми було сформовано послідовність процесів та дій які необхідно виконати для забезпечення захисту інформації організаціями які є критичними для держави та суспільства. Виконання вимог ЦПБ забезпечує захист критичної інфраструктури для уникнення катастрофічних наслідків та неможливості держави виконувати свої базові функції. Запровадження системи авторизації надає гнучкості моделі управління кіберризиками на основі стандартів NIST. Для сфери інформаційних технологій це означає впровадження динамічного моніторингу замість статичного оцінювання. Створення та виконання ЦПБ дозволить оптимізувати архітектуру захисту сучасних хмарних та сервісних рішень (SaaS/PaaS - Software-as-a-Service (програмне забезпечення як послуга)/Platform-as-a-Service (платформа як послуга), що використовуються в державному секторі, забезпечуючи їхню стійкість до актуальних кіберзагроз.

1. Закон України «Про критичну інфраструктуру».
2. Постанова Кабінету Міністрів України від 18 червня 2025 р. № 712 «Про затвердження Порядку авторизації з безпеки інформаційних, електронних комунікаційних, інформаційно-комунікаційних, технологічних систем».
3. Наказ Адміністрації Держспецзв'язку від 30.06.2025 № 409 "Про затвердження базового профілю безпеки системи, де обробляється відкрита або конфіденційна інформація"
4. Наказ Адміністрації Держспецзв'язку від 02.07.2025 № 419 "Про затвердження базового профілю безпеки системи, де обробляється службова інформація".

Стебловський Г.В. Земляков С.О.
Студент групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ВАЖЛИВІСТЬ ЗАХИСТУ КІНЦЕВИХ ПРИСТРОЇВ НА ОСНОВІ ESET INSPECT & PROTECT

У роботі розглядається критична важливість захисту кінцевих пристроїв у бізнес-середовищі, що трансформувалося через масовий перехід до віддаленої роботи. Проаналізовано ключові кіберзагрози, вразливості віддалених робочих місць та зміну парадигми безпеки на підхід «Post-breach». Особливу увагу приділено комплексним рішенням ESET класу EPP та EDR та їх основним інструментам. Обґрунтовано ефективність використання платформ ESET PROTECT та ESET Inspect для глибокого моніторингу, своєчасного виявлення інцидентів і швидкого відновлення даних, що гарантує безперервність бізнесу.

Ключові слова: захист кінцевих пристроїв, кібербезпека, ESET, EPP, EDR, кіберзагроза, шкідливе програмне забезпечення

Зміни у моделях ведення бізнесу змусили компанії переглянути свої підходи до корпоративної безпеки. Поняття захисту кінцевих пристроїв (Endpoint Protection) охоплює процес моніторингу та убезпечення від кіберзагроз різноманітних пристроїв, до яких належать настільні комп’ютери, ноутбуки, смартфони, планшети тощо [1]. Якщо раніше ці пристрої знаходилися здебільшого всередині корпоративного офісу, де діяв надійний периметральний захист, то сьогодні ситуація змінилася [1, 2].

Пандемія та масовий перехід до віддаленого або гібридного форматів роботи сприяли тому, що працівники можуть виконувати свої обов'язки з будь-якого місця: з дому, з готелю або кав’ярні [2]. Водночас політика Bring Your Own Device (BYOD) дозволила співробітникам підключати власні пристрої до корпоративної мережі [1]. Як наслідок, кінцеві пристрої опинилися за межами захищеного корпоративного периметра та перетворилися на першу лінію оборони бізнесу, а також на головну ціль для зловмисників [1, 2].

Традиційна архітектура IT-безпеки передбачала, що головний офіс компанії є найбільш захищеним: тут встановлено системи логування, моніторингу, фаєрволи, а також ключові ресурси, такі як корпоративна пошта та CRM-системи. Філіали зазвичай отримують захищений доступ до цих ресурсів через спеціальне обладнання. Проте віддалені локації (де працюють співробітники поза офісом) стали найменш захищеними зонами підприємства [2].

Основними викликами для безпеки кінцевих точок є [2]:

«Білі плями». Це сліпі зони в поводженні з робочими комп’ютерами, які компанія не може контролювати. Наприклад, співробітник може встановити шкідливу програму, використати заборонені інструменти на корпоративному ноутбуці або витягнути жорсткий диск та зробити копію всіх доступних робочих систем [2].

Витік конфіденційної інформації. Співробітник може випадково або навмисне надіслати конфіденційні дані компанії конкурентам. Це можуть бути

навіть найпростіші документи на кшталт цінових пропозицій або фінансового звіту, компрометація яких вже завдасть шкоди бізнесу [2].

«Shadow IT». Це будь-які додатки, які використовують співробітники, але які не контролює IT-підрозділ компанії. Наприклад, різноманітні месенджери або програми для ведення задач (task management), візуалізації даних тощо. Загроза полягає у тому, що такі продукти можуть мати вразливості «нульового дня», а працівник вноситиме туди частково конфіденційні дані компанії [2].

При цьому вартість організації кібератак є вкрай низькою, що робить їх масовими та доступними для зловмисників. Віруси-шифрувальники (Ransomware) можна купити на тіньовому ринку всього за 66 доларів (або за 30% від прибутку) [2]. Атаки типу Denial of Service (DoS), які блокують доступ до ресурсів, коштують у середньому близько 102 доларів на день [2]. Водночас найдорожчою є атака через вразливість «нульового дня» (0days) — її вартість може сягати 350 000 доларів, оскільки для неї ще не створено механізмів захисту [2].



Рис. 1. Вартість організації кібератак [2]

Для ефективної протидії сучасним загрозам програмне забезпечення має працювати на різних рівнях. На ринку існують кілька класів рішень [1]:

Endpoint Protection Platform (EPP). Платформи для базового захисту, що запобігають атакам та виявляють шкідливу активність.

Endpoint Detection and Response (EDR). Рішення, які в режимі реального часу відстежують та збирають дані про кінцеві точки для миттєвого реагування на інциденти.

Mobile Threat Defense (MTD). Важливий інструмент для захисту смартфонів та планшетів, актуальність якого зростає через BYOD.

Advanced Threat Protection (ATP). Захист від прогресивних, багатоступеневих загроз.

Сучасна система безпеки повинна включати низку компонентів: поведінкову аналітику для виявлення загроз без використання відомих сигнатур, контроль додатків та фаєрвол для сегментації мережі, повне шифрування дисків (FDE) для захисту даних від несанкціонованого доступу, а також «пісочницю»

(Sandbox) для ізольованої перевірки підозрілих файлів [1]. Для боротьби з фішингом необхідна URL-фільтрація, а для убезпечення віддалених працівників — безпечний віддалений доступ (VPN) [1].

Окрім того, парадигма кібербезпеки змінилася з «Pre-breach» (спроби виключно не допустити хакерів у мережу) на «Post-breach» (дослідження та аналіз атак, які вже відбулися, для запобігання їхньому подальшому поширенню) [2]. Для посилення захисту від витоку даних компанії також застосовують технології DLP (Data Loss Prevention) та брокерів безпеки хмарного доступу (CASB), які разом з EDR створюють поведінкові моделі користувачів і виявляють аномалії [2].

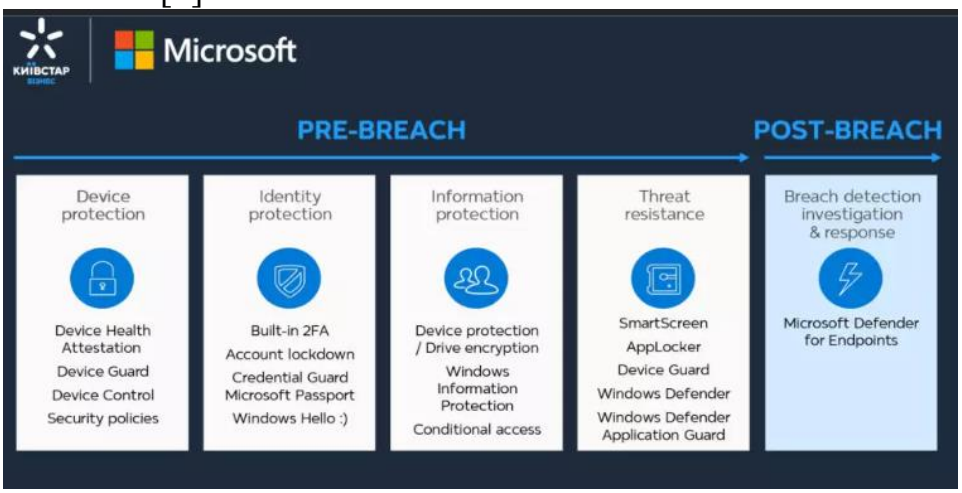


Рис. 2. Парадигма кібербезпеки [2]

Технології ESET пропонують потужний захист кінцевих точок завдяки багаторівневому підходу, який використовує можливості глобальної мережі виявлення загроз та машинного навчання [3]. Шкідливе програмне забезпечення аналізується та блокується на трьох етапах життєвого циклу загрози: до виконання (pre-execution), під час виконання (during execution) та після виконання (post-execution) [3].

Управління всією інфраструктурою безпеки, включно з мобільними пристроями, здійснюється з єдиної централізованої консолі ESET PROTECT [3].

Для реалізації підходу «Post-breach» (аналіз вже здійснених атак та складних загроз) ESET пропонує інструмент ESET Inspect [2, 4]. Головне меню платформи ESET Inspect містить ключові інструменти для моніторингу та розслідування інцидентів [4]:

Панель інструментів (Dashboard) для моніторингу загального статусу системи [4].

Комп’ютери (Computers) для відстеження стану кожної окремої кінцевої точки [4].

Виявлені об’єкти (Detections) та Інциденти (Incidents) для детального вивчення загроз і керування реакцією на них [4].

Виконувані файли (Executables) та Сценарії (Scripts) для глибокого аналізу процесів і коду, що запускаються в інфраструктурі (важливо для виявлення прихованих загроз) [4].

Пошук (Search) та Сповіщення (Notifications) дозволяють оперативно знаходити аномалії та реагувати на них [4].

В умовах розділеного корпоративного периметра та постійного зростання кількості і якості кібератак, кінцеві пристрої є найвразливішим елементом IT-інфраструктури підприємства [1, 2]. Стандартні інструменти захисту вже не здатні забезпечити належний рівень безпеки від цілеспрямованих атак, програм-вимагачів або експлойтів «нульового дня» [2, 3].

Впровадження комплексних платформ класу EPP та EDR на основі рішень ESET дозволяє підприємствам закрити «білі плями» у захисті віддалених працівників [2, 3]. Завдяки багаторівневому підходу, модулям відновлення після атак, захисту від безфайлового шкідливого ПЗ [3] та глибокому моніторингу через платформи ESET PROTECT і ESET Inspect [3, 4], бізнес отримує змогу не лише запобігати інцидентам, а й ефективно розслідувати їх [2, 4]. Інвестиція в комплексний Endpoint Security сьогодні — це гарантія безпеки та стабільності бізнесу завтра.

Перелік використаних джерел

1. What is Endpoint Protection? - Check Point Software. Check Point Software. URL: https://www.checkpoint.com/cyber-hub/threat-prevention/what-is-endpoint-security/ (дата звернення: 20.10.2026).

2. Endpoint Protection: як вберегти робочі місця від кіберзагроз | Kyivstar Business Hub. Kyivstar Business Hub – корпоративний блог для бізнесу. URL: https://hub.kyivstar.ua/articles/endpoint-protection-yak-vberegty-robochi-misczya-vid-kiberzagroz (дата звернення: 20.10.2026).

3. Modern Endpoint protection - Multilayered defense for company | ESET. ESET. URL: https://www.eset.com/us/business/solutions/endpoint-protection/?srsltid=AU7gw4XgJwOgp6TJ-zkwgxgT2DADwjD5Cm3FTmyjb-fk-SFTJ_35q13X (дата звернення: 20.10.2026).

4. ESET Inspect. ESET Online Help. URL: https://help.eset.com/ei_cloud/uk-UA/main_menu.html?index.html (дата звернення: 20.10.2026).

Гуленов І.В.
Студент групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ПІДВИЩЕННЯ БЕЗПЕКИ VPN ЗА ДОПОМОГОЮ БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ

У роботі розглядається проблема забезпечення безпеки віртуальних приватних мереж за допомогою багатофакторної автентифікації в умовах віддаленої роботи. Проаналізовано основні принципи функціонування VPN та визначено ключові вразливості систем, що покладаються виключно на традиційну парольну автентифікацію, зокрема загрози викрадення облікових даних, соціальної інженерії та фішингу. Обґрунтовано необхідність впровадження MFA як критично важливого рівня захисту, який комплексно перевіряє особу користувача за допомогою кількох незалежних факторів. Досліджено різні методи MFA та їх вплив на захист корпоративних мереж. На основі опрацьованих джерел доведено, що інтеграція MFA та VPN створює надійну архітектуру безпеки, що блокує більшість автоматизованих атак і гарантує безпечний віддалений доступ.

Ключові слова: VPN, MFA, 2FA, безпека мережі, віддалений доступ, кібербезпека, шифрування, облікові дані, фішинг, SASE, ZTNA, біометрія, апаратні токени.

Із переходом до гібридних та віддалених моделей роботи, організації стикаються з серйозними викликами у сфері кібербезпеки. Статистика свідчить, що значна частина працівників працює поза межами традиційного офісу, отримуючи доступ до критично важливих корпоративних даних з домашніх мереж, публічних точок Wi-Fi та особистих пристроїв [1]. У цьому контексті віртуальні приватні мережі (далі — VPN) стали фундаментальним інструментом для забезпечення захищеного з'єднання. Вони створюють зашифрований тунель між віддаленим користувачем та корпоративною мережею [2]. Однак використання лише VPN з традиційною автентифікацією (логін і пароль) вже не відповідає сучасним вимогам безпеки. Зловмисники постійно вдосконалюють методи викрадення облікових даних. Саме тому інтеграція VPN із багатофакторною автентифікацією (MFA/2FA) стала стандартом захисту віддаленого доступу [3, 4].

Захищений віддалений доступ — це комплекс технологій, політик і контролів, що дозволяють користувачам підключатися до мережі організації з-поза її фізичних меж, захищаючи дані від перехоплення [2]. Цей процес базується на трьох основних стовпах: автентифікації, тунелюванні та шифруванні.

Автентифікація перевіряє, ким є користувач. Тунелювання логічно відокремлює корпоративний трафік від загального інтернет-трафіку. Шифрування (переважно з використанням надійних алгоритмів, таких як AES-256) гарантує, що навіть у разі перехоплення пакетів даних зловмисники побачать лише набір випадкових символів [2].

Віддалені VPN-з'єднання історично були основною технологією такого доступу. Вони змінюють IP-адресу пристрою таким чином, що користувач нібито знаходиться всередині офісної мережі [1]. Сьогодні, окрім традиційних VPN, компанії все частіше звертаються до сучасніших концепцій, таких як Zero Trust Network Access (ZTNA) та Secure Access Service Edge (SASE). ZTNA реалізує принцип «ніколи не довіряй, завжди перевіряй», надаючи доступ лише

до конкретних додатків, а не до всієї мережі, тоді як SASE об'єднує мережеві функції та функції безпеки в хмарну архітектуру [1, 2]. Незважаючи на це, VPN залишається основою для більшості корпоративних рішень.

Основна проблема VPN полягає в тому, що він створює безпечний тунель, але якщо «ключі» від цього тунелю потрапляють у чужі руки, вся система опиняється під загрозою [4]. VPN-шлюз перевіряє правильність введених даних, але без додаткових факторів не може гарантувати, що їх вводить легітимний працівник.

Звіт Verizon Data Breach Investigations Report зазначає, що 31% порушень безпеки пов'язані з використанням викрадених облікових даних [2]. Хакери застосовують фішинг, підроблені електронні листи, атаки на підбір паролів (brute-force) та використання вразливостей протоколів віддаленого робочого столу (RDP) [1]. За відсутності MFA, один скомпрометований пароль дозволяє зловмиснику не лише увійти в мережу, але й вільно переміщуватися всередині неї (lateral movement), оскільки стандартний VPN надає широкий рівень доступу до корпоративних ресурсів [2, 4].

Для усунення ризиків, пов'язаних зі слабкими або вкраденими паролями, застосовується багатофакторна або двофакторна автентифікація. MFA вимагає від користувача підтвердження своєї особи щонайменше двома незалежними способами [3].

Ці фактори класифікуються за трьома основними категоріями:

Те, що ви знаєте (фактор знання) — це пароль, PIN-код або відповідь на секретне запитання.

Те, що ви маєте (фактор володіння) наприклад смартфон, апаратний токен (YubiKey), смарт-картка або додаток-автентифікатор.

Те, ким ви є (фактор властивості) біометричні показники, такі як відбиток пальця, розпізнавання обличчя або сканування сітківки ока [3].

На сьогодні існує широкий спектр методів реалізації MFA. До них належать одноразові паролі (OTP) через SMS або електронну пошту, спеціальні додатки-автентифікатори (Google Authenticator, Authy, Microsoft Authenticator), push-повідомлення на мобільний пристрій, а також використання біометрії та апаратних токенів [3]. Згідно з даними компанії Microsoft, використання багатофакторної автентифікації здатне заблокувати до 99,9% автоматизованих атак на облікові записи [2].

Поєднання MFA та VPN формує потужну лінію оборони. Під час спроби встановлення VPN-з'єднання процес проходить у кілька етапів: спочатку користувач вводить базові облікові дані (логін/пароль), після чого VPN-шлюз (часто через інтеграцію з RADIUS або SAML) надсилає запит на другий фактор автентифікації [2, 4]. Лише після підтвердження на мобільному пристрої або через токен генерується зашифрований тунель [4].

Таке поєднання має низку переваг:

MFA зменшує поверхню атаки — при успішній фішинговій атаці не дозволить зловмиснику отримати доступ без фізичного пристрою працівника [3].

Відповідність нормативним вимогам, таким як GDPR, HIPAA або PCI DSS 4.0, які прямо вимагають впровадження шифрування та суворих засобів контролю доступу (включно з MFA) для віддалених підключень [2].

Сучасні системи MFA дозволяють враховувати контекст (розташування користувача, час доби, стан безпеки пристрою). Якщо пристрій не має оновленого антивірусу або підключення відбувається з підозрілої локації, доступ може бути заблоковано [1, 3].

Однак розгортання MFA для VPN має свої виклики. Головний з них — баланс між безпекою та користувацьким досвідом. Надто складні процедури входу можуть викликати опір співробітників [3, 4]. Для вирішення цієї проблеми компаніям варто впроваджувати зручні методи, такі як push-повідомлення або біометрична автентифікація. Також фахівці рекомендують переходити на phishing-resistant MFA (стійку до фішингу автентифікацію), таку як ключі стандарту FIDO2, які захищають навіть від складних MitM-атак, де хакери намагаються перехопити одноразовий код [4].

Отже, віртуальні приватні мережі є надійною основою для шифрування даних у транзиті, проте вони беззахисні перед сучасними методами компрометації облікових даних, якщо використовуються як самостійний інструмент.

Таким чином інтеграція MFA з VPN є не просто рекомендацією, а критичною необхідністю для будь-якої організації. Вона перетворює процес входу з простої перевірки пароля на комплексну верифікацію особи користувача. Вибираючи правильні та зручні методи MFA, компанії можуть ефективно захистити конфіденційні дані, відповідати регуляторним вимогам і гарантувати безперебійну та безпечну роботу своїх віддалених команд.

Перелік використаних джерел

1. The Various Types of Secure Remote Access Solutions - Check Point Software. Check Point Software. URL: https://www.checkpoint.com/cyber-hub/network-security/types-of-secure-remote-access-solutions/ (дата звернення: 20.10.2026).

2. What Is Secure Remote Access and VPN? | FatPipe. FatPipe SD-WAN & Cloud | 99.98% WAN Reliability & VPN Failover. URL: https://www.fatpipeinc.com/resources/secure-remote-access-and-vpn (дата звернення: 20.10.2026).

3. Multi-Factor Authentication and VPNs: A Winning Combination. Check Point SASE. URL: https://sase.checkpoint.com/blog/network/multi-factor-authentication-and-vpn (дата звернення: 20.10.2026).

4. How to Secure VPN Access With 2FA: Benefits & Methods. miniOrange Security Software: IAM, PAM, IGA, MDM, DLP, Data Privacy & AI Security. URL: https://www.miniorange.com/blog/vpn-security-how-2fa-helps-to-secure-your-vpn/ (дата звернення: 20.10.2026).

Завгородня Євгенія Ярославівна
студентка 6 курсу, групи УБДМ-61

КІБЕРСТІЙКІСТЬ ОБ'ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ УКРАЇНИ: УРОКИ АТАКИ INDUSTROYER2 ТА ПІДХІД НА ОСНОВІ NIST CSF 2.0

Проаналізовано динаміку кібератак на критичну інфраструктуру України та розглянуто спробу атаки на електропідстанції енергопостачальника із застосуванням шкідливого програмного забезпечення Industroyer2 як приклад цілеспрямованої операції проти промислових систем керування. На основі аналізу виявлено етапи атаки, на яких її можна було б виявити, та запропоновано підхід до підвищення кіберстійкості операторів критичної інфраструктури, що спирається на шість функцій NIST CSF 2.0 і вимоги Закону України «Про критичну інфраструктуру».

Критична інфраструктура – енергетика, електронні комунікації, фінансові послуги, охорона здоров'я та інші сфери – забезпечує безперервність життєдіяльності суспільства, а кібернетичне ураження її об'єктів може спричинити каскадні наслідки для економіки, безпеки та населення. Закон України «Про критичну інфраструктуру» від 16.11.2021 № 1882-IX заклав правові засади її захисту та покладає на операторів обов'язки щодо захисту об'єктів і повідомлення про інциденти [2]. В умовах повномасштабної війни ці вимоги щодня перевіряються на практиці, тому аналіз реальних атак і пошук практичних підходів до підвищення кіберстійкості операторів є актуальним завданням. Метою роботи є аналіз тактики зловмисників на прикладі спроби атаки на електропідстанції та формування системи заходів кіберстійкості для операторів критичної інфраструктури на основі NIST CSF 2.0.

Проблема має системний характер. За даними CERT-UA, у 2024 році було опрацьовано 4315 кіберінцидентів, що на 69,8 % більше, ніж у 2023 році, коли їх було 2541; фахівці відзначають стійку тенденцію до зростання кібератак передусім на критично важливу інфраструктуру [3]. Отже, ризик для операторів не є гіпотетичним, і підготовка до таких атак має бути частиною повсякденної діяльності, а не реакцією на окремі події.

Показовим прикладом цілеспрямованої операції проти промислових систем є спроба вимкнення електропостачання у квітні 2022 року. За результатами спільного розслідування ESET і CERT-UA група Sandworm намагалася застосувати шкідливе програмне забезпечення Industroyer2 проти високовольтних електричних підстанцій в Україні, а дослідники з високим ступенем впевненості приписують атаку саме цій групі [1]. Первинне проникнення в мережу енергопостачальника відбулося не пізніше лютого 2022 року, а деструктивні дії були заплановані на 8 квітня 2022 року [1]. Industroyer2 реалізує лише протокол IEC-104 для зв'язку з промисловим обладнанням, а одночасно використовувалися деструктивні програми CaddyWiper, ORCSHRED, SOLOSHRED та AWFULSHRED, орієнтовані на системи Windows, Linux і Solaris [1]. За оцінкою ESET, Industroyer2 створено на основі коду Industroyer, який використовували для порушення енергопостачання в Україні ще у 2016 році [1]. Після запуску Industroyer2 на тому самому комп'ютері передбачалося виконання

CaddyWiper для знищення слідів [1]. Атаку вдалося відвернути завдяки спільним діям CERT-UA та ESET [1].

Хронологія завершальної фази атаки показує, наскільки стислим було вікно для реагування. 8 квітня 2022 року о 14:58 за UTC на частині комп'ютерів із Windows, а також на системах Linux і Solaris було розгорнуто деструктивне програмне забезпечення, о 15:02 оператор створив заплановане завдання для запуску Industroyer2, о 16:10 розпочалося його виконання з метою відключення електропостачання в одному з регіонів України, а о 16:20 на тому самому комп'ютері було запущено CaddyWiper для стирання слідів [1]. Отже, від створення завдання до запуску минуло трохи більше години, і спиратися на ручне виявлення на цьому етапі недоцільно: ефективніше виявляти зловмисників на ранніх стадіях, коли вони лише закріплюються в мережі.

Аналіз інциденту дозволяє зробити кілька висновків для практики захисту. По-перше, між початковою компрометацією та запланованим моментом деструктивних дій минуло щонайменше кілька тижнів, тобто в оператора існувало вікно для виявлення присутності зловмисників. По-друге, атака поєднувала інструменти для IT- та OT-середовищ: засоби взаємодії з промисловим обладнанням і програми знищення даних. По-третє, Industroyer2 є гнучко конфігурованим, тому для кожного нового середовища зловмисникам потрібно перекомпілювати його [1]; це означає, що суто сигнатурні засоби виявлення недостатні, а захист має спиратися також на моніторинг поведінки та мережевого трафіку промислових протоколів. По-четверте, повторне використання коду 2016 року свідчить, що інструменти, раніше застосовані проти української енергетики, не зникають, а розвиваються, тому попередні інциденти слід розглядати як джерело вимог до систем моніторингу та реагування.

Структурувати відповідні заходи доцільно за NIST Cybersecurity Framework 2.0, випущеним у лютому 2024 року. Він містить шість функцій: Govern (управління), Identify (ідентифікація), Protect (захист), Detect (виявлення), Respond (реагування) та Recover (відновлення), причому нова функція Govern охоплює стратегію управління кіберризиками, ролі, політики та нагляд [4]. Для оператора критичної інфраструктури, який може зіткнутися з атаками, подібними до розглянутої, пропонується така конкретизація цих функцій. Слід зазначити, що NIST CSF 2.0 є добровільною рамкою і не замінює вимог національного законодавства, тому його положення доцільно застосовувати разом з нормами Закону № 1882-IX [2, 4].

Govern: закріпити відповідальність керівництва за кіберризики технологічних систем і узгодити внутрішні політики з вимогами Закону № 1882-IX щодо захисту об'єктів і повідомлення про інциденти [2]. Identify: провести інвентаризацію IT- та OT-активів, зокрема підстанційного обладнання, і зв'язків між сегментами мережі, оскільки невідомий актив неможливо ні захистити, ні відстежувати. Protect: сегментувати корпоративну та технологічну мережі, обмежити доступ з першої до другої, вимагати багатофакторної автентифікації

для віддаленого доступу та дозволяти передавання команд за протоколом IEC-104 лише з визначених вузлів.

Detect: вести моніторинг мережевого трафіку промислових протоколів, створення нових запланованих завдань на серверах і нетипової активності облікових записів, щоб скоротити час перебування зловмисників у мережі. Respond: розробити та регулярно відпрацьовувати плани реагування, що передбачають оперативний обмін інформацією з CERT-UA; розглянутий випадок показує цінність такої взаємодії, оскільки саме спільна робота оператора, національної команди реагування та дослідників дала змогу відвернути атаку [1]. Recover: зберігати резервні копії в ізольованому від робочої мережі середовищі та підготувати процедури ручного керування обладнанням на випадок втрати автоматизованих систем.

З практичного погляду заходи доцільно впроваджувати поетапно. На першому етапі, що не потребує значних витрат, варто завершити інвентаризацію IT- та OT-активів, переглянути правила доступу між корпоративною та технологічною мережами, запровадити багатофакторну автентифікацію для віддаленого доступу та призначити відповідальних за кіберризики. На другому етапі розгортаються засоби моніторингу промислового трафіку й централізованого збирання журналів, а також створюються ізольовані резервні копії. На третьому етапі проводяться навчання зі сценаріями, подібними до розглянутої атаки, за участі технологічного персоналу та служби безпеки. Для оцінювання прогресу доцільно відстежувати частку проінвентаризованих активів і середній час виявлення нетипової активності.

Таким чином, спроба атаки Industroyer2 демонструє, що сучасні операції проти критичної інфраструктури планують заздалегідь, вони поєднують IT- та OT-компоненти і залишають оператору час для виявлення, якщо моніторинг налагоджено. Запропонована конкретизація функцій NIST CSF 2.0 може бути використана операторами як основа для планування заходів кіберстійкості та узгодження їх із вимогами національного законодавства. Подальші дослідження доцільно спрямувати на розроблення показників оцінювання зрілості кіберстійкості операторів у різних секторах.

Список використаних джерел

1. Industroyer2: Industroyer reloaded. ESET Research. 12.04.2022. URL: https://blog.eset.ie/2022/04/12/industroyer2-industroyer-reloaded (дата звернення: 21.10.2026).

2. Про критичну інфраструктуру: Закон України від 16.11.2021 № 1882-IX. URL: https://zakon.rada.gov.ua/laws/show/1882-20#Text (дата звернення: 21.10.2026).

3. В Україні за рік кількість кібератак зросла на 70 %. Forbes Ukraine. 08.01.2025. URL: https://forbes.ua/news/v-ukraini-za-rik-kilkist-kiberatak-zrosla-na-70-nayposhirenishi-tipi-intsidentiv-i-golovni-tsili-khakeriv-08012025-26137 (дата звернення: 21.10.2026).

4. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. Gaithersburg: National Institute of Standards and Technology, 2024. DOI: https://doi.org/10.6028/NIST.CSWP.29.

Шимчук Г. В.,
старший викладач каф. КН,
ТНТУ ім. І. Пулюя,
Тернопіль, Україна
Яцків П.О.,
студент групи КТ-12,
ТНТУ ім. І. Пулюя,
Тернопіль, Україна

КОНТЕКСТНА МОДЕЛЬ КЕРУВАННЯ ДОСТУПОМ ДЛЯ ПІДВИЩЕННЯ БЕЗПЕКИ ХМАРНИХ СЕРЕДОВИЩ

Запропоновано контекстну модель керування доступом до хмарних ресурсів, що поєднує принципи Zero Trust, оцінювання ризику ідентичності, стану пристрою або робочого навантаження, поведінкових відхилень, мережевого контексту та чутливості ресурсу. Модель передбачає чотири реакції: дозвіл, посилену автентифікацію, обмежений доступ і відмову. На модельному наборі зі 120 запитів 59 із 60 ризикових сценаріїв активували додаткові засоби контролю, тоді як серед нормальних запитів лише два потребували повторної перевірки.

Ключові слова: Cloud Security Alliance, Zero Trust, API, IAM, MFA

Поширення хмарних і мультихмарних архітектур істотно змінює модель захисту інформаційних систем. Традиційне припущення, що ресурс усередині корпоративного периметра є більш довіреним, поступово втрачає практичну цінність: користувачі працюють віддалено, сервіси взаємодіють між різними хмарами, а значна частина доступу виконується не людьми, а сервісними обліковими записами та програмними навантаженнями. У звіті Cloud Security Alliance серед найкритичніших хмарних ризиків визначено помилки конфігурації, недоліки керування ідентичностями та доступом, небезпечні API, недостатню видимість середовища і неконтрольований спільний доступ [1]. Тому захист хмарної системи доцільно будувати не лише на статичних ролях, а на безперервному аналізі контексту кожного запиту.

Архітектура Zero Trust формалізує підхід, за якого мережеве розташування саме по собі не є підставою для довіри. NIST SP 800-207 вимагає виконувати автентифікацію й авторизацію до встановлення сесії та розглядати ресурс, користувача і пристрій як окремі об’єкти політики [2]. Для cloud-native і multi-cloud середовищ NIST SP 800-207A додатково акцентує ідентичність застосунків і сервісів, використання API-шлюзів, service mesh і точок примусового застосування політик [3]. Практична проблема полягає в тому, що класична рольова модель RBAC здебільшого відповідає на питання «чи має суб’єкт право?», але недостатньо враховує «за яких умов це право використовується саме зараз».

Метою роботи є розроблення компактної контекстної моделі прийняття рішення про доступ, придатної для інтеграції з IAM, API Gateway, service mesh або хмарним policy engine. Для кожного запиту пропонується формувати п’ять нормованих показників у діапазоні від 0 до 1. Перший характеризує ризик ідентичності: наявність MFA, тип облікового запису, надмірність ролі, вік і

спосіб отримання облікових даних. Другий відображає ризик пристрою або робочого навантаження: стан оновлень, відповідність політикам, наявність підтвердженої workload identity та цілісність середовища виконання. Третій описує поведінкове відхилення: нетиповий час входу, різку зміну частоти запитів, звернення до незвичних ресурсів. Четвертий враховує мережевий контекст: репутацію адреси, географічну аномалію, використання анонімізаторів, відкритість точки входу. П’ятий показник задає чутливість ресурсу.

Інтегральний ризик визначається як зважена сума чотирьох динамічних складових: 30 % припадає на ризик ідентичності, 25 % – на стан пристрою або workload, 25 % – на поведінкове відхилення, 20 % – на мережевий контекст. Отримане значення додатково масштабується коефіцієнтом чутливості ресурсу, щоб однаковий контекст приводив до жорсткішого рішення для секретів, адміністративних API чи баз даних. Запропоновано чотири інтервали реакції. За низького ризику запит дозволяється відповідно до принципу найменших привілеїв, за помірного виконується step-up перевірка, наприклад повторна MFA, за підвищеного видається короткоживучий токен з обмеженням операцій до читання або окремого API, за високого ризику доступ блокується, а подія передається до SIEM/SOC.

Запропонована схема не скасовує RBAC або ABAC, а доповнює їх контекстним рівнем. Спочатку policy engine перевіряє базове право суб’єкта на ресурс. Лише після цього виконується оцінювання поточного ризику і, за необхідності, звужується вже надане право. Такий порядок зменшує небезпеку «розростання привілеїв» і компрометації легітимних облікових записів. Окремо передбачено використання короткоживучих ідентифікаційних даних для сервісів, регулярну ротацію секретів, журналювання всіх рішень policy engine та кореляцію з CloudTrail/Azure Activity Log або аналогічними джерелами. Це забезпечує можливість ретроспективної перевірки, чому конкретний запит було дозволено, обмежено чи відхилено.

Для первинної перевірки поведінки моделі проведено відтворюваний модельний експеримент із фіксованим генератором випадкових чисел. Сформовано 120 запитів чотирьох типів: 60 нормальних, 20 з аномальним контекстом, 20 із ознаками компрометації ідентичності та 20 від недовіреного робочого навантаження. Для кожного запиту генерувалися п’ять нормованих параметрів відповідно до типового профілю сценарію. Модельний набір не є заміною журналів реальної хмарної платформи, його завданням було перевірити логіку порогових рішень та оцінити співвідношення між чутливістю контролю і кількістю зайвих перевірок.

Середнє значення інтегрального ризику для нормальних запитів становило 0,133, для аномального контексту – 0,407, для компрометованої ідентичності – 0,546, для недовіреного workload – 0,521. Із 60 нормальних запитів 58 були дозволені без додаткових дій, а два спрямовані на посилену автентифікацію, жоден нормальний запит не був обмежений або заблокований. Серед 60 ризикових сценаріїв лише один запит потрапив до інтервалу безумовного

дозволу. Для 17 запитів застосовано step-up автентифікацію, для 40 – обмежений режим, для двох – повну відмову. Отже, 98,3 % модельних ризикових запитів активували додатковий контроль, а частка зайвих перевірок серед нормальних становила 3,3 %.

Наведені результати слід інтерпретувати як перевірку алгоритмічної логіки, а не як оцінку промислового продукту. У реальній інфраструктурі ваги та пороги мають адаптуватися до типу сервісу, профілю користувача і вартості помилкового блокування. Сучасні дослідження динамічного Zero Trust доступу підтверджують доцільність поєднання атрибутів, оцінки довіри та контекстних правил для хмарних середовищ [4].

Практичне впровадження моделі може виконуватися поетапно. На першому етапі збираються журнали IAM, API-шлюзів, мережевих засобів і засобів керування кінцевими пристроями без автоматичного блокування. На другому етапі policy engine працює у режимі рекомендацій, що дає змогу відкоригувати ваги та пороги. Після стабілізації профілів активуються step-up перевірка і обмеження токенів, а повна відмова залишається для явно критичних комбінацій ознак. Такий перехід зменшує ризик масових хибних блокувань і дозволяє інтегрувати контекстний контроль у чинні процеси DevSecOps.

Окремим завданням є захист самого механізму прийняття рішень. Телеметрія, що використовується для розрахунку ризику, повинна надходити з автентифікованих джерел, мати узгоджені часові мітки та бути захищеною від підміни. Журнали рішень доцільно зберігати окремо від робочих журналів застосунків, щоб компрометація одного сервісу не дозволяла приховати зміну політики. Для критичних сервісів необхідно визначити режим роботи під час недоступності policy engine: «fail closed» для адміністративних операцій і контрольований короткочасний кеш для операцій, де повне блокування створює неприйнятний ризик доступності. Важливим є також контроль версій policy-as-code, обов’язковий review змін і автоматичне тестування політик перед розгортанням.

Таким чином, для хмарних середовищ доцільно доповнювати статичне призначення ролей контекстним ризик-орієнтованим рішенням для кожної сесії. Запропонована модель узгоджується з принципами Zero Trust, враховує ідентичність, стан workload, поведінкові та мережеві ознаки й чутливість ресурсу. Модельний експеримент показав можливість відокремлювати нормальні запити від ризикових без значного збільшення кількості зайвих блокувань.

Перелік посилань:

1. Cloud Security Alliance. Top Threats to Cloud Computing 2024. 2024. URL: https://cloudsecurityalliance.org/artifacts/top-threats-to-cloud-computing-2024.

2. Rose S., Borchert O., Mitchell S., Connelly S. Zero Trust Architecture. NIST Special Publication 800-207. 2020. https://doi.org/10.6028/NIST.SP.800-207.

3. Chandramouli R., Butcher Z. A Zero Trust Architecture Model for Access Control in Cloud-Native Applications in Multi-Location Environments. NIST Special Publication 800-207A. 2023. https://doi.org/10.6028/NIST.SP.800-207A.

4. Wang R., Li C., Zhang K., Tu B. Zero-trust based dynamic access control for cloud computing. Cybersecurity. 2025. Vol. 8. Art. 12. https://doi.org/10.1186/s42400-024-00320-x.

Шимчук Г. В.,
старший викладач каф. КН,
ТНТУ ім. І. Пулюя,
Тернопіль, Україна
Яцків П.О.,
студент групи КТ-12,
ТНТУ ім. І. Пулюя,
Тернопіль, Україна

АДАПТИВНИЙ ЗАХИСТ ІОТ-МЕРЕЖ НА ОСНОВІ ОЦІНКИ ДОВІРИ ПРИСТРОЇВ ТА ДИНАМІЧНОЇ МІКРОСЕГМЕНТАЦІЇ

Запропоновано адаптивний підхід до захисту IoT-мереж, який поєднує інвентаризацію пристроїв, оцінювання їхнього поточного ризику та динамічну мікросегментацію. Для кожного пристрою враховуються стан автентифікації, актуальність програмного забезпечення, зовнішня експозиція, поведінкові відхилення та критичність функції. У модельному експерименті з 60 вузлами динамічне обмеження зв'язків та ізоляція підозрілих пристроїв зменшили середню кількість скомпрометованих вузлів з 30,87 до 3,25, тобто приблизно на 89,5 %.

Ключові слова: шлюз, мережа, IoT, SDN-контролер, MQTT-брокер, VLAN-сегментація

Захист Internet of Things ускладнюється значною неоднорідністю пристроїв, довгими життєвими циклами, обмеженими обчислювальними ресурсами та залежністю від хмарних сервісів і шлюзів. У межах однієї мережі можуть одночасно працювати камери, сенсори, контролери, шлюзи та пристрої попередніх поколінь із різним рівнем підтримки криптографії й оновлень. Компрометація одного вузла створює ризик латерального переміщення до інших компонентів, особливо якщо мережа побудована як великий довірений сегмент. CISA окремо звертає увагу на небезпеку публічно доступних пристроїв, конфігурацій за замовчуванням, застарілого ПЗ та зайвої експозиції сервісів в Інтернеті [1].

Базова рекомендація NIST для IoT полягає в тому, що продукт повинен мати набір технічних можливостей, які дозволяють організації ідентифікувати пристрій, керувати його конфігурацією, захищати дані, обмежувати логічний доступ, безпечно оновлювати ПЗ та отримувати інформацію про стан кібербезпеки [2]. У редакції NIST IR 8259 Rev. 1, опублікованій у 2026 році, додатково підкреслено необхідність закладати кібербезпеку в життєвий цикл IoT-продукту до його введення в експлуатацію [3]. Однак навіть за наявності цих можливостей статична VLAN-сегментація не реагує на зміну стану конкретного вузла: пристрій після компрометації часто продовжує мати ті самі мережеві права, що і до інциденту.

Метою роботи є розроблення адаптивного механізму, в якому мережеві дозволи IoT-пристрою визначаються не лише його постійною належністю до сегмента, а й актуальним рівнем довіри. Для кожного вузла формується профіль, що включає п'ять груп ознак: надійність автентифікації та унікальність облікових даних, актуальність прошивки й наявність відомих вразливостей, ступінь зовнішньої експозиції портів і сервісів, відхилення мережевої поведінки

від звичного профілю, критичність функції пристрою. Високий ризик за будь-якою критичною ознакою повинен мати пріоритет над середнім інтегральним балом, наприклад використання заводського пароля або несанкціоноване відкриття керуючого порту одразу переводить вузол у режим обмеження.

Запропоновано три стани політики. У стані «довірений» пристрій взаємодіє лише з визначеними сервісами: MQTT-брокером, локальним шлюзом, DNS/NTP та сервером оновлень. У стані «обмежений» забороняється peer-to-peer взаємодія з однотипними IoT-вузлами, скорочується перелік дозволених протоколів і підвищується деталізація журналювання. У стані «карантин» зберігається лише керований канал до системи моніторингу та, за потреби, сервера відновлення. Перехід між станами виконується policy engine на основі телеметрії, а реалізація правил можлива засобами SDN, ACL на комутаторах, host firewall або політиками програмно визначеної мережі.

Ключовою властивістю підходу є мікросегментація за мінімально необхідними потоками, а не за фізичним місцем пристрою. Камера не повинна мати прямого доступу до інших камер лише тому, що вони розміщені в одній VLAN, сенсор має передавати телеметрію брокеру, але не ініціювати з'єднання з контролером, контролер повинен приймати команди лише від авторизованого сервісу керування. Така структура обмежує радіус наслідків компрометації. Дослідження гнучкої Zero Trust архітектури для IIoT також демонструє практичну доцільність поєднання мікросегментації та програмно визначеного керування політиками для гетерогенних середовищ [4].

Для кількісної первинної оцінки виконано модельний експеримент. Побудовано мережу з 60 IoT-вузлів, поділених на чотири функціональні групи по 15 пристроїв. У статичному варіанті вузли одного сегмента могли безпосередньо взаємодіяти між собою, а шлюзи забезпечували міжсегментні зв'язки. В адаптивному варіанті кожен вузол мав лише необхідні службові зв'язки з локальним шлюзом і обмеженим набором сусідів, після виявлення підозрілої поведінки скомпрометований вузол ізолювався через два модельні такти. Моделювалося поширення компрометації впродовж десяти тактів із ймовірністю успішного переходу через дозволений зв'язок 0,1. Проведено 500 повторів із фіксованим початковим зерном генератора.

За статичної сегментації середня кількість вузлів, що встигали бути скомпрометованими, становила 30,87 із 60, стандартне відхилення – 13,45, в окремих прогонах компрометація охоплювала всі 60 пристроїв. За динамічної мікросегментації з автоматичною ізоляцією середнє значення становило 3,25 вузла, стандартне відхилення – 3,51, а максимальне – 19. Відносне зменшення середньої кількості скомпрометованих вузлів становило приблизно 89,5 %. Результат пояснюється двома механізмами: меншою кількістю дозволених ребер між пристроями та обмеженим часом, протягом якого ризиковий вузол залишається активним.

Модельний характер експерименту накладає суттєві обмеження: ймовірність поширення задавалася штучно, не враховувалися різні класи експлойтів, затримки реального SDN-контролера та можливі помилки

детектування. Тому числовий результат не слід переносити безпосередньо на виробничу IoT-мережу. Водночас експеримент показує структурний ефект мікросегментації: навіть недосконалий механізм виявлення має значно менше потенційних напрямів латерального переміщення, якщо взаємодія між пристроями спочатку обмежена принципом least privilege.

Для практичного впровадження пропонується починати з пасивної інвентаризації: визначити всі IoT-вузли, виробника, версію прошивки, відкриті порти, типові адреси призначення та нормальний обсяг трафіку. Далі формуються дозволені комунікаційні шаблони для кожного класу пристроїв. На другому етапі вводиться оцінка довіри без автоматичного карантину, щоб визначити частоту помилкових спрацювань. Після калібрування в автоматичний режим переводяться лише однозначні події: поява забороненого сервісу, зміна MAC/IP без реєстрації, спроби сканування, масові з'єднання з новими адресами, використання небезпечних облікових даних. Для legacy-пристроїв, які не підтримують сучасну автентифікацію, компенсувальним контролем має бути жорсткіша мережна ізоляція.

Для різних типів загроз механізм використовує різні комбінації ознак. Масове сканування або спроба поширення ботнету проявляються зростанням кількості нових адрес призначення та нетипових портів, такий вузол доцільно негайно відокремлювати від peer-to-peer потоків. Поява невідомого пристрою з підробленою або незареєстрованою ідентичністю повинна блокувати доступ до брокерів керування ще до накопичення поведінкової статистики. Для пристрою з простроченою прошивкою ризик може підвищуватися поступово: спочатку обмежується вихід в Інтернет, потім, якщо вразливість стає критичною, дозволяється лише канал оновлення. Таким чином, адаптивна політика поєднує реактивний контроль подій із профілактичним зменшенням поверхні атаки.

Отже, ефективний захист IoT потребує переходу від статичної сегментації до моделі, де довіра до пристрою змінюється відповідно до його поточного стану. Запропонований підхід поєднує інвентаризацію, поведінковий контроль, мінімально необхідні мережеві зв'язки та автоматичний карантин. Модельний експеримент показав суттєве зменшення потенційного масштабу компрометації порівняно зі статичною схемою. Подальші дослідження доцільно спрямувати на перевірку механізму в Mininet/Containernet або фізичному IoT-стенді, оцінювання затримки прийняття рішень і розроблення адаптивних порогів для різних класів пристроїв.

Перелік посилань:

1. Cybersecurity and Infrastructure Security Agency. Internet Exposure Reduction Guidance. 2025. URL: https://www.cisa.gov/resources-tools/resources/exposure-reduction.

2. Fagan M., Megas K., Scarfone K., Smith M. IoT Device Cybersecurity Capability Core Baseline. NIST IR 8259A. 2020. https://doi.org/10.6028/NIST.IR.8259A.

3. Fagan M., Megas K., Cuthill B., Marron J., Hoehn B. Foundational Cybersecurity Activities for IoT Product Manufacturers. NIST IR 8259 Rev. 1. 2026. https://doi.org/10.6028/NIST.IR.8259r1.

4. Zanasi C., Russo S., Colajanni M. Flexible zero trust architecture for the cybersecurity of industrial IoT infrastructures. Ad Hoc Networks. 2024. Vol. 156. Art. 103414. https://doi.org/10.1016/j.adhoc.2024.103414.

Онопрієнко М.Ю
Студент УБДМ-61, ННІКБЗІ
Київ,Україна

МЕТОДИ І ТЕХНОЛОГІЇ КІБЕРАТАК, ОРГАНІЗОВАНИХ ЗА ПІДТРИМКИ НАЦІОНАЛЬНИХ ДЕРЖАВ ТА ЗАСОБИ ПРОТИДІЇ ЇМ

Анотація. У роботі розглянуто основні методи та технології кібератак, організованих за підтримки національних держав. Проаналізовано характерні етапи проведення цілеспрямованих кібероперацій, зокрема розвідку, отримання початкового доступу, закріплення, підвищення привілеїв, переміщення мережею та викрадення інформації. Розглянуто можливості застосування мережевих і операційних засобів захисту для виявлення та протидії таким атакам. Запропоновано комплексний підхід до виявлення кіберзагроз на різних етапах їх реалізації з використанням моніторингу мережі, контролю кінцевих пристроїв та централізованого аналізу подій безпеки.

Ключові слова: кібербезпека, кібератака, кібероперація, національна держава, APT, MITRE ATT&CK, мережева безпека, операційна безпека.

Сучасні інформаційні системи відіграють важливу роль у функціонуванні державних установ, підприємств та організацій. Одночасно зі зростанням залежності від цифрової інфраструктури збільшується кількість і складність кіберзагроз. Особливу увагу привертають атаки, пов’язані з діяльністю державних або пов’язаних із державами угруповань. Такі операції можуть мати розвідувальний, політичний, економічний або деструктивний характер і часто передбачають тривале перебування зловмисника в інформаційній системі. Актуальність дослідження таких атак також підтверджується сучасними оцінками кіберзагроз. За даними ENISA, діяльність state-nexus угруповань залишається одним із чинників формування кіберзагроз у Європі. Зокрема, у звітному періоді спостерігалося посилення операцій, пов’язаних із кіберрозвідкою, а державні угруповання продовжували використовувати вразливості та фішингові кампанії для отримання доступу до цільових систем.

Особливістю таких кібератак є те, що вони, як правило, не обмежуються використанням одного інструменту або однієї вразливості. Атака може складатися з декількох послідовних етапів, кожен з яких виконує окрему функцію. Для систематизації подібних дій доцільно використовувати базу знань MITRE ATT&CK, яка описує тактики та техніки зловмисників на основі реальних спостережень. Поточна Enterprise-матриця ATT&CK охоплює такі етапи, як розвідка, початковий доступ, виконання, закріплення, підвищення привілеїв, приховування діяльності, отримання облікових даних, виявлення ресурсів, переміщення мережею, збір та виведення інформації.

Першим етапом цілеспрямованої атаки зазвичай є збір інформації про потенційну жертву. Зловмисник може визначати доменні імена, IP-адреси, мережеві сервіси, типи операційних систем, програмне забезпечення та іншу інформацію, яка може бути використана для подальшого проникнення. Для цього можуть застосовуватися методи OSINT, аналіз відкритих джерел, сканування мережевої інфраструктури та дослідження інформації про працівників організації.

Після збору необхідної інформації здійснюється пошук способу отримання початкового доступу. Серед найбільш поширених варіантів можна виділити фішинг, компрометацію облікових даних та експлуатацію вразливостей програмного забезпечення. За даними ENISA, фішинг та експлуатація вразливостей залишаються серед основних шляхів отримання початкового доступу до інформаційних систем .

Метод кібератаки	**Основна мета**	**Засоби виявлення та протидії**
Розвідка	Збір інформації про цільову інфраструктуру	Моніторинг мережі, аналіз зовнішніх сервісів
Фішинг	Отримання облікових даних або початкового доступу	Антифішингові засоби, MFA, аналіз поштового трафіку
Експлуатація вразливостей	Отримання доступу до системи	Сканери вразливостей, IDS/IPS, своєчасне оновлення ПЗ
Підвищення привілеїв	Отримання розширених прав	Контроль привілеїв, аудит облікових записів
Переміщення мережею	Розширення контролю над інфраструктурою	Сегментація мережі, SIEM, моніторинг автентифікації
Викрадення даних	Отримання та передача конфіденційної інформації	DLP, аналіз мережевого трафіку, SIEM

Таблиця 1. Основні методи кібератак та засоби їх виявлення

Важливою особливістю цільових атак є використання соціальної інженерії. На відміну від масових фішингових розсилок, цільові повідомлення можуть бути адаптовані під конкретну організацію або працівника. Використання інформації, отриманої на етапі розвідки, дозволяє зробити повідомлення більш правдоподібним та збільшити ймовірність взаємодії користувача зі шкідливим вмістом.

Після отримання доступу до окремого пристрою зловмисник може намагатися закріпитися в системі та розширити свої можливості. Одним із напрямів є отримання підвищених привілеїв. Надалі може здійснюватися пошук інших пристроїв, серверів та облікових записів у корпоративній мережі. Отримання доступу до додаткових вузлів дозволяє поступово розширювати контроль над інформаційною системою.

Особливу небезпеку становить використання легітимних інструментів операційної системи. У такому випадку діяльність зловмисника може бути схожою на звичайні адміністративні дії, що ускладнює її автоматичне виявлення. Подібний підхід відображений у MITRE ATT&CK як використання штатних можливостей системи для виконання дій без необхідності застосування окремого шкідливого програмного забезпечення.

Показовим прикладом використання складного ланцюга технік є діяльність APT28, яку MITRE пов’язує з Головним управлінням Генерального штабу Збройних сил росії. Для цього угруповання описано використання різних технік, серед яких отримання інфраструктури, маніпуляції обліковими записами, підвищення привілеїв та інші способи розвитку атаки.

З огляду на багаторівневий характер таких атак, протидія їм повинна передбачати використання декількох взаємопов’язаних механізмів захисту. На мережевому рівні доцільним є використання міжмережевих екранів, систем виявлення та запобігання вторгненням, сегментації мережі та моніторингу мережевого трафіку. На рівні кінцевих пристроїв важливими є контроль виконання процесів, своєчасне встановлення оновлень, захист облікових записів та обмеження адміністративних привілеїв.

Не менш важливим є централізований збір та аналіз подій інформаційної безпеки. Поєднання журналів із мережевого обладнання, серверів, робочих станцій та інших компонентів дозволяє виявляти взаємопов’язані події, які окремо можуть не виглядати підозрілими. Наприклад, послідовність мережевого сканування, невдалих спроб автентифікації, отримання підвищених привілеїв та доступу до внутрішніх ресурсів може свідчити про розвиток цілеспрямованої атаки.

Такий підхід відповідає загальній логіці NIST Cybersecurity Framework 2.0, яка орієнтована на управління кіберризиками та передбачає комплексний підхід до ідентифікації, захисту, виявлення, реагування та відновлення.

Отже, кібератаки, організовані за підтримки національних держав, характеризуються комплексним використанням різних методів та технологій і можуть тривалий час залишатися непоміченими. Основною проблемою для систем захисту є не тільки виявлення окремих шкідливих дій, а встановлення взаємозв’язку між подіями, що відбуваються на різних рівнях інформаційної системи. Тому ефективна протидія таким загрозам потребує поєднання мережевих та операційних засобів захисту, централізованого моніторингу та аналізу поведінки зловмисника.

Подальше дослідження доцільно спрямувати на моделювання окремих сценаріїв кібератак у тестовому середовищі та оцінювання можливостей мережевих і операційних засобів щодо виявлення різних етапів кібероперації. Це дозволить визначити найбільш ефективні точки контролю та сформувати практичні рекомендації щодо підвищення стійкості корпоративних інформаційних систем.

Література

1. **ENISA Threat Landscape 2025**. European Union Agency for Cybersecurity, 2025. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025
2. **MITRE ATT&CK®**. Adversarial Tactics, Techniques, and Common Knowledge. URL: https://attack.mitre.org/
3. **APT28**. MITRE ATT&CK, Group G0007. https://attack.mitre.org/groups/G0007/?trk=public_post_comment-text/
4. **NIST Cybersecurity Framework (CSF) 2.0**. National Institute of Standards and Technology, 2024. https://csrc.nist.gov/pubs/sp/1299/

Ланде Дмитро Володимирович, доктор технічних наук, професор, завідувач кафедри інформаційної безпеки
Кузьмінський Віктор Сергійович, аспірант кафедри інформаційної безпеки
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ

ВИЛУЧЕННЯ ТРИПЛЕТІВ ПРО КІБЕРЗАГРОЗИ З ТЕКСТОВИХ ПОВІДОМЛЕНЬ ІЗ ГАРАНТОВАНОЮ МЕЖЕЮ ЧАСТКИ ПОМИЛОК

Анотація. Запропоновано метод вилучення триплетів «суб'єкт — зв'язок — об'єкт» про кіберзагрози з текстових повідомлень, який приймає лише кандидатів з оцінкою, не нижчою за поріг, перевірений на розмічених текстах джерела. Гарантовано, що середня по текстах частка хибних серед прийнятих з тексту триплетів не перевищує заданого рівня α з імовірністю щонайменше $1 - \delta$. Під нове джерело поріг перекалібровують без перенавчання великих моделей. На англомовних звітах CTINexus і українських публікаціях CERT-UA рівень перевищувався не більш як у 1,3 % повторів, а влучність за $\alpha = 0{,}3$ становила 80–81 % проти 45–74 % у мовної моделі без фільтра.

Графи знань про кіберзагрози дедалі частіше наповнюють автоматично: велика мовна модель (LLM) читає звіт і записує з нього триплети «суб'єкт — зв'язок — об'єкт», наприклад «UAC-0050 — використовує — RemcosRAT». Скільки серед них хибних, наперед невідомо, і ця частка залежить від джерела: у моделі Ministral 3 8B вона становила 26 % на англомовних звітах CTINexus [1] і 49 % на українських публікаціях CERT-UA. Чимало помилок стосуються не зв'язку, а самих сутностей: замість назви діяча чи шкідливої програми модель ставить загальне слово або уламок назви. Власна впевненість LLM у задачах про кіберзагрози погано узгоджується з часткою її правильних відповідей [2], а поріг, підібраний на розмічених текстах без статистичної перевірки, у наших дослідах перевищував заданий рівень помилок у 38–54 % випадкових поділів текстів.

Мета роботи — метод, який приймає з тексту лише ті триплети, для яких середня частка хибних не перевищує заданого аналітиком рівня α, і зберігає цю властивість під час переходу до нового джерела повідомлень.

Кандидати і оцінки. Кандидати бувають двох видів. Перші — пари сутностей: компактна модель GLiNER2 знаходить згадки сутностей, LLM відповідає, чи варта кожна з них вузла графа, а для пари — чи є між сутностями зв'язок, якого він типу і яка з них суб'єкт. Другі — триплети, які LLM виписує з тексту сама. Дві логістичні регресії зводять ознаки кандидатів до оцінок від 0 до 1, і про всіх кандидатів вирішує один спільний поріг (рис. 1).

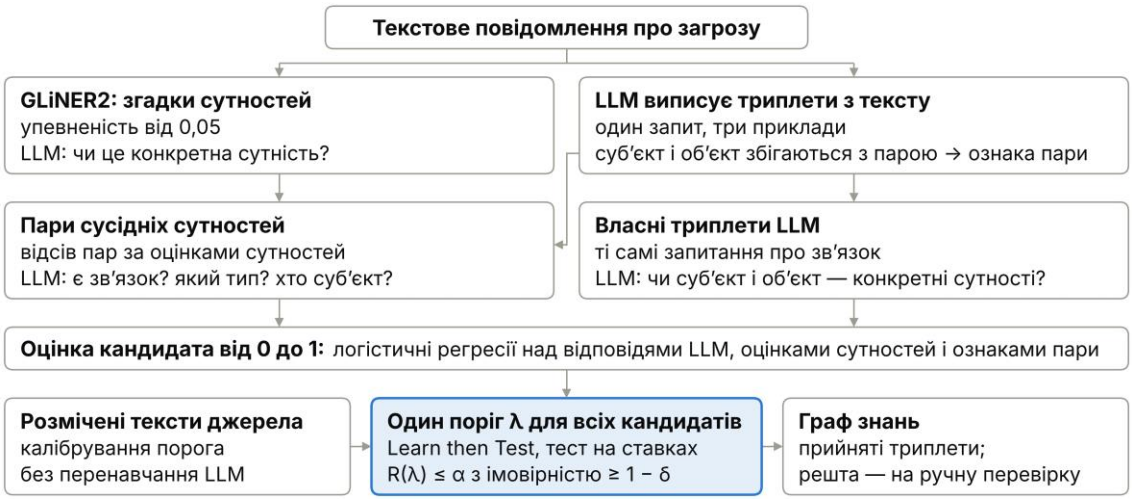


Рис. 1. Схема методу: один перевірений поріг для двох видів кандидатів

Строге означення правильного триплета. Триплет правильний, якщо текст прямо стверджує зв'язок між суб'єктом і об'єктом і обидва вони є конкретними сутностями, тобто називають річ, яку аналітик лишив би вузлом графа. Поріг перевіряють саме за такими мітками, тому гарантія охоплює і зв'язок, і обидві сутності, а оцінки сутностей лише допомагають прийняти більше правильних триплетів. Для порівняння, на CTINexus два послідовні фільтри (спершу сутності, потім зв'язки) з тим самим рівнем для кожного перевищували рівень для триплетів у 58 % повторів, а оцінки сутностей у складі оцінки пари збільшували відносну повноту (частку знайдених зі спільного переліку правильних зв'язків) в 1,3–4 рази.

Вибір і перевірка порога. Одиницею гарантії є текст, бо помилки в межах одного звіту залежні: у них спільні сутності, контекст і запит до моделі. Ризик порога — сподівана частка хибних серед прийнятих з тексту триплетів, усереднена по текстах, з яких щось прийнято. Розмічені калібрувальні тексти ділять на дві частини: чверть пропонує кілька порогів, решта перевіряє їх тестом на основі ставок [3] за схемою Learn then Test [4]. Якщо тексти незалежні й однаково розподілені, ймовірність прийняти поріг із ризиком, не меншим за α, не перевищує заданого δ; якщо перевірки не пройшов жоден поріг, метод нічого не приймає. Пройти перевірку можна лише за достатньої кількості текстів, з яких щось прийнято: за $\delta = 0{,}1$ це щонайменше 88, 43, 20 і 13 текстів для $\alpha = 0{,}05$; 0,1; 0,2; 0,3. Ця межа задає найменший обсяг розмітки під нове джерело.

Адаптація до джерела. Великі моделі під нове джерело не перенавчають: досить заново відкалібрувати поріг на розмічених текстах цього джерела, а щоб приймати більше, перенавчити й логістичні регресії. Без перекалібрування гарантія не переноситься: у постановці з еталонними сутностями пороги, перевірені на CTINexus, на текстах CERT-UA перевищували рівень у 55–71 % повторів, а після перекалібрування на текстах CERT-UA — не більш як в 1 % повторів.

Результати. Метод перевірено на англомовному наборі CTINexus і на корпусі зі 158 публікацій CERT-UA за 2018–2026 роки, зібраному в цій роботі, з двома LLM на 8 млрд параметрів (Ministral 3 8B і Llama 3.1 8B), що працюють локально; на CERT-UA — за протоколом, записаним наперед. Середня по тестових текстах частка хибних перевищувала α не більш як у 1,3 % із 300 випадкових поділів (з Llama на CERT-UA — у 4,3 %). За α = 0,3 влучність методу (частка правильних серед прийнятих) становила 80–81 % на обох наборах проти 73–74 % на CTINexus і 51 % (Ministral) чи 45 % (Llama) на CERT-UA у LLM без фільтра. Перевага не пояснюється лише більшими обчисленнями: самоперевірка і самоузгодженість LLM за зіставних витрат часу мали на CERT-UA з Ministral на 19–23 пункти нижчу влучність. Рівень α задає компроміс між влучністю і відносною повнотою: за α = 0,3 на CERT-UA з Ministral метод знаходить удвічі менше зв’язків, ніж LLM без фільтра, а за α = 0,5 — і більше зв’язків, і з вищою влучністю. Мітки для калібрування поставила мовна модель Claude за письмовим правилом; за мітками незалежного фахівця з кібербезпеки на вибірці текстів CERT-UA середня частка хибних за α = 0,3 становить 0,218 (верхня одностороння 95-відсоткова межа 0,265), тобто нижча за α.

Висновки. Метод дає аналітикові один керований параметр — допустиму частку помилок — і триплети для графа знань із передбачуваною середньою часткою хибних. Під нове джерело потрібні лише розмічені тексти цього джерела, а конвеєр вилучення працює локально, тож тексти не залишають організацію. Код, мітки й протоколи відкрито (https://github.com/datch1588/cti-gate-experiment). Подальші дослідження: більші LLM і послідовна перевірка, що вказує, коли потрібне перекалібрування.

Список використаних джерел

1. Cheng Y., Bajaber O., Tsegai S. A., Song D., Gao P. CTINexus: Automatic cyber threat intelligence knowledge graph construction using large language models. *Proceedings of the 10th IEEE European Symposium on Security and Privacy (EuroS&P)*. 2025. URL: https://arxiv.org/abs/2410.21060.

2. Mezzi E., Massacci F., Tuma K. Large language models are unreliable for cyber threat intelligence. *Availability, Reliability and Security (ARES 2025)*. LNCS, vol. 15993. Springer, 2025. P. 343–364. DOI: 10.1007/978-3-032-00627-1_17.

3. Waudby-Smith I., Ramdas A. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B*. 2024. Vol. 86, no. 1. P. 1–27. DOI: 10.1093/jrsssb/qkad009.

4. Angelopoulos A. N., Bates S., Candès E. J., Jordan M. I., Lei L. Learn then test: Calibrating predictive algorithms to achieve risk control. *The Annals of Applied Statistics*. 2025. Vol. 19, no. 2. P. 1641–1662. DOI: 10.1214/24-AOAS1998.

Четвертуха Н.Р. аспірант
Отенко В.І. к.т.н., доцент
Національний університет «Львівська політехніка», Львів, Україна

ВИБІР ІНСТРУМЕНТУ ДИНАМІЧНОГО АНАЛІЗУ ДЛЯ ШКІДЛИВИХ ПРОГРАМ, ЯКІ ПРИХОВУЮТЬ СВОЮ ПОВЕДІНКУ

Шкідливі програми з техніками протидії аналізу можуть приховувати поведінку під час аналізу й поводитися як безпечні. У роботі з'ясовано, якими перевірками такі зразки помічають кожен із трьох інструментів динамічного аналізу, різних за принципом роботи. Для цього створено 210 тестових програм для Linux на основі п'яти типових перевірок середовища. Їх виконано під засобом відстеження ptrace, емулятором qemu-user і засобом динамічного бінарного інструментування valgrind. Засіб відстеження й емулятор помітні для двох перевірок із п'яти, valgrind для трьох. Жодна перевірка не помітила всіх трьох інструментів. Для якісного аналізу досить знайти перевірки в коді зразка, щоб до його запуску обрати інструмент, якого вони не помічають.

Ключові слова: динамічний аналіз, поведінкова інваріантність, протидія аналізу.

Постановка проблеми. Моделі машинного навчання для виявлення шкідливого програмного забезпечення (ШПЗ) здебільшого навчають на поведінці зразків в ізольованому середовищі для контрольованого виконання програм і запису їхніх дій (пісочниці). Навчання на таких записах має сенс тоді, коли зразок у пісочниці поводиться так само, як на комп'ютері користувача. Однак зразки з техніками протидії аналізу цієї умови не виконують [1, 2]. Зразок із такою технікою спершу перевіряє, чи не стежить за ним відлагоджувач, емулятор або інший засіб аналізу. Якщо такий засіб знайдено, шкідлива функціональність лишається прихованою. Моделі, навчені на даних із пісочниці, розпізнають понад 90 % шкідливих зразків на таких самих даних, але лише близько 17 % на даних із реальних комп'ютерів [3]. Однією з причин цієї відмінності є техніки протидії аналізу. Мета роботи полягає в тому, щоб з'ясувати, якими перевірками шкідлива програма помічає кожен із трьох інструментів динамічного аналізу, різних за принципом роботи.

Поведінкова інваріантність. Запропонована авторами фасетна таксономія технік обфускації ШПЗ [4] описує кожну техніку за кількома незалежними вимірами. Один із них, поведінкова інваріантність, показує, чи поводиться програма під наглядом засобу аналізу так само, як на комп'ютері користувача. Значення «Повна» мають техніки, які лише ускладнюють будову коду й не перевіряють середовища, тому поведінка програми не залежить від наявності системи аналізу. Значення «Часткова» мають техніки протидії аналізу, тобто програма з такою технікою змінює поведінку, коли помічає засоби аналізу. Значення «Відсутня» мають техніки, що запускають шкідливу частину лише на заздалегідь обраній машині, поведінка такої програми може відрізнятись навіть між декількома звичайними комп'ютерами.

Методика експерименту. Для експерименту згенеровано 210 тестових програм для Linux. Вони відтворюють логіку ШПЗ із технікою протидії аналізу, але не містять шкідливого навантаження. Кожна програма спершу перевіряє середовище, а потім виконує одну з двох гілок. Якщо ознак аналізу немає, програма виконує «шкідливу» гілку, яка імітує шкідливу функціональність.

Якщо ознаки є, виконується підставна гілка, яка поводиться як безпечна програма. Кожна гілка виводить власну текстову мітку, тож за нею видно, яка гілка виконалася. Набір складається з 30 базових програм, кожна в семи варіантах. Платформу Linux обрано тому, що інструменти аналізу для неї відкриті, а експеримент легко відтворити.

Контрольний варіант не містить перевірок і завжди виконує «шкідливу» гілку. П'ять наступних варіантів містять по одній типовій перевірці. Дві з них з'ясовують, чи стежить за процесом інша програма, наприклад відлагоджувач. Перша читає поле TracerPid у файлі /proc/self/status, а друга виконує виклик PTRACE_TRACEME. Ще дві шукають сторонні бібліотеки в адресному просторі процесу. Одна читає карту пам'яті у файлі /proc/self/maps, а інша перевіряє змінну середовища LD_PRELOAD. П'ята перевірка помічає уповільнення виконання. Вона сім разів вимірює час виконання ділянки коду й бере найменше значення, щоб зменшити вплив фонового навантаження. Сьомий варіант поєднує всі п'ять перевірок.

Кожну програму запускали без засобів аналізу та під трьома інструментами, що працюють за різними принципами. Засіб відстеження приєднується до процесу через системний виклик ptrace. Емулятор режиму користувача qemu-user перекладає інструкції програми в інструкції процесора комп'ютера, де його запущено. Засіб динамічного бінарного інструментування valgrind вставляє власний код між інструкціями програми. Без засобів аналізу всі програми виконують «шкідливу» гілку, і цей результат слугує еталоном. Останній рядок табл. 1 показує, яка частка з 210 програм помітила інструмент і перейшла до підставної гілки.

Результати. Зміну поведінки програм під інструментами подано в табл. 1.

Таблиця 1. Зміна поведінки програм під інструментами аналізу

Перевірка середовища	**Засіб відстеження ptrace**	**Емулятор qemu-user**	**valgrind (DBI)**
відсутня (контроль)	–	–	–
поле TracerPid	+	–	–
виклик PTRACE_TRACEME	+	+	–
час виконання	–	+	+
карта пам'яті процесу	–	–	+
змінна LD_PRELOAD	–	–	+
усі п'ять перевірок	+	+	+
Помітили інструмент	**0,43**	**0,43**	**0,57**

Знак «+» у таблиці означає, що поведінка, записана під інструментом, не відповідає поведінці на комп'ютері користувача, а «–» що вона не змінилася. Засіб відстеження помічають обидві перевірки стеження за процесом. Часова перевірка засіб відстеження не помічає, бо він не зупиняє процес на ділянці, час якої вимірюють. Емулятор помічають і часова перевірка, бо переклад інструкцій уповільнює виконання, і виклик PTRACE_TRACEME. Обидві перевірки стеження за процесом valgrind не помічають, але загалом він найпомітніший із трьох, оскільки його викривають часова перевірка, перевірка карти пам'яті, де видно його бібліотеки, і перевірка змінної LD_PRELOAD, через яку valgrind

підвантажує ці бібліотеки в процес. Також показано, що одна перевірка може помітити два інструменти одразу (PTRACE_TRACEME і часова перевірка). Однак жодна перевірка не помічає всіх трьох.

Правило вибору інструменту. Табл. 1 дає змогу обирати інструмент для кожного зразка окремо. Правило можна описати тим, що зразок помічає інструмент лише тими перевірками, які в ньому є. Якщо жодна з них інструмент не помічає, зразок під ним поводиться так само, як на комп'ютері користувача, тож записана поведінка придатна для навчання моделі.

Спершу в зразку знаходять перевірки середовища. Для цього досить статичного аналізу, тобто дослідження програми без її запуску, оскільки перевірки є частиною коду програми. Потім, використовуючи табл. 1, можна проаналізувати рядки знайдених перевірок і обрати стовпець, у якому в усіх цих рядках стоїть «–».

Якщо перевірок не знайдено, це не обов'язково означає, що їх немає, оскільки статичний аналіз міг їх пропустити, наприклад через обфускацію коду. Тоді зразок запускають під двома інструментами, які помічають різні перевірки, наприклад під засобом відстеження та valgrind. За табл. 1 кожна з п'яти перевірок помічає один із цих двох інструментів. Тому зразок з однією перевіркою під одним інструментом виконає «шкідливу» гілку, а під другим підставну. Така розбіжність дозволяє зробити висновок про використання технік протидії аналізу. Якщо ж поведінка під обома інструментами збігається, то перевірок у зразку або немає, або їх кілька і разом вони помічають обидва інструменти, що вимагає ручного аналізу.

Висновки. Поведінка шкідливої програми, записана під інструментом динамічного аналізу, не завжди збігається з її поведінкою на комп'ютері користувача. Це залежить від двох чинників. Перший чинник стосується самої техніки. Програма, яка не перевіряє середовища, поводиться однаково під будь-яким інструментом, і це випливає з означення поведінкової інваріантності. Другий чинник стосується інструмента, оскільки кожен інструмент має свої ознаки, за якими його можна помітити. Програма помічає інструмент, лише якщо перевіряє саме ці ознаки. Жоден із трьох інструментів не залишився непомітним для всіх перевірок, однак жодна перевірка не помітила всіх трьох використаних інструментів. З цього можна зробити висновок, що інструмента, придатного для будь-якого зразка, немає. Обирати його слід для кожного зразка окремо, тобто знайти в ньому перевірки й узяти інструмент, якого вони не помічають.

Перелік посилань:

1. Afianian A., Niksefat S., Sadeghiyan B. et al. Malware dynamic analysis evasion techniques: a survey. *ACM Computing Surveys*. 2019. Vol. 52, No. 6. Art. 126. DOI: 10.1145/3365001.

2. Galloro N., Polino M., Carminati M. et al. A systematical and longitudinal study of evasive behaviors in Windows malware. *Computers & Security*. 2022. Vol. 113. Art. 102550. DOI: 10.1016/j.cose.2021.102550.

3. Kaya Y., Chen Y., Botacin M. et al. ML-based behavioral malware detection is far from a solved problem. *2025 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*. Copenhagen : IEEE, 2025. P. 921–940. DOI: 10.1109/SaTML64287.2025.00056.

4. Четвертуха Н. Р., Отенко В. І. Розроблення багатовимірної фасетної таксономії технік обфускації шкідливого програмного забезпечення. *Кібербезпека: освіта, наука, техніка*. 2026. № 1(33). С. 683–696. DOI: 10.28925/2663-4023.2026.33.1254.

Опірський І.Р.
завідувач кафедри ЗІ, НУ «Львівська політехніка», Львів, Україна
Стефанків А.В.
аспірант кафедри ЗІ, НУ «Львівська політехніка», Львів, Україна

МЕХАНІЗМИ ВІДБОРУ ОЗНАК СИГНАЛІВ У СИСТЕМАХ РАДІОЧАСТОТНОГО ФІНГЕРПРИНТИНГУ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Розглянуто механізми відбору ознак у системах радіочастотного фінгерпринтингу IoT-пристроїв на основі нейронних мереж. Запропоновано класифікацію за двома незалежними ознаками — наявністю множини іменованих кандидатів і суб'єктом відбору (дослідник або навчання), що дає чотири механізми: відбір рукотворних ознак, відбір через інваріантності конвеєра обробки, вбудований відбір і навчання на необробленому сигналі. Для кожного механізму визначено характерні помилки та засоби контролю. Сформульовано критерій допустимості перетворень конвеєра та показано, що в наскрізних системах відповідальність за відбір ознак переходить до планування експерименту.

Ключові слова: радіочастотний фінгерпринтинг, відбір ознак, нейронна мережа, IoT, інваріантність, аугментація даних, навчання представлень.

Метою ідентифікації за радіочастотними відбитками (Radio Frequency Fingerprint Identification, RFFI) є встановлення ідентичності бездротового пристрою за особливостями його сигналу, зумовленими виробничими відхиленнями компонентів аналогового тракту передавача. На відміну від протокольних ознак (адрес та ідентифікаторів), підміна яких є тривіальною, такий відбиток визначається апаратурою і значно складніше піддається програмній підробці.

Розглянемо модель прийнятого сигналу $r = R(H(T_d(x)) + n)$, де x — еталонний сигнал протоколу, T_d — апаратні спотворення передавача d (розбаланс квадратур, витік гетеродина, нелінійність підсилювача, фазовий шум, зсув несучої) [1], H — канал, R — тракт приймача, n — адитивний шум (Рис. 1).



Рис. 1. Стадії формування результуючого сигналу та їх вплив на формування відбитку пристрою.

Отже, шукана функція класифікатора має залежати від T_d і бути інваріантною до H, R, n та вмісту x.

Класифікація механізмів відбору ознак. У загальному випадку механізми відбору ознак поділяють на явні та неявні. Явний відбір обирає підмножину з наперед заданої множини іменованих кандидатів (ознак, представлень, ділянок сигналу), і його результат можна переглянути; неявний відбір такої множини не

має. Доцільно ввести другу вісь класифікації — суб'єкт відбору: дослідник або модель під час навчання. Таким чином отримуємо чотири класи механізмів (Таблиця 1).

Таблиця 1. Класифікація механізмів відбору ознак

	Відбір дослідником	**Відбір навчанням**
Явний відбір	Відбір рукотворних ознак, вибір ділянки пакета та виду представлення	Вбудовані методи: L1-регуляризація, маски
Неявний відбір	Відбір через інваріантності конвеєра	Навчання на основі необробленого сигналу

Явний відбір дослідником. Механізм явного відбору дослідником передбачає попередній аналіз сигналів, обчислення ознак із фізичним змістом і відбір тих, що найбільше впливають на точність класифікації, фільтровими або обгортковими методами [2]. Модель з явно сконструйованим вектором ознак передбачуваніша та прозоріша, оскільки ключові параметри визначено наперед; для навчання потрібно значно менше даних; дослідник може свідомо виключити нестабільні ознаки.

Водночас застосування підходу обмежене змодельованими чи виявленими спотвореннями та залежить від точності протокольно-специфічних оцінювачів. Специфічною вадою явного відбору в системах RFF є те, що розрізнювальна здатність, виміряна в межах одного сеансу запису, переоцінює нестабільні ознаки. Наприклад, класифікація пристроїв за зсувом несучої частоти ефективна лише в межах часового вікна, у якому температурний дрейф опорного генератора несуттєвий [3]. Тому критерієм відбору має бути відношення дисперсії ознаки між пристроями до її дисперсії для того самого пристрою між сеансами.

Відбір через інваріантності конвеєра. Відбір через інваріантності полягає в тому, що модель роблять нечутливою до певних класів змін сигналу, а отже, позбавляють можливості використовувати їх як ознаки. Інваріантність функції розпізнавання $f(r)$ до сімейства перетворень G означає $f\big(g(r)\big) \approx f(r)$, де $g \in G$. Її досягають нормалізацією, компенсацією та аугментацією даних. Нормалізація та компенсація приводять сигнал до канонічної форми: наприклад, ділення на середньоквадратичне значення робить модель інваріантною до рівня прийнятого сигналу й повністю знищує цю ознаку, а компенсація зсуву опорної частоти виключає можливість тренування моделі на цій неточній характеристиці сигналу, зміна якої може бути викликана великою кількістю факторів протягом незначних часових періодів. Аугментація вносить штучну варіативність у властивості, щодо яких потрібна інваріантність, і навчає модель їх ігнорувати; у системах RFF це рівень адитивного шуму, багатопроменеві відбиття, доплерівський зсув частоти тощо.

Перетворення конвеєра має задовольняти критерій допустимості: воно допустиме, лише якщо сімейство змін, яке воно усуває, породжується каналом H або приймачем R і не породжується спотвореннями передавача T_d. Складність у тому, що деякі сімейства спільні: частотний зсув породжують як доплерівський ефект, так і похибка опорного генератора, сталу складову сигналу — і приймач, і витік гетеродина. Їх розділяють за масштабом: для нерухомих IoT-пристроїв

доплерівський зсув не перевищує десятків герц [3], тоді як похибка кварцового генератора сягає кілогерців, тому рандомізувати частоту під час аугментації допустимо лише в доплерівському діапазоні; інакше модель стає нечутливою до однієї з найсильніших апаратних ознак.

Явний відбір навчанням. Явний відбір навчанням реалізують вбудованими методами [2]: L1-регуляризацією, груповою L1-регуляризацією та масками над множиною кандидатів (ознак, представлень сигналу, ділянок пакета). До функції втрат додається штраф за використання кожного кандидата, і під час мінімізації сумарних втрат неефективні кандидати втрачають вагу аж до відключення, тоді як ефективні виправдовують накладений штраф. Штраф оптимізує точність на навчальних даних, тому вада оцінювання в межах одного сеансу переходить і сюди: силу штрафу слід обирати на валідації з інших сеансів запису.

Навчання на основі необробленого сигналу. У цьому випадку множина кандидатів відсутня: мережа самостійно формує ознаки з необробленого сигналу [1] і визначає, до яких його властивостей стає чутливою. Перевагу отримують властивості, що корелюють з міткою в навчальних даних і легко вивчаються. Тому мережа схильна до навчання на «коротких шляхах» — простих побічних ознаках (канал, тракт приймача, вміст пакета, умови сеансу), які корелюють з міткою лише в навчальних даних. Так, при тестуванні на записах іншого дня точність може падати майже до рівня випадкового вгадування, а щоб виключити навчання на адресі пристрою, застосовують контрольні експерименти з побітово ідентичними передачами [4]. Отже, у наскрізній парадигмі відповідальність за відбір ознак переходить від моделювання до планування експерименту.

Висновки

Поділ механізмів відбору ознак на явні та неявні не збігається з поділом на класичні та нейромережеві методи: явний відбір можливий і в наскрізній системі через вибір ділянки сигналу й представлення або вбудовані методи. Запропонована класифікація за двома осями дає змогу визначити для кожного механізму характерну помилку та засіб контролю: міжсеансовий критерій стабільності для явного відбору, критерій допустимості перетворень для конвеєра обробки, тести на «короткі шляхи» для навчання на необробленому сигналі. Для задач захисту інформації доцільна гібридна схема, у якій нейромережа навчається на сигналі після компенсації нестабільних складових, а коректність відбору підтверджується протоколом оцінювання, а не лише точністю класифікації.

Список використаних джерел:

1. No Radio Left Behind: Radio Fingerprinting Through Deep Learning of Physical-Layer Hardware Impairments / K. Sankhe et al. IEEE Transactions on Cognitive Communications and Networking. 2020. Vol. 6, no. 1. P. 165–178. URL: https://doi.org/10.1109/tccn.2019.2949308.

2. Guyon I., Elisseff A. An introduction to variable and feature selection. The Journal of Machine Learning Research. 2003. Vol. 3. P. 1157–1182. URL: https://doi.org/10.5555/944919.944968.

3. Towards Scalable and Channel-Robust Radio Frequency Fingerprint Identification for LoRa / G. Shen et al. New York : arXiv, 2021. 12 p. (Preprint. IEEE ; 2107.02867). URL: https://doi.org/10.48550/arXiv.2107.02867.

4. Deep Learning for RF Fingerprinting: A Massive Experimental Study / T. Jian et al. IEEE Internet of Things Magazine. 2020. Vol. 3, no. 1. P. 50–57. URL: https://doi.org/10.1109/iotm.0001.1900065.

Афонін М.В.
студент групи ІКБМ-51, ННІКБЗІ ДУІКТ, Київ, Україна

АКТУАЛЬНІ ПРОБЛЕМИ БЕЗПЕКИ ІНСТРУМЕНТІВ ВІДДАЛЕНОГО ДОСТУПУ В КОРПОРАТИВНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ

Розглянуто ризики, які створюють засоби віддаленого доступу та адміністрування (VPN, RDP, RMM-системи) у корпоративних інформаційних системах, зокрема під час надання IT-підтримки клієнтським організаціям. Проаналізовано типові вектори атак на ці засоби та обмеження периметрової моделі захисту. Запропоновано набір організаційно-технічних заходів на основі принципів нульової довіри: багатофакторна автентифікація, мінімальні привілеї, сегментація, контроль інструментів, журналювання сеансів. Окреслено напрями подальших досліджень.

Ключові слова: віддалений доступ, нульова довіра, RMM, VPN, корпоративні інформаційні системи.

Постановка проблеми. Поширення дистанційної роботи, хмарних сервісів та аутсорсингу IT-підтримки призвело до того, що значну частину адміністративних дій у корпоративних інформаційних системах (КІС) виконують віддалено. Засоби віддаленого доступу, а саме VPN-шлюзи, служби віддалених робочих столів (RDP), RMM-системи та інструменти віддаленої підтримки, забезпечують оперативність обслуговування, але водночас створюють додаткову поверхню атаки. Скомпрометований віддалений канал надає зловмиснику привілейований доступ до робочих станцій і серверів, часто в обхід решти засобів захисту. Особливо гостро проблема постає для постачальників IT-послуг, які обслуговують багато організацій: компрометація їхніх інструментів може спричинити каскадний вплив на всіх клієнтів [1]. Метою роботи є систематизація ризиків інструментів віддаленого доступу та формування набору заходів захисту для організацій, які надають IT-підтримку.

Аналіз підходів. Використано порівняльний аналіз настанов NIST і CISA та узагальнення практики IT-підтримки клієнтських організацій. Рекомендації щодо захисту віддаленого доступу викладено в настанові NIST SP 800-46 Rev. 2, яка розглядає безпекові аспекти різних типів рішень для віддаленого доступу (тунелювання, портали, віддалені робочі столи, прямий доступ до застосунків) [2, с. 4–11] та радить формувати відповідні політики безпеки [2, с. 30]. Концепція нульової довіри (Zero Trust), описана в NIST SP 800-207, виходить з того, що розташування в мережі саме по собі не означає довіри, а доступ до кожного ресурсу надається окремо для кожного сеансу [3, с. 7]; захист зміщується з периметра на окремі ресурси, що відповідає зростанню кількості віддалених користувачів і хмарних активів поза межами корпоративної мережі. У 2023 р. агентство CISA оприлюднило план кіберзахисту для RMM-програмного

забезпечення, у якому наголошено, що доступ до RMM-інструментів може дати зловмисникам плацдарм у середовищах постачальників керованих послуг і, відповідно, у їхніх клієнтів [1].

Основні загрози. Аналіз зазначених джерел дає змогу згрупувати типові ризики для віддаленого доступу в КІС: 1) компрометація облікових даних (повторне використання паролів, фішинг, відсутність багатофакторної автентифікації); 2) публікація служб віддаленого доступу в мережі Інтернет із невиправленими вразливостями або слабкою конфігурацією; 3) зловживання легітимними засобами віддаленого адміністрування, присутність яких на пристрої не викликає підозр і ускладнює виявлення атаки; 4) надмірні привілеї та плоска мережева топологія, за якої один обліковий запис відкриває доступ до багатьох систем; 5) спільні облікові записи та ключі, які використовуються в середовищах різних клієнтів; 6) відсутність журналювання та запису адміністративних сеансів, через що розслідування інциденту стає неможливим.

Пропоновані заходи. З урахуванням принципів нульової довіри [3, с. 7] і рекомендацій щодо захисту віддаленого доступу [2, с. 4–11] запропоновано шість груп заходів. По-перше, єдина контрольована точка входу: заборона прямого доступу до RDP з мережі Інтернет, доступ лише через VPN або шлюз віддалених робочих столів. По-друге, обов’язкова багатофакторна автентифікація для всіх віддалених сеансів, передусім адміністративних. По-третє, принцип мінімальних привілеїв: персональні облікові записи замість спільних, окремі облікові записи для адміністрування, надання доступу на обмежений час і лише до необхідних систем. По-четверте, сегментація та ізоляція клієнтських середовищ: окремі тунелі, облікові дані й правила доступу для кожного клієнта, щоб компрометація одного середовища не поширювалась на інші. По-п’яте, інвентаризація та контроль інструментів віддаленого доступу: затверджений перелік дозволеного програмного забезпечення, блокування неузгоджених засобів, своєчасне оновлення. По-шосте, централізоване журналювання та моніторинг: збір журналів сеансів, сповіщення про аномальну активність, запис адміністративних сеансів, періодичний перегляд прав доступу. Ці заходи доповнюються навчанням персоналу та формалізованою політикою віддаленого доступу [2, с. 30].

Особливості застосування в ІТ-аутсорсингу. Для організацій, що надають послуги підтримки багатьом клієнтам, до загальних заходів додаються специфічні вимоги. Кожне клієнтське середовище доцільно підключати через окремий VPN-тунель із власними обліковими даними та правилами доступу, а інженерів підтримки допускати лише до тих клієнтів і систем, які закріплено за ними. Інструменти віддаленої підтримки варто налаштовувати так, щоб неконтрольований постійний доступ до пристроїв був вимкнений, а підключення до робочої станції користувача підтверджувалося самим користувачем або відбувалося в межах зареєстрованої заявки. Важливим є життєвий цикл доступу:

оперативне надання прав під час прийому інженера на роботу та їх негайне відкликання після звільнення чи переведення.

Контроль ефективності. Щоб заходи не залишалися формальними, для них доцільно визначити події, на які система моніторингу реагує в першу чергу: вхід поза робочим часом або з нетипового розташування, серія невдалих спроб автентифікації, встановлення нового засобу віддаленого керування на робочій станції, адміністративний сеанс без пов’язаної заявки. Результати перевірок рекомендовано фіксувати та використовувати під час внутрішнього аудиту інформаційної безпеки, а виявлені відхилення оформлювати як коригувальні дії з визначеними відповідальними та строками виконання.

Обмеження та етапність впровадження. Заходи ускладнюють роботу інженерів і потребують додаткового адміністрування, тому впроваджувати їх доцільно поетапно: спершу прибрати прямий доступ із мережі Інтернет і спільні облікові записи, далі запровадити багатофакторну автентифікацію для привілейованих користувачів і лише потім переходити до детального моніторингу.

Висновки. Безпеку віддаленого доступу слід розглядати не як окремий технічний засіб, а як складову загальної політики безпеки КІС, що охоплює автентифікацію, управління привілеями, мережеву сегментацію та моніторинг. Запропонований набір заходів може бути застосований як у корпоративному середовищі, так і в організаціях, що надають послуги ІТ-підтримки багатьом клієнтам. Починати доцільно з найризикованіших елементів: публічно доступних служб і привілейованих облікових записів. Перспективним напрямом подальших досліджень є розроблення методики оцінювання зрілості захисту віддаленого доступу з кількісними показниками, зокрема часткою сеансів під багатофакторною автентифікацією, часом відкликання доступу та охопленням журналюванням.

Список літератури

1. CISA Releases JCDC Remote Monitoring and Management (RMM) Cyber Defense Plan. Cybersecurity and Infrastructure Security Agency. 2023. URL: https://cisa.gov/news-events/alerts/2023/08/16/cisa-releases-jcdc-remote-monitoring-and-management-rmm-cyber-defense-plan (дата звернення: 21.10.2026).

2. Souppaya M., Scarfone K. Guide to Enterprise Telework, Remote Access, and Bring Your Own Device (BYOD) Security : NIST Special Publication 800-46 Rev. 2 / M. Souppaya, K. Scarfone. Gaithersburg : NIST, 2016. DOI: https://doi.org/10.6028/NIST.SP.800-46r2.

3. Zero Trust Architecture : NIST Special Publication 800-207 / S. Rose, O. Borchert, S. Mitchell, S. Connelly. Gaithersburg : NIST, 2020. 59 p. DOI: https://doi.org/10.6028/NIST.SP.800-207.

Гончаренко М.
студент групи ІКБм-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТА КЛАСИФІКАЦІЇ МЕРЕЖЕВИХ АТАК У ПРОМИСЛОВИХ МЕРЕЖАХ MODBUS TCP НА ОСНОВІ АЛГОРИТМУ XGBOOST

Розглянуто проблему захисту промислових мереж із протоколом Modbus TCP, який не має механізмів автентифікації та шифрування, на прикладі атаки шкідливого програмного забезпечення FrostyGoop на систему теплопостачання м. Львова. Запропоновано технологію виявлення та класифікації мережевих атак на основі алгоритму XGBoost, доповненого автоенкодером для виявлення раніше невідомих атак. Описано етапи технології, методику її експериментальної перевірки на наборі даних CIC Modbus 2023 та сформульовано рекомендації щодо впровадження на об'єктах критичної інфраструктури.

Ключові слова: *промислові системи керування, Modbus TCP, виявлення вторгнень, машинне навчання, XGBoost, автоенкодер, критична інфраструктура.*

Промислові системи керування (ПСК) об'єктів енергетики, тепло- та водопостачання значною мірою досі використовують протокол Modbus TCP, створений без механізмів автентифікації та шифрування. Будь-який вузол, що має мережевий доступ до контролера, може прочитати або змінити значення його регістрів, і пристрій виконає таку команду як легітимну. В умовах повномасштабної війни ця вразливість перестала бути теоретичною: у січні 2024 року шкідливе програмне забезпечення FrostyGoop, надсилаючи звичайні команди Modbus TCP на контролери, доступні з мережі Інтернет, порушило роботу комунального теплоенергетичного підприємства у Львові, унаслідок чого понад 600 багатоквартирних будинків майже дві доби залишалися без опалення за мінусових температур [1]. FrostyGoop не експлуатує вразливостей програмного забезпечення, а лише зловживає штатними функціями протоколу, тому сигнатурні засоби захисту його не розпізнавали.

Відповідно, захист таких мереж потребує не пошуку відомих сигнатур, а аналізу поведінки трафіку на прикладному рівні протоколу. Методи машинного навчання дають змогу автоматизувати такий аналіз, однак більшість досліджень оцінює моделі на тих самих типах атак, що були присутні в навчальній вибірці. За таких умов класифікатори демонструють високу точність, але реальна атака з використанням нового інструменту, подібного до FrostyGoop, може бути віднесена до нормального трафіку. Отже, актуальним є завдання побудови технології, яка не лише класифікує відомі атаки, а й сигналізує про раніше невідомі.

Для розв'язання цього завдання запропоновано технологію виявлення та класифікації мережевих атак, ядром якої є алгоритм градієнтного бустингу XGBoost [2], доповнений модулем виявлення аномалій (рис. 1). Вибір XGBoost зумовлений високою точністю на табличних даних, стійкістю до незбалансованих вибірок і малим часом обробки одного запису, що

підтверджують, зокрема, дослідження гібридних ансамблевих моделей виявлення зловмисної активності, виконані на кафедрі [3].

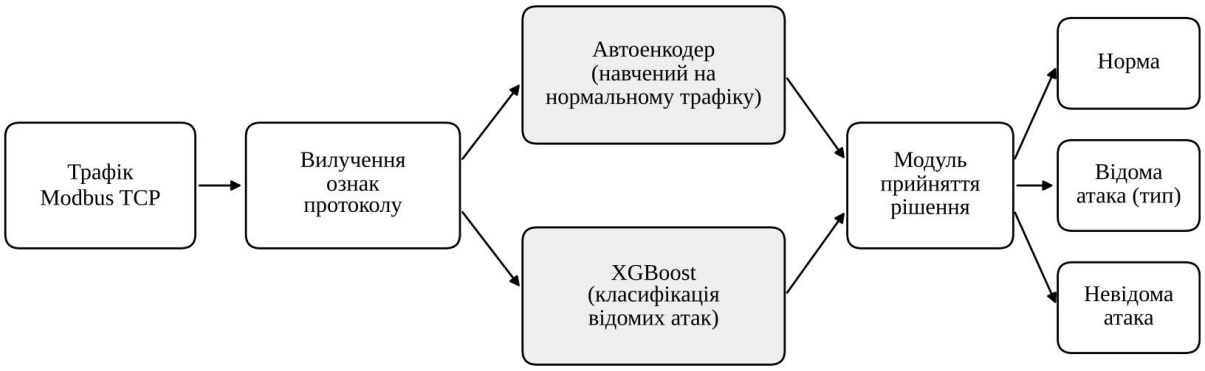


Рис. 1. Схема технології виявлення та класифікації мережевих атак

Технологія передбачає таку послідовність дій. На першому етапі з трафіку Modbus TCP вилучаються ознаки прикладного рівня протоколу: код функції (читання чи запис), адреса та кількість регістрів, записувані значення, довжина кадру, а також часові ознаки - інтервали між запитами та частота звернень кожного вузла до контролера. На другому етапі ознаки паралельно подаються на два модулі. Автоенкодер, навчений виключно на нормальному трафіку, відтворює вхідний вектор; якщо помилка відтворення перевищує поріг, обчислений за розподілом помилок на нормальних даних, запис вважається аномальним. Класифікатор XGBoost, навчений на розмічених прикладах, визначає, чи належить запис до одного з відомих типів атак. На третьому етапі модуль прийняття рішення поєднує результати: запис, який XGBoost із високою впевненістю відносить до певного класу, позначається як відома атака відповідного типу, а запис, визнаний аномальним автоенкодером, але не розпізнаний класифікатором, - як невідома атака, що потребує уваги оператора.

Для експериментальної перевірки обрано відкритий набір даних CIC Modbus 2023, сформований Канадським інститутом кібербезпеки на моделі електричної підстанції, розгорнутій у середовищі Docker [4]. Набір містить записи нормального обміну та трафік низки атак, зокрема розвідки, переповнення запитами, ін’єкції хибних даних і повторного відтворення пакетів. Ефективність оцінюватиметься за показниками точності, повноти, F1-міри та частки хибних спрацювань у двох режимах: на атаках, присутніх у навчальній вибірці, та за схемою «виключення одного типу атаки» (leave-one-attack-out), коли модель тестується на атаці, якої вона не бачила під час навчання. Саме другий режим відображає реальну ситуацію появи нової кіберзброї. Особливістю промислового трафіку є значний дисбаланс класів: під час атак контролери продовжують штатне циклічне опитування, тому нормальних записів набагато більше, ніж шкідливих. Для врахування цього під час навчання XGBoost застосовуватиметься зважування класів, а поріг автоенкодера визначатиметься лише за нормальними даними, що взагалі не потребує прикладів атак.

За результатами аналізу сформульовано такі рекомендації щодо впровадження технології. По-перше, аналізатор доцільно підключати пасивно -

через дзеркальний порт комутатора (SPAN) або мережевий відгалужувач, що виключає вплив на технологічний процес. По-друге, поріг автоенкодера слід налаштовувати з урахуванням допустимого рівня хибних спрацювань, адже надмірна кількість тривог знижує довіру персоналу до системи. По-третє, події «невідома атака» мають передаватися черговому оператору або аналітику SOC для ручної перевірки, а підтверджені випадки - використовуватися для донавчання класифікатора. Водночас технологія не замінює базових заходів: контролери з протоколом Modbus TCP не повинні бути доступні з мережі Інтернет, а промислова мережа має бути відокремлена від корпоративної.

Таким чином, запропонована технологія поєднує точну класифікацію відомих атак алгоритмом XGBoost зі здатністю автоенкодера сигналізувати про раніше невідомі загрози, що є критично важливим для захисту об’єктів критичної інфраструктури України в умовах воєнних кіберзагроз. Подальші дослідження спрямовані на програмну реалізацію технології та експериментальне порівняння її ефективності з окремими класифікаторами на наборі даних CIC Modbus 2023.

Перелік посилань:

1. Kovacs E. FrostyGoop ICS Malware Left Ukrainian City’s Residents Without Heating. SecurityWeek. 2024. 23 July. URL: https://www.securityweek.com/frostygoop-ics-malware-left-ukrainian-citys-residents-without-heating/ (дата звернення: 23.10.2026).

2. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. P. 785-794. DOI: https://doi.org/10.1145/2939672.2939785

3. 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

4. Boakye-Boateng K., Ghorbani A. A., Lashkari A. H. Implementation of a Trust-Based Framework for Substation Defense in the Smart Grid. Smart Cities. 2024. Vol. 7, No. 1. P. 99-140. DOI: https://doi.org/10.3390/smartcities7010005

Ступченко А.О
студентка групи ІКБмз-11, ДУІКТ,
Київ, Україна

ПОБУДОВА ЦЕНТРУ ОПЕРАЦІЙНОЇ БЕЗПЕКИ (SOC) ЯК ІНСТРУМЕНТ ПІДВИЩЕННЯ КІБЕРСТІЙКОСТІ ОРГАНІЗАЦІЇ

У тезах розглянуто особливості побудови центру операційної безпеки (Security Operations Center, SOC) як складової системи кіберзахисту організації. Визначено основні функції SOC, охарактеризовано технологічні та організаційні компоненти його функціонування. Розглянуто роль централізованого моніторингу подій безпеки, систем SIEM та процедур реагування на кіберінциденти. Обґрунтовано необхідність комплексного підходу до створення SOC, який передбачає поєднання технічних засобів, кваліфікованого персоналу й регламентованих процесів.

Ключові слова: кібербезпека, SOC, SIEM, моніторинг безпеки, кіберінцидент, реагування на інциденти.

Зі зростанням кількості кіберзагроз та ускладненням інформаційних систем організаціям необхідно забезпечувати постійний контроль стану інформаційної безпеки. Використання окремих захисних засобів не завжди дає змогу своєчасно виявляти складні атаки, які можуть проявлятися через сукупність подій на різних вузлах інфраструктури. За таких умов актуальним є створення центру операційної безпеки (SOC), який забезпечує централізований моніторинг, аналіз подій та координацію реагування на кіберінциденти.

Відповідно до рекомендацій Агентства Європейського Союзу з кібербезпеки (ENISA), створення SOC потребує послідовного підходу, що охоплює оцінювання готовності організації, проєктування, впровадження, експлуатацію та вдосконалення відповідних процесів [1].

SOC - це організаційно-технічна структура, призначена для моніторингу подій інформаційної безпеки, виявлення потенційних загроз, аналізу підозрілої активності та координації реагування на інциденти. Його основне завдання полягає у своєчасному виявленні небезпечних подій та зменшенні можливих наслідків кібератак. Побудова SOC передбачає узгодження трьох основних компонентів: технологічної інфраструктури, персоналу та операційних процесів. Технологічна складова забезпечує отримання, зберігання й аналіз даних. Персонал здійснює моніторинг, перевіряє сповіщення, досліджує інциденти та координує відповідні дії. Операційні процеси визначають порядок оброблення подій, пріоритизацію загроз, ескалацію інцидентів і взаємодію з іншими підрозділами організації.

Важливим елементом технологічної архітектури SOC є система управління інформацією та подіями безпеки (SIEM). Вона дає змогу централізовано збирати журнали подій із серверів, мережевого обладнання, робочих станцій, засобів захисту та інших інформаційних ресурсів. Завдяки кореляції подій SIEM допомагає виявляти підозрілі закономірності, які складно встановити під час аналізу окремих журналів. Водночас ефективність SIEM залежить від повноти джерел даних, якості налаштування правил кореляції та здатності фахівців відрізняти реальні загрози від хибнопозитивних сповіщень. Тому впровадження технологічних засобів має супроводжуватися розробленням процедур аналізу подій і регулярним переглядом правил виявлення.

Значення централізованого моніторингу узгоджується з положеннями NIST Cybersecurity Framework 2.0, де функція Detect передбачає своєчасне виявлення й аналіз потенційних кібератак та компрометації інформаційних систем [2]. У контексті діяльності SOC це означає необхідність не лише реєструвати події, а й визначати їхню значущість з урахуванням критичності активів, характеру загрози та можливого впливу на діяльність організації. Іншим важливим аспектом побудови SOC є організація процесу реагування на інциденти. Відповідно до рекомендацій NIST, реагування необхідно інтегрувати в загальну систему управління кіберризиками організації. Такий підхід передбачає підготовку до інцидентів, їх виявлення, реагування, відновлення та постійне вдосконалення захисних заходів [3].

Під час створення SOC необхідно також визначити модель його функціонування. Організація може сформувати власну команду, залучити зовнішнього постачальника послуг моніторингу або застосувати комбіновану модель. Вибір залежить від розміру інфраструктури, критичності інформаційних активів, доступного бюджету, вимог до безперервності моніторингу та наявності кваліфікованих спеціалістів.

Для оцінювання ефективності SOC доцільно використовувати показники, що характеризують швидкість виявлення та оброблення інцидентів, частку хибнопозитивних сповіщень, дотримання встановлених строків реагування й повноту охоплення інформаційних активів моніторингом. Аналіз таких показників дає змогу визначати проблемні ділянки процесу та обґрунтовувати подальший розвиток центру. Варто наголосити, що створення SOC не обмежується придбанням SIEM-системи або організацією робочих місць аналітиків. Для досягнення практичного результату необхідні чіткий розподіл відповідальності, документовані процедури, належна якість телеметрії, постійне навчання персоналу та регулярне вдосконалення механізмів виявлення загроз.

Висновки. Побудова центру операційної безпеки є комплексним завданням, яке передбачає поєднання технологічних засобів моніторингу, компетентного персоналу та формалізованих процесів управління кіберінцидентами. Централізований збір і кореляція подій дають змогу підвищити видимість стану інформаційної інфраструктури, а регламентовані процедури реагування сприяють своєчасному виявленню та локалізації загроз. Ефективність SOC залежить від відповідності його архітектури потребам організації, якості налаштування засобів моніторингу й систематичного вдосконалення процесів безпеки. Отже, SOC доцільно розглядати не як окремий програмно-технічний комплекс, а як інтегровану складову системи управління кібербезпекою організації.

Список використаних джерел

1. European Union Agency for Cybersecurity (ENISA). How to Set Up CSIRT and SOC. 2020. URL: https://www.enisa.europa.eu/publications/how-to-set-up-csirt-and-soc.

2. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. 2024. URL: https://doi.org/10.6028/NIST.CSWP.29.

3. Nelson A., Rekhi S., Scarfone K., Souppaya M. 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.

Шапко Олег Олександрович
студент групи ІКБм-61, ННІКБЗІ ДУІКТ,
Київ, Україна

ПРОАКТИВНЕ ВИЯВЛЕННЯ ЗАГРОЗ НА КІНЦЕВИХ ТОЧКАХ ОРГАНІЗАЦІЇ З ВИКОРИСТАННЯМ ПОВЕДІНКОВИХ МОДЕЛЕЙ У СЕРЕДОВИЩІ WAZUH

Розглянуто проблему пізнього виявлення атак на кінцеві точки організації, під час яких зловмисник використовує легітимні засоби операційної системи. Запропоновано технологію проактивного виявлення загроз, що доповнює відкриту платформу Wazuh поведінковими моделями користувачів і вузлів та автоматизованим реагуванням. Описано етапи збору телеметрії, побудови моделей, статистичного аналізу відхилень і реагування, а також порядок поетапного впровадження. Визначено показники для оцінювання ефективності технології на віртуальному стенді організації.

Ключові слова: кінцева точка, проактивне виявлення загроз, поведінкова модель, UEBA, Wazuh, Active Response, кореляція подій.

Атака на інформаційну систему організації майже завжди проходить через кінцеву точку: робочу станцію працівника, сервер або віртуальну машину. Саме там зловмисник отримує перший доступ, закріплюється й звідти просувається мережею далі. Від стану захисту кінцевих точок значною мірою залежить, на якому етапі атаку буде помічено і скільки шкоди вона встигне завдати.

Типовий набір засобів захисту в українських організаціях середнього розміру складається з антивірусу та системи збору журналів подій. Обидва засоби працюють за принципом збігу: файл порівнюється з базою відомих зразків, подія — з описаним наперед шаблоном. Проти поширеного шкідливого коду цього достатньо. Значно гірше такий підхід спрацьовує, коли нападник не приносить власних програм, а користується тим, що вже є в системі: PowerShell, інструментарієм WMI, планувальником завдань, засобами віддаленого адміністрування. Кожна подібна дія зовні не відрізняється від роботи системного адміністратора, тож сигнатури для неї не існує. На практиці такі атаки нерідко виявляють лише під час шифрування даних або після їх витоку, коли збитків уже завдано [1].

Проактивне виявлення виходить з іншого запитання. Замість перевірки того, чи відомий певний файл, система оцінює, чи звично поводиться конкретний користувач або вузол. Поведінка в корпоративному середовищі доволі стабільна: бухгалтер працює в робочі години з обмеженим набором програм, сервер бази даних не звертається до зовнішніх сайтів, офісний редактор не запускає командних оболонок. Формалізований опис такої звичної поведінки в роботі названо поведінковою моделлю, а суттєве відхилення від неї розглядається як підстава для перевірки. Перевага підходу полягає в тому, що він реагує на дії нападника, а не на його інструменти, і тому здатний помітити атаку ще під час закріплення в системі чи підвищення привілеїв [2].

Подібну логіку реалізують комерційні платформи класу XDR з модулем аналітики поведінки користувачів і сутностей (UEBA), однак для багатьох вітчизняних організацій їхня вартість ліцензування неприйнятна. Відкрита

платформа Wazuh поєднує функції SIEM і XDR: має агент для Windows, Linux і macOS, централізований механізм аналізу подій, модулі контролю цілісності файлів і виявлення вразливостей, а також засіб автоматичного реагування Active Response [3]. Окремого модуля поведінкової аналітики в ній немає, проте наявні механізми дозволяють відтворити основну частину його функцій. На цьому ґрунтується запропонована технологія, структуру якої наведено на рис. 1.



Рис. 1. Структура технології проактивного виявлення загроз на кінцевих точках у середовищі Wazuh

На першому етапі збирається детальна телеметрія. Стандартних журналів Windows для поведінкового аналізу недостатньо, оскільки вони не показують, яка програма запустила процес або з якою адресою встановлено з’єднання. Тому на вузлах Windows разом з агентом Wazuh встановлюється системний монітор Sysmon, а на вузлах Linux налаштовується підсистема аудиту auditd. Агент передає на сервер події запуску процесів, мережевої активності, змін файлів і реєстру, входів до системи.

На другому етапі на сервері Wazuh будуються поведінкові моделі у вигляді власних правил кореляції. Застосовано моделі трьох типів. Модель інтенсивності фіксує незвично велику кількість однотипних подій за короткий проміжок часу, наприклад серію невдалих входів з однієї адреси або масову зміну файлів, характерну для шифрувальника. Контекстна модель враховує обставини події: час доби, вузол, з якого виконано вхід, програму, що породила процес. Вхід адміністратора о третій ночі з комп’ютера бухгалтерії або запуск PowerShell текстовим редактором технічно можливі, але нетипові для конкретної організації. Модель послідовності об’єднує кілька окремих подій в один ланцюжок: запуск підозрілого процесу, звернення до пам’яті процесу автентифікації, нове мережеве з’єднання. Кожна з цих подій окремо має низький рівень критичності, а ланцюжок загалом отримує високий.

Правила добре працюють там, де межу норми можна задати заздалегідь. Для числових показників, які залежать від ролі вузла (кількість процесів, з’єднань чи входів за годину), фіксовані пороги дають або пропуски, або надлишок тривог. Тому третій етап виконується в індексаторі Wazuh,

побудованому на OpenSearch. Його модуль виявлення аномалій використовує алгоритм Random Cut Forest [4], який навчається на історії кожного вузла окремо й для кожного нового значення обчислює ступінь аномальності без порогів, заданих вручну.

Четвертий етап передбачає автоматичне реагування. Для кожного класу загроз заздалегідь визначено дію, яку виконує агент: блокування IP-адреси засобами брандмауера, тимчасове вимкнення облікового запису, завершення процесу або мережеву ізоляцію вузла. Дії, здатні зачепити легітимну роботу, задаються як тимчасові й скасовуються автоматично через встановлений інтервал. Кожне правило прив'язане до техніки бази знань MITRE ATT&CK, що дає змогу оцінити повноту покриття та узгодити дії з життєвим циклом реагування на інциденти, описаним у NIST SP 800-61r3 [1].

Роботу моделей зручно показати на прикладі атаки шифрувальника. Працівник відкриває документ із поштового вкладення, і текстовий редактор запускає PowerShell. Контекстна модель позначає цю подію як нетипову, хоча сама по собі вона ще не доводить атаки. Через кілька хвилин той самий процес звертається до зовнішньої адреси й завантажує файл, а модель послідовності пов'язує обидві події в ланцюжок і підвищує рівень сповіщення. Коли завантажена програма починає масово змінювати файли в мережевій папці, спрацьовує модель інтенсивності, і Active Response завершує процес та ізолює робочу станцію від мережі. Аналітик отримує одне сповіщення з повною хронологією замість кількох розрізнених подій, які довелося б зіставляти вручну. Без поведінкових моделей перша тривога в такому сценарії з'явилася б лише після появи зашифрованих файлів.

Впроваджувати технологію доцільно поетапно. Протягом перших двох–трьох тижнів правила працюють лише в режимі сповіщення. За цей час уточнюються переліки довірених вузлів і облікових записів, з розгляду виключаються регулярні службові завдання, а модуль виявлення аномалій накопичує історію. Автоматичне реагування вмикається спочатку для сценаріїв з низьким ризиком помилки, як-от підбір пароля, і лише потім для ізоляції вузлів.

Перевірку технології передбачено на віртуальному стенді, що моделює організацію: контролер домену Active Directory, робоча станція Windows, вебсервер Linux і вузол нападника. Відтворюються шість сценаріїв: підбір пароля, виконання закодованих команд PowerShell, викрадення облікових даних із пам'яті, закріплення через реєстр і планувальник завдань, емуляція шифрувальника та нетиповий вхід привілейованого облікового запису. Для кожного сценарію фіксуються факт виявлення, рівень сповіщення, час від початку атаки до виявлення і до реагування. Ці показники разом із часткою хибних спрацювань порівнюються з базовою конфігурацією Wazuh без поведінкових моделей.

Технологія має свої обмеження. Поведінкові моделі описуються вручну й потребують перегляду після змін в інфраструктурі, а нападник, який діє повільно й у межах звичного робочого графіка, може залишатися в межах норми. Водночас для організації без бюджету на комерційну UEBA-платформу

запропонований підхід дає змогу виявляти атаки без відомих сигнатур на ранніх етапах і автоматично стримувати їх, використовуючи лише відкрите програмне забезпечення.

Перелік використаної літератури

1. Nelson A., Rekhi S., Souppaya M., Scarfone K. Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile : NIST Special Publication 800-61r3. Gaithersburg, MD : NIST, 2025. URL: https://doi.org/10.6028/NIST.SP.800-61r3 (дата звернення: 24.10.2026).

2. Roddie M., Deyalsingh J., Katz G. J. Practical Threat Detection Engineering: A hands-on guide to planning, developing, and validating detection capabilities. Birmingham : Packt Publishing, 2023. 328 p.

3. Gupta R. Security Monitoring with Wazuh: A hands-on guide to effective enterprise security using real-life use cases in Wazuh. Birmingham : Packt Publishing, 2024. 322 p.

4. Guha S., Mishra N., Roy G., Schrijvers O. Robust Random Cut Forest Based Anomaly Detection on Streams. Proceedings of the 33rd International Conference on Machine Learning. New York, NY, 2016. Vol. 48. P. 2712–2721.

Харькевич Дмитро Олегович
магістрант групи ІКБм-61, ННІКБЗІ, ДУІКТ
Київ, Україна

АВТОМАТИЗОВАНИЙ КОНТРОЛЬ КОДУ В DEVSECOPS ІЗ ВИКОРИСТАННЯМ SONARQUBE QUALITY GATE

Анотація. Розглянуто інтеграцію статичного аналізу коду в DevSecOps-процес із використанням SonarQube Quality Gate. Запропоновано послідовність автоматизованої перевірки, за якої пакування застосунку залежить від виконання визначених критеріїв. На тестовому JavaScript-проєкті зіставлено базову версію, версію з контрольованим порушенням правила та виправлену версію. За незмінних тестів і політики отримано статуси OK - ERROR - OK та підтверджено блокування пакування. Визначено умови відтворюваності перевірки й межі застосування результатів.

Безперервне внесення змін до програмного забезпечення потребує перевірок, результати яких впливають на рішення про подальший випуск. Функціональні тести встановлюють відповідність поведінки програми заданим сценаріям, проте не охоплюють усіх вимог до безпеки коду. У NIST Secure Software Development Framework передбачено включення практик безпечної розробки в наявний життєвий цикл програмного забезпечення [1]. Для реалізації такого підходу необхідно визначити склад перевірок, критерії приймання та реакцію конвеєра на невідповідність.

У DevSecOps статичний аналіз є одним із засобів раннього зворотного зв’язку щодо програмних змін. Його впровадження потребує узгодження інструментів із практиками команди та умовами виконання. Систематичний огляд R. N. Rajapakse та співавторів розглядає проблеми впровадження DevSecOps за напрямами людей, практик, інструментів та інфраструктури [2]. Отже, наявність аналізатора в конвеєрі слід доповнювати визначеним порядком використання його результату. Звіт, після якого пакування відбувається незалежно від знайдених проблем, не забезпечує виконання обов’язкової політики контролю.

Метою дослідження є перевірка можливості автоматично зупиняти пакування застосунку за невиконання критерію безпеки SonarQube та відновлювати успішне проходження після виправлення коду. Дослідницьке питання стосується зв’язку між повідомленням аналізатора, результатом Quality Gate і поведінкою наступного етапу CI. Оцінюється працездатність цього механізму на контрольованому прикладі, а не повнота виявлення всіх вразливостей програмного продукту.

SonarQube Quality Profile визначає набір активних правил аналізу, тоді як Quality Gate оцінює прийнятність отриманих показників. Умови Gate можуть застосовуватися до нового або всього коду; кожна умова містить метрику, оператор порівняння та поріг [3]. Для інтеграції запропоновано послідовність: отримання визначеної ревізії, підготовка середовища, виконання тестів, формування звіту покриття, статичний аналіз, очікування рішення Gate та пакування лише за успіху перевірок. Докази виконання потрібно зберігати також за відмови.

Тестове середовище побудовано на JavaScript HTTP API керування завданнями. Проєкт містить модулі сховища, оброблення HTTP-запитів та перевірки авторизації. Передбачено отримання, створення й завершення завдань; для операції завершення потрібне значення авторизації. Підготовлено 12 автоматизованих тестів, які перевіряють роботу сховища, авторизацію та HTTP-інтерфейс. Невеликий обсяг проєкту дозволяє локалізувати експериментальну зміну й зіставити її з конкретним повідомленням аналізатора.

Автоматизовані запуски виконано в GitHub Actions на Ubuntu 24.04 із Node.js 24.19.0, SonarQube Community Build 26.10.0.133078 та SonarScanner CLI 8.1.0.6389. Сервер розгортався в контейнері окремо для кожного запуску. Використано режим Multi-Quality Rule Mode та профіль Sonar way core. Покриття імпортувалося зі звіту LCOV після виконання тестів. Тимчасова інсталяція з вбудованою H2 призначена для навчального експерименту; готовність до промислової експлуатації в цій постановці не оцінюється.

Для порівняння прийнято чотири умови відмови Gate: покриття нижче 80%, дублювання вище 3%, рейтинг безпеки гірше A та рейтинг надійності гірше A. Умови застосовано до всього аналізованого коду. Це дозволяє зіставляти однакову область невеликого проєкту без залежності від кількості нових рядків. Порогові значення є параметрами досліду; їхню оптимальність для інших програмних систем не встановлено.

Сканер запускається з параметром sonar.qualitygate.wait=true. Відповідно до документації цей режим передбачає очікування рішення сервера та невдале завершення кроку за відмови Gate [4]. У реалізації час очікування обмежено 300 секундами. Пакування залежить від успішності попереднього кроку. Окремо зберігаються журнали тестів і сканера, показники, повідомлення про проблеми та рішення політики. Сам факт передавання звіту серверу не використовується як підтвердження завершення перевірки.

Методика охоплює три стани коду за незмінних тестів, профілю й умов Gate. Спочатку перевіряється базова версія, далі вноситься контрольоване порушення активного правила, після чого виконується виправлення й повторний аналіз. Основні спостережувані результати - статус Gate, рейтинг безпеки, успішність тестів і фактичний статус пакування. Конфігурацію правил та умов зіставлено за збереженими експортами. Виправлена версія має той самий вміст файлів, що й базова, та окрему ревізію в історії проєкту.

У базовому запуску пройшли всі 12 тестів. Покриття за даними SonarQube становило 97,7%, дублювання - 0%, рейтинги безпеки та надійності - A. Quality Gate отримав OK, сканер завершився з кодом 0, пакування та передавання артефакту виконалися успішно. Цей стан використано як основу для перевірки реакції на локальну зміну.

У другій версії до обчислення дайджесту отриманого рядка внесено зайвий виклик eval(JSON.stringify(providedPassword)) замість безпосереднього використання providedPassword. Наявні функціональні тести зберегли успішність. Водночас аналізатор виявив проблему javascript:S1523 із впливом на безпеку High. Рейтинг безпеки змінився на D, що порушило прийняту вимогу A.

Інші три критерії залишилися успішними. Підсумковий Gate отримав ERROR, Linux-сканер повернув код 3, а кроки пакування та передавання пакета були пропущені. Матеріали перевірки збережено.

Виправлення полягало у видаленні зайвого динамічного виконання та безпосередньому передаванні рядка до обчислення дайджесту. Після повторного запуску всі тести пройшли, рейтинг безпеки повернувся до A, Gate отримав OK, пакування виконалося успішно. Покриття 97,7%, дублювання 0% і рейтинг надійності A залишилися однаковими в усіх трьох станах. Послідовність OK - ERROR - OK підтверджує реакцію конвеєра на досліджувану конструкцію за незмінних критеріїв.

Для відтворюваності результат пов’язано з ревізією коду, запуском CI та ідентифікатором завершеного аналізу. Ці дані дозволяють перевірити, який саме звіт зумовив відмову. Коди завершення сканера потрібно розглядати разом зі статусами наступних кроків: отримання ERROR без блокування пакування свідчило б про невиконання вимоги до інтеграції. Окреме збереження доказів забезпечує можливість дослідити причину зупинки після завершення тимчасового сервера.

Технічна відмова потребує окремої діагностики. Недоступність сервера або перевищення часу очікування не доводить наявності дефекту коду, проте має зупиняти пакування зі збереженням причини. Перед перенесенням рішення до іншого проєкту доцільно повторити позитивний і негативний сценарії та перевірити залежність пакування від результату.

Отриманий результат має визначені межі. JSON.stringify екранує переданий рядок, тому контрольований приклад підтверджує спрацювання правила й політики, але не доводить можливості ін’єкційної атаки. Один негативний сценарій не дозволяє визначити точність або повноту роботи аналізатора. DAST, аналіз сторонніх залежностей і автоматичне розгортання не виконувалися. Для ширшого оцінювання потрібні додаткові контрольовані приклади та перевірка поведінки інтеграції за технічних відмов.

Таким чином, перевірено механізм застосування SonarQube Quality Gate як умови переходу до пакування в DevSecOps-процесі. На тестовому проєкті підтверджено блокування цього етапу за порушення критерію безпеки та відновлення після виправлення без послаблення політики. Практичне застосування підходу потребує перевірки області аналізу, погодження порогів і збереження результатів кожного запуску. Подальший розвиток дослідження пов’язаний із розширенням набору правил і проєктів, перевіркою промислової інфраструктури та поєднанням статичного контролю з іншими методами безпеки.

Використана література

1. Souppaya M., Scarfone K., Dodson D. Secure Software Development Framework (SSDF) Version 1.1. NIST SP 800-218. 2022. DOI: https://doi.org/10.6028/NIST.SP.800-218.
2. Rajapakse R. N., Zahedi M., Babar M. A., Shen H. Challenges and solutions when adopting DevSecOps: A systematic review. Information and Software Technology. 2022. Vol. 141. Article 106700. DOI: https://doi.org/10.1016/j.infsof.2021.106700.
3. SonarSource. Understanding quality gates. SonarQube Community Build. URL: https://docs.sonarsource.com/sonarqube-community-build/quality-standards-administration/managing-quality-gates/introduction-to-quality-gates (дата звернення: 22.10.2026).
4. SonarSource. CI integration overview. SonarQube Community Build. URL: https://docs.sonarsource.com/sonarqube-community-build/analyzing-source-code/ci-integration/overview (дата звернення: 22.10.2026).

Курінний О.С.
студент групи УБДМ-51, ННІКБЗІ ДУІКТ
Київ, Україна

ОЦІНЮВАННЯ ТА ЗНИЖЕННЯ РИЗИКІВ ЛЮДСЬКОГО ФАКТОРА ПІД ЧАС ПРОТИДІЇ СОЦІАЛЬНІЙ ІНЖЕНЕРІЇ

Розглянуто людський фактор як складову ризиків реалізації атак соціальної інженерії. Обґрунтовано доцільність переходу від оцінювання теоретичних знань працівників до аналізу їхньої фактичної поведінки під час контрольованих сценаріїв кіберзагроз. Запропоновано підхід до оцінювання поведінкових ризиків на основі показників небезпечних дій, своєчасності повідомлення про загрози, швидкості реагування та дотримання вимог інформаційної безпеки. Визначено можливості використання отриманих результатів для сегментації ризиків і планування цільових навчальних заходів.

Ключові слова: соціальна інженерія, людський фактор, поведінкові ризики, кібербезпека, фішинг, культура кібербезпеки.

Розвиток сучасних кіберзагроз характеризується активним використанням методів соціальної інженерії, спрямованих на маніпулювання поведінкою працівників організацій. Цільовий фішинг, підроблення особистості керівництва, шахрайські телефонні дзвінки та використання технологій штучного інтелекту для створення переконливих повідомлень ускладнюють своєчасне виявлення зловмисних дій. За даними звіту Європейського агентства з кібербезпеки ENISA Threat Landscape 2025, соціальна інженерія залишається одним із провідних початкових векторів атак, а на фішинг припадало близько 60 % випадків у проаналізованому масиві кіберінцидентів [1]. Це свідчить про необхідність розглядати поведінку персоналу як важливий чинник системи управління кібербезпекою організації.

Традиційні заходи протидії соціальній інженерії передбачають проведення інструктажів, ознайомлення працівників із політиками інформаційної безпеки та перевірку їхніх теоретичних знань. Проте успішне засвоєння навчального матеріалу не гарантує правильних дій у ситуації реального психологічного тиску, дефіциту часу або використання авторитету керівництва. Крім того, результати одноразового тестування не забезпечують достатнього розуміння того, як змінюється поведінка працівників упродовж тривалого періоду. У зв’язку з цим доцільним є перехід до систематичного оцінювання поведінкових ризиків, яке враховує практичні дії персоналу та дає змогу визначати найбільш уразливі аспекти організаційної кібербезпеки.

Метою роботи є обґрунтування підходу до оцінювання та зниження ризиків людського фактора на основі поведінкових показників персоналу під час протидії соціальній інженерії. Основна ідея запропонованого підходу полягає у використанні результатів контрольованих перевірок поведінки працівників для визначення рівня ризику, виявлення типових помилок і планування цільових заходів підвищення обізнаності та практичної готовності до протидії кіберзагрозам.

Основою такого оцінювання є сукупність поведінкових показників, які характеризують різні аспекти взаємодії працівників із потенційними загрозами. Доцільно враховувати частку працівників, які виконують небезпечні дії під час контрольованих сценаріїв фішингу, своєчасність повідомлення про підозрілі повідомлення, час реагування на виявлені ознаки атаки, а також правильність подальших дій відповідно до встановлених процедур. Додатковим показником може бути рівень дотримання вимог щодо використання облікових даних, розмежування доступу та виконання внутрішніх регламентів інформаційної безпеки. Сукупне оцінювання цих показників дає змогу отримати більш повне уявлення про поведінкові особливості персоналу, ніж перевірка виключно теоретичних знань.

Збирання відповідних даних може здійснюватися за допомогою контрольованих симуляцій фішингових атак, навчальних сценаріїв телефонного шахрайства, перевірок установлених процедур повідомлення про підозрілі події та аналізу результатів внутрішніх перевірок дотримання політик безпеки. Такі заходи мають проводитися в контрольованих умовах, із дотриманням вимог законодавства та внутрішніх правил захисту інформації. Для забезпечення об’єктивності результати необхідно оцінювати за єдиними критеріями, враховуючи специфіку посадових обов’язків, рівень доступу до інформаційних ресурсів і характер потенційних загроз.

Наступним етапом є сегментація персоналу за рівнями поведінкового ризику. Вона передбачає визначення груп із відносно низьким, середнім або високим рівнем ризику на основі результатів оцінювання та заздалегідь установлених критеріїв. Працівникам, які систематично демонструють небезпечні реакції або несвоєчасно повідомляють про підозрілі події, доцільно призначати додаткові практичні заняття, моделювання відповідних ситуацій і повторне оцінювання. Для працівників, які демонструють належний рівень реагування, доцільно передбачати періодичне контрольне тестування та підтримувальні навчальні заходи. Такий підхід дає змогу спрямовувати ресурси на ті напрями, де існує найбільша потреба у зниженні ризиків, замість застосування однакової навчальної програми до всього персоналу.

Важливим елементом запропонованого підходу є регулярність оцінювання. Повторне проведення контрольованих перевірок дає змогу відстежувати зміни поведінкових показників, визначати типові помилки та оцінювати доцільність застосованих навчальних заходів. Така циклічність відповідає рекомендаціям Національного інституту стандартів і технологій США, викладеним у документі NIST SP 800-50 Revision 1, щодо системного розвитку програм навчання з кібербезпеки, оцінювання їхньої ефективності та вдосконалення з урахуванням змін організаційних потреб [2]. При цьому результати оцінювання доцільно використовувати насамперед для попередження інцидентів і розвитку компетентностей працівників, забезпечуючи належний захист персональних даних та уникаючи необґрунтованих висновків щодо надійності окремих осіб на підставі одного тестування.

Отже, оцінювання поведінкових ризиків персоналу є перспективним напрямом удосконалення системи протидії соціальній інженерії. Запропонований підхід передбачає поєднання контрольованих перевірок, аналізу поведінкових показників, сегментації ризиків і цільових навчальних заходів. Його застосування може сприяти своєчасному виявленню типових помилок персоналу, раціональному розподілу ресурсів кібербезпеки та формуванню стійкої корпоративної культури безпеки. Подальші дослідження доцільно спрямувати на практичну перевірку запропонованого підходу, обґрунтування критеріїв визначення рівнів ризику та оцінювання змін поведінки працівників у результаті повторного навчання.

Перелік посилань:

1. European Union Agency for Cybersecurity. ENISA Threat Landscape 2025. 2025. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025 (дата звернення: 24.10.2026).
2. Merritt M., Hansche S., Ellis B., Sanchez-Cherry K., Snyder J. N., Walden D. Building a Cybersecurity and Privacy Learning Program. Gaithersburg, MD: National Institute of Standards and Technology, 2024. (NIST Special Publication 800-50r1). DOI: https://doi.org/10.6028/NIST.SP.800-50r1 (дата звернення: 24.10.2026).

Литвинюк Владислав Вадимович
Студент групи ІКБм-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ СЕГМЕНТАЦІЇ КОРПОРАТИВНОЇ МЕРЕЖІ ДЛЯ РЕАЛІЗАЦІЇ КОНЦЕПЦІЇ ZERO TRUST НА ОБЛАДНАННІ CISCO

Традиційні підходи до захисту корпоративних мереж на основі периметрової безпеки виявляються недостатньо ефективними перед сучасними кіберзагрозами. Концепція Zero Trust пропонує нову парадигму, що вимагає безперервної верифікації та суворого контролю доступу. У роботі досліджується застосування макро- та мікросегментації мережі для реалізації архітектури нульової довіри з використанням систем Cisco ISE та технології Cisco TrustSec, що дозволяє ізолювати критичні активи та блокувати горизонтальне переміщення зловмисників.

Ключові слова: корпоративна мережа, концепція Zero Trust, макросегментація, мікросегментація, Cisco ISE, Cisco TrustSec, інформаційна безпека, контроль доступу.

Сучасне корпоративне середовище стикається з безпрецедентним зростанням кількості та складності кіберзагроз, що робить традиційну модель безпеки, орієнтовану виключно на захист периметра, застарілою. Стрімкий перехід до гібридних хмарних середовищ, масове впровадження пристроїв Інтернету речей (IoT) та поширення моделей віддаленої роботи призвели до розмиття або повного зникнення чітких меж корпоративної мережі. Як наслідок, у разі подолання зовнішнього ешелону захисту, наприклад через фішинг чи компрометацію VPN-з'єднання, зловмисники отримують можливість вільно переміщуватися всередині мережі (lateral movement), що неминуче призводить до витоку або шифрування критично важливих даних підприємства [1].

Ефективною та безальтернативною відповіддю на ці виклики є впровадження концепції Zero Trust (нульової довіри), яка базується на фундаментальному принципі «ніколи не довіряй, завжди перевіряй». Одним із головних та найбільш дієвих механізмів практичної реалізації цієї архітектури є багаторівнева сегментація мережі. На відміну від класичного, статичного розділення на VLAN чи використання громіздких списків контролю доступу (ACL), які важко масштабувати, сучасний підхід вимагає впровадження динамічної мікросегментації. У такій моделі політики доступу формуються не на основі IP-адрес, а на основі комплексної ідентифікації користувача, його пристрою, контексту підключення та поточного рівня ризику [2].

Для побудови такої відмовостійкої архітектури доцільно використовувати передові рішення компанії Cisco, зокрема платформу Cisco Identity Services Engine (ISE) та технологію Cisco TrustSec. Система Cisco ISE виступає централізованим ядром управління політиками (Policy Decision Point). Вона забезпечує глибоку ідентифікацію (через інтеграцію з Active Directory), автентифікацію за протоколом 802.1X та авторизацію кожної кінцевої точки. Крім того, ISE виконує профілювання пристроїв (зокрема IoT) та перевірку їхнього стану (Posture Assessment) на відповідність корпоративним стандартам безпеки перед наданням доступу [3].

Ключовою складовою запропонованої технології є мікросегментація за допомогою Cisco TrustSec. Ця технологія дозволяє призначати користувачам, пристроям та серверам спеціальні мітки груп безпеки (Security Group Tags – SGT)

безпосередньо під час їх підключення до мережі. Завдяки цьому комутатори, маршрутизатори та міжмережеві екрани (Policy Enforcement Points) застосовують політики контролю доступу на основі цих логічних міток. Для передачі міток через ділянки мережі, що не підтримують їх нативну інкапсуляцію в апаратних заголовках, використовується протокол SXP (SGT Exchange Protocol), що забезпечує зворотну сумісність із застарілим обладнанням [3].

Логіка управління політиками доступу будується навколо централізованої матриці TrustSec (TrustSec Matrix), яка визначає дозволену взаємодію між різними SGT-мітками. Замість адміністрування тисяч рядків статичних IP-орієнтованих ACL на кожному пристрої, адміністратор налаштовує SGACL (Security Group Access Control Lists) безпосередньо в Cisco ISE. Ці політики динамічно завантажуються на мережеві вузли, що дозволяє застосовувати мікросегментацію на рівні порту доступу.

У рамках дослідження також розглядається посилення архітектури Zero Trust за допомогою інтеграції Cisco ISE з системами аналізу мережевої поведінки (наприклад, Cisco Secure Network Analytics). Такий симбіоз забезпечує безперервний моніторинг телеметрії: при виявленні аномальної активності рівень довіри до кінцевої точки автоматично знижується. Система миттєво призначає пристрою карантинну SGT-мітку, ізолюючи загрозу в реальному часі. Розроблено моделі застосування такої мікросегментації як у традиційних мережах, так і в сучасних фабриках SD-Access (Software-Defined Access).

Процес практичного впровадження розробленої технології має відбуватися поетапно. На першому етапі система працює в режимі моніторингу (Monitor Mode), що дозволяє вивчити легітимні потоки трафіку та уникнути порушення бізнес-процесів. На наступних етапах впроваджується логічне розділення та повне блокування несанкціонованих з'єднань (Enforcement Mode). На макрорівні забезпечується відокремлення гостьового трафіку, IoT та корпоративних сервісів, а на рівні мікросегментації створюються ізольовані зони для окремих серверних додатків, що мінімізує поверхню потенційної атаки (blast radius).

Таким чином, розроблена технологія макро- та мікросегментації на базі Cisco ISE та TrustSec є високоефективним інструментом для практичної реалізації концепції Zero Trust у корпоративних мережах. Відмова від традиційних ACL на користь SGT-міток значно спрощує адміністрування, зменшує навантаження на мережеве обладнання та забезпечує гнучкість системи безпеки при масштабуванні інфраструктури. Запропоновані архітектурні рішення гарантують проактивний захист та високий рівень стійкості інформаційного середовища підприємства в умовах сучасного ландшафту кіберзагроз.

Перелік посилань:

1. Rose S., Borchert O., Mitchell S., Connelly S. Zero Trust Architecture. NIST Special Publication 800-207. 2020. URL: https://csrc.nist.gov/pubs/sp/800/207/final

2. Garbis J., Chapman J. W. Zero Trust Security: An Enterprise Guide. Apress, 2021. 312 p. URL: https://link.springer.com/book/10.1007/978-1-4842-6702-8

3. Cisco TrustSec Switch Configuration Guide. Cisco Systems. URL: https://www.cisco.com/c/en/us/td/docs/switches/lan/trustsec/configuration/guide/trustsec/config.html

Виноградов К.В.
Студент групи ІКБм-61 ННІКБЗІ ДУІКТ
Київ, Україна

ТЕХНОЛОГІЯ АВТОМАТИЗОВАНОГО ВИЯВЛЕННЯ ТА РЕАГУВАННЯ НА КІБЕРІНЦИДЕНТИ В SENTINELONE

Розглянуто підхід до автоматизації опрацювання кіберінцидентів у SentinelOne на основі штатних механізмів виявлення та зовнішнього модуля взаємодії через API. Запропоновано послідовність перевірки контексту загрози, результатів захисних дій і умов завершення інциденту. На прикладі спрацювань, пов'язаних зі списком блокування, визначено критерії встановлення вердикту True Positive та статусу Resolved. Обґрунтовано необхідність журналювання, обмеження повноважень і захисту від повторного виконання. Окреслено методику перевірки коректності рішень та оцінювання часу опрацювання спрацювань.

Ключові слова: SentinelOne, кіберінцидент, автоматизація, EDR, API, реагування.

Автоматизація опрацювання спрацювань у центрі операцій безпеки (SOC) потребує узгодження швидкості реагування з достовірністю прийнятих рішень. Повторювані перевірки стану загроз і результатів захисних дій збільшують навантаження на аналітика. Водночас механічне закриття повідомлень може приховати незавершене реагування. У рекомендаціях NIST SP 800-61 Rev. 3 виявлення, реагування та відновлення розглядаються як складові управління кіберризиками [1]. Тому автоматизований сценарій має забезпечувати перевірюваність рішення та збереження матеріалів для подальшого розслідування.

Метою роботи є обґрунтування технології автоматизованого виявлення та реагування на кіберінциденти в SentinelOne з додатковим опрацюванням спрацювань через програмний інтерфейс (API). Запропоновано поєднати штатне виявлення загроз на кінцевих пристроях із зовнішнім модулем, який застосовує визначені правила, перевіряє результат реагування та фіксує рішення. Предметом автоматизації є насамперед повторювані операції аналітика з підтвердженими спрацюваннями, для яких можна встановити однозначні умови оброблення.

SentinelOne Singularity Complete поєднує засоби запобігання загрозам із виявленням і реагуванням на кінцевих пристроях (EDR). Виробник описує використання статичних і поведінкових моделей для ідентифікації шкідливих ознак, автоматичне реагування відповідно до політик та можливість блокування мережевої активності пристрою [2]. Технологія Storyline пов'язує споріднені події в контекст атаки [3]. Такий контекст доцільно враховувати під час визначення меж інциденту, оскільки окреме спрацювання може бути лише частиною ширшої послідовності дій.

У запропонованій технології виявлення виконує агент SentinelOne, а зовнішній модуль організовує подальший робочий процес. Це дає змогу відокремити рішення про наявність загрози від адміністративних операцій з її записом у консолі. Зміна вердикту аналітика або статусу інциденту сама по собі не підтверджує припинення шкідливої активності. Підставою для завершення сценарію має бути перевірений стан захисних дій та відповідність спрацювання умовам затвердженого правила.

Архітектуру модуля пропонується побудувати з клієнта API, механізму перевірки правил, виконавця дій і журналу оброблення. Клієнт отримує записи в межах дозволених сайтів і часових інтервалів. Механізм правил зіставляє атрибути загрози з визначеними умовами. Виконавець надсилає дозволені запити, після чого повторно перевіряє стан запису. Журнал зберігає ідентифікатор загрози, версію правила, підставу рішення, час виконання та результат. Формати полів, доступні операції й обмеження запитів необхідно узгоджувати з документацією API конкретної консолі SentinelOne.

Базовим сценарієм є оброблення загрози, пов’язаної зі списком блокування (Blocklist). Модуль перевіряє, чи належить спрацювання до дозволеної області, чи підтверджено зв’язок із відповідним механізмом виявлення та чи збігається хеш із погодженим записом списку блокування. Самого збігу назви файлу або процесу недостатньо. Додатково перевіряються завершення необхідних захисних дій, відсутність їхніх помилок і невирішених пов’язаних ознак компрометації. За неповного контексту або суперечливих даних запис передається аналітику.

Лише після виконання всіх умов сценарій передбачає встановлення вердикту True Positive та переведення інциденту в стан Resolved із пояснювальним коментарем. Вердикт означає підтвердження загрози, а завершення опрацювання має спиратися на окрему перевірку результатів реагування. Правило також повинно враховувати ситуацію повторної появи того самого об’єкта: нове спрацювання потребує перевірки поточного стану пристрою, навіть якщо попередній запис уже закрито. Це запобігає автоматичному перенесенню застарілого рішення на нову активність.

Надійність оброблення залежить від повноти отримання даних і стійкості до повторних запусків. Пропонується використовувати посторінкове отримання записів, контрольну часову позначку та перекриття сусідніх часових вікон. Повторно отримані записи перевіряються за ідентифікатором, актуальним станом і версією правила. Перед зміною статусу модуль перечитує запис, щоб урахувати рішення іншого аналітика. У разі втрати відповіді на запит спочатку перевіряється фактичний результат операції. Помилка зв’язку не повинна автоматично трактуватися як успішне реагування.

Окремої уваги потребує розмежування типів записів у консолі. Для загроз і повідомлень від правил можуть застосовуватися різні структури даних та операції API. Тому модуль має спочатку визначати тип об’єкта, а потім використовувати відповідний обробник. Внутрішня модель може уніфікувати ідентифікатор пристрою, час виявлення, вердикт і стан реагування, проте повинна зберігати вихідні значення. Невідоме значення поля або відсутність обов’язкового атрибута є підставою зупинити автоматичну зміну цього запису. Такий підхід дає змогу уникнути застосування правила до об’єкта, семантика якого відрізняється від очікуваної.

Для безпечного застосування пропонується окремий службовий обліковий запис із мінімально необхідними правами та обмеженою областю керування. Секрети автентифікації мають зберігатися поза кодом і журналами, а

передавання даних здійснюватися захищеним з’єднанням. На початковому етапі модуль працює в режимі спостереження: формує рішення без зміни записів. Після перевірки правил дозволяється автоматичне завершення визначених сценаріїв. Дії з істотним впливом на роботу пристрою, зокрема мережева ізоляція, потребують окремих умов застосування та врахування критичності активу.

Керування правилами пропонується здійснювати через інтерфейс аналітика з переглядом умов, попередніх рішень і причин відхилення автоматизації. Зміна правила повинна створювати нову версію, щоб надалі можна було відтворити логіку конкретного рішення. Перед увімкненням оновлене правило перевіряється на збереженому наборі прикладів, включно з випадками, які не мають завершуватися автоматично. Доцільно передбачити зупинення сценарію при збільшенні частоти помилок або неочікуваному обсязі змін. Межі прийнятного обсягу визначаються під час апробації з урахуванням звичайного потоку спрацювань.

Методика перевірки передбачає відтворення контрольованих подій у тестовому середовищі й порівняння ручного та автоматизованого опрацювання однакового набору спрацювань. До набору слід включити підтверджену заблоковану загрозу, незавершене реагування, помилку захисної дії, невідповідний хеш, повторну подію та недоступність API. Для кожного випадку заздалегідь задається очікуваний результат: завершення сценарію, передавання аналітику або повторна перевірка. Окремо перевіряється, чи зберігаються всі записи після перезапуску модуля та чи не дублюються вже виконані зміни.

Ефективність пропонується оцінювати за тривалістю опрацювання спрацювання, часткою коректно автоматизованих рішень і кількістю помилково закритих записів. Час блокування загрози агентом потрібно вимірювати окремо від часу адміністративного завершення інциденту. Скорочення другого показника не доводить прискорення виявлення або локалізації атаки. До проведення випробувань очікуване зменшення ручних операцій слід розглядати як гіпотезу, а кількісний ефект визначати лише за фактичними результатами порівняння.

Отже, запропонована технологія доповнює штатні механізми SentinelOne правилами контрольованого опрацювання спрацювань через API. Її практична цінність полягає в можливості автоматизувати повторювані рішення зі збереженням доказової основи та історії виконання. Визначальними умовами застосування є перевірка успішності реагування, захист від повторних змін і передавання неоднозначних випадків аналітику.

Список використаних джерел

1. 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: 25.10.2026).

2. SentinelOne. Singularity Complete. URL: https://www.sentinelone.com/platform/singularity-complete/ (дата звернення: 25.10.2026).

3. SentinelOne. Singularity Endpoint Protection Platform. URL: https://www.sentinelone.com/platform/endpoint-protection-platform/ (дата звернення: 25.10.2026).

Довгошея Д. О.
студент групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ МОНІТОРИНГУ СТАНУ ДЖЕРЕЛ ДАНИХ ТА СПОВІЩЕННЯ АДМІНІСТРАТОРА У SPLUNK ENTERPRISE SECURITY

Розглянуто технологію моніторингу стану джерел даних у Splunk Enterprise Security (ES) та автоматичного сповіщення адміністратора про їх несправності. Описано принцип виявлення «мовчазних» джерел на основі аналізу часу надходження останньої події, реалізацію контролю за допомогою планових кореляційних пошуків та механізми реагування: формування notable events і надсилання повідомлень. Запропоновано підхід до визначення порогових значень для різних типів джерел.

Ключові слова: Splunk Enterprise Security, SIEM, моніторинг джерел даних, кореляційний пошук, сповіщення.

Ефективність центру моніторингу безпеки (SOC) безпосередньо залежить від повноти та актуальності даних, що надходять до SIEM-системи. Якщо джерело (міжмережевий екран, контролер домену, EDR-агент, веб-сервер) припиняє надсилати події, кореляційні правила перестають спрацьовувати, а атака залишається непоміченою. Тому постійний контроль стану джерел даних є необхідною умовою надійної роботи системи [1].

Splunk Enterprise Security — це SIEM-рішення, побудоване на платформі Splunk Enterprise. Дані надходять через універсальні forwarder, HTTP Event Collector, syslog та API сторонніх сервісів, зберігаються в індексах і нормалізуються відповідно до моделі Common Information Model (CIM) [1]. Типовими проблемами на цьому шляху є зупинка forwarder, збій мережевого з’єднання, зміна формату журналів після оновлення програмного забезпечення та переповнення черг індексації. Усі вони проявляються однаково: нові події від джерела перестають з’являтися в індексах.

Базовий принцип технології полягає в обчисленні для кожної комбінації index, sourcetype та host часу надходження останньої події. Для швидкого отримання цієї інформації використовується команда tstats, яка працює з метаданими індексів і не сканує сирі події [2]. Типовий пошук має вигляд: | tstats latest(_time) AS lastTime WHERE index=* BY index, sourcetype, host. Відставання джерела обчислюється як різниця між поточним часом та часом останньої отриманої події.

Джерело вважається неактивним, якщо це відставання перевищує задане порогове значення. Додатково доцільно контролювати обсяг подій за інтервал часу: різке зменшення кількості повідомлень порівняно зі звичним рівнем також свідчить про часткову втрату даних.

Порогове значення не може бути однаковим для всіх джерел. Для контролерів домену та міжмережевих екранів допустима пауза складає лише кілька хвилин, тоді як для щоденних звітів або резервних систем вона може сягати доби. Тому параметри контролю (індекс, sourcetype, поріг, відповідальна особа, критичність) доцільно зберігати в таблиці пошуку (lookup), яку пошук

зіставляє з результатами tstats. Такий підхід дозволяє змінювати пороги та додавати нові джерела без редагування самого пошуку.

У Splunk ES перевірка оформлюється як кореляційний пошук (correlation search), що виконується за розкладом, наприклад, кожні 15 хвилин. Якщо умова виконується, система створює notable event, який з'являється на панелі Incident Review і може бути призначений відповідальному аналітику чи адміністратору [1]. Щоб уникнути повторних повідомлень про одне й те саме джерело, налаштовується режим throttling, який блокує повторне спрацювання протягом заданого часу. Структуру процесу наведено на рис. 1.



Рис. 1. Схема моніторингу джерел даних та сповіщення адміністратора

Для безпосереднього сповіщення адміністратора до кореляційного пошуку підключаються адаптивні дії та дії сповіщення: надсилання електронного листа, виклик webhook до системи обліку заявок або месенджера, запуск скрипта автоматичного перезапуску служби [1]. Стан самої платформи Splunk (черги індексації, активність forwarder, використання ресурсів) додатково контролюється засобами Monitoring Console та Splunk Health Report [3], що дозволяє розрізнити проблему на стороні джерела та несправність інфраструктури збору даних.

Повний цикл роботи пошуку складається з трьох кроків. Спочатку tstats повертає перелік джерел із часом останньої події. Потім результати об'єднуються з таблицею пошуку через команду lookup, після чого обчислюється відставання та порівнюється з порогом. Нарешті, до результату потрапляють лише ті записи, для яких поріг перевищено, і саме вони перетворюються на notable events. Для кожного такого запису в повідомлення включаються назва джерела, час останньої події, величина відставання та контакт відповідальної особи, що суттєво спрощує початкову діагностику.

Практичне застосування описаної технології скорочує час виявлення збоїв у збиранні журналів, зменшує кількість «сліпих зон» у системі безпеки та підвищує довіру до результатів кореляційного аналізу. Для ефективної експлуатації рекомендується: класифікувати джерела за критичністю, закріпити за кожним відповідальну особу, періодично переглядати порогові значення та перевіряти не лише наявність подій, а й коректність розбору полів відповідно до CIM.

Література

1. Splunk Enterprise Security Documentation. URL: https://docs.splunk.com/Documentation/ES (дата звернення: 23.10.2026)
2. Splunk Enterprise Search Reference: tstats. URL: https://docs.splunk.com/Documentation/Splunk/latest/SearchReference/Tstats (дата звернення: 23.10.2026)
3. Splunk Enterprise: Monitoring Splunk Enterprise (Monitoring Console). URL: https://docs.splunk.com/Documentation/Splunk/latest/DMC/DMCoverview (дата звернення: 23.10.2026)

Чуєва С.О.
студентка групи БІКСм-25-1, ХНУРЕ,
Харків, Україна

ЗАСТОСУВАННЯ ЛОКАЛЬНИХ МАЛИХ МОВНИХ МОДЕЛЕЙ ДЛЯ ПОЯСНЮВАЛЬНОГО АНАЛІЗУ КІБЕРЗАГРОЗ В ІЗОЛЬОВАНИХ МЕРЕЖЕВИХ СЕРЕДОВИЩАХ

Розглянуто проблему забезпечення конфіденційності та прозорості під час розслідування кіберінцидентів у повністю ізольованих мережевих середовищах. Обґрунтовано доцільність використання локальних малих мовних моделей (Small Language Models, SLM) як інтерпретаторів пояснювального штучного інтелекту (Explainable AI, XAI) для подолання проблеми «чорної скрині» наявних засобів захисту. Описано принцип функціонування модуля семантичної підготовки даних, який конвертує розширений масив статичних індикаторів у структурований текстовий контекст. Показано, що детерміноване зіставлення індикаторів із базою знань MITRE ATT&CK за допомогою вбудованих програмних алгоритмів у поєднанні з мовною моделлю дає змогу формувати аргументовані висновки природною мовою та покрокові сценарії реагування (Playbook).

Ключові слова: ізольоване мережеве середовище (air-gapped network), малі мовні моделі (SLM), пояснювальний штучний інтелект (XAI), семантична підготовка даних, сценарій реагування (Playbook), статичний аналіз коду.

Сучасні засоби захисту інформації та комплексні рішення розширеного виявлення та реагування застосовують широкий спектр методів виявлення загроз – від евристичного та поведінкового моніторингу до алгоритмів машинного навчання. При цьому більшість наявних інструментів функціонує за принципом «чорної скрині», надаючи користувачу або аналітику лише двозначний висновок (шкідливий чи безпечний) або формальний код загрози без розкриття причинно-наслідкової логіки прийняття рішень [1].

Така непрозорість створює високе когнітивне навантаження на спеціалістів центрів моніторингу безпеки (Security Operations Center, SOC), вимагаючи додаткового тривалого ручного сортування складних індикаторів компрометації [1]. Водночас спроби отримати розгорнуту текстову аналітику за допомогою хмарних сервісів супроводжуються критичними ризиками витоку конфіденційної інформації, що є неприпустимим для державних установ, об’єктів критичної інфраструктури та підприємств із повністю ізольованими мережевими середовищами [2].

Для розв’язання проблеми «чорної скрині» та забезпечення абсолютної автономності аналізу запропоновано застосування локальних SLM. Сучасні малі мовні моделі за умови подачі чітко структурованого контексту здатні ефективно вирішувати завдання аналізу безпеки та інтерпретації загроз, не поступаючись за функціональністю великим хмарним аналогам [2]. Головною перевагою локальних моделей є можливість їх виконання безпосередньо на робочих станціях операторів із помірними апаратними ресурсами без зовнішніх мережевих запитів.

Передавання двійкових даних у мовну модель є неефективним через обмеженість контекстного вікна та ризик генерації неправдивих фактів (галюцинацій) [3]. Для усунення цього обмеження розроблено модуль

семантичної підготовки даних. Цей модуль виконує первинне вилучення статичних індикаторів: заголовків PE-файлів, таблиці імпортованих функцій, посекційної ентропії Шеннона, сигнатур YARA в оперативній пам’яті, цифрових підписів, нечітких хешів TLSH, результатів поведінкового аналізу CAPA та лексичних характеристик URL-адрес. Отримані технічні параметри конвертуються у структурований текстовий опис, який відображає логічні взаємозв’язки між виявленими аномаліями.

Зіставлення виявлених індикаторів із тактиками та техніками бази знань MITRE ATT&CK здійснюється детерміновано на рівні алгоритмів комплексу та поведінкового компонента CAPA. Сформований контекст передається на вхід локальній мовній моделі, яка виконує роль пояснювального інтерпретатора XAI. Модель не замінює функцій захисних засобів, а трансформує накопичені технічні факти у зв’язний експертний висновок природною мовою, пояснюючи причини віднесення об’єкта до категорії шкідливих, підозрілих чи безпечних, та доповнює його покроковим сценарієм Playbook [4].

Результати дослідження підтверджують ефективність поєднання статичного аналізу з локальними малими мовними моделями для розв’язання проблеми «чорної скрині». Запропонований підхід забезпечує прозору аналітику загроз та автоматизацію сценаріїв реагування на інциденти в повністю ізольованому середовищі без ризику витоку даних.

Перелік посилань:

1. Sharma A., Rani S., Shabaz M. A comprehensive review of explainable AI in cybersecurity: Decoding the black box. ICT Express. 2025. URL: https://www.sciencedirect.com/science/article/pii/S2405959525001584 (дата звернення: 24.10.2026).

2. Schmidt H., Kaiser K., Spiekermann D. Language Models in Cybersecurity: A Comparative Approach to Task-Driven Model Assessment. SPAIML@ ECAI. 2025. P. 43–54. URL: https://ceur-ws.org/Vol-4154/paper5.pdf (дата звернення: 24.10.2026).

3. Gao C., Fan G., Chong C. Y. [et al.]. A systematic literature review of code hallucinations in LLMs: Characterization, mitigation methods, challenges, and future directions for reliable AI. arXiv preprint arXiv:2511.00776. 2025. URL: https://arxiv.org/abs/2511.00776 (дата звернення: 24.10.2026).

4. Jaffal N. O., Alkhanafseh M., Mohaisen D. Large Language Models in Cybersecurity: A Survey of Applications, Vulnerabilities, and Defense Techniques. AI. 2025. Vol. 6, No. 9. P. 216. DOI: https://doi.org/10.3390/ai6090216.

Середін Андрій Валерійович
Студент групи ІКБМ-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ УПРАВЛІННЯ ІДЕНТИФІКАЦІЄЮ ТА БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ ІНФОРМАЦІЙНИХ СИСТЕМ НА БАЗІ KEYCLOAK

Інформаційні системи потребують ефективного управління обліковими записами та надійної перевірки особи користувача. Використання лише пароля створює ризики несанкціонованого доступу через фішинг, витоки даних, повторне використання паролів і підбір слабких комбінацій. Тому важливими завданнями кібербезпеки є централізоване управління ідентифікацією, розмежування прав доступу та застосування багатофакторної автентифікації. Метою роботи є розгляд основних проблем управління ідентифікацією користувачів, методів автентифікації та можливостей сучасних засобів їх централізованого управління.

Ключові слова: ідентифікація користувачів, автентифікація, авторизація, багатофакторна автентифікація, інформаційна безпека, Keycloak, контроль доступу.

Управління ідентифікацією користувачів охоплює створення, зміну, блокування та видалення облікових записів, а також підтримання актуальності даних і призначення прав доступу. У цьому процесі важливо розрізняти три поняття: ідентифікація визначає, ким заявляє себе користувач; автентифікація перевіряє цю заяву; авторизація визначає, які дії дозволені користувачу після входу. Недоліки на будь-якому з цих етапів можуть призвести до надмірних привілеїв або доступу сторонніх осіб до інформаційних ресурсів [1; 3].

Одним із поширених засобів підтвердження особи є пароль, однак його надійність залежить від складності, унікальності та захищеного зберігання. Додатковий рівень захисту забезпечує багатофакторна автентифікація (Multi-Factor Authentication, MFA), яка використовує щонайменше два різні фактори: знання (наприклад, пароль), володіння (наприклад, смартфон або апаратний ключ) чи біометричну характеристику. Поєднання різних факторів зменшує ймовірність компрометації облікового запису в разі викрадення лише пароля [3; 4].

Під час вибору засобу управління ідентифікацією враховують спосіб розгортання, можливості інтеграції, адміністрування користувачів і ролей, а також підтримку стандартних протоколів. Служби каталогів забезпечують централізоване зберігання облікових записів у корпоративному середовищі, хмарні платформи ідентифікації пропонують керування доступом до онлайн-сервісів, а рішення класу Identity and Access Management поєднують функції

управління користувачами, автентифікації та контролю доступу. Вибір рішення залежить від вимог організації, інфраструктури й доступних ресурсів.

Одним із засобів централізованого управління ідентифікацією є Keycloak — програмна платформа з відкритим вихідним кодом. Вона дає змогу створювати користувачів і групи, призначати ролі, налаштовувати політики входу та інтегрувати застосунки за допомогою OpenID Connect і SAML. Підтримка єдиного входу (Single Sign-On, SSO) дозволяє користувачам автентифікуватися для сумісних застосунків через централізований механізм [1].

Для посилення захисту облікових записів у Keycloak може налаштовуватися багатофакторна автентифікація з використанням одноразових кодів TOTP (Time-Based One-Time Password). Код формується на основі спільного секретного ключа та поточного часу, а користувач вводить його після основного пароля [2]. Такий підхід дозволяє перевірити не лише знання пароля, а й володіння пристроєм, на якому працює застосунок-автентифікатор. Водночас належний рівень захисту потребує коректного налаштування політик, безпечного зберігання облікових даних і відповідального адміністрування прав доступу [1; 3].

Для подальшого дослідження Keycloak доцільно розглядати як середовище, у якому можна послідовно перевірити основні сценарії управління користувачами. До них належать створення тестового облікового запису, призначення ролі, налаштування способу входу та активація TOTP. Під час перевірки важливо встановити, чи може користувач увійти з правильним паролем і одноразовим кодом, та чи відхиляється вхід за неправильного або відсутнього коду. Така перевірка дозволяє оцінити не лише наявність відповідної функції в системі, а й правильність її налаштування. Алгоритм TOTP визначає одноразовий код на основі секретного ключа та часу, тому для його коректної роботи важливі захищене зберігання секрету й узгодженість часу на пристроях [1; 2].

Застосування централізованої платформи допомагає узгодити процес входу до кількох інформаційних систем. Наприклад, інтеграція через OpenID Connect або SAML дає змогу делегувати перевірку облікових даних спеціалізованому сервісу ідентифікації, а не реалізовувати окремий механізм у кожному застосунку. Єдиний вхід SSO може спростити роботу користувачів і адміністраторів, однак потребує належного захисту самої платформи, її адміністративних облікових записів і конфігурації інтегрованих застосунків. Отже, централізація має супроводжуватися обмеженням адміністративних прав, контролем налаштувань і регулярним переглядом доступів [1].

Під час аналізу методів автентифікації важливо враховувати не лише зручність входу, а й стійкість до типових атак. Пароль може бути викрадений через фішингову сторінку, повторно використаний після витоку іншого сервісу

або підібраний у разі недостатньої складності. MFA створює додатковий бар’єр, оскільки для входу потрібно підтвердити особу ще одним фактором. Водночас фактори мають бути різних типів: використання двох паролів саме по собі не є багатофакторною автентифікацією. Організація також повинна визначити порядок відновлення доступу, адже ненадійна процедура відновлення може послабити весь механізм захисту [3; 4].

Окремою проблемою є життєвий цикл облікового запису. Під час прийняття працівника або створення нового користувача необхідно надати лише ті права, які потрібні для виконання його завдань. У разі зміни обов’язків права слід переглядати, а після припинення співпраці обліковий запис необхідно своєчасно блокувати або видаляти. Якщо ці дії виконуються вручну в різних системах, зростає ризик помилок, появи невикористовуваних облікових записів і збереження надлишкових привілеїв. Централізоване управління дає змогу упорядкувати такі процеси та застосовувати єдині правила доступу [1; 3].

Отже, управління ідентифікацією та багатофакторна автентифікація є важливими складовими захисту інформаційних систем. Розмежування ідентифікації, автентифікації й авторизації допомагає визначити основні вимоги до управління доступом, а застосування MFA знижує залежність захисту від одного пароля. Keycloak поєднує централізоване управління користувачами, ролями та підтримку стандартних протоколів із можливістю налаштування TOTP. Це робить платформу доречною для подальшого розгляду та перевірки налаштувань багатофакторної автентифікації в тестовому середовищі.

Джерела:

1. Keycloak. Server Administration Guide. URL: https://www.keycloak.org/documentation (дата звернення: 19.10.2026).

2. M’Raihi D., Machani S., Pei M., Rydell J. TOTP: Time-Based One-Time Password Algorithm. RFC 6238. Internet Engineering Task Force, 2011. URL: https://www.rfc-editor.org/rfc/rfc6238 (дата звернення: 19.10.2026).

3. National Institute of Standards and Technology. Digital Identity Guidelines: Authentication and Authenticator Management. NIST SP 800-63B-4. 2025. URL: https://csrc.nist.gov/pubs/sp/800/63/b/4/final (дата звернення: 19.10.2026).

4. OWASP Foundation. Multifactor Authentication Cheat Sheet. URL: https://cheatsheetseries.owasp.org/cheatsheets/Multifactor_Authentication_Cheat_Sheet.html (дата звернення: 19.10.2026).

Стожок М.Р., студент групи ІКБм-61
Науковий керівник: Борсуковський Ю.В
Державний університет інформаційно-комунікаційних технологій, м. Київ

ТЕХНОЛОГІЯ АВТОМАТИЗОВАНОГО ВИЯВЛЕННЯ ТА РИЗИК-ОРІЄНТОВАНОЇ ПРІОРИТИЗАЦІЇ ВРАЗЛИВОСТЕЙ МЕРЕЖЕВИХ СЕРВІСІВ КОРПОРАТИВНОЇ МЕРЕЖІ

Анотація. Розглянуто проблему пріоритизації вразливостей мережевих сервісів корпоративної мережі, виявлених засобами автоматизованого сканування. Запропоновано технологію, що поєднує інвентаризацію сервісів, виявлення відомих вразливостей та їх ранжування за інтегральним показником, який враховує базову оцінку CVSS, імовірність експлуатації за моделлю EPSS, наявність вразливості в каталозі CISA KEV та критичність активу. Описано тестовий стенд для експериментальної перевірки технології та прототип засобу її автоматизації мовою Python.

Корпоративна мережа сучасної організації містить десятки та сотні мережевих сервісів: вебсервери, поштові сервери, служби віддаленого доступу (SSH, RDP), файлові сервіси (SMB, FTP), системи керування базами даних тощо. Кожен такий сервіс є потенційною точкою входу для зловмисника, а застарілі версії програмного забезпечення та помилки конфігурації залишаються одними з найпоширеніших причин успішних атак. Щорічно в базі CVE публікуються десятки тисяч нових записів про вразливості, тому навіть невелика організація регулярно стикається з сотнями знахідок після кожного сканування.

Традиційний підхід до обробки результатів сканування ґрунтується на базовій оцінці CVSS (Common Vulnerability Scoring System) [1]. Однак CVSS характеризує технічну тяжкість вразливості, а не ймовірність її реальної експлуатації. Дослідження, на яких ґрунтується система EPSS [2], показують, що реально експлуатується лише невелика частка опублікованих вразливостей. Унаслідок цього значна кількість знахідок отримує «високий» або «критичний» рівень, а дійсно небезпечні вразливості можуть залишатися неусуненими.

Метою роботи є розроблення технології автоматизованого виявлення вразливостей мережевих сервісів корпоративної мережі з ризик-орієнтованою пріоритизацією, яка дозволяє визначити першочергові заходи з їх усунення. Технологія ґрунтується на відкритих інструментах і може застосовуватися в організаціях малого та середнього бізнесу. Об’єктом дослідження є процес виявлення та оцінювання вразливостей мережевих сервісів, предметом – методи та засоби їх автоматизованої ризик-орієнтованої пріоритизації.

Запропонована технологія включає п’ять етапів: 1) інвентаризація хостів і мережевих сервісів з визначенням їх версій (Nmap) [4]; 2) виявлення відомих вразливостей (CVE) за версіями сервісів (скрипт NSE vulners, шаблони Nuclei); 3) збагачення даних про кожну вразливість – базова оцінка CVSS, ймовірність експлуатації за моделлю EPSS та наявність у каталозі відомих експлуатованих вразливостей CISA KEV [3]; 4) розрахунок інтегрального показника пріоритету з урахуванням критичності активу; 5) формування звіту з ранжованим переліком вразливостей (рис. 1).

Для визначення черговості усунення вразливостей пропонується інтегральний показник пріоритету R, що поєднує технічну тяжкість, імовірність експлуатації та цінність активу:

$$R = (w_1 \cdot \frac{S_{cvss}}{10} + w_2 \cdot P_{epss} + w_3 \cdot K_{kev}) \cdot W_a, \quad (1)$$

де S_{cvss} — базова оцінка CVSS (від 0 до 10) [1]; P_{epss} — імовірність експлуатації вразливості протягом 30 днів за EPSS (від 0 до 1) [2]; K_{kev} — ознака наявності вразливості в каталозі CISA KEV (1 або 0) [3]; w_1, w_2, w_3 — вагові коефіцієнти ($w_1 + w_2 + w_3 = 1$, попередньо 0,4; 0,3; 0,3); W_a — коефіцієнт критичності активу (1,0 — звичайний вузол, 1,25 — важливий сервер, 1,5 — критичний сервіс).

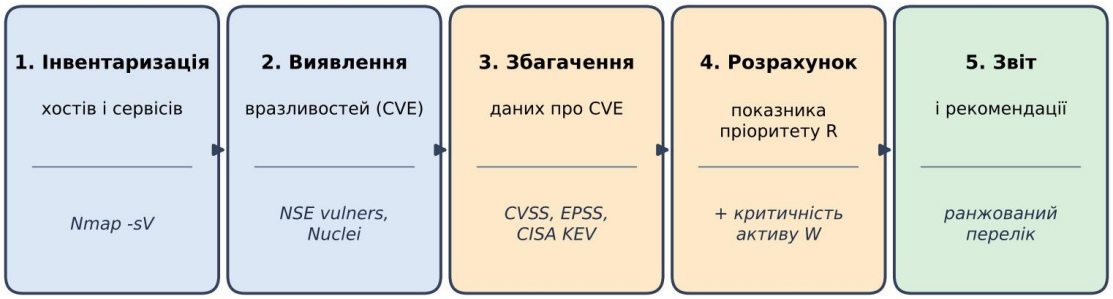


Рис. 1. Етапи технології автоматизованого виявлення та пріоритизації вразливостей

Для експериментальної перевірки технології спроєктовано ізольований тестовий стенд у середовищі VirtualBox, що моделює сегмент корпоративної мережі: вузол-сканер під керуванням Linux та навмисно вразливі вузли (Metasploitable 2 та інші) з типовими сервісами FTP, SSH, Telnet, SMB, HTTP, MySQL, PostgreSQL. Прототип засобу автоматизації мовою Python запускає сканування, отримує актуальні дані EPSS та CISA KEV, розраховує показник R за формулою (1) і формує звіт. Ефективність технології планується оцінювати порівнянням ранжування вразливостей за «чистою» оцінкою CVSS та за показником R. Як критерії ефективності розглядаються частка вразливостей із каталогу CISA KEV та з високим значенням EPSS серед перших позицій звіту, кількість вразливостей, віднесених до першочергових, та час формування звіту.

Таким чином, запропонована технологія поєднує автоматизоване виявлення вразливостей мережевих сервісів з ризик-орієнтованою пріоритизацією на основі відкритих джерел даних і дозволяє адміністраторам зосередити обмежені ресурси на вразливостях, що мають високу ймовірність експлуатації або вже використовуються в реальних атаках.

Література:

1. FIRST. Common Vulnerability Scoring System v4.0: Specification Document. URL: https://www.first.org/cvss/v4-0/specification-document (дата звернення: 20.10.2026).
2. Jacobs J., Romanosky S., Edwards B., Adjerid I., Roytman M. Exploit Prediction Scoring System (EPSS) // Digital Threats: Research and Practice. 2021. Vol. 2, No. 3. Article 20.
3. CISA. Known Exploited Vulnerabilities Catalog. URL: https://www.cisa.gov/known-exploited-vulnerabilities-catalog (дата звернення: 20.10.2026).
4. Lyon G. Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning. URL: https://nmap.org/book/ (дата звернення: 20.10.2026).

Архипенко Дмитро Євгенович
студент групи ІКБм-62, ННІЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ МЕРЕЖЕВИХ АНОМАЛІЙ ІЗ ЗАСТОСУВАННЯМ МАШИННОГО НАВЧАННЯ НА БАЗІ ZEEK

Розглянуто, як систему моніторингу мережі Zeek і алгоритми машинного навчання можна поєднати для виявлення мережевих аномалій. Пояснено, які дані дає Zeek і чим відрізняються навчання на розмічених прикладах та пошук відхилень без позначок. Окрему увагу приділено якості даних: відомі помилки популярного набору CIC-IDS2017 можуть завищувати оцінки систем. Сформульовано мету дослідження, яке можна виконати на безкоштовному програмному забезпеченні в умовах невеликого SOC.

Традиційні системи захисту шукають у трафіку відомі ознаки атак, унаслідок цього нові або змінені атаки можуть залишитися непоміченими. Тому дедалі більше уваги приділяють виявленню аномалій, тобто пошуку всього, що відрізняється від звичайної поведінки мережі. Машинне навчання варто розглядати не як заміну сигнатурних систем, а як їх доповнення, особливо для невеликих центрів моніторингу безпеки (SOC), яким потрібне недороге й відкрите рішення.

Zeek - система моніторингу мережевої активності з відкритим кодом. Вона не блокує трафік, а уважно «слухає» його й веде докладні журнали, схожі на нотатки про кожну розмову в мережі. У журналі з'єднань (conn.log) для кожного з'єднання записано, хто з ким спілкувався, скільки це тривало, скільки даних передано, який протокол використано та чим усе завершилося. Такі журнали зручно перетворювати на таблицю для алгоритмів машинного навчання (МН), і їх уже використовують у наукових наборах даних, зокрема UWF-ZeekData22 [3].

Для виявлення аномалій застосовують два підходи. У першому моделі показують приклади звичайного й шкідливого трафіку з готовими позначками, і вона вчиться їх розрізняти; так працюють Random Forest і Gradient Boosting. Такий підхід може давати кращі результати, якщо розмічених даних достатньо і вони репрезентативні. У другому позначок немає: модель, наприклад Isolation Forest, вивчає, як виглядає звичайний трафік, і вважає підозрілим усе, що від нього сильно відрізняється. Однак відхилення від норми не завжди є атакою, тому такі сповіщення потребують перевірки.

Для навчання моделей потрібні дані, і тут виникає проблема. Багато досліджень перевіряє моделі на наборі CIC-IDS2017 і повідомляє про дуже високу точність. Однак аналіз показав, що в цьому наборі є помилки: у способі створення трафіку, побудові потоків і розмітці, тобто в позначках «атака» чи «норма» [1]. Після виправлення набору результати перевірених алгоритмів у більшості показників покращилися, а деякі прості моделі стали конкурентними, що дає підстави переглянути вибір алгоритмів [2]. Крім того, у дослідженні [4] моделі майже ідеально працювали всередині одного набору, але на іншому

наборі якість різко падала. Отже, висока точність на «сирому» наборі ще не доводить, що система добре працюватиме в реальній мережі.

Мета роботи - дослідити вплив якості вихідних даних і способу перенесення міток на результати виявлення аномалій за ознаками з журналів Zeek. Складність у тому, що CIC-IDS2017 створено програмою CICFlowMeter, яка обчислює ознаки інакше, ніж Zeek, тому розмітку набору потрібно коректно зіставити з'єднаннями Zeek.

Для оцінки таких систем використовують Recall (частка знайдених атак), Precision (частка правильних сповіщень серед усіх), F1 (баланс між ними) та FPR (частка звичайних з'єднань, помилково названих атаками). Загальної точності недостатньо: атак у реальному трафіку значно менше, ніж звичайних з'єднань, тому модель, що не помічає жодної атаки, все одно матиме високу точність. Також важлива абсолютна кількість хибних тривог, бо вона визначає навантаження на аналітиків SOC.

Перелік посилань:

1. Engelen G., Rimmer V., Joosen W. Troubleshooting an Intrusion Detection Dataset: the CICIDS2017 Case Study. 2021 IEEE Security and Privacy Workshops (SPW). IEEE, 2021. P. 7–12. DOI: 10.1109/SPW53761.2021.00009

2. Rosay A., Cheval E., Carlier F., Leroux P. Network Intrusion Detection: A Comprehensive Analysis of CIC-IDS2017. Proc. of the 8th Int. Conf. on Information Systems Security and Privacy (ICISSP). SCITEPRESS, 2022. P. 25–36.

3. Bagui S. S., Mink D., Bagui S. C. et al. Detecting Reconnaissance and Discovery Tactics from the MITRE ATT&CK Framework in Zeek Conn Logs Using Spark's Machine Learning in the Big Data Framework. Sensors. 2022. Vol. 22, No. 20. Art. 7999.

4. Cantone M., Marrocco C., Bria A. On the Cross-Dataset Generalization of Machine Learning for Network Intrusion Detection. arXiv:2402.10974. 2024.

Гавриленко Д.П.
студент групи ІКБм-61, ННІКБЗІ, ДУІКТ, Київ, Україна

МЕТОДИ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ОПТИМІЗАЦІЇ ПРОЦЕСІВ МОНІТОРИНГУ ТА РЕАГУВАННЯ НА ІНЦИДЕНТИ В СУЧАСНОМУ SOC

Розглянуто застосування штучного інтелекту для моніторингу та реагування на інциденти в центрі операцій кібербезпеки. Метою дослідження є визначення відповідності методів завданням SOC та умов їх доцільного використання. На основі порівняльного аналізу підходів систематизовано можливості виявлення аномалій, класифікації сповіщень і підтримки рішень аналітика. Запропоновано поєднання спеціалізованих моделей із мовною моделлю та визначено критерії оцінювання такого підходу. Окреслено обмеження донавчання, ризики автоматизації та напрями подальшої перевірки дослідницької гіпотези.

Ключові слова: штучний інтелект, кібербезпека, SOC, моніторинг, реагування на інциденти, машинне навчання, мовні моделі.

Центр операцій кібербезпеки (Security Operations Center, SOC) має перетворювати потік сповіщень на обґрунтовані рішення щодо загроз. NIST SP 800-61 Rev. 3 розглядає виявлення, реагування та відновлення в контексті управління кіберризиками [1]. Тому оптимізацію SOC доцільно пов’язувати з якістю рішень і наслідками інцидентів, а не лише зі швидкістю оброблення сповіщень.

Метою дослідження є визначення напрямів застосування методів штучного інтелекту в SOC та обґрунтування умов їх доцільності. Використано аналіз наукових і нормативних джерел та порівняльне зіставлення підходів за призначенням, вимогами до даних, пояснюваністю й ресурсними витратами. Досліджується розподіл функцій між моделями та аналітиком зі збереженням контролю над ризиком пропуску загрози.

Для моніторингу пропонується розрізняти виявлення відхилень і розпізнавання відомих класів активності. Методи виявлення аномалій придатні для пошуку незвичної поведінки користувачів або вузлів, однак відхилення саме по собі не доводить наявності атаки. Навчання з учителем доцільно розглядати для класифікації та пріоритизації сповіщень, коли доступні достовірні попередні оцінки аналітиків. Узагальнювальна здатність таких моделей залежить від відповідності навчальних прикладів середовищу застосування; методологічні помилки можуть створювати завищене уявлення про якість [2].

Мовні моделі доцільно досліджувати насамперед як засіб роботи з текстовим контекстом: пояснення сповіщень, узагальнення пов’язаних подій, пошуку відомостей у внутрішніх інструкціях і підготовки варіантів подальшої перевірки. Згенероване пояснення потребує перевірки: переконливість викладу не є підтвердженням правильності висновку.

Для підтримки реагування пропонується поєднувати відомості про інцидент із затвердженими сценаріями дій організації. Пошук релевантних

документів із подальшим включенням їх до контексту запиту (RAG) дає змогу надавати моделі актуальні інструкції без зміни її ваг. Донавчання слід розглядати окремо — для адаптації до термінології, категорій інцидентів і формату відповідей. Метод LoRA зберігає базові ваги та навчає низькорангові матриці оновлення, скорочуючи кількість навчуваних параметрів [3]. Це не гарантує достовірності відповідей або переваги над простішими методами.

Робоча гіпотеза дослідження полягає в тому, що розподіл функцій між спеціалізованими моделями та мовною моделлю може бути доцільнішим за покладання всіх завдань на одну універсальну систему. У запропонованій схемі засоби моніторингу формують сповіщення, моделі класифікації визначають черговість аналізу, а мовна модель узагальнює доступний контекст і пропонує наступні кроки. Аналітик перевіряє висновки та погоджує дії, здатні вплинути на інформаційні ресурси. Ефективність цієї концептуальної схеми потребує перевірки.

Критерії доцільності мають охоплювати точність класифікації, повноту виявлення загроз, частоту хибних спрацювань, затримку та сукупні витрати. Для мовної моделі додатково важливі фактична обґрунтованість відповіді й відповідність запропонованих дій регламентам. Ресурсну оцінку слід доповнювати витратами на донавчання, підтримання бази знань і перевірку відповідей людиною. Вплив на тривалість розслідування потребує оцінювання роботи аналітика.

Окремим обмеженням є безпека самого інтелектуального компонента. Текст сповіщення або зовнішнього документа може містити інструкції, що змінюють поведінку мовної моделі. OWASP визначає це як ризик ін’єкції інструкцій; RAG і донавчання не усувають його повністю [4]. Тому в запропонованій схемі потрібні розмежування довірених інструкцій і зовнішнього вмісту, мінімальні права доступу та погодження критичних дій. Конфіденційність даних також має впливати на вибір між локальним розгортанням і зовнішнім сервісом.

Отже, результатом теоретичного дослідження є систематизація ролей методів ШІ та концептуальна схема їх поєднання в SOC. Доцільність застосування визначається відповідністю конкретному завданню, ціною помилки й витратами на перевірку результату. Перспективним напрямом подальшого дослідження є порівняння початкової та донавченої мовних моделей зі спеціалізованим класифікатором для первинного аналізу інцидентів.

Список використаних джерел

1. 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. DOI: 10.6028/NIST.SP.800-61r3. URL: https://csrc.nist.gov/pubs/sp/800/61/r3/final (дата звернення: 25.10.2026).
2. Arp D. et al. Dos and Don’ts of Machine Learning in Computer Security. 31st USENIX Security Symposium. 2022. P. 3971–3988. URL: https://www.usenix.org/conference/usenixsecurity22/presentation/arp (дата звернення: 25.10.2026).
3. Hu E. J. et al. LoRA: Low-Rank Adaptation of Large Language Models. 2021. arXiv:2106.09685. DOI: 10.48550/arXiv.2106.09685. URL: https://arxiv.org/abs/2106.09685 (дата звернення: 25.10.2026).
4. OWASP GenAI Security Project. LLM01:2025 Prompt Injection. URL: https://genai.owasp.org/llmrisk/llm01-prompt-injection/ (дата звернення: 25.10.2026).

Душник Володимир Володимирович
Студент групи ІКБМ-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАСТОСУВАННЯ SPLUNK ДЛЯ МОНІТОРИНГУ ТА ВИЯВЛЕННЯ КІБЕРІНЦИДЕНТІВ У ДІЯЛЬНОСТІ SOC

У тезі розглянуто технологію застосування платформи Splunk для моніторингу подій інформаційної безпеки та виявлення кіберінцидентів у діяльності центру операційної безпеки (SOC). Проаналізовано основні етапи збирання, індексації та аналізу журналів подій, використання мови пошукових запитів SPL, кореляції подій і налаштування автоматизованих сповіщень. Визначено практичне значення Splunk для своєчасного виявлення підозрілої активності, підвищення ефективності роботи аналітиків SOC та вдосконалення процесів реагування на кіберзагрози.

Ключові слова: Splunk, SOC, кібербезпека, кіберінциденти, моніторинг, кореляція подій, виявлення загроз.

Технологія застосування Splunk у діяльності SOC передбачає послідовне виконання декількох взаємопов’язаних етапів: збирання даних, їх нормалізацію та індексацію, пошук закономірностей, виявлення підозрілих подій і формування сповіщень для аналітиків безпеки.

На першому етапі здійснюється збирання інформації з різноманітних джерел, зокрема журналів операційних систем, мережевого обладнання, міжмережевих екранів, серверів, систем автентифікації та засобів захисту кінцевих пристроїв. Отримані дані передаються до Splunk за допомогою відповідних механізмів збору, зокрема Splunk Universal Forwarder, або інших підтримуваних способів інтеграції. Централізоване накопичення інформації дає змогу аналізувати події з різних компонентів інфраструктури в межах єдиного середовища.

На другому етапі виконується індексація даних, що забезпечує можливість оперативного пошуку та дослідження подій. Для оброблення інформації використовується мова пошукових запитів Splunk Search Processing Language (SPL). Вона дає змогу фільтрувати журнали, групувати події за визначеними ознаками, обчислювати статистичні показники та встановлювати взаємозв’язки між подіями з різних джерел.

Наступним етапом є виявлення потенційних кіберінцидентів. За допомогою пошукових запитів і правил кореляції аналітики SOC можуть виявляти численні невдалі спроби автентифікації, нетипові входи до облікових записів, підозрілі мережеві з’єднання, аномальну активність користувачів та інші ознаки можливого компрометування інформаційної системи. Наприклад, групування невдалих спроб входу за IP-адресами та обліковими записами дає змогу виявляти ознаки атак методом перебору паролів (brute-force).

Для автоматизації моніторингу в Splunk можуть створюватися заплановані пошукові запити та правила сповіщення (alerts), які активуються за визначених умов. Результати аналізу можуть відображатися на інтерактивних панелях моніторингу (dashboards), що допомагають аналітикам оцінювати поточний стан інформаційної безпеки, відстежувати тенденції та визначати пріоритетність подій.

Важливим аспектом використання Splunk є кореляція подій із різних джерел. Окремий запис у журналі не завжди свідчить про наявність кіберінциденту, проте сукупність взаємопов'язаних подій може вказувати на підготовку або реалізацію атаки. Наприклад, поєднання численних невдалих спроб автентифікації, подальшого успішного входу та підозрілої активності після нього може бути ознакою компрометування облікового запису. Такий підхід дає змогу підвищити точність виявлення загроз і зменшити кількість хибнопозитивних спрацювань за умови належного налаштування правил.

У практичній діяльності SOC Splunk використовується для підтримки процесів аналізу кіберінцидентів, дослідження першопричин подій і підготовки даних для реагування. Інтеграція з іншими засобами кіберзахисту та платформами автоматизації реагування, зокрема рішеннями класу SOAR, дає змогу організувати передачу сповіщень, збагачення інформації про інциденти й автоматизацію окремих процедур реагування. Водночас ефективність платформи залежить від повноти джерел даних, якості налаштування правил виявлення, актуальності сценаріїв атак і кваліфікації аналітиків.

Важливим напрямом удосконалення використання Splunk у діяльності SOC є розроблення та постійне оновлення правил виявлення кіберзагроз з урахуванням змін у тактиках і методах діяльності зловмисників. Особливу увагу доцільно приділяти кореляції подій, аналізу нетипової поведінки користувачів і мережевих пристроїв, а також зменшенню кількості хибнопозитивних сповіщень.

Перспективи подальших досліджень полягають у розширенні можливостей Splunk шляхом інтеграції з іншими засобами кіберзахисту, використання автоматизованих механізмів реагування та впровадження інтелектуальних методів аналізу даних. Поєднання централізованого моніторингу, автоматизації оброблення сповіщень і вдосконалення аналітичних підходів сприятиме підвищенню рівня захищеності інформаційних систем.

Крім того, результативність застосування Splunk значною мірою залежить від правильного налаштування системи, повноти зібраних даних і професійної підготовки фахівців з кібербезпеки. Регулярний перегляд результатів моніторингу, перевірка ефективності правил виявлення та аналіз попередніх інцидентів дають змогу вдосконалювати процеси забезпечення інформаційної безпеки. Важливим є також визначення чітких процедур ескалації інцидентів і розподілу відповідальності між учасниками SOC. Реалізація цих заходів сприятиме формуванню комплексного підходу до управління кіберризиками, підвищенню стійкості інформаційної інфраструктури та забезпеченню безперервності функціонування організації в умовах зростання кількості кіберзагроз.

Перелік посилань:

About Splunk Enterprise Security // Splunk Documentation. URL: https://help.splunk.com/en/splunk-enterprise-security-8/user-guide/8.7/introduction/about-splunk-enterprise-security (дата звернення: 13.10.2026).

Фесенко Д.С.
Студент групи ІКБм-22, ННІКБЗІ, ДУІКТ
Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТЕХНІК ОБХОДУ ЗАСОБІВ ЗАХИСТУ КІНЦЕВИХ ПРИСТРОЇВ НА ОСНОВІ MICROSOFT DEFENDER XDR

У роботі розглянуто технологію виявлення технік обходу засобів захисту кінцевих пристроїв на основі платформи Microsoft Defender XDR. Проаналізовано можливості інтеграції Microsoft Defender for Endpoint, Microsoft Defender XDR, Microsoft Sentinel та механізмів розширеного пошуку загроз для виявлення підозрілих дій, спрямованих на вимкнення захисних механізмів, обхід антивірусного контролю, використання легітимних системних інструментів і приховування шкідливої активності. Визначено основні етапи технології: збір телеметрії, кореляція подій, поведінковий аналіз, виявлення індикаторів компрометації та автоматизоване реагування на інциденти. Обґрунтовано доцільність використання централізованого аналізу подій для підвищення ефективності виявлення складних атак і скорочення часу реагування на кіберзагрози.

Ключові слова: кібербезпека, Microsoft Defender XDR, захист кінцевих пристроїв, виявлення загроз, обхід засобів захисту, EDR, XDR, поведінковий аналіз, реагування на інциденти.

Сучасні інформаційні системи функціонують в умовах постійного зростання кількості кіберзагроз, спрямованих на компрометацію робочих станцій, серверів і корпоративних інформаційних ресурсів. Однією з актуальних проблем кібербезпеки є використання зловмисниками технік обходу засобів захисту кінцевих пристроїв (Defense Evasion), які дозволяють зменшити ймовірність виявлення шкідливої активності традиційними антивірусними системами. Такі техніки можуть передбачати зміну поведінки процесів, використання легітимних системних утиліт, маніпуляції з налаштуваннями безпеки, приховування артефактів виконання та спроби обмеження реєстрації подій. Унаслідок цього виникає потреба в застосуванні комплексних технологій, здатних аналізувати не лише окремі файли або сигнатури, а й послідовності подій у контексті роботи всієї інформаційної системи.

Перспективним підходом до розв’язання цієї проблеми є використання платформи Microsoft Defender XDR, яка забезпечує інтеграцію засобів захисту кінцевих пристроїв, ідентичностей, електронної пошти, хмарних застосунків та інших компонентів корпоративного середовища. Платформа об’єднує сигнали безпеки з різних джерел, установлює взаємозв’язки між подіями та формує узгоджене представлення інцидентів. Це дає змогу виявляти атаки, які окремо можуть виглядати як звичайна адміністративна активність, але в сукупності свідчать про спробу компрометації системи [1].

Основним компонентом технології виявлення технік обходу захисту є Microsoft Defender for Endpoint -рішення класу Endpoint Detection and Response (EDR). Воно забезпечує збір телеметрії кінцевих пристроїв, аналіз активності процесів, контроль підозрілих дій і формування сповіщень про потенційні загрози. Зокрема, аналізуються відомості про створення та завершення процесів,

запуск команд, файлові операції, мережеві з'єднання й зміни конфігурації системи. На основі цих даних можна виявляти ознаки спроб вимкнення захисних механізмів, використання нетипових способів виконання програм і маскування шкідливих процесів під легітимну активність [2].

Важливе значення має зіставлення виявлених подій із таксономією MITRE ATT&CK, яка систематизує тактики та техніки, що застосовуються під час кіберзлочинних операцій. Для досліджуваної проблеми особливо актуальною є тактика Defense Evasion, що охоплює методи обходу механізмів виявлення та контролю. Застосування цієї таксономії дає змогу класифікувати спостережувану активність, формувати правила виявлення та оцінювати повноту покриття загроз засобами захисту. Водночас відповідність окремої події певній техніці MITRE ATT&CK не є самостійним доказом атаки, тому результати аналізу необхідно оцінювати з урахуванням контексту, поведінки користувача та інших подій [3].

Технологія виявлення технік обходу засобів захисту на основі Microsoft Defender XDR передбачає послідовне виконання кількох взаємопов'язаних етапів. На першому етапі здійснюється збір телеметрії з кінцевих пристроїв та інших підключених джерел. На другому етапі виконується первинний аналіз подій, спрямований на виявлення відхилень від типової поведінки системи. На третьому етапі відбувається кореляція сигналів, що дає змогу об'єднати пов'язані дії в єдиний ланцюг атаки. На четвертому етапі проводиться перевірка виявлених ознак із використанням правил виявлення, аналітичних запитів і контексту інциденту. Завершальний етап передбачає оцінювання ризику та, за наявності відповідних підстав і налаштувань, автоматизоване або ручне реагування.

Одним із ключових механізмів реалізації цієї технології є розширений пошук загроз (Advanced Hunting), доступний у Microsoft Defender XDR. Він дає змогу виконувати запити до телеметрії безпеки за допомогою мови Kusto Query Language (KQL). Використовуючи таблиці подій, зокрема DeviceProcessEvents, DeviceFileEvents, DeviceRegistryEvents і DeviceNetworkEvents, фахівці можуть досліджувати запуск процесів, файлову активність, зміни реєстру та мережеві взаємодії. Наприклад, поєднання інформації про незвичний запуск системної утиліти, зміну параметрів безпеки та подальше мережеве з'єднання може свідчити про підозрілий сценарій, що потребує додаткового розслідування. Ефективність такого підходу залежить від доступності відповідної телеметрії, налаштувань аудиту, ліцензування та конфігурації середовища [1; 4].

Окремим напрямом є виявлення спроб втручання в роботу самих засобів захисту. До таких дій можуть належати несанкціоновані зміни параметрів безпеки, спроби вимкнення компонентів захисту, втручання в процеси моніторингу або використання облікових записів із надмірними привілеями. Для виявлення цих загроз доцільно поєднувати сигнали Microsoft Defender for Endpoint із журналами аудиту, інформацією про ідентичності та контекстом адміністративної активності. Такий підхід дає змогу відрізняти легітимні дії

системних адміністраторів від підозрілих операцій, виконаних унаслідок компрометації облікового запису або запуску шкідливого процесу.

Платформа Microsoft Defender XDR також забезпечує кореляцію подій між різними доменами безпеки. Наприклад, підозріла активність на кінцевому пристрої може бути пов’язана з аномальним входом користувача, фішинговим повідомленням або незвичною поведінкою в хмарному середовищі. Завдяки цьому аналіз здійснюється не ізольовано в межах одного пристрою, а з урахуванням ширшого контексту інциденту. Такий механізм підвищує ймовірність виявлення багатоступеневих атак, у яких зловмисник послідовно використовує кілька векторів проникнення та приховування активності [1].

Для зменшення часу реагування доцільно застосовувати автоматизовані сценарії реагування, що підтримуються Microsoft Defender XDR. Залежно від конфігурації середовища вони можуть передбачати ізоляцію скомпрометованого пристрою, блокування виявлених загроз, запуск розслідування, збирання додаткових даних або передачу інциденту відповідальному фахівцю. Водночас автоматизація має ґрунтуватися на визначених критеріях ризику та правилах управління інцидентами. Надмірно агресивні сценарії можуть призводити до хибнопозитивних спрацювань і переривання легітимних бізнес-процесів, тому перед широким упровадженням необхідно перевіряти їх у контрольованому середовищі.

Таким чином, технологія виявлення технік обходу засобів захисту кінцевих пристроїв на основі Microsoft Defender XDR ґрунтується на комплексному використанні телеметрії, поведінкового аналізу, кореляції подій, розширеного пошуку загроз і автоматизованого реагування. Її основною перевагою є можливість виявляти підозрілі послідовності дій, які можуть залишатися непомітними під час ізольованого аналізу окремих подій. Поєднання EDR-можливостей Microsoft Defender for Endpoint, аналітичних інструментів Microsoft Defender XDR і методології MITRE ATT&CK створює основу для систематизованого виявлення технік обходу захисту та підвищення стійкості корпоративної інфраструктури до кіберзагроз. Подальший розвиток цього підходу передбачає вдосконалення правил виявлення, адаптацію аналітики до особливостей організації, перевірку ефективності на контрольованих сценаріях і постійне оновлення механізмів реагування відповідно до актуальних моделей атак.

Перелік посилань

1. Microsoft. *Microsoft Defender XDR documentation*. Microsoft Learn. URL: https://learn.microsoft.com/en-us/defender-xdr/ (дата звернення: 25.10.2026).
2. Microsoft. *Microsoft Defender for Endpoint documentation*. Microsoft Learn. URL: https://learn.microsoft.com/en-us/defender-endpoint/ (дата звернення: 25.10.2026).
3. MITRE. *Enterprise ATT&CK: Defense Evasion*. MITRE ATT&CK. URL: https://attack.mitre.org/tactics/TA0005/ (дата звернення: 25.10.2026).
4. Microsoft. *Advanced hunting overview*. Microsoft Learn. URL: https://learn.microsoft.com/en-us/defender-xdr/advanced-hunting-overview (дата звернення: 25.10.2026).

Піскунов Костянтин Валерійович
студент групи ІКБм-61, ННІЗІ ДУІКТ,
Київ, Україна

ТЕХНОЛОГІЯ ЗАПОБІГАННЯ ФІШИНГОВИМ АТАКАМ, СПРЯМОВАНИМ НА ВИКРАДЕННЯ ОБЛІКОВИХ ДАНИХ КОРИСТУВАЧІВ

Запропоновано технологію захисту корпоративних облікових даних, яка поєднує фільтрацію повідомлень, контроль вебресурсів, фішингостійку автентифікацію та реагування на компрометацію. Визначено порядок застосування захисних механізмів, обмеження підходу й показники оцінювання його ефективності. Очікуваний результат полягає у зниженні ризику компрометації облікових записів завдяки перевірці повідомлень і сторінок до введення даних та прив'язці автентифікації до легітимного сервісу.

Ключові слова: фішинг, облікові дані, багатофакторна автентифікація, FIDO2, WebAuthn, контроль доступу.

Фішингові повідомлення спонукають користувачів переходити на підроблені сторінки та розкривати паролі й одноразові коди. У дослідженні D. Lain, K. Kostiainen та S. Čapkun за участю понад 14 тисяч працівників упродовж 15 місяців показано, що поширений формат навчання під час симуляцій не гарантує підвищення стійкості до фішингу [1]. Додатковою проблемою є передавання викраденого коду справжньому сервісу в реальному часі. За NIST SP 800-63B-4, автентифікація з ручним введенням одноразового коду не вважається фішингостійкою [2]. Це обґрунтовує необхідність технічного захисту, який зменшує залежність від уважності користувача.

Метою роботи є обґрунтування технології запобігання викраденню та зловмисному використанню облікових даних шляхом узгодженого застосування засобів виявлення фішингу, контролю автентифікації та припинення підозрілих сеансів. Технологію запропоновано для організації з керованими робочими пристроями та централізованим постачальником ідентичності. Модель загроз охоплює підроблені сторінки входу, посилання в листах і QR-кодах, ретрансляцію одноразових кодів та використання викрадених сеансових токенів.

Перший етап захисту передбачає перевірку повідомлень. Поштовий шлюз аналізує відправника, вкладення й посилання, зокрема адреси, закодовані в QR-кодах. Перевірки SPF, DKIM і DMARC застосовують для протидії підміні домену відправника. DMARC узгоджує домен у полі From з доменом, підтвердженим SPF або DKIM [3]. Водночас успішна перевірка не підтверджує безпечність змісту листа: повідомлення може надійти зі скомпрометованого облікового запису або схожого домену.

На наступному етапі здійснюється контроль вебресурсу. У керованому браузері перевіряють повне доменне ім'я, перенаправлення, репутацію адреси та відповідність форми входу переліку погоджених доменів. Підтверджені

фішингові ресурси блокують; невідомі сторінки з ознаками імітації корпоративного входу передають на додаткову перевірку. Наявність HTTPS сама по собі не є підставою дозволяти введення даних. Контроль має спрацьовувати до їх введення, оскільки шкідливий сценарій може зчитувати символи ще до натискання кнопки надсилання.

Для корпоративного входу застосовують FIDO2/WebAuthn з локальною перевіркою користувача за PIN-кодом або біометричною ознакою. Автентифікатор створює криптографічну відповідь, прив'язану до домену легітимного сервісу, без передавання секретного ключа вебсайту. Сервер перевіряє одноразовий виклик, походження запиту, ідентифікатор сервісу та цифровий підпис [2]. Для відновлення доступу передбачають резервний автентифікатор та перевірену процедуру, яка запобігає обходу захисту через слабший спосіб входу.

Взаємодію компонентів запропонованої технології наведено на рис. 1. Перевірки повідомлень і вебресурсів доповнюються фішингостійкою автентифікацією, а їхні події використовуються для виявлення інцидентів та оновлення індикаторів загроз.

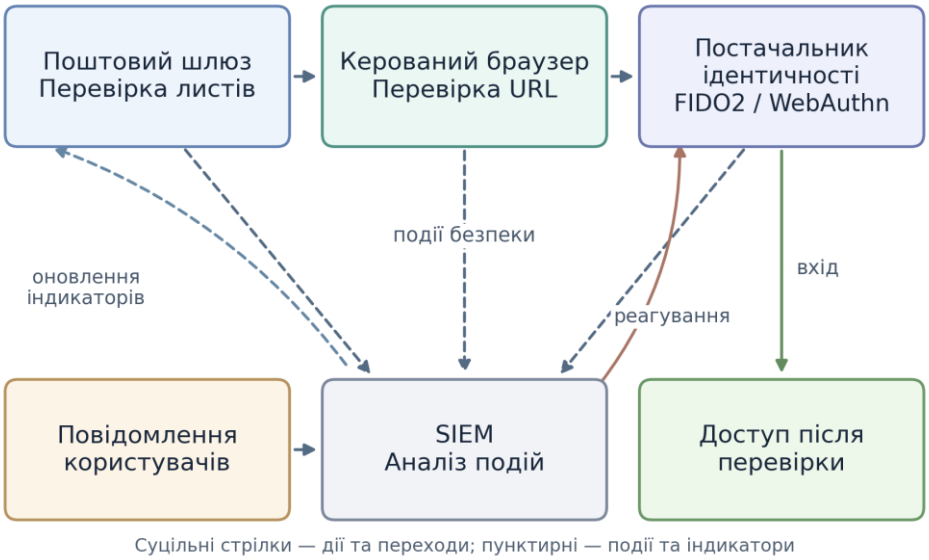


Рис. 1. Взаємодія компонентів технології запобігання фішингу

Події поштового шлюзу, браузера й постачальника ідентичності передають до системи керування інформацією та подіями безпеки (SIEM). За поєднання переходу на підтверджений фішинговий ресурс і підозрілого входу запускають перевірку інциденту. У разі підтвердження компрометації припиняють сеанси постачальника ідентичності та окремих застосунків, відкликають підтримувані системою токени й відновлюють захищений доступ [2]. Дії виконують за узгодженими процедурами реагування.

У поштовому клієнті передбачають кнопку повідомлення про підозрілий лист. Після перевірки такого повідомлення нові індикатори передають засобам блокування. Дослідження [1] підтверджує практичну придатність залучення працівників до колективного виявлення фішингу. Навчання зосереджують на розпізнаванні підозрілих запитів і швидкому повідомленні про них.

Правило ухвалення рішення розділяє підтверджені загрози та евристичні ознаки. Збіг із перевіреним індикатором зумовлює блокування; лише підозріла структура адреси потребує додаткової перевірки. Сам факт нового пристрою не повинен автоматично вважатися компрометацією. Особливість запропонованого порядку полягає в поєднанні блокування фішингової сторінки до введення даних із перевіркою права на вхід незалежно від того, чи була сторінка вчасно розпізнана.

Для оцінювання ефективності технології передбачено порівняння базової фільтрації, фільтрації з одноразовими кодами та запропонованого комплексу на однакових сценаріях і тестових облікових записах. Окремо враховують частку виявлених фішингових ресурсів серед усіх фішингових ресурсів, частку помилково заблокованих легітимних ресурсів, затримку перевірки та частку успішних захоплень облікового запису. Додатково вимірюють час від підтвердження інциденту до припинення доступу. Числові значення цих показників мають бути отримані під час випробувань.

Впровадження технології пропонується починати з пілотної групи користувачів. Для неї реєструють основний і резервний автентифікатори, визначають дозволені домени входу та перевіряють надходження подій до SIEM. На тестових облікових записах відтворюють перехід за QR-кодом, імітацію сторінки входу, ретрансляцію коду та втрату автентифікатора. Для кожного сценарію фіксують етап зупинення атаки, спрацювання контролів і доступність легітимних сервісів. За результатами пілотування уточнюють правила блокування та порядок відновлення доступу. Перед розгортанням у всій організації перевіряють сумісність сервісів, визначають відповідальних за обробку повідомлень користувачів і документують порядок припинення скомпрометованих сеансів.

Межі застосування визначаються станом кінцевого пристрою та повнотою інтеграції. Фішингостійка автентифікація не усуває загрозу викрадення вже створеного сеансу шкідливим програмним забезпеченням. Тому її доповнюють захистом пристроїв, моніторингом сеансів і контрольованим відновленням доступу.

Отже, запропонована технологія охоплює доставлення фішингового повідомлення, відкриття сторінки, автентифікацію та подальше використання сеансу. Її практична цінність полягає в узгодженні захисних дій на цих етапах. Очікуване зменшення кількості компрометацій потребує підтвердження випробуваннями з одночасним контролем помилкових блокувань.

Перелік посилань:

1. Phishing in Organizations: Findings from a Large-Scale and Long-Term Study [Електронний ресурс] / D. Lain, K. Kostiainen, S. Čapkun // 2022 IEEE Symposium on Security and Privacy (SP). 2022. — Режим доступу: https://doi.org/10.1109/SP46214.2022.9833766 (дата звернення: 23.10.2026).
2. Digital Identity Guidelines: Authentication and Authenticator Management. SP 800-63B-4 [Електронний ресурс] // National Institute of Standards and Technology. 2025. — Режим доступу: https://doi.org/10.6028/NIST.SP.800-63B-4 (дата звернення: 23.10.2026).
3. Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489 [Електронний ресурс] / M. Kucherawy, E. Zwicky. 2015. — Режим доступу: https://doi.org/10.17487/RFC7489 (дата звернення: 23.10.2026).

Гриньов Тимофій Олександрович
студент групи ІКБм-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ТА БЛОКУВАННЯ РОЗПОДІЛЕНИХ МЕРЕЖЕВИХ АТАК У КОРПОРАТИВНІЙ МЕРЕЖІ ОРГАНІЗАЦІЇ НА БАЗІ CROWDSEC

Розглянуто технологію виявлення та автоматичного блокування розподілених мережевих атак у корпоративній мережі організації на основі open-source системи CrowdSec. Описано архітектуру рішення, принципи роботи сценаріїв виявлення, використання профілів рішень та спільнотного списку репутації IP-адрес. Запропоновано схему розгортання на периметрі мережі та показники оцінювання ефективності. Визначено переваги й обмеження підходу порівняно з Fail2Ban.

Корпоративні мережі постійно зазнають розподілених атак: DDoS [1], перебору паролів, сканування портів, експлуатації вразливостей. Через множину джерел засоби зі статичними порогами для окремих IP-адрес спрацьовують із запізненням, а зловмисник може розподілити запити так, щоб жодне джерело не перевищило поріг. Системи виявлення та запобігання вторгненням (IDPS) [2] фіксують підозрілу активність і намагаються її зупинити, проте для розподілених атак потрібна ще й колективна розвідка про загрози.

CrowdSec — система з відкритим кодом (мова Go), що аналізує журнали сервісів і HTTP-запити та блокує зловмисників [3]. Її архітектура (рис. 1): Security Engine зчитує журнали (Nginx, SSH, міжмережевий екран тощо), розбирає їх парсерами та перевіряє сценаріями виявлення; локальний API (LAPI) приймає сповіщення, на основі профілів ухвалює рішення (ban, captcha) і передає їх компонентам виконання (remediation components, bouncers), а також обмінюється даними з центральним API CrowdSec; компоненти виконання блокують джерела на міжмережевому екрані (iptables), reverse-proxy або вебсервері.

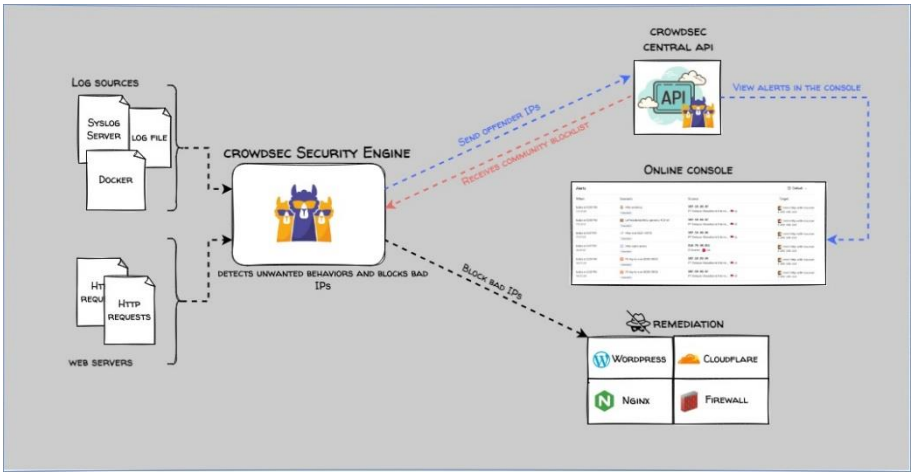


Рис. 1. Архітектура рішення на базі CrowdSec [3]

Кожен сценарій виявлення описується у файлі YAML і працює за принципом «дірявого відра» (leaky bucket): події певного типу надходять у відро окремо для кожного джерела (зазвичай за IP-адресою), а його місткість і швидкість «витікання» визначають, чи станеться переповнення. Після переповнення формується сповіщення й ухвалюється рішення про блокування. Додатково Security Engine завантажує спільнотний список перевірених зловмисників (community blocklist), що дозволяє блокувати їх ще до першої спроби атаки на організацію [3].

Тип і тривалість рішення задаються профілями (profiles.yaml): умова, дія (ban, captcha) і строк блокування. Хибні спрацювання зменшують білі списки (whitelists) довірених адрес — адміністраторів, систем моніторингу, партнерських підмереж; блокування тимчасові, тому помилково заблокована адреса згодом розблоковується [3].

У разі розподіленого перебору паролів кожне джерело робить лише кілька спроб, тому локальні пороги не спрацьовують. Спільнотний список частково вирішує проблему: адреси ботнету, виявлені іншими учасниками, блокуються раніше, ніж атака досягне організації. Типові сценарії групують події за адресою джерела, тому повільну атаку з багатьох адрес можна виявити лише за спільнотним списком або сценаріями з іншою ознакою групування, наприклад іменем облікового запису [3].

Запропоновано таку схему розгортання: агенти CrowdSec встановлюються на серверах із публічними сервісами (вебсервер, SSH, поштовий сервіс), централізований LAPI приймає сповіщення від усіх агентів, а компоненти виконання на периметрі (міжмережевий екран і reverse-proxy) отримують від нього рішення. Завдяки єдиному LAPI джерело, виявлене на одному сервері, блокується на всьому периметрі.

Ефективність планується оцінити на тестовому стенді у віртуальному середовищі (вебсервер, SSH, пошта, периметральний вузол, вузол атакувальника). Атаки імітуватимуться інструментами Hydra (перебір паролів SSH), Nmap (сканування портів) і wrk (HTTP-флуд). Показники: час від початку атаки до блокування, частка заблокованих шкідливих запитів, кількість хибних спрацювань, навантаження на систему.

На відміну від Fail2Ban [4], що змінює правила міжмережевого екрана зазвичай лише на окремому хості, CrowdSec підтримує розподілену архітектуру та обмін репутацією IP-адрес. Працюючи з журналами та HTTP-запитами, він підходить для виявлення атак рівня застосунків (L7) і розподіленого перебору, але не замінює захист від об'ємних L3/L4 DDoS-атак на боці провайдера чи CDN.

Отже, CrowdSec дозволяє автоматизувати виявлення та блокування розподілених атак у корпоративній мережі без витрат на ліцензування; подальші дослідження — експериментальна перевірка схеми, власні сценарії виявлення та інтеграція зі SIEM-системою.

Перелік посилань:

1. Mirkovic J., Reiher P. A taxonomy of DDoS attack and DDoS defense mechanisms // ACM SIGCOMM Computer Communication Review. – 2004. – Vol. 34, No. 2. – P. 39–53. DOI: 10.1145/997150.997156.
2. Scarfone K., Mell P. Guide to Intrusion Detection and Prevention Systems (IDPS): NIST SP 800-94. – Gaithersburg: NIST, 2007. URL: https://csrc.nist.gov/pubs/sp/800/94/final (дата звернення: 15.10.2026).
3. CrowdSec Documentation. URL: https://docs.crowdsec.net (дата звернення: 15.10.2026).
4. Fail2Ban. URL: https://github.com/fail2ban/fail2ban (дата звернення: 15.10.2026).

Шавшин Є.О
Студент групи ІКБмз-11, ННІКБЗІ ДУІКТ, Київ, Україна

КОМПЛЕКСНИЙ ПІДХІД ДЛЯ ПІДВИЩЕННЯ РІВНЯ ЗАХИЩЕНОСТІ ІНФОРМАЦІЙНО-ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ

У роботі досліджено проблему трансформації сучасних кіберзагроз в умовах розмиття класичного мережевого периметра та стрімкої автоматизації векторів атак. Проаналізовано обмеження традиційних сигнатурних методів захисту перед цільовими APT-атаками (Advanced Persistent Threats) та техніками прихованої присутності в системі. Запропоновано комплексний підхід до підвищення рівня захищеності інформаційно-телекомунікаційних систем, що базується на впровадженні архітектури нульової довіри (Zero Trust), поведінковому аналізі за матрицею MITRE ATT&CK та використанні відкритих систем класу SIEM/XDR для автоматизації реагування на інциденти на кінцевих точках.

Ключові слова: кібербезпека, захист інформації, APT-загрози, Zero Trust, поведінковий аналіз, SIEM, XDR, захист кінцевих точок, MITRE ATT&CK.

Сучасний етап розвитку інформаційних технологій характеризується фундаментальним парадоксом: попри постійне вдосконалення криптографічних протоколів, міжнародних стандартів безпеки та зростання витрат на кіберзахист, кількість успішних кібератак на об’єкти критичної інфраструктури, державні реєстри та корпоративний сектор продовжує зростати [4, с. 18]. Причина цього явища полягає у зміні самої парадигми протистояння в кіберпросторі. Класичні моделі захисту, які ще кілька років тому вважалися еталонними, сьогодні втрачають ефективність через зміну архітектури корпоративних мереж та еволюцію тактик зловмисників.

Протягом тривалого часу корпоративна безпека будувалася за захистом зовнішнього контуру: існувала чітка межа локальної мережі, захищена міжмережевими екранами (Firewall) та системами виявлення вторгнень (IDS/IPS). Проте масовий перехід до хмарних обчислень, гібридних форматів роботи та використання мобільних пристроїв фактично змінив поняття фізичного периметра [1, с. 4].

Паралельно з цим кардинально змінилася тактика проведення цільових довготривалих атак - APT [4, с. 112]. Сучасні хакерські угруповання мінімізують використання класичного шкідливого програмного забезпечення, яке легко ідентифікується антивірусними рішеннями за відомими сигнатурами чи хеш-сумами файлів. Натомість домінуючою стала концепція Living off the Land (LotL) - використання зловмисниками легітимних системних утиліт та вбудованих засобів адміністрування (таких як PowerShell, Windows Management Instrumentation, системні планувальники завдань чи штатні утиліти віддаленого керування) [3, с. 8].

З точки зору традиційного сигнатурного антивіруса, виконання команди у PowerShell є штатною роботою операційної системи. Однак у контексті атаки

саме через ці легітимні інструменти відбувається закріплення в системі, підвищення привілеїв та горизонтальне переміщення мережею [3, с. 14]. Таким чином, захист, що спирається виключно на статичні індикатори компрометації, виявляється неефективним перед сучасними безфайловими (fileless) атаками [3, с. 5].

Другим вагомим викликом є індустріалізація кіберзлочинності та інтеграція генеративного штучного інтелекту в інструментарій нападників. Хоча алгоритми ШІ не створили принципово нових класів вразливостей у мережевих протоколах, вони критично знизили поріг входу в кіберзлочинність та масштабували відомі вектори атак.

Зокрема, автоматизація дозволяє генерувати поліморфний шкідливий код, який динамічно змінює свою структуру при кожному розгортанні, непомітно для стандартного антивірусу [4, с. 125]. Крім того, поєднання автоматизованого OSINT-аналізу (розвідки на основі відкритих джерел) із великими мовними моделями вивело соціальну інженерію та цільовий фішинг (spear-phishing) на безпрецедентний рівень переконливості [4, с. 164].

Третій виклик лежить в операційній площині функціонування центрів моніторингу безпеки (SOC). В галузі зберігається фундаментальна асиметрія: зловмиснику для успішного проникнення достатньо знайти одну вразливість або помилку конфігурації, тоді як фахівці з кібербезпеки змушені безперервно контролювати всю поверхню атаки [2, с. 15].

Збільшення кількості підключених вузлів та інформаційних систем призводить до генерації величезних масивів реєстраційних даних (логів). При неоптимальному налаштуванні правил кореляції виникає Alert Fatigue (виснаження від сповіщень), коли аналітики безпеки отримують тисячі сповіщень щодня, переважна більшість із яких є хибними спрацьовуваннями (false positives) [2, с. 29].

Водночас занадто висока ціна платних рішень систем класу SIEM (Security Information and Event Management) та EDR/XDR (Extended Detection and Response) від світових лідерів ринку часто залишає бюджетні установи та невеликі підприємства без належного інструментарію моніторингу, формуючи хибне уявлення про недоступність побудови власного центру реагування на інциденти.

Аналіз окреслених загроз свідчить про те, що екстенсивне нарощування кількості ізольованих засобів захисту є неефективним. Необхідною є якісна зміна підходів до побудови систем кібербезпеки за чотирма ключовими напрямами:

1. Перехід від сигнатурного детектування до поведінкового аналізу. Сучасний захист кінцевих точок і телекомунікаційних мереж має базуватися на безперервному зборі телеметрії, моніторингу цілісності системних файлів та аналізі поведінкових патернів. Замість пошуку відомих хеш-сум файлів, система безпеки має виявляти аномалії у взаємодії процесів (наприклад, нетипове

породження командного рядка процесом текстового редактора або аномальну мережеву активність у неробочий час).

2. Впровадження концепції Zero Trust (Нульова довіра) та принципу Assume Breach (Припущення компрометації). Згідно зі стандартом NIST SP 800-207, архітектура безпеки повинна проєктуватися з припущенням, що мережевий периметр вже скомпрометовано і зловмисник потенційно присутній у внутрішньому сегменті [1, с. 5]. Це вимагає відмови від поняття «довіреної внутрішньої мережі», впровадження жорсткої мікросегментації, обов'язкової багатофакторної автентифікації (MFA) при зверненні до будь-якого ресурсу та суворого дотримання принципу найменших привілеїв (Least Privilege) [1, с. 9-11]. За таких умов навіть успішна компрометація окремої робочої станції не дозволить атакувальнику здійснити горизонтальне переміщення до критичних серверів або резервних копій [1, с. 17].

3. Використання рішень із відкритим вихідним кодом (Open Source) для побудови SIEM/XDR. Обмеженість фінансових ресурсів не повинна бути перешкодою для впровадження ешелонованого захисту. Сучасний рівень розвитку платформ із відкритим кодом, зокрема систем класу Wazuh, дозволяє реалізувати повноцінний функціонал централізованого збору та кореляції логів (SIEM) у поєднанні з розширеним захистом кінцевих точок (XDR). Інтеграція подібних рішень із зовнішніми базами аналітики загроз (Threat Intelligence) та правилами виявлення загроз за методологією MITRE [3, с. 24] дозволяє ефективно ідентифікувати спроби закріплення в системі, експлуатації вразливостей та порушення політик конфігурації без витрат на дорогі комерційні ліцензії.

4. Автоматизація активного реагування (Active Response) та розвиток культури кібербезпеки. Для подолання проблеми Alert Fatigue та скорочення часу реакції на інцидент (MTTR - Mean Time to Respond) типові вектори атак мають блокуватися в автоматичному режимі [2, с. 36]. Налаштування сценаріїв активного реагування безпосередньо на кінцевих точках (наприклад, автоматичне блокування IP-адреси на рівні хостового брандмауера при виявленні перебору паролів або ізоляція вузла при спрацюванні критичного правила кореляції) дозволяє зупинити атаку в перші секунди та вивільнити ресурс аналітиків для проактивного пошуку загроз (Threat Hunting) [2, с. 38].

Підсумовуючи проведене дослідження, слід зазначити, що сучасна кібербезпека остаточно перестала бути статичним станом інфраструктури, перетворившись на безперервний адаптивний процес. Виклик сучасних загроз - від прихованих APT-кампаній до автоматизованих атак із залученням штучного інтелекту - вимагає відмови від застарілих периметральних моделей захисту. Поєднання архітектурних принципів Zero Trust [1, с. 5], глибокого поведінкового аналізу за методологією MITRE ATT&CK [3, с. 19] та раціонального впровадження відкритих платформ моніторингу й активного реагування класу

SIEM/XDR дозволяє побудувати стійку систему захисту інформації, здатну діяти на випередження навіть в умовах обмежених ресурсів та високої інтенсивності кіберзагроз.

Список використаних джерел

1. Rose S., Borchert O., Mitchell S., Connelly S. Zero Trust Architecture : NIST Special Publication 800-207. Gaithersburg, MD : National Institute of Standards and Technology, 2020. 59 p. DOI: https://doi.org/10.6028/NIST.SP.800-207

2. Cichonski P., Millar T., Grance T., Scarfone K. Computer Security Incident Handling Guide : NIST Special Publication 800-61 Rev. 2. Gaithersburg, MD : National Institute of Standards and Technology, 2012. 79 p. DOI: https://doi.org/10.6028/NIST.SP.800-61r2

3. Strom B. E., Applebaum A., Miller D. P., Nickels K. C., Pennington A. G., Thomas C. B. MITRE ATT&CK: Design and Philosophy. McLean, VA : The MITRE Corporation, 2020. 37 p. URL: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf

4. Бурячок В. Л., Толубко В. Б., Хорошко В. О., Толюпа С. В. Інформаційна та кібербезпека: соціотехнічний аспект : підручник / за заг. ред. д-ра техн. наук, проф. В. Б. Толубка. Київ : ДУТ, 2015. 288 с.

Твердохліб Я.М.
студентка групи ІКБм-62, ННІКБЗІ ДУІКТ,
Київ, Україна

ЗАСТОСУВАННЯ СТАТИЧНОГО АНАЛІЗУ КОДУ ДЛЯ ПІДВИЩЕННЯ БЕЗПЕКИ ВЕБ-ЗАСТОСУНКІВ ОРГАНІЗАЦІЇ

Безпека веб-застосунків є важливою складовою захисту інформаційних ресурсів організації, оскільки вразливості програмного коду можуть призводити до несанкціонованого доступу та витоку даних. Одним зі способів запобігання таким загрозам є статичний аналіз коду (SAST), який дозволяє виявляти потенційні вразливості ще на етапі розробки. Використання платформи SonarQube та її інтеграція в процеси CI/CD забезпечують регулярну перевірку вихідного коду, своєчасне виявлення помилок і контроль дотримання вимог безпеки. Такий підхід допомагає зменшити ризики кібератак, підвищити якість програмного забезпечення та вдосконалити процес розроблення веб-застосунків.

Ключові слова: статичний аналіз коду, SAST, кібербезпека, безпека веб-застосунків, SonarQube, CI/CD, Quality Gates.

Веб-застосунки залишаються одним із головних каналів взаємодії організації з клієнтами та партнерами, а отже й однією з найпривабливіших цілей для кіберзлочинців. Значна частина інцидентів не завжди в серверах чи мережі: нерідко слабке місце сидить у вихідному коді. Чим пізніше вразливість виявлено, тим дорожчим буде її усунення, тому сучасні організації переносять перевірки безпеки на ранні етапи життєвого циклу розробки за підходом «shift left». Його суть полягає в тому, щоб виявляти і усунути потенційні проблеми безпеки як можна раніше – ще під час написання та перевірки коду, а не після завершення розробки або впровадження застосунку.

Одним із засобів реалізації цього підходу є статичний аналіз коду (SAST, Static Application Security Testing) – це перевірка вихідного коду без його виконання. Інструмент будує модель програми та зіставляє її з набором правил, що описують типові помилки й небезпечні конструкції. Перевагами такого підходу є виявлення проблем ще на етапі написання коду, аналіз усієї кодової бази, зокрема гілок, які не завжди виконуються під час ручного тестування, автоматичність і повторюваність перевірок, а також конкретні рекомендації щодо виправлення для розробника. Це дозволяє організаціям усувати вразливості до їх розгортання у виробництві, зменшуючи ризик інцидентів безпеки та мінімізуючи витрати на усунення [1, с.121]. Практики безпечного розроблення, зокрема перевірку та аналіз коду, передбачає і документ NIST SP 800-218 (SSDF) [2]. Водночас статичний аналіз має певні обмеження: інструменти можуть повідомляти про проблеми, яких насправді немає в коді, а також не виявляти вразливості, що проявляються лише під час виконання застосунку. Тому SAST

доцільно поєднувати з динамічним тестуванням безпеки (DAST) та аналізом залежностей.

Для практичного впровадження статичного аналізу організація може використовувати платформу SonarQube, яка допомагає систематично перевіряти код і контролювати його якість. Залежно від версії платформи та мови програмування результати аналізу можуть охоплювати помилки, вразливості, проблеми підтримуваності коду й Security Hotspots – ділянки, які потребують додаткової перевірки з погляду безпеки [3]. Такий розподіл дає змогу визначати пріоритети під час опрацювання результатів і зосереджувати увагу розробників на проблемах, які можуть найбільше впливати на надійність і захищеність застосунку. Водночас виявлені попередження необхідно аналізувати з урахуванням контексту, оскільки автоматизована перевірка не замінює експертної оцінки.

Для забезпечення систематичного контролю якості та безпеки коду SonarQube використовує Quality Gates – визначені критерії, яким має відповідати проєкт. Якщо результати аналізу не задовольняють встановлені вимоги, наприклад через виявлення критичних проблем у новому коді, перехід до наступного етапу розробки може бути заблокований. Найбільшої ефективності цей механізм досягає за умови інтеграції автоматизованого аналізу в процес розробки: засоби на кшталт Jenkins, GitLab CI та GitHub Actions запускають перевірку під час внесення змін або подання їх на розгляд. Розробники отримують результати перевірки на ранньому етапі, а фахівці з безпеки можуть відстежувати зміни показників якості коду в єдиному інтерфейсі. Для організації це забезпечує єдині вимоги до всіх веб-застосунків і спрощує контроль результатів перевірок. У такий спосіб безпека стає невіддільною частиною щоденного процесу розробки, а не окремим етапом наприкінці роботи.

Отже, застосування статичного аналізу коду є критично важливим елементом захисту веб-застосунків, що дозволяє перенести акцент із ліквідації наслідків інцидентів на їх запобігання на етапі створення програмного забезпечення. Системна інтеграція платформи SonarQube у конвеєри CI/CD із застосуванням механізмів Quality Gates забезпечує безперервний контроль захищеності вихідного коду, мінімізує вплив людського фактору та запобігає появі критичних вразливостей у продуктивних середовищах. Регулярне проведення таких перевірок також дозволяє своєчасно виявляти недоліки в нових версіях програмного забезпечення та підтримувати належний рівень його захищеності протягом усього життєвого циклу. У результаті організація отримує ефективний інструмент, оптимізації витрат на безпеку, зниження репутаційних ризиків та побудови культури безпечної розробки, яка повністю відповідає сучасним стандартам захисту інформаційних систем.

1. Sehgal V. V. Front cover image for Implementing DevSecOps practices : understand application security testing and secure coding by integrating SAST and DAST Implementing DevSecOps practices : understand application security testing and secure coding by integrating SAST and DAST. Birmingham : Packt Publishing Ltd., 2023.
2. NIST Special Publication 800-218. Secure Software Development Framework (SSDF) Version 1.1: Recommendations for Mitigating the Risk of Software Vulnerabilities. National Institute of Standards and Technology, 2022.
3. SonarSource. SonarQube Documentation: Security-related rules. URL: https://docs.sonarsource.com/sonarqube-server/quality-standards-administration/managing-rules/security-related-rules

Мурзін І. А., студент групи ІКБм-51
Науковий керівник: Гайдур Г. І., д.т.н., професор
Державний університет інформаційно-комунікаційних технологій

ВИЯВЛЕННЯ ПОВТОРНИХ АТАК І ДОВГОТРИВАЛОГО ДОСТУПУ ЗЛОВМИСНИКІВ ЗАСОБАМИ XDR

Анотація. За даними CERT-UA, у другому півріччі 2025 року зловмисники дедалі частіше прагнуть не разового викрадення даних, а довготривалого доступу до інфраструктури та повертаються до вже скомпрометованих систем. Розглянуто, чому такі атаки погано виявляються окремими засобами захисту, і як системи XDR завдяки кореляції телеметрії з кінцевих точок, облікових записів, мережі та хмари допомагають помітити закріплення і повторний вхід. Запропоновано практичні заходи для SOC після завершення реагування на інцидент.

Ключові слова: XDR, EDR, SOC, закріплення (persistence), повторна компрометація, MITRE ATT&CK.

За даними CERT-UA, у другому півріччі 2025 року зафіксовано 2909 кіберінцидентів проти 3018 у першому, тобто 5927 за рік, причому інтенсивність атак лишалася стабільно високою [1]. Важливіша за кількість зміна тактики: зловмисники відходять від підходу «Steal & Go» (швидко викрасти дані й піти) і розглядають первинний доступ як плацдарм для наступних етапів. Вони використовують легітимні сервіси віддаленого доступу, а через деякий час повертаються до скомпрометованих систем, щоб перевірити, чи доступ ще працює, чи не закрито вразливості і чи дійсні викрадені облікові дані. CERT-UA окремо наголошує, що відновлення сервісів без усунення першопричин суттєво підвищує ризик повторної, часто більш руйнівної атаки [1].

У термінах MITRE ATT&CK така поведінка відповідає насамперед тактиці «Закріплення» (Persistence), мета якої – зберегти присутність у системі після перезавантаження, зміни паролів чи інших перешкод: заплановані завдання, перехоплення потоку виконання, механізми до завантаження ОС тощо [4]. Повторний вхід часто виконується з дійсними обліковими записами, тож формально виглядає як звичайна робота користувача. Головна складність виявлення полягає в тому, що встановлення механізму закріплення і його спрацювання рознесені в часі – іноді на тижні. Liu та співавтори показують, що саме зв'язування цих двох фаз у графі подій дає змогу точно виявляти закріплення, а запропонований ними підхід зменшує кількість хибних спрацювань у середньому на 93 % порівняно з наявними методами [3].

Окремі засоби захисту бачать лише свою частину картини: антивірус – файл, поштовий фільтр – лист, система керування доступом – успішний вхід. Системи XDR (Extended Detection and Response) збирають телеметрію з кінцевих точок, мережі, пошти та хмарних сервісів і корелюють її в єдиний інцидент.

Опірський, Олійник і Василишин на прикладі фішингового сценарію показують, що XDR відстежує повний ланцюжок атаки, тоді як EDR фіксує лише окремі його фрагменти; водночас для нових і невідомих загроз автори вважають необхідним експертний аналіз фахівців DFIR [2].

Розглянемо типовий сценарій повторної атаки. Після первинного інциденту організація відновила роботу, але не змінила пароль одного сервісного облікового запису і не перевірила автозавантаження на сервері. Через два місяці під цим обліковим записом відбувається вхід через легітимний засіб віддаленого адміністрування з нової IP-адреси, а на сервері спрацьовує запланований завдання, створене ще під час першої атаки. Кожна подія окремо не є підозрілою, але XDR може зіставити їх між собою та з індикаторами первинного інциденту і підняти тривогу до того, як зловмисник перейде до дій щодо цілей.

Щоб це працювало, SOC після закриття інциденту має виконати кілька кроків: зберігати телеметрію довше, ніж типовий інтервал повернення зловмисника (кілька місяців, а не тижнів); додати індикатори та техніки ATT&CK, використані в інциденті, до правил виявлення; змінити всі облікові дані, які могли бути скомпрометовані, включно з сервісними; окремо контролювати появу нових механізмів закріплення (заплановані завдання, служби, ключі автозапуску) і використання засобів віддаленого доступу; встановити період посиленого моніторингу після інциденту. Обмеженнями підходу є вартість тривалого зберігання даних, залежність «нативних» XDR від продуктів одного виробника та потреба у кваліфікованих аналітиках.

Ефективність такого підходу варто вимірювати не лише кількістю виявлених атак. Для SOC корисними показниками є частка інцидентів, після яких повторної компрометації не було протягом кількох місяців; середній час від повторного входу зловмисника до його виявлення; частка інцидентів, для яких підтверджено усунення першопричини, а не лише відновлення сервісів; кількість хибних спрацювань нових правил, доданих після інциденту. Саме ці показники відображають, чи навчається система захисту на попередніх атаках, чи лише реагує на кожну з них окремо. Для порівняння конкретних XDR-рішень за такими параметрами доцільно застосовувати метод експертних оцінок.

Отже, перехід зловмисників до довготривалого доступу і повторних атак робить завершення реагування на інцидент не кінцем, а початком нового етапу моніторингу. Системи XDR завдяки кореляції подій з різних джерел і в часі є зручною основою для виявлення закріплення та повторного входу, однак їхня ефективність залежить від налаштування правил, строку зберігання телеметрії та роботи аналітиків SOC.

Список використаних джерел

1. Кіберзагрози: Україна. Аналітика за II півріччя 2025 / Держспецзв'язку, CERT-UA. Київ, 2026. URL: https://cip.gov.ua/services/cm/api/attachment/download?id=74645 (дата звернення: 18.10.2026).

2. Опірський І. Р., Олійник А. В., Василишин С. І. Дослідження ефективності XDR рішень для виявлення та усування загроз. Сучасний захист інформації. 2025. № 2(62). С. 141–157. DOI: 10.31673/2409-7292.2025.027390.

3. Liu Q., Shoaib M., Rehman M. U., Bao K., Hagenmeyer V., Hassan W. U. Accurate and Scalable Detection and Investigation of Cyber Persistence Threats. arXiv:2407.18832. 2024. URL: https://arxiv.org/abs/2407.18832 (дата звернення: 18.10.2026).
4. Strom B. E., Applebaum A., Miller D. P. et al. MITRE ATT&CK: Design and Philosophy. The MITRE Corporation, 2020. URL: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf (дата звернення: 18.10.2026).

Тамбієва В.Г.
студентка групи БСДм-61, ННІКБЗІ ДУІКТ
Київ, Україна

ПРОГРАМНА СИСТЕМА КОМПЛЕКСНОГО ЗАХИСТУ БАЗ ДАНИХ ВІД НЕСАНКЦІОНОВАНОГО ДОСТУПУ

У роботі представлено розробку програмної системи комплексного захисту баз даних від несанкціонованого доступу, побудованої за клієнт-серверною архітектурою з використанням мови програмування Python. Система складається з модулів шифрування та розшифрування даних, їх передавання, взаємодії з базою даних, а також клієнтської та серверної частин. Реалізовано автентифікацію користувачів, розмежування прав доступу та шифрування інформації за алгоритмом AES-256-GCM перед передаванням і зберіганням. Проведено експериментальне тестування основних функцій, зокрема перевірку підключення, передавання даних і виявлення помилок під час розшифрування. Розроблена система орієнтована на автономне використання в локальній інфраструктурі підприємства без постійної залежності від зовнішніх сервісів.

Ключові слова: програмна система, захист баз даних, клієнт-серверна архітектура, автентифікація, розмежування прав доступу, AES-256-GCM, шифрування даних.

Перехід підприємств до цифрового зберігання й оброблення інформації підвищує вимоги до захисту баз даних (БД), що містять конфіденційні відомості. Несанкціонований доступ (НСД) може спричинити розголошення, модифікацію або знищення даних, порушуючи безперервність бізнес-процесів. Ефективний захист БД передбачає поєднання автентифікації користувачів і пристроїв, розмежування прав доступу, шифрування та контролю цілісності інформації.

Тому актуальною є розробка програмної системи, що працює в локальній мережі підприємства, забезпечує безпечне передавання даних без постійного підключення до Інтернету та дає змогу зберігати інформацію в зашифрованому вигляді.

Підходи до захисту баз даних охоплюють організаційні, програмно-технічні та криптографічні заходи. У роботах [1, 2] наголошується на необхідності врахування моделі порушника, можливих цілей атаки та рівня його технічних можливостей. На практиці це означає, що система має протидіяти не тільки зовнішнім спробам підключення, а й помилкам конфігурації, зловживанню привілеями та спробам доступу з пристроїв, які не належать до переліку довірених.

Серед розглянутих програмних рішень: HashiCorp Vault, Acra Database Protection Suite та SAP SQL Anywhere. Вони забезпечують керування секретами й захист даних, проте відрізняються моделями розгортання, ліцензування та доступністю вихідного коду. Це обґрунтовує доцільність створення локальної системи, у якій підприємство самостійно контролює передавання, шифрування та зберігання інформації. Криптографічний компонент розробки використовує

AES-256 у режимі GCM, що забезпечує конфіденційність і дає змогу виявляти модифікацію даних [3]. Для обміну інформацією між клієнтом і сервером застосовано сокети, які забезпечують мережеву взаємодію компонентів системи.

Метою роботи є розроблення програмної системи комплексного захисту баз даних від несанкціонованого доступу, яка забезпечує автентифікацію користувачів, розмежування доступу за рівнями допуску, наскрізне шифрування даних і захищене передавання в межах локальної мережі без необхідності залучення зовнішніх сервісів.

Розроблену систему побудовано за клієнт-серверною моделлю з використанням мови Python і сокетів для мережевого обміну. Серверна частина забезпечує взаємодію з БД, перевірку користувачів і надання даних відповідно до їхніх прав, а клієнтська – підключення, автентифікацію, завантаження файлів і надсилання оновлень. Графічний інтерфейс реалізовано за допомогою бібліотеки Tkinter.

Захист передбачає перевірку дозволених пристроїв, автентифікацію користувачів і розмежування доступу за рівнями допуску, призначеними адміністратором. Це обмежує можливості перегляду та зміни даних відповідно до повноважень користувача. Також передбачено завершення сесії після відключення клієнта або закриття застосунку.

Для захисту інформації застосовано наскрізне шифрування AES-256-GCM: файл шифрується на клієнтському пристрої перед передаванням, а розшифрування виконується отримувачем за допомогою секретного ключа. Випадковий 96-бітний вектор ініціалізації використовується під час шифрування та розшифрування, а режим GCM дає змогу виявляти неправильний ключ або модифікацію зашифрованих даних [6].

Система складається з модулів data_cyphering.py для шифрування, data_transfer.py для мережевого обміну, db_interaction.py для роботи з БД, а також secure_db_server.py і secure_db_client.py, що реалізують серверну та клієнтську логіку. Така модульна структура спрощує супровід і подальше вдосконалення програмного засобу.

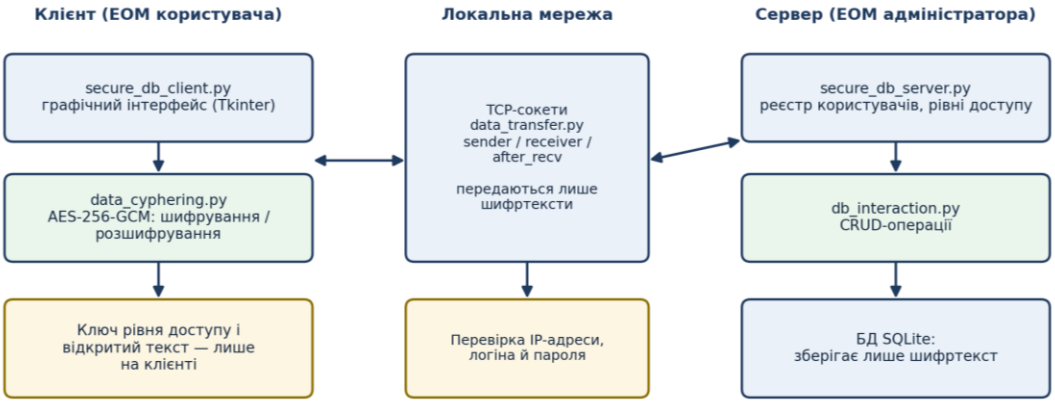


Рис. 1. Архітектура програмної системи захисту БД від НСД

У ході роботи було проведено експериментальне дослідження розробленої програмної системи, перевірено працездатність основних функцій системи, зокрема підключення клієнта до сервера, автентифікацію користувачів, розмежування прав доступу, передавання, шифрування та розшифрування файлів, а також оброблення помилок у разі використання неправильного ключа або модифікації зашифрованих даних. Результати дослідження підтвердили працездатність реалізованих функцій і здатність системи коректно обробляти передбачені помилкові ситуації. Також було проведено порівняння з наявними програмними рішеннями, яке засвідчило доцільність використання розробки для автономного захисту даних у локальній інфраструктурі підприємства без постійної залежності від зовнішніх сервісів.

Висновки

У роботі розроблено клієнт-серверну систему, що поєднує ідентифікацію пристрою, автентифікацію користувача, розмежування доступу за рівнями допуску, передавання даних через сокети та шифрування AES-256-GCM. Експериментально перевірено основні сценарії роботи: підключення, вхід користувача, отримання та надсилання файлів, а також реакцію на неправильний ключ або зміну зашифрованих даних. Отримані результати свідчать про працездатність реалізованих механізмів у дослідженому середовищі.

Практична цінність системи полягає в можливості локального розгортання, відсутності потреби в постійному Інтернет-з’єднанні для обміну даними всередині підприємства, модульній структурі та можливості подальшої інтеграції до комплексної системи захисту інформації. Подальший розвиток доцільно спрямувати на розширення журналювання подій безпеки, формалізоване тестування контролю доступу, перевірку стійкості до мережевих атак і вдосконалення процедур керування та ротації ключів.

Список використаних джерел

1. Погребняк А. В. Технології комп’ютерної безпеки: монографія. Рівне: МЕГУ, 2011. 117 с. URL: https://studfile.net/preview/7435414/page:11/ (дата звернення: 22.10.2026).

2. Колесник Д. Модель порушника, загроз, безпеки для постквантового цифрового підпису. Наука онлайн : міжнародний електронний науковий журнал. 2020. № 11. URL: https://nauka-online.com/ua/publications/informatsionnye-tehnologii/2020/11/model-porushnika-zagroz-bezpeki-dlya-postkvantovogo-tsifrovogo-pidpisu/ (дата звернення: 22.10.2026).

3. Dworkin M. Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC. NIST Special Publication 800-38D. Gaithersburg: National Institute of Standards and Technology, 2007. URL: https://csrc.nist.gov/pubs/sp/800/38/d/final (дата звернення: 22.10.2026).

Мазур Олександра
студентка групи БСДм-61, ННІКБЗІ ДУІКТ
Київ, Україна

СИСТЕМА ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ ВІД ШКІДЛИВОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ НА ОСНОВІ ПРОАКТИВНОГО АНАЛІЗУ

У роботі представлено розробку системи захисту інформаційних систем від шкідливого програмного забезпечення на основі проактивного аналізу. Розглянуто архітектуру системи, що передбачає сканування мережі та системи, аналіз мережевої активності, аудит безпеки й пошук потенційно небезпечних об'єктів. Основну увагу приділено поєднанню сигнатурного аналізу, зіставлення характеристик програмних об'єктів, ізольованого виконання в середовищі Sandbox і виявлення експлойтів. Описано основні етапи роботи системи: оцінювання загроз, запобігання, моніторинг і формування звітності. Проведено експериментальне тестування з використанням тестового об'єкта. За результатами перевірки вдосконалений засіб забезпечив виявлення та видалення об'єкта, який не було виявлено під час попереднього сканування базовим антивірусним засобом. Отримані результати підтверджують працездатність розробки в межах проведеного експерименту та обґрунтовують доцільність подальшого дослідження ефективності проактивного аналізу на ширшому наборі загроз.

Ключові слова: інформаційна безпека, шкідливе програмне забезпечення, проактивний аналіз, сигнатурний аналіз, поведінковий аналіз, Sandbox, виявлення експлойтів.

Сучасний розвиток інформаційних технологій супроводжується зростанням кількості кіберзагроз, серед яких значну небезпеку становить шкідливе програмне забезпечення (ШПЗ), здатне порушувати конфіденційність, цілісність і доступність інформації. Традиційні антивірусні засоби переважно використовують сигнатурний аналіз, ефективний для виявлення відомих загроз, проте обмежений у разі появи нових або модифікованих шкідливих програм. Тому важливим напрямом удосконалення захисту є проактивний аналіз, що передбачає дослідження характеристик і поведінки програмних об'єктів для виявлення потенційних загроз до настання негативних наслідків.

Поєднання проактивного та сигнатурного аналізу, технологій ізольованого виконання програм (Sandbox) і виявлення експлойтів дає змогу створити багаторівневу систему захисту. Отже, актуальність роботи полягає в розробленні системи на основі проактивного аналізу, спрямованої на своєчасне виявлення підозрілих об'єктів, запобігання виконанню небезпечних операцій і підвищення рівня інформаційної безпеки.

Проблематика захисту інформаційних систем від шкідливого програмного забезпечення розглядається в сучасних працях, присвячених інформаційній безпеці, виявленню шкідливих програм і вдосконаленню засобів протидії їм. У цих дослідженнях висвітлено актуальні загрози, методи їх виявлення та підходи до підвищення ефективності захисту інформаційних систем [1, 2]. Одним із традиційних методів виявлення ШПЗ є сигнатурний аналіз, який дає змогу

ефективно ідентифікувати відомі загрози, проте має обмеження під час виявлення нових і модифікованих програм. Для подолання цих обмежень застосовують проактивні методи аналізу, зокрема дослідження поведінки програмних об’єктів і перевірку їхньої активності в ізольованому середовищі Sandbox. У дослідженні [3] розглянуто підходи до вдосконалення систем протидії шкідливому програмному забезпеченню, що є важливим напрямом розвитку засобів інформаційного захисту. Комплексний аналіз шкідливих програм, представлений у праці [4], також підкреслює актуальність використання різних підходів до дослідження загроз і вдосконалення методів їх виявлення.

Отже, аналіз джерел підтверджує доцільність комбінування сигнатурного та поведінкового аналізу для підвищення ефективності захисту. Результативність конкретної системи при цьому необхідно підтверджувати експериментальним тестуванням.

Метою роботи є розроблення та вдосконалення системи захисту інформаційних систем від шкідливого програмного забезпечення на основі проактивного аналізу для підвищення ефективності виявлення потенційно небезпечних об’єктів, своєчасного реагування на загрози та зменшення ризику негативного впливу шкідливих програм на роботу інформаційної системи.

Основою розробленої системи є комплексний підхід до захисту інформаційних систем, що поєднує різні механізми виявлення та запобігання загрозам. Архітектура містить модулі сканування мережі, аналізу мережевої активності, перевірки безпеки, аудиту паролів і системи, а також пошуку об’єктів на дисках. Основним компонентом є підсистема проактивного аналізу, яка передбачає використання сигнатурного аналізу, DNA Matching, Sandbox та Exploit Detection. Сигнатурний аналіз забезпечує виявлення відомих загроз шляхом порівняння файлів із наявними зразками. DNA Matching дає змогу зіставляти характерні ознаки програмних об’єктів, Sandbox забезпечує ізольоване виконання та аналіз поведінки підозрілих програм, а Exploit Detection спрямований на виявлення ознак використання вразливостей програмного забезпечення. Зазначені механізми доповнюють один одного: сигнатурний аналіз орієнтований на відомі загрози, тоді як поведінковий аналіз та виявлення експлойтів дають змогу досліджувати потенційно небезпечні дії. Ефективність системи залежить від реалізації та налаштувань кожного компонента.

Функціонування системи передбачає чотири основні етапи: оцінювання, запобігання, моніторинг і формування звітності. На етапі оцінювання аналізуються програмні об’єкти та визначаються ознаки загроз. У разі їх виявлення система може блокувати підозрілий код або видаляти відомі шкідливі об’єкти. Моніторинг забезпечує реєстрацію подій і надсилання сповіщень через електронну пошту, SNMP, SMS, Syslog або журнал подій операційної системи. На завершальному етапі формуються звіти для подальшого аналізу інцидентів.

Алгоритм роботи передбачає перевірку вхідних даних і мережевого трафіку, аналіз об’єктів та прийняття рішення залежно від результатів. Якщо загрозу виявлено, система блокує відповідний об’єкт або фіксує подію для подальшого реагування.

Систему реалізовано мовами C/C++, що дає змогу розробляти компоненти для роботи з файлами, системними ресурсами та механізмами операційного середовища. Така архітектура створює основу для комплексного захисту інформаційних систем від шкідливого програмного забезпечення.

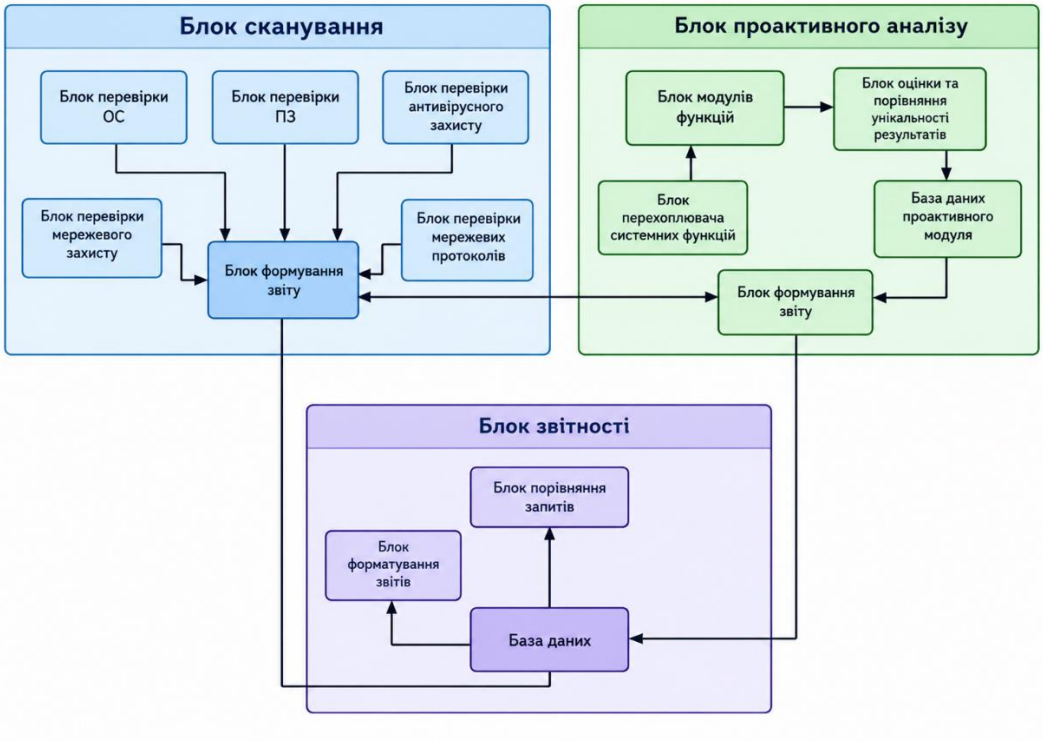


Рис. 1. Структура системи

Для перевірки розробленої системи проведено експеримент із тестовим об’єктом Alarm-Worm.HCF2R. Під час першої перевірки базовий антивірусний засіб без удосконаленого модуля не виявив загрозу. Після встановлення вдосконаленого засобу повторне сканування забезпечило виявлення та видалення тестового об’єкта. Це підтверджує працездатність механізму в межах проведеного експерименту, проте для оцінювання загальної ефективності потрібне тестування на ширшому наборі загроз. Практичне значення розробки полягає в поєднанні перевірки програмних об’єктів, контролю мережевої активності, аудиту системи та реєстрації подій. Подальше вдосконалення передбачає розширення набору тестових зразків, перевірку поведінкового аналізу, оцінювання хибнопозитивних спрацювань і дослідження стійкості системи до спроб обходу механізмів захисту.

Висновки

У роботі розроблено систему захисту інформаційних систем від шкідливого програмного забезпечення на основі проактивного аналізу, що поєднує сигнатурний аналіз, DNA Matching, Sandbox та Exploit Detection. Визначено структуру системи, яка включає модулі сканування, перевірки безпеки, аудиту та моніторингу, а також етапи виявлення загроз, реагування й формування звітності.

Експериментальне дослідження з тестовим об’єктом Alarm-Worm.HCF2R показало, що базовий антивірусний засіб не виявив загрозу, тоді як після встановлення вдосконаленого засобу об’єкт було виявлено та видалено. Також оцінено час запуску системи, тривалість сканування та використання системних ресурсів. Отримані результати підтверджують працездатність розробки в межах проведеного експерименту. Подальше дослідження передбачає тестування на ширшому наборі шкідливих об’єктів і порівняння з іншими антивірусними рішеннями для об’єктивнішого оцінювання ефективності проактивного аналізу.

Список використаних джерел

1. Педяш В., Ледовський Є., Ткач В. Сучасні методи виявлення шкідливого програмного забезпечення. Вісник Хмельницького національного університету. Технічні науки. 2024. Т. 333, № 2. С. 389–392. DOI: https://doi.org/10.31891/2307-5732-2024-333-2-60.

2. Піддубний Д., Шабала Є. Аналіз сучасного шкідливого програмного забезпечення та методи боротьби з ним. Build-Master-Class-2024: International scientific-practical conference of young scientists. Київ: КНУБА, 2024. С. 465–466. URL: https://repositary.knuba.edu.ua/handle/123456789/15932.

3. Сундучков А., Аносов А. Удосконалення системи протидії впливу шкідливого програмного забезпечення. Безпека інформаційно-комунікаційних систем 2025. 2025. URL: https://elibrary.kubg.edu.ua/id/eprint/54214.

4. Кіцель Н., Волканін Є., Борисенко О., Матвєєв В., Мальований В. Комплексний аналіз шкідливих програм: підходи, виклики та перспективи. Інформаційні технології та суспільство. 2025. DOI: https://doi.org/10.32689/maup.it.2025.4.14.

Сокульський О. Є.
доцент, к.т.н., ВІТІ імені Героїв Крут, Київ, Україна
Топольськов Є.О.
доцент, к.т.н., КНУ імені Тараса Шевченка, Україна
Москаленко Н. В.
асистент, КНУ імені Тараса Шевченка, Україна

ГОМОМОРФНЕ ШИФРУВАННЯ ТА ХМАРНІ ТЕХНОЛОГІЇ

Шифрування — це процес перетворення даних у спеціальний код, який можна зберігати, передавати та обробляти з забезпеченням необхідної конфіденційності, цілісності та безпеки. З розвитком інформаційних технологій зростає потреба у підвищенні якості роботи та захисті інформації. Особливо чутливими галузями є медицина, фінансовий сектор, електронне голосування, де використання гомоморфного шифрування є досить ефективним.

Ключові слова. Шифрування, хмарні технології, унікальний ключ.

З появою та розвитком хмарних технологій розробники та користувачі мають ширші можливості використовувати ресурси без їхньої фізичної наявності. Дані можуть перебувати в стані спокою (Data at Rest), під час передачі (Data in Transit) та під час використання (Data in Use). Для кожного з цих станів існують як загальні, так і власні загрози. Особливе значення набувають проблеми контролю доступу та ідентифікації (IAM), шифрування даних, моніторингу й аудиту, а також резервного копіювання [1]. Шифрування даних з використанням хмарних технологій можна розділити на стандартне (Server-side encryption), наскрізне (E2EE – End-to-End Encryption) та BYOK (Bring Your Own Encryption).

Платформа Dropbox реалізує симбіоз всіх трьох типів шифрування даних, забезпечуючи цим кілька рівнів безпеки [2]. Найбільш складним і ефективним захистом інформації в хмарі є використання гомоморфного шифрування (Homomorphic Encryption), яке дає змогу обробляти зашифровану інформацію в хмарі без попереднього дешифрування. Залежно від потреб можна реалізувати один з типів гомоморфного шифрування:

- Частково гомоморфне (PHE).
- Майже гомоморфне (SWHE).
- Повністю гомоморфне(FHE).

Використання гомоморфного шифрування, попри наявні переваги, має й свої недоліки. Це низька швидкість обробки даних, збільшення обсягів пам’яті та іноді досить висока складність інтеграції, особливо під час роботи зі звичайними базами даних (SQL/NoSQL). Найширше цей тип шифрування використовується в системах із високими вимогами до приватності (FinTech, BioTech, AI). Успішне використання демонструє Google Cloud (Confidential Computing). Microsoft Azure, IBM Cloud, Amazon Web Services.

Перелік посилань:

1. Ю. Є. Яремчук, О. В. Салієва, І. О. Бондаренко. Основи криптографічного захисту інформації. ВНТУ, 2024 140 с. [Електронний ресурс]. Режим доступу: https://pdf.lib.vntu.edu.ua/books/2024/Yaremchuk_2024_139.pdf

2. Повністю гомоморфне шифрування. [Електронний ресурс]. Режим доступу: https://www.gate.com/uk/learn/articles/fully-homomorphic-encryption-introduction-and-use-cases/3876

Головненко Н.Р.
Студент групи ІКБм-62 ННІКБЗІ ДУІКТ
Київ, Україна

РОЗРОБЛЕННЯ ТЕХНОЛОГІЇ ВИЯВЛЕННЯ КІБЕРАТАК НА ОСНОВІ КОРЕЛЯЦІЇ ПОДІЙ SPLUNK ENTERPRISE SECURITY ТА ТЕЛЕМЕТРІЇ SENTINELONE

Ключові слова: кібербезпека, SIEM, EDR, Splunk Enterprise Security, SentinelOne, кореляція подій, виявлення кібератак.

Актуальність дослідження зумовлена зростанням складності кібератак, під час яких зловмисники послідовно використовують кілька способів проникнення, закріплення та поширення в інформаційній системі. Окремий журнал мережевого обладнання, запис операційної системи або спрацювання засобу захисту кінцевої точки часто відображає лише частину такої активності. За великої кількості повідомлень це ускладнює встановлення взаємозв'язків між подіями, підвищує навантаження на аналітиків центру моніторингу безпеки (SOC) і створює ризик несвоєчасного реагування. Тому практичного значення набуває об'єднання засобів SIEM та EDR для аналізу сукупності ознак компрометації.

Метою роботи є розроблення технології виявлення кібератак шляхом кореляції подій, що обробляються у Splunk Enterprise Security, із телеметрією SentinelOne, для підвищення контекстності сповіщень і обґрунтованості їх пріоритизації. Об'єктом дослідження є процес виявлення кібератак в інформаційній системі організації, а предметом — методи нормалізації, часової та поведінкової кореляції подій SIEM і EDR. Основне завдання полягає не лише в об'єднанні журналів, а у формуванні відтворюваної послідовності дій від отримання даних до встановлення підозрілого сценарію та створення інциденту для перевірки.

Як технологічну основу обрано Splunk Enterprise Security — рішення для централізованого аналізу подій безпеки, створення правил виявлення та оцінювання ризиків — і SentinelOne, що забезпечує захист кінцевих пристроїв та формує інформацію про виявлені загрози й, залежно від ліцензії та доступних функцій, розширену телеметрію активності процесів. Офіційні засоби інтеграції SentinelOne зі Splunk передбачають передавання даних у середовище аналітики [1, 2]. Водночас сам факт інтеграції не гарантує доступності кожного типу подій: склад полів, повнота історії та частота надходження визначаються конфігурацією джерела, правами доступу та способом експорту.

Запропонована технологія охоплює п'ять взаємопов'язаних етапів: збирання, нормалізацію, контекстне збагачення, кореляцію та оцінювання результатів. На етапі збирання до Splunk надходять сповіщення SentinelOne, доступні події про процеси, файли й мережеві з'єднання, а також журнали

автентифікації, DNS, проксі-серверів і міжмережевих екранів. Передавання даних EDR можливе через підтримувані інтеграційні компоненти або API, а конкретну схему отримання слід обирати з урахуванням доступних функцій середовища. Для контролю якості необхідно перевіряти затримку доставки, дублікати, пропуски записів і коректність часових позначок.

На етапі нормалізації різнорідні записи приводяться до узгоджених атрибутів: часу події, ідентифікатора пристрою, користувача, адреси джерела та призначення, типу дії, імені процесу, командного рядка й контрольної суми файлу — якщо відповідні дані присутні в події. Для узгодження семантики доцільно використовувати Common Information Model (CIM) і пов’язані моделі Endpoint, Authentication та Network Traffic [3]. Особливого значення набуває відповідність імен пристроїв, локальних та доменних облікових записів і часових зон: помилки зіставлення можуть спричиняти як хибні зв’язки, так і втрату важливого контексту.

Контекстне збагачення полягає у приєднанні відомостей про критичність активу, роль користувача, репутацію індикаторів компрометації та відому активність, дозволену внутрішніми правилами організації. При цьому адресу IP або ім’я хоста не варто розглядати як безумовно унікальний ідентифікатор: для динамічних адрес, VPN і спільних робочих станцій потрібне врахування часових меж та додаткових атрибутів. Кореляційні правила мають зберігати посилання на первинні події, щоб аналітик міг відтворити ланцюг активності та відрізнити спостереження від аналітичного висновку.

Основним елементом технології є поєднання часової та поведінкової кореляції. Часова кореляція групує записи, пов’язані з одним пристроєм або користувачем, у визначеному інтервалі; поведінкова — перевіряє послідовність або поєднання дій, характерних для сценарію атаки. Наприклад, сигналом підвищеного ризику може бути сукупність невдалих входів з подальшою успішною автентифікацією, запуском незвичного інтерпретатора команд на цій самій кінцевій точці та підозрілим мережевим з’єднанням. Кожна з цих ознак окремо може мати легітимне пояснення, проте їх послідовна поява підсилює підстави для перевірки.

Для реалізації правил у Splunk Enterprise Security пропонується використовувати пошуки мовою SPL з обмеженням за інтервалом часу, агрегацією за ідентифікаторами сутностей і перевіркою умов сценарію. Під час їх побудови доцільно застосовувати таксономію MITRE ATT&CK для зв’язування спостережуваної поведінки з тактиками та техніками порушника [4]. Це дозволяє оцінювати не тільки наявність окремих індикаторів, але й охоплення етапів атаки, а також підтримувати простежуваність логіки виявлення. Водночас позначення техніки ATT&CK не є доказом успішної реалізації атаки та потребує підтвердження контекстом.

Щоб зменшити кількість ізольованих сповіщень, запропоновано підхід до пріоритизації за ризиком. Первинні правила формують проміжні результати з урахуванням достовірності ознаки, критичності активу, джерела даних і повторюваності подій. Підсумковий ризик оцінюється для визначеної сутності за сукупністю різнорідних сигналів у часовому вікні. Функціональні можливості ризик-орієнтованого виявлення у Splunk Enterprise Security забезпечують агрегування відповідних подій та формування інцидентів для подальшого розгляду [5]. Порогові значення та ваги мають налаштовуватися на основі характеристик конкретної мережі, а не використовуватися як універсальні константи.

Типовим сценарієм перевірки технології є моделювання активності, що включає запуск сценаріїв PowerShell з нетиповими аргументами, появу підозрілого дочірнього процесу, звернення до зовнішнього ресурсу та реакцію EDR. У межах тестування слід зіставити часову шкалу подій Windows, мережевих журналів і SentinelOne, оцінити доступність спільних ідентифікаторів та перевірити коректність спрацьовування кореляційного правила. Окремо потрібно перевіряти легітимні адміністративні дії зі схожими ознаками, щоб правила не ототожнювали використання стандартних інструментів системи зі шкідливою діяльністю.

Методика оцінювання передбачає контрольований експеримент на ізольованому тестовому пристрої з агентом SentinelOne та підключенням до дозволеного контуру моніторингу Splunk. Необхідно заздалегідь визначити набір сценаріїв, очікувані спостереження, контрольні часові межі та еталонну розмітку подій. Для порівняння результатів окремого аналізу SIEM, окремих EDR-сповіщень і запропонованої кореляції використовуються частка виявлених тестових сценаріїв, кількість хибнопозитивних спрацювань, точність сповіщень та затримка між подією і формуванням сигналу. Виробничі дані доцільно знеособлювати, а активні випробування не виконувати на робочих серверах.

Очікуваний практичний результат полягає в підвищенні інформативності сповіщення завдяки поєднанню мережевого, системного та кінцевого контексту, що може скоротити обсяг ручного зіставлення журналів під час розслідування. Водночас ефективність залежатиме від якості нормалізації, повноти телеметрії, синхронізації часу та правильності добору сценаріїв. Отже, запропонована технологія створює основу для керованого процесу виявлення багатокрокових кібератак засобами Splunk Enterprise Security і SentinelOne, а підтвердження її переваг потребує експериментального порівняння з початковими підходами за визначеними метриками. Для відтворюваності підходу потрібно документувати версії правил, часові вікна, джерела даних, відповідність полів та умови винятків.

Список використаних джерел

1. SentinelOne. SentinelOne App For Splunk. Splunkbase. URL: https://splunkbase.splunk.com/app/5433 (дата звернення: 26.10.2026).

2. SentinelOne. Does SentinelOne integrate with my SIEM? FAQ. URL: https://www.sentinelone.com/nl/faq/ (дата звернення: 26.10.2026).
3. Splunk. Use the CIM to normalize data at search time. URL: https://help.splunk.com/en/splunk-enterprise/common-information-model/6.4/using-the-common-information-model/use-the-cim-to-normalize-data-at-search-time (дата звернення: 26.10.2026).
4. MITRE. ATT&CK Matrix for Enterprise. URL: https://attack.mitre.org/matrices/enterprise/ (дата звернення: 26.10.2026).
5. Splunk. Risk scoring in Splunk Enterprise Security. URL: https://help.splunk.com/en/splunk-enterprise-security-8/administer/8.5/risk-based-alerting/risk-scoring-in-splunk-enterprise-security (дата звернення: 26.10.2026).

Журавель Олександр Олегович
студент групи ІКБм-61, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ АТАК ТИПУ BRUTE FORCE ІЗ ЗАСТОСУВАННЯМ МЕТОДУ МАШИННОГО НАВЧАННЯ RANDOM FOREST

Атаки типу Brute Force є одним зі способів отримання несанкціонованого доступу до інформаційних систем шляхом багаторазового підбору паролів. Для їх виявлення можна застосувати методи машинного навчання, які допомагають розрізняти звичайну активність і спроби атаки. Одним із таких методів є Random Forest, який дає змогу класифікувати дані за певними ознаками та може бути використаний для виявлення атак типу Brute Force.

Ключові слова: Brute Force, машинне навчання, Random Forest, кібербезпека, класифікація, виявлення атак.

Особливості атак типу Brute Force

Залежно від способу підбору облікових даних розрізняють декілька різновидів атак Brute Force. Класична атака передбачає перевірку різних комбінацій символів, тоді як словникова використовує заздалегідь підготовлений список поширених паролів. Гібридний підхід поєднує ці способи, наприклад шляхом додавання цифр або спеціальних символів до відомих слів. Окремим різновидом є розпилення паролів (Password Spraying), за якого один або декілька поширених паролів перевіряються для багатьох облікових записів. Успішне проведення таких атак може призвести до отримання несанкціонованого доступу, викрадення конфіденційної інформації та подальшої компрометації системи [1].

Застосування методу Random Forest для виявлення Brute Force атак

Random Forest – це метод машинного навчання, який використовує сукупність дерев рішень для класифікації даних. Під час роботи алгоритму (рис. 1) кожне дерево формує власний прогноз, після чого результати об’єднуються за принципом голосування більшості. Для побудови дерев застосовуються випадкові вибірки даних і підмножини ознак, що дає змогу врахувати різні закономірності в досліджуваних даних [2, 3, с. 5-6].

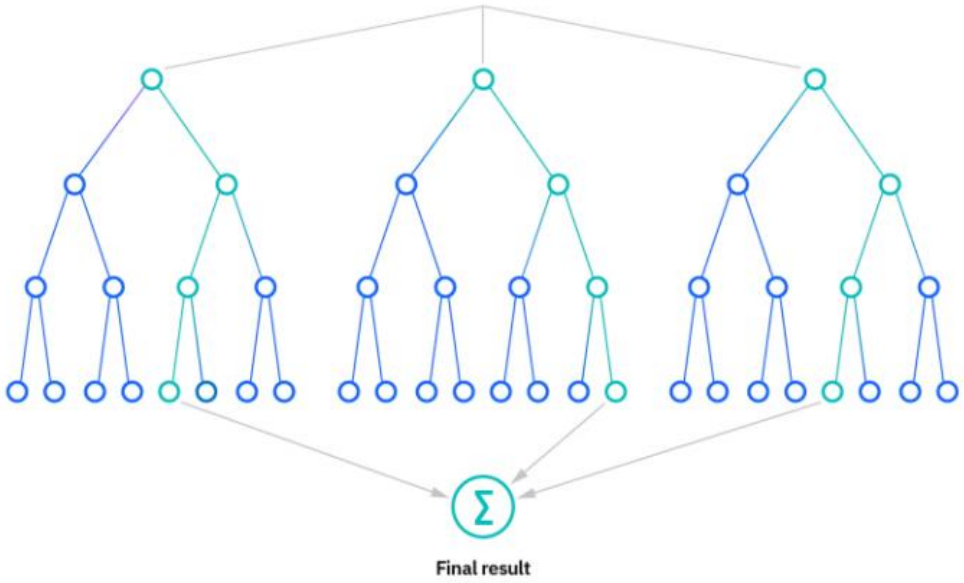


Рис. 1 – Принцип роботи алгоритму Random Forest [2]

Для виявлення атак типу Brute Force метод Random Forest можна застосувати для класифікації записів набору даних на два класи: нормальна активність (benign) та активність, пов'язана з атакою. З цією метою необхідно підготувати дані, визначити відповідні ознаки та навчити модель на прикладах із відомими класами. Після навчання модель зможе класифікувати нові записи відповідно до виявлених закономірностей. Якість роботи моделі можна оцінити за допомогою метрик точність (precision), повнота (recall) та F1-показник (F1-score). Вони дають змогу оцінити правильність визначення атак, повноту їх виявлення та збалансованість цих показників. Оцінювання за зазначеними метриками дасть змогу охарактеризувати якість класифікації нормальної активності та атак типу Brute Force [4].

Проблеми забезпечення ефективного виявлення атак типу Brute Force

Одним із викликів під час розроблення технологій виявлення атак типу Brute Force є обмеженість наборів даних, які використовуються для навчання та оцінювання моделей машинного навчання. За даними Канадського інституту з кібербезпеки (Canadian Institute for Cybersecurity) [5] використання наборів даних для дослідження кіберзагроз ускладнюється їх обмеженою доступністю та невідповідністю характеристик сучасного мережевого трафіку. Це ускладнює оцінювання здатності моделі виявляти атаки в умовах, наближених до реальних умов експлуатації.

У контексті застосування Random Forest зазначена проблема підкреслює важливість вибору репрезентативних даних для навчання та перевірки моделі. Тому під час розроблення технології виявлення Brute Force атак необхідно врахувати характеристики набору даних і перевірити якість класифікації на прикладах, які не використовуються під час навчання.

Отже, використання методу Random Forest дозволяє сформувати основу для автоматизованого виявлення підозрілої активності. Практична цінність такого підходу полягає в можливості класифікувати події автентифікації та визначати спроби несанкціонованого доступу. Доцільність застосування моделі в системах захисту інформації визначається, зокрема, її здатністю виявляти атаки та мінімізувати кількість хибнопозитивних результатів.

Перелік посилань:

1. SoSafe. Brute-Force Attack. URL: https://sosafe-awareness.com/glossary/brute-force-attack/ (дата звернення: 20.10.2026).

2. IBM. What is Random Forest? URL: https://www.ibm.com/think/topics/random-forest (дата звернення: 23.10.2026).

3. Breiman L. Random Forests. Machine Learning. 2001. Vol. 45. P. 5–32. URL: https://doi.org/10.1023/A:1010933404324 (дата звернення: 23.10.2026).

4. Scikit-learn. Classification metrics. URL: https://scikit-learn.org/stable/modules/model_evaluation.html (дата звернення: 24.10.2026).

5. Canadian Institute for Cybersecurity. CSE-CIC-IDS2018 on AWS. URL: https://www.unb.ca/cic/datasets/ids-2018.html (дата звернення: 25.10.2026).

Хорольський Костянтин Андрійович
Студент групи ІКБм-62, ННІКБЗІ ДУІКТ, Київ, Україна

ЗАСТОСУВАННЯ ГРАДІЄНТНОГО БУСТИНГУ ДЛЯ ВИЯВЛЕННЯ ШКІДЛИВИХ ДОКУМЕНТНИХ ВКЛАДЕНЬ ФІШИНГОВИХ ЕЛЕКТРОННИХ ЛИСТІВ

Розглянуто проблему виявлення шкідливих документних вкладень фішингових електронних листів методами статичного аналізу. Проаналізовано набір даних CIC-Trap4Phish 2025 та результати його базової оцінки, виявлено нижчу ефективність класифікації HTML-вкладень і ризик завищення метрик для документів Office через різне походження класів. Запропоновано схему технології виявлення з окремими моделями градієнтного бустингу для кожного формату та поясненням рішень на основі значень SHAP, а також методику експериментального порівняння реалізацій XGBoost, LightGBM і CatBoost з контролем ознак, пов'язаних із джерелом даних.

Ключові слова: фішинг, шкідливі вкладення, градієнтний бустинг, машинне навчання, статичний аналіз, SHAP, CIC-Trap4Phish.

Фішингові електронні листи залишаються одним із провідних векторів початкового доступу до корпоративних інформаційних систем: за даними звіту Verizon DBIR 2026, людський чинник присутній у 62 % витоків даних, а фішинг є вектором первинного доступу у 16 % випадків [4]. Особливу небезпеку становлять шкідливі вкладення – документи Word, Excel, PDF та HTML-сторінки, у які зловмисники вбудовують макроси, скрипти або посилання на фішингові ресурси. Для України ця загроза має прикладний характер: CERT-UA регулярно фіксує кампанії проти державних органів та об'єктів критичної інфраструктури, у яких шкідливе програмне забезпечення доставляється через вкладення листів. Сигнатурні засоби захисту недостатньо ефективні проти нових та обфускованих зразків, а динамічний аналіз у «пісочниці» потребує значних обчислювальних ресурсів, що зумовлює актуальність застосування методів машинного навчання для статичного аналізу вкладень.

Більшість наявних досліджень зосереджена на окремих форматах файлів (PDF, документи Office або HTML), що ускладнює побудову та порівняння моделей для різнорідних вкладень [1]. Цю проблему частково усуває набір даних CIC-Trap4Phish 2025, розроблений Канадським інститутом кібербезпеки, який містить по 20 000 збалансованих зразків для форматів Word, Excel, PDF та HTML, а також від 40 до 48 статичних ознак для кожного формату [1]. Автори набору оцінили його за допомогою класифікаторів Decision Tree, Random Forest та XGBoost: для Word, Excel і PDF значення F1-міри перевищили 0,99, тоді як для HTML становили близько 0,94 [1]. Водночас залишаються невирішеними такі питання: нижча ефективність виявлення шкідливих HTML-вкладень; можливе завищення результатів для документів Office через різне походження класів (легітимні зразки отримано веб-краулінгом або синтетичною генерацією,

шкідливі – з платформи MalwareBazaar); відсутність порівняння сучасних реалізацій градієнтного бустингу (XGBoost, LightGBM, CatBoost) з налаштуванням гіперпараметрів.

Метою роботи є обґрунтування підходу до виявлення шкідливих документних вкладень фішингових електронних листів на основі градієнтного бустингу (ГБ), який забезпечує високу точність класифікації, стійкість до артефактів збору даних та пояснюваність рішень для аналітика центру моніторингу безпеки (SOC).

Статичні ознаки вкладень мають табличну структуру та є неоднорідними: вони поєднують бінарні індикатори (наявність макросів, DDE-команд, автоматичних дій OpenAction у PDF), лічильники (кількість OLE-об'єктів, потоків, тегів, форм) та неперервні характеристики (ентропія вмісту, розмір файлу) [1]. Для таких даних ансамблі дерев рішень, побудовані методом градієнтного бустингу, є одним із найефективніших класів моделей: кожне наступне дерево навчається на градієнті функції втрат попереднього ансамблю, що дає змогу моделювати нелінійні залежності та взаємодії ознак без масштабування і складної попередньої обробки даних [2]. Додатковими перевагами для поштових шлюзів є висока швидкість класифікації та вбудована обробка пропущених значень, що виникають під час аналізу пошкоджених файлів.

За результатами відбору ознак, виконаного авторами набору, найінформативнішими для документів Word є наявність OLE-об'єктів, макросів VBA, DDE-команд і підозрілих ключових слів VBA (CreateObject, PowerShell) та ентропія вмісту; для Excel – ентропія тексту комірок, метрики обфускації коду макросів (зокрема кількість викликів Chr/ChrW) та посилання на віддалений шаблон; для PDF – кількість об'єктів, потоків і фільтрів кодування, ентропія потоків та коректність заголовка файлу; для HTML – кількість тегів, форм і вбудованих скриптів, ентропія коду JavaScript, характеристики посилань та частка пробільних символів [1]. Серед найінформативніших ознак HTML переважають структурні та статистичні характеристики, а не явні індикатори зловмисної поведінки, що може пояснювати нижчу якість класифікації цього формату та обґрунтовує доцільність пошуку ефективніших моделей саме для нього.

Запропонована технологія (рис. 1) передбачає такі етапи: вилучення вкладення з електронного листа; визначення формату файлу за сигнатурою, а не за розширенням, оскільки зловмисники часто маскують тип вкладення; статичне вилучення ознак спеціалізованим парсером без відкриття чи виконання файлу; класифікацію окремою моделлю градієнтного бустингу, навченою для відповідного формату; формування вердикту з оцінкою ймовірності шкідливості та поясненням рішення.

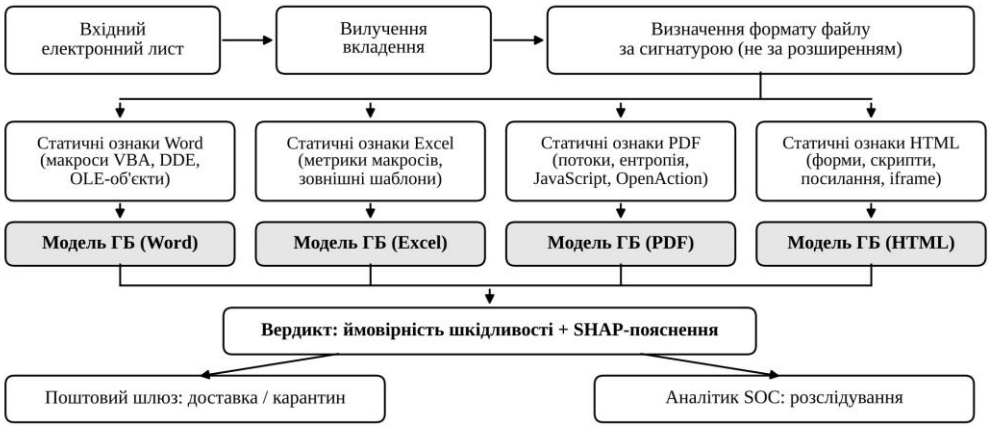


Рис. 1. Схема технології виявлення шкідливих вкладень

Розділення моделей за форматами обґрунтоване тим, що набори ознак для різних типів файлів не збігаються, а окремі моделі можна незалежно оновлювати й налаштовувати для кожної з них поріг ймовірності, за яким вкладення поміщається в карантин.

Для пояснення рішень застосовано метод SHAP (SHapley Additive exPlanations), який на основі значень Шеплі з теорії кооперативних ігор оцінює внесок кожної ознаки в конкретний прогноз моделі [3]. На відміну від авторів набору даних, які використовували SHAP лише для відбору ознак [1], у запропонованій технології пояснення формується для кожного вердикту: аналітик SOC отримує не лише мітку «шкідливе», а й перелік чинників, що її зумовили (наприклад, наявність DDE-команди, висока ентропія потоків PDF чи прихований iframe у HTML-сторінці). Це полегшує розслідування інцидентів і виявлення хибних спрацювань.

Експериментальну перевірку заплановано у три етапи. На першому етапі для HTML-вкладень, де базова F1-міра становить близько 0,94 [1], порівнюються реалізації XGBoost, LightGBM та CatBoost на повному наборі з 40 ознак і на скороченому наборі з 13 ознак, відібраних авторами набору, з підбором гіперпараметрів (кількість дерев, глибина, темп навчання) методом k-кратної перехресної перевірки; для запобігання завищенню оцінок дублікати зразків вилучаються до розділення даних на навчальну та тестову вибірки. На другому етапі для документів Word і Excel перевіряється гіпотеза про вплив ознак, пов'язаних із джерелом даних: легітимні Word-документи збережено у форматі .docx, який за специфікацією не підтримує макроси, тому ознака їх наявності може тривіально розділяти класи. Моделі навчаються після вилучення таких ознак, і оцінюється зміна якості класифікації. На третьому етапі аналізуються SHAP-пояснення для правильно та хибно класифікованих зразків. Для оцінювання використовуються точність, повнота, F1-міра та частка хибнопозитивних спрацювань, критична для поштових систем, де блокування легітимного листа порушує бізнес-процеси.

Таким чином, набір даних CIC-Trap4Phish 2025 створює основу для порівняльних досліджень виявлення шкідливих вкладень, проте високі базові результати для документів Office потребують критичної перевірки, а виявлення шкідливих HTML-вкладень залишається найменш вирішеною задачею. Запропонована технологія поєднує окремі моделі градієнтного бустингу для кожного формату з поясненням рішень на основі SHAP, що робить її придатною для інтеграції в поштові шлюзи та процеси SOC. Подальші дослідження передбачають проведення описаного експерименту та оцінку узагальнювальної здатності моделей на зразках, зібраних поза межами набору даних.

Список використаних джерел

1. Nejati F., Rabbani M., Eskandarian M. et al. CIC-Trap4Phish: A Unified Multi-Format Dataset for Phishing and Quishing Attachment Detection. arXiv:2602.09015. 2026. URL: https://arxiv.org/abs/2602.09015

2. Ke G., Meng Q., Finley T. et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017. P. 3146–3154.

3. Lundberg S. M., Lee S.-I. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017. P. 4765–4774.

4. Verizon. 2026 Data Breach Investigations Report. 2026. URL: https://www.verizon.com/business/resources/reports/dbir/

Савченко В.В., студент групи ІКБм-61
Науковий керівник: Борсуковський Ю.В.
Державний університет інформаційно-комунікаційних технологій, м. Київ

ТЕХНОЛОГІЯ ТЕСТУВАННЯ ЗАХИЩЕНОСТІ API ДЛЯ ВИЯВЛЕННЯ ВРАЗЛИВОСТЕЙ АВТЕНТИФІКАЦІЇ

Анотація. Розглянуто проблему виявлення вразливостей автентифікації та контролю доступу в API вебзастосунків, які складно виявити традиційними сканерами безпеки. Запропоновано технологію тестування захищеності API, що включає картування та розвідку API, тестування автентифікації та покрокову перевірку авторизації, а також модель оцінювання ризику виявлених вразливостей. Наведено результати експериментальної перевірки технології в лабораторному середовищі на базі OWASP crAPI та OWASP Juice Shop і рекомендації щодо підвищення захищеності API.

Стрімкий розвиток сучасних вебзастосунків та мікросервісних архітектур зумовив широке впровадження програмних інтерфейсів API (Application Programming Interface), зокрема REST, GraphQL та gRPC. Проте прискорені цикли розробки часто призводять до недоліків у перевірці прав доступу, роблячи API одним із головних векторів кібератак. Значна частка сучасних атак на вебзастосунки спрямована саме на програмні інтерфейси API через наявність критичних вразливостей у механізмах автентифікації та розмежування доступу.

Аналіз загроз безпеки показує, що найнебезпечнішими вразливостями API за класифікацією OWASP API Security Top 10 [1] є порушення авторизації на рівні об’єктів (BOLA/IDOR), несправна автентифікація (Broken Authentication) та порушення авторизації на рівні функцій (BFLA). Такі дефекти складно виявити традиційними сканерами безпеки та міжмережевими екранами (WAF), оскільки це вимагає глибшого розуміння логіки роботи вебзастосунку, структури JWT-токена та контексту запиту конкретного користувача.

Метою даної роботи є розроблення та практична апробація комплексної технології тестування захищеності API вебзастосунків, яка дозволяє системно виявляти вразливості автентифікації та контролю доступу на ранніх етапах аудиту безпеки. Запропонована технологія ґрунтується на методології OWASP Web Security Testing Guide [2] і поєднує автоматизоване сканування із комбінованим фаззингом та ручною перевіркою бізнес-логіки.

Технологія включає три основні етапи (рис. 1): 1) картування та розвідка API (аналіз специфікацій OpenAPI/Swagger, збір кінцевих точок); 2) автоматизоване та напівавтоматизоване тестування автентифікації (фаззинг токенів JWT/OAuth, перевірка алгоритмів підпису, аналіз стійкості сесій); 3) покрокове тестування авторизації та контролю доступу (перевірка BOLA, BFLA, маніпулювання ідентифікаторами користувачів та ролями).

Для кількісної оцінки критичності виявлених вразливостей в API пропонується використовувати адаптовану математичну модель розрахунку

ризику, яка базується на метриках CVSS, але враховує специфіку архітектури API. Загальний показник ризику вразливості R обчислюється за формулою:

$$R = (I_{conf} + I_{int} + I_{avail}) \cdot P_{exploit} \cdot W_{api}, \quad (1)$$

де I_{conf}, I_{int}, I_{avail} — вплив на конфіденційність, цілісність та доступність даних відповідно (від 0 до 10); $P_{exploit}$ — ймовірність успішної експлуатації вразливості через API (від 0 до 1); W_{api} — ваговий коефіцієнт критичності конкретного ендпоінту (наприклад, для платіжного шлюзу W_{api} = 1,5).

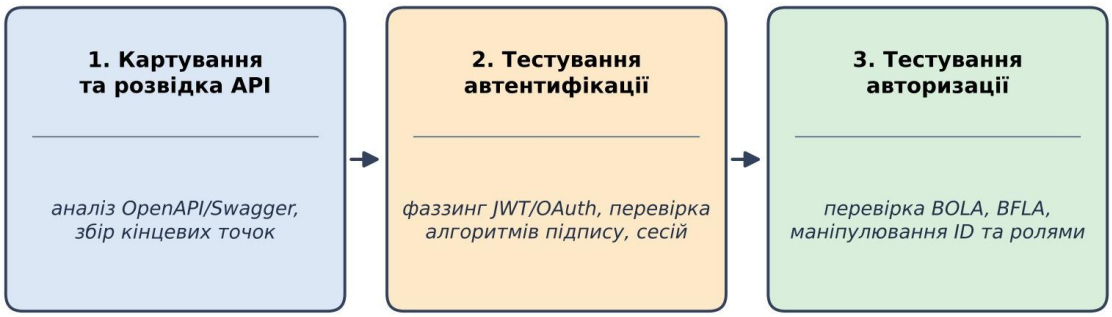


Рис. 1. Етапи комплексної технології тестування захищеності API вебзастосунків

Для практичної реалізації та експериментальної перевірки розробленої технології було розгорнуто ізольоване лабораторне середовище на базі спеціалізованих вразливих вебзастосунків OWASP crAPI (completely ridiculous API) [3] та OWASP Juice Shop. У ході проведення серії тестових атак із використанням інструментів Burp Suite Professional, OWASP ZAP, Postman та спеціалізованих Python-скриптів вдалося виявити критичні вразливості обходу авторизації, підробки підпису JWT-токенів та несанкціонованого доступу до персональних даних інших користувачів.

Окрему увагу було приділено перевірці токенів доступу (JWT): відсутність валідації алгоритму підпису (зміна RS256 на HS256 або на «none») дозволяє зловмиснику генерувати валідні токени для будь-якого користувача й отримувати адміністративний доступ до ресурсів API.

За результатами тестування розроблено рекомендації щодо підвищення захищеності API: контроль доступу на основі ролей та атрибутів (RBAC/ABAC) на кожному рівні обробки запиту; централізована перевірка підпису токенів із жорстко заданим переліком допустимих алгоритмів [4]; валідація вхідних даних для запобігання ін’єкціям; обмеження частоти запитів (Rate Limiting) для захисту від брутфорс-атак.

Таким чином, запропонована технологія дозволяє підвищити рівень безпеки сучасних вебзастосунків, а її впровадження в процес розробки та аудиту (DevSecOps) дає змогу усувати вразливості API до моменту їх експлуатації зловмисниками.

Література:

1. OWASP API Security Project. OWASP API Security Top 10 – 2023. URL: https://owasp.org/API-Security/ (дата звернення: 18.10.2026).
2. OWASP Foundation. Web Security Testing Guide (WSTG) v4.2. URL: https://owasp.org/www-project-web-security-testing-guide/ (дата звернення: 18.10.2026).
3. OWASP crAPI: completely ridiculous API. URL: https://github.com/OWASP/crAPI (дата звернення: 18.10.2026).
4. Sheffer Y., Hardt D., Jones M. JSON Web Token Best Current Practices. RFC 8725. IETF, 2020. URL: https://www.rfc-editor.org/rfc/rfc8725 (дата звернення: 18.10.2026).

Караєв Владислав Русланович,
студент, групи КНДм-61 ННІІТ ДУІКТ,
Київ, Україна
Науковий керівник Катков Юрій Ігорович,
д.т.н., доцент, ДУІКТ,
Київ, Україна

ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ОБРОБКИ ДАНИХ У ВЕБ-СИСТЕМАХ АВТОМАТИЗОВАНОГО ПІДБОРУ КОМПЛЕКТУЮЧИХ ПК

У роботі розглянуто забезпечення безпеки обробки даних у веб-системі автоматизованого підбору комплектуючих персональних комп'ютерів на базі Node.js, Express, Sequelize та MySQL. Визначено ризики, пов'язані з передачею некоректних даних через API, SQL-ін'єкціями та міжсайтовим скриптингом (XSS). Запропоновано багаторівневий захист серверної логіки: сувору валідацію вхідних параметрів за допомогою проміжних обробників (middleware), параметризацію запитів через ORM, обмеження частоти запитів, контроль доступу (CORS), шифрування з'єднань (HTTPS) і журналювання критичних помилок. Загрози зіставлено із засобами захисту.

Ключові слова: кібербезпека, веб-застосунок, валідація даних, SQL-ін'єкція, XSS, Node.js, Express, Sequelize, ORM, CORS, HTTPS, middleware.

Веб-сервіси електронної комерції та автоматизованого вибору техніки дедалі частіше містять інтелектуальні системи підтримки прийняття рішень, які за бюджетом і вимогами користувача формують конфігурацію персонального комп'ютера (ПК). Така система приймає від клієнта параметри запиту, обробляє їх на сервері та звертається до бази даних каталогу комплектуючих, тому серверна частина стає привабливою ціллю для зловмисників. За даними OWASP, ін'єкції, до яких віднесено й міжсайтовий скриптинг (XSS), залишаються серед десяти найкритичніших ризиків безпеки вебзастосунків [1]. Отже, в таких сервісах важливою є не лише зручність інтерфейсу, а й надійність обробки даних користувача, а забезпечення безпеки цієї обробки є актуальним завданням кібербезпеки.

Основні ризики під час роботи подібних застосунків пов'язані з передачею некоректних даних через API, спробами реалізації SQL-ін'єкцій або XSS під час формування запитів на підбір заліза. Від'ємне значення ціни, нечислове значення бюджету, сторонні символи у фільтрах характеристик чи лавина автоматизованих запитів можуть призвести до маніпуляції серверною логікою, несанкціонованого зчитування каталогу, перевантаження системних ресурсів і перехоплення даних у каналі між клієнтом і сервером. Практичні рекомендації щодо захисту вебзастосунків на Node.js та Express [2, 3] і вимоги до безпечної розробки програмного забезпечення [4] мають загальний характер, тому проблемою залишається їх поєднання в єдиний контур захисту конкретної системи підбору комплектуючих.

Метою роботи є підвищення безпеки обробки даних у веб-системі автоматизованого підбору комплектуючих ПК шляхом побудови багаторівневого захисту серверної логіки та каналу взаємодії між клієнтом і базою даних.

Для розв'язання проблеми в розробленій системі на базі Node.js та Express із базою даних MySQL застосовано послідовні рівні захисту, які проходить кожен запит (табл. 1). Підхід узгоджується з практиками безпечної розробки [2–4].

Перший рівень — сувора серверна валідація вхідних параметрів бюджету та характеристик компонентів. Її виконують спеціалізовані проміжні обробники (middleware): вони перевіряють JSON-структуру запиту, тип і діапазон значень та відхиляють запити, що не відповідають очікуваній схемі. Зокрема, від'ємні значення ціни й сторонні символи у фільтрах не потрапляють до логіки підбору. Перевірка виконується саме на сервері, оскільки обмеження на стороні клієнта користувач може обійти.

Другий рівень — використання ORM-бібліотеки Sequelize. Усі операції з базою даних MySQL параметризуються, що дозволяє повністю виключити пряме виконання SQL-запитів, сформованих із даних користувача. Це запобігає SQL-ін'єкціям і несанкціонованому зчитуванню каталогу заліза за умови, що «сирі» запити з підстановкою рядків не застосовуються.

Третій рівень — захист від масових автоматизованих запитів і журналювання. Обмеження кількості звернень від одного клієнта за визначений інтервал часу знижує ризик перевантаження системних ресурсів, а логування критичних помилок серверної частини дає змогу виявляти підозрілу активність і розслідувати інциденти.

Четвертий рівень — контроль доступу та захист каналу. Механізм CORS обмежує перелік зовнішніх джерел, яким дозволено звертатися до API, а шифрування з'єднань за протоколом HTTPS захищає дані, що передаються між клієнтом і сервером, від перехоплення та підміни. CORS є механізмом браузера й не замінює серверну валідацію, тому використовується разом із нею.

Типовий запит обробляється так: клієнт надсилає параметри підбору (бюджет, вимоги до компонентів) через HTTPS; сервер перевіряє джерело запиту (CORS) і частоту звернень; middleware перевіряє структуру та значення параметрів; лише після цього логіка підбору через Sequelize параметризовано звертається до MySQL; критичні помилки фіксуються в журналі. Якщо запит містить від'ємну ціну чи сторонні символи, його відхиляють ще до звернення до бази даних.

Для наочного подання результатів у таблиці 1 зіставлено основні загрози для веб-системи підбору комплектуючих із засобами захисту та компонентами системи, що їх реалізують. Перелік загроз сформовано на основі типових ризиків вебзастосунків [1] та рекомендацій щодо захисту застосунків на Express і Node.js [2, 3]. Таке зіставлення показує, що кожну з розглянутих загроз перекрито

принаймні одним рівнем захисту, а ризик ін’єкцій знижується одразу двома рівнями — валідацією вхідних даних і параметризацією запитів. (Табл. 1)

Таблиця 1

Загрози для веб-системи підбору комплектуючих і засоби захисту

Загроза	Засіб захисту	Компонент
Некоректні вхідні дані (від’ємна ціна, сторонні символи) [2, 3]	Серверна валідація JSON-структур, типів і діапазонів значень	Middleware (Express)
SQL-ін’єкція [1]	Параметризація запитів через ORM; відмова від прямих SQL-запитів	Sequelize, MySQL
Міжсайтовий скриптинг (XSS) [1]	Фільтрація вхідних даних; відхилення сторонніх символів	Middleware (Express)
Масові автоматизовані запити [2]	Обмеження частоти запитів від одного клієнта	Middleware (Express)
Несанкціоновані зовнішні запити [2]	Політика CORS	Express (CORS)
Перехоплення даних у каналі [2, 3]	Шифрування з’єднань за протоколом HTTPS	Транспортний рівень
Непомічені збої та спроби атак [4]	Журналювання критичних помилок сервера	Серверна частина

Отже, запропоновано багаторівневий підхід до забезпечення безпеки обробки даних у веб-системі автоматизованого підбору комплектуючих ПК, що поєднує серверну валідацію вхідних параметрів, параметризацію запитів через ORM, обмеження частоти запитів, контроль доступу за допомогою CORS, шифрування з’єднань HTTPS і журналювання помилок. Такий підхід сприяє збереженню цілісності логіки підбору конфігурацій, захищає системні ресурси від перевантаження та створює надійну основу для подальшого розширення функціоналу системи. Наступний етап — кількісне оцінювання ефективності захисту: тестування на проникнення, вимірювання частки відхилених некоректних запитів і поведінки системи під навантаженням (табл. 1).

Перелік використаної літератури:

1. OWASP Foundation. (2025) OWASP Top 10:2025. 2025. URL: https://owasp.org/Top10/2025/ (дата звернення: 25.10.2026).owasp
2. Express. Production (2024) Best Practices: Security. OpenJS Foundation. URL: https://expressjs.com/en/advanced/best-practice-security.html (дата звернення: 25.10.2026).expressjs
3. Goldberg Y. та ін. (2025) Node.js Best Practices: Security Best Practices. GitHub. URL: https://github.com/goldbergyoni/nodebestpractices#6-security-best-practices (дата звернення: 25.10.2026).github

4. Souppaya M., Scarfone K., Dodson D. (2022) Secure Software Development Framework (SSDF) Version 1.1: Recommendations for Mitigating the Risk of Software Vulnerabilities. Gaithersburg, MD : National Institute of Standards and Technology, 2022. 36 p. NIST Special Publication 800-218. DOI: https://doi.org/10.6028/NIST.SP.800-218. URL: https://csrc.nist.gov/pubs/sp/800/218/final (дата звернення: 25.10.2026).csrc.nist+1

Габорець О. А.,
доктор філософії, доцент, доцент кафедри оперативно-розшукової діяльності та інформаційної безпеки ННІПФПКП імені Е. О. Дідоренка,
Донецький державний університет внутрішніх справ,
Кропивницький, Україна

Хоменко О. М.,
курсант 2 курсу ННІПФПКП імені Е. О. Дідоренка, рядовий поліції,
Донецький державний університет внутрішніх справ
Кропивницький, Україна

ДЕСТРУКТИВНИЙ ІНФОРМАЦІЙНИЙ ВПЛИВ НА ДИТИНУ В ЦИФРОВОМУ СЕРЕДОВИЩІ: ПОНЯТТЯ, ОЗНАКИ ТА СУЧАСНІ ФОРМИ

Розглянуто міжособистісний деструктивний інформаційний вплив на дитину в цифровому середовищі. Уточнено робоче визначення та запропоновано матрицю відповідності між змістовими, комунікативними ознаками, використанням даних і спрямованістю впливу. Розмежовано приниження, сексуальну експлуатацію та спонукання до небезпечних дій; синтетичні матеріали визначено як технологічний засіб. Окреслено правові орієнтири оцінювання, вікові аспекти й обмеження запропонованої систематизації. Практичне значення полягає в упорядкуванні ознак для контекстного аналізу цифрової взаємодії та визначення потреби в захисті дитини.

Ключові слова: деструктивний інформаційний вплив, цифрове середовище, безпека дитини, кібербулінг, онлайн-грумінг, сексторшен, інформаційні маніпуляції, синтетичний контент.

Цифрова комунікація може використовуватися для приниження, сексуальної експлуатації та втягнення дитини в небезпечні дії. За звітом UNICEF 2026 року, майже кожна п’ята дитина віком 12–17 років, яка користується інтернетом, у 21 досліджуваній країні протягом року зазнала сексуальної експлуатації або насильства із застосуванням цифрових технологій [1]. Ці дані не характеризують Україну; у цій роботі національних репрезентативних оцінок поширеності досліджуваних форм не наведено.

Правовими орієнтирами є положення ст. 10 Закону України «Про охорону дитинства» та Загального коментаря № 25 Комітету ООН з прав дитини [2; 3]. Мета дослідження – конкретизувати ознаки міжособистісного деструктивного інформаційного впливу та пов’язати їх із формами небезпечної цифрової взаємодії. Елемент наукової новизни полягає в авторській матриці відповідності між рівнями ознак і функціональною спрямованістю впливу. Методами є аналіз та узагальнення наукових публікацій і документів. Масове поширення

шкідливого контенту й договірні ризики залишаються поза предметом дослідження.

О. Невельська-Гордєєва та В. Нечитайло пов’язують інформаційно-психологічний вплив із дією на свідомість і поведінку, а маніпуляцію розглядають як прихований вплив в інтересах маніпулятора [4]. У пропонованому підході «інформаційний» означає здійснення впливу через зміст повідомлень, образ співрозмовника чи контроль над відомостями; «деструктивний» – спрямованість на приниження гідності, сексуальну експлуатацію або нав’язування небезпечних дій. Отже, робоча категорія вужча за інформаційно-психологічний вплив загалом.

Деструктивний інформаційний вплив у межах цього дослідження – цілеспрямоване використання цифрових повідомлень або відомостей про дитину для її приниження, сексуальної експлуатації чи спонукання до небезпечних дій через обман, залякування або експлуатацію довіри. Для його виявлення необхідно встановити зв’язок між інформаційним засобом, способом тиску та конкретною загрозою; самого неприємного повідомлення недостатньо. Це аналітична категорія для зіставлення механізмів, а не самостійний склад правопорушення. Кібербулінг зберігає ознаки повторюваної агресії [5], грумінг – сексуально спрямованого контакту, сексторшен – шантажу інтимними матеріалами [6; 7]. Варіативність оцінок кібергрумінгу [8] потребує врахування методики дослідження. Модель 4Cs розрізняє ризики змісту, контактів, поведінки та договірних відносин [9]. Ризик означає можливість негативного наслідку, а шкода – фактично заподіяні негативні зміни; сам контакт не доводить їх настання.

За функціональною спрямованістю розрізняємо приниження, сексуальну експлуатацію та нав’язування небезпечних дій. Матрицю сформовано шляхом тематичного узагальнення описів кібербулінгу [5], грумінгу й сексторшену [6; 7] та механізму обману й примушування у повідомленні СБУ [10]. Описані дії згруповано за змістом повідомлень, способом взаємодії та використанням відомостей про дитину. Формулювання клітинок є авторськими індикаторами, а не дослівними положеннями джерел; вони можуть перетинатися й не доводять правопорушення.

Таблиця 1. Відповідність рівнів ознак і спрямованості впливу

Рівень ознак	***Приниження***	***Сексуальна експлуатація***	***Небезпечні дії***
Зміст	Образи; погрози осоромлення	Сексуалізовані вимоги; погрози	Небезпечні доручення; залякування
Комунікація	Повторюваний тиск; групове виключення	Удавана близькість; секретність	Удаваний авторитет; обіцянки вигоди

Рівень ознак	*Приниження*	*Сексуальна експлуатація*	*Небезпечні дії*
Використання даних	Оприлюднення приватних відомостей	Контроль над інтимними матеріалами	Збір геолокації; погрози викриття

Синтетичні зображення, відео й аудіозаписи є технологічним засобом у кожній із трьох груп. Їх використання оцінюється за відповідними клітинками матриці: підміна співрозмовника – як обман, погроза поширення підробленого матеріалу – як контроль над даними. Ознаки повторюваності та багатоканальності наративів [11] тут запозичено як аналітичну схему для вивчення послідовності повідомлень. Це не означає підтвердження її ефективності в аналізі комунікації з дітьми.

Вік у методичному аналізі слід враховувати разом зі здатністю дитини розпізнавати обман, розуміти наслідки передавання даних і звертатися по допомогу [3]. Жорсткі вікові межі для механізмів впливу не встановлюємо: удаваний авторитет, тиск групи й репутаційні погрози перевіряються з урахуванням індивідуального розвитку. Правові вікові пороги застосовуються окремо.

Правова оцінка проводиться окремо. Стаття 173-4 КУпАП стосується булінгу учасника освітнього процесу; цькування поза відповідними відносинами потребує іншої оцінки. Особиста адміністративна відповідальність настає з 16 років (ст. 12 КУпАП). Діти до 14 років її не несуть; спеціальні положення ч. 3–4 ст. 173-4 передбачають відповідальність батьків або осіб, які їх замінюють, за діяння малолітніх, а також неповнолітніх віком від 14 до 16 років [12]. Сексуальний шантаж не охоплюється автоматично ст. 156-1 КК: вимога грошей під погрозою розголошення таємних відомостей може оцінюватися за ст. 189, а примушування особи без її добровільної згоди до здійснення акту сексуального характеру з іншою особою – за ст. 154 за наявності всіх відповідних ознак. Визначення дитячої порнографії міститься у примітці до ст. 156-1 і застосовується також для ст. 301-1; воно охоплює змодельовані образи та поведінку [13]. Генерація зображення сама по собі не визначає кваліфікації.

Стаття 304 КК України стосується втягнення неповнолітніх у протиправну діяльність [13]. Повідомлення СБУ від 8 жовтня 2026 року про фейкові акаунти та шантаж підлітків [10] використовуємо лише як ілюстрацію міжособистісного механізму обману й примушування у третій групі, без прирівнювання вербування до грумінгу. За цим повідомленням, 17-річному оголошено підозру за ч. 2, 3 ст. 15 та ст. 194 КК. Зазнання маніпулятивного впливу не надає автоматично статусу потерпілого і не виключає відповідальності за власне діяння; кожний процесуальний статус та умови відповідальності встановлюються окремо.

Практичне значення матриці полягає у встановленні того, яке повідомлення, спосіб тиску або використання даних створює конкретну загрозу

дитині. Висновок про вплив потребує оцінювання контексту, віку й послідовності взаємодії. Подальша перевірка матриці має спиратися на українські емпіричні дані. Пріоритетами реагування залишаються безпека дитини, збереження приватності та доступ до підтримки.

Список використаних джерел

1. UNICEF. Through Children's Eyes: How digital technologies enable child sexual abuse. 2026. URL: https://www.unicef.org/innocenti/reports/through-childrens-eyes (дата звернення: 21.10.2026).
2. Про охорону дитинства : Закон України від 26.04.2001 № 2402-III. URL: https://data.rada.gov.ua/ogd/zak/perv/text/d82435.htm (дата звернення: 15.10.2026).
3. Committee on the Rights of the Child. General comment No. 25 (2021) on children's rights in relation to the digital environment. CRC/C/GC/25. 02.03.2021. URL: https://docs.un.org/en/CRC/C/GC/25 (дата звернення: 21.10.2026).
4. Невельська-Гордєєва О. П., Нечитайло В. О. Маніпуляції як засіб інформаційно-психологічного впливу у інформаційній війні. *Вісник НЮУ імені Ярослава Мудрого. Серія: Філософія, філософія права, політологія, соціологія*. 2021. № 3 (50). С. 71–83. DOI: https://doi.org/10.21564/2663-5704.50.235389
5. Бедан В. Б. Проблема кібербулінгу у молодіжному середовищі: теоретичний аналіз. *Дніпровський науковий часопис публічного управління, психології, права.* 2022. № 1. С. 42–47. DOI: https://doi.org/10.51547/ppp.dp.ua/2022.1.6.
6. Копйова І. А. Кримінально-правова кваліфікація дитячого секстингу, сексторшену і онлайн-грумінгу. *Вісник Асоціації кримінального права України*. 2021. № 2 (16). С. 209–222. DOI: https://doi.org/10.21564/2311-9640.2021.16.244349.
7. Алексійчук О. Домагання дітей для сексуальних цілей як об'єкт кримінально-правового регулювання та криміналістичного дослідження. *Наука і правоохорона*. 2025. Вип. 4 (70). С. 61–70. DOI: https://doi.org/10.33270/05257004.6
8. Schittenhelm C., Kops M., Moosburner M., Fischer S. M., Wachs S. Cybergrooming Victimization Among Young People: A Systematic Review of Prevalence Rates, Risk Factors, and Outcomes. *Adolescent Research Review.* 2025. Vol. 10. P. 169–200. DOI: https://doi.org/10.1007/s40894-024-00248-w
9. Livingstone S., Stoilova M. The 4Cs: Classifying Online Risk to Children. CO:RE Short Report Series on Key Topics. Hamburg : Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI) ; CO:RE – Children Online: Research and Evidence, 2021. 14 p. DOI: https://doi.org/10.21241/ssoar.71817
10. Служба безпеки України. СБУ викрила спробу рф вчинити масштабну диверсію на Київщині руками неповнолітніх (відео). 08.10.2026. URL: https://www.ssu.gov.ua/novyny/sbu-vykryla-sprobu-rf-vchynyty-masshtabnu-dyversiiu-na-kyivshchyni-rukamy-nepovnolitnikh-video (дата звернення: 12.10.2026).
11. Габорець О. А. Аналіз деструктивних інформаційних наративів у цифровому просторі як складова кримінального аналізу інформаційних загроз. *Кібербезпека: освіта, наука, техніка*. 2026. № 1 (33). С. 320–329. DOI: https://doi.org/10.28925/2663-4023.2026.33.1178
12. Кодекс України про адміністративні правопорушення від 07.12.1984 № 8073-X. URL: https://zakon.rada.gov.ua/laws/show/8073-10 (дата звернення: 12.10.2026).
13. Кримінальний кодекс України від 05.04.2001 № 2341-III. URL: https://zakon.rada.gov.ua/laws/show/2341-14 (дата звернення: 12.10.2026).

Дубик Артур Русланович,
студент, групи КНДм-61 ННІІТ ДУІКТ,
Київ, Україна
Науковий керівник Катков Юрій Ігорович,
д.т.н., доцент, ДУІКТ,
Київ, Україна

АРХІТЕКТУРНА МОДЕЛЬ ЗАХИЩЕНОЇ ІНТЕГРАЦІЇ AI-АГЕНТІВ У КОРПОРАТИВНІ ІНФОРМАЦІЙНІ СИСТЕМИ НА ЗАСАДАХ ZERO TRUST

У роботі розглянуто безпечну інтеграцію AI-агентів у корпоративні інформаційні системи (КІС) в умовах зростання кількості машинних ідентичностей. Запропоновано референсну архітектуру на основі Zero Trust (NIST SP 800-207) з окремим шлюзом для агентів (agent gateway), мікросегментацією, безперервним моніторингом і автоматизованим реагуванням. Загрози для AI-агентів — ін'єкцію запитів, надмірні повноваження, викрадення токенів, поширення атак мережею та витік даних — зіставлено із засобами захисту, що обмежують повноваження агентів і наслідки компрометації.

Ключові слова: інформаційна безпека, Zero Trust, AI-агенти, штучний інтелект, корпоративні інформаційні системи, кіберзахист, agent gateway, машинні ідентичності, prompt injection, мікросегментація.

В умовах зростання кіберзагроз та активного впровадження AI-агентів у корпоративні інформаційні системи актуальним є розроблення референсної архітектури їх безпечної інтеграції спираючись на концепцію нульової довіри (Zero Trust) з метою забезпечення високого рівня захищеності цифрових ресурсів.

Корпоративні інформаційні системи (КІС) — це програми, якими користується організація: ERP, CRM, хмарні сервіси. Їх дедалі частіше з'єднують через API та доповнюють інтелектуальною автоматизацією: від RPA-ботів до AI-агентів — програм на основі штучного інтелекту, які самостійно виконують дії в інших системах. Тому доступ до даних мають не лише люди, а й машинні ідентичності — цифрові «посвідчення» сервісів, ботів і агентів. Периметровий захист значною мірою спирається на довіру до внутрішньої мережі, чого для агентів недостатньо. По-перше, агента можна ввести в оману шкідливими інструкціями в документі чи на вебсторінці (опосередкована підміна інструкцій є критичною вразливістю у сферах безпеки штучного інтелекту) [1, 2]. По-друге, агент із надмірними повноваженнями може виконати дію, недоступну самому зловмиснику (проблема «confused deputy») [3]. Крім того, VPN і великі мережеві сегменти часто надають доступ до мережі загалом, а не до окремих ресурсів, і самі по собі не враховують контексту запиту. Тому базові моделі Zero Trust [4] потребують адаптації до агентів: делегування повноважень, контролю викликів інструментів і обмеження наслідків ін'єкції запитів. У межах розглянутих джерел недостатньо конкретизовано, як поєднати ідентичність агента, делегування повноважень і контроль кожного виклику інструмента в єдиному контурі Zero Trust.

Для розв’язання окресленої проблеми необхідно розробити концептуальну референсну архітектуру Zero Trust для безпечної інтеграції AI-агентів із корпоративними інформаційними системами та встановити відповідність між основними загрозами агентній взаємодії й засобами їх пом’якшення.

Запропонована референсна (еталонна) архітектура безпечної автоматизації та інтеграції КІС на основі Zero Trust («нульова довіра», NIST SP 800-207 [4]), за якою кожен запит перевіряється незалежно від джерела. До базової моделі додається окремий шлюз для агентів (agent gateway), а загрози для AI-агентів зіставляються з конкретними засобами захисту. Архітектура спирається на три принципи Zero Trust: не довіряти за замовчуванням, надавати дозвіл лише в обсязі, необхідному для виконання завдань (least privilege) і виходити з того, що компрометація можлива (assume breach).(Рис. 1):

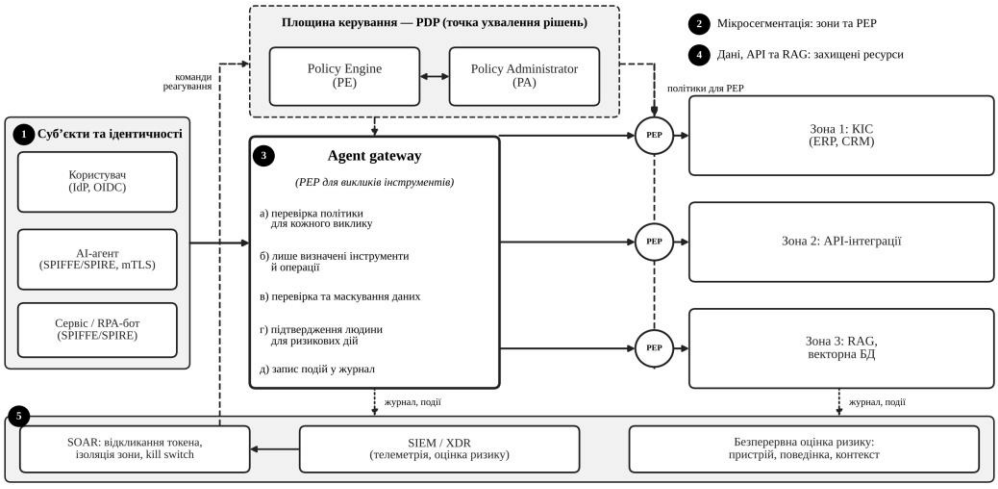


Рис. 1. Референсна архітектура безпечної інтеграції AI-агентів у КІС на основі Zero Trust

На рис. 1 наведено п’ять функціональних груп архітектури: суб’єкти та ідентичності; мікросегментація й PEP; шлюз агентів; захищені дані, API та RAG; моніторинг і реагування. PDP показано окремо як площину керування, що взаємодіє з PEP. Policy Engine (PE) формує рішення щодо доступу, а Policy Administrator (PA) реалізує його в площині керування; разом PE і PA утворюють PDP. Policy Enforcement Point (PEP) забезпечує виконання політики для контрольованого запиту.

Користувачі автентифікуються через IdP (OIDC), а сервіси й агенти отримують власну криптографічну ідентичність (SPIFFE/SPIRE, mTLS) і короткострокові токени. Повноваження агента обмежуються правами користувача, дозволами типу агента й умовами задачі, а остаточний доступ визначає політика авторизації. OAuth 2.0 Token Exchange (RFC 8693) дає змогу обмінювати токени для делегування повноважень; токени захищаються перевіркою призначення, коротким строком дії, прив’язкою до клієнта (DPoP, mTLS) і відкликанням. Мережу й застосунки поділяють на невеликі ізольовані

зони, щоб обмежити «радіус ураження»; взаємодія між зонами проходить лише через PEP, який виконує рішення PDP [4].

Агенти взаємодіють із системами лише через шлюз, що виконує функції PEP для викликів інструментів. Він перевіряє політику для кожного виклику; дозволяє тільки визначені інструменти й операції; перевіряє та маскує дані; запитує підтвердження людини для ризикових дій; записує події в журнал. Шлюз не запобігає самій ін'єкції запитів, але зменшує її наслідки.

Дані шифруються, чутлива інформація маскується, кожна API-інтеграція є окремим ресурсом з явною авторизацією. У RAG-системах (генерації відповідей із пошуком у базі документів) доступ до документів визначається правами користувача, від імені якого виконується запит; векторні бази розміщуються в окремій зоні. Ризик (з урахуванням стану пристрою) повторно оцінюється протягом сеансу; дії всіх ідентичностей відстежують SIEM/XDR, а SOAR може відкликати токен, ізолювати зону або зупинити агента (kill switch).

Типовий виклик виглядає так: користувач ставить задачу агенту; агент через Token Exchange отримує вузькоспрямований короткостроковий токен; шлюз зіставляє запит із рішенням PDP і лише тоді пропускає його до потрібної зони. Якщо агент прочитав шкідливу інструкцію в документі й спробував викликати інструмент, якого немає в дозволеному переліку, шлюз заблокує виклик, а подія потрапить до SIEM/XDR. Архітектура узгоджується з напрямами моделі CISA Zero Trust Maturity Model (ідентичності, мережі, застосунки, дані, аналітика й автоматизація); стан пристроїв ураховується через оцінювання ризику.

Таблиця 1

Загрози для AI-агентів і засоби захисту

Загроза	Засіб захисту	Компонент
Ін'єкція запитів (шкідливі інструкції) [1, 2]	Дозволені лише перелічені інструменти й операції; перевірка даних; підтвердження людиною ризикових дій	Agent gateway
Надмірні повноваження агента (confused deputy) [1, 3]	Повноваження звужено до задачі та прав користувача (Token Exchange); короткоживучі токени	Ідентичності
Викрадення токена агента	Прив'язка токена до клієнта (DPoP, mTLS); короткий строк дії; відкликання через SOAR	Ідентичності; SOAR
Поширення атаки мережею	Поділ на ізольовані зони; доступ лише через PEP	Мікросегментація
Витік даних через RAG або відповіді агента [1]	Права доступу до документів; маскування чутливих даних	Дані; шлюз
Непомічена підозріла поведінка	Безперервна оцінка ризику; моніторинг; kill switch	Перевірка та реагування

Отже, розроблено концептуальну архітектуру Zero Trust для інтеграції AI-агентів із КІС, що передбачає окрему ідентичність агента, обмежене делегування повноважень, контроль викликів інструментів через шлюз і сегментацію ресурсів. Ефективність архітектури експериментально не оцінювалася;

наступний етап — прототипування та вимірювання покриття сценаріїв атак, радіуса ураження й часу реагування. (рис. 1, табл. 1).

Перелік використаної літератури:

1. OWASP GenAI Security Project. OWASP GenAI LLM Top 10 2026. OWASP Foundation, 2026. URL: https://genai.owasp.org/resource/owasp-genai-llm-top-10-2026/.

2. Greshake K. та ін. Not What You've Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. arXiv, 2023. URL: https://arxiv.org/abs/2302.12173.

3. Hardy N. The Confused Deputy (or why capabilities might have been invented). ACM SIGOPS Operating Systems Review, 1988. Vol. 22, No. 4. P. 36–38.

4. Rose S. та ін. Zero Trust Architecture: NIST Special Publication 800-207. NIST, 2020. DOI: https://doi.org/10.6028/NIST.SP.800-207.

Михайленко Юлія Іванівна
Студентка групи ІКБм-62, ННІКБЗІ ДУІКТ,
Київ, Україна

ПІДХІД ДО ВИЯВЛЕННЯ DDOS SYN FLOOD-АТАК У МЕРЕЖАХ ІНТЕРНЕТУ РЕЧЕЙ НА ОСНОВІ АЛГОРИТМУ XGBOOST

Стрімкий розвиток Інтернету речей (IoT) супроводжується зростанням кількості мережевих пристроїв і підвищенням ризику реалізації кібератак. Однією з актуальних проблем є своєчасне виявлення шкідливої мережевої активності, зокрема розподілених атак на відмову в обслуговуванні. Перспективним напрямом розв'язання цієї проблеми є застосування алгоритмів машинного навчання, здатних автоматично аналізувати характеристики мережевого трафіку.

Ключові слова: Інтернет речей, IoT, кібербезпека, мережеві атаки, машинне навчання, XGBoost, DDoS, SYN Flood, CICIoT2023.

Розвиток Інтернету речей (IoT) супроводжується збільшенням кількості взаємопов’язаних пристроїв і підвищенням вимог до захисту мережевої інфраструктури. Національний інститут стандартів і технологій США (NIST) визначає базові можливості кіберзахисту IoT-пристроїв, зокрема керування конфігурацією, контроль доступу, захист даних та оновлення програмного забезпечення [1].

Однією з актуальних загроз є розподілені атаки на відмову в обслуговуванні (DDoS), зокрема TCP SYN Flood. Їхній механізм передбачає інтенсивне надсилання SYN-запитів, що може спричиняти виснаження ресурсів цільової системи та порушення доступності мережевих сервісів. Це зумовлює необхідність дослідження ефективних методів автоматизованого виявлення таких атак на основі аналізу статистичних і протокольних характеристик мережевого трафіку.

Одним із перспективних підходів до розв’язання зазначеної задачі є застосування алгоритму XGBoost (Extreme Gradient Boosting), який ґрунтується на градієнтному бустингу дерев рішень. Алгоритм формує ансамбль дерев, кожне наступне з яких коригує попередні прогнози, а їхнє об’єднання забезпечує побудову класифікаційної моделі [2].

Використання XGBoost дозволяє досліджувати можливість розпізнавання DDoS SYN Flood-атак за сукупністю характеристик мережевого трафіку. Схематичне представлення принципу роботи алгоритму наведено на рис. 1.

Мета роботи — дослідження підходу до автоматизованого виявлення DDoS SYN Flood-атак у мережах Інтернету речей на основі алгоритму XGBoost шляхом аналізу характеристик мережевого трафіку та експериментального оцінювання результатів класифікації.

.

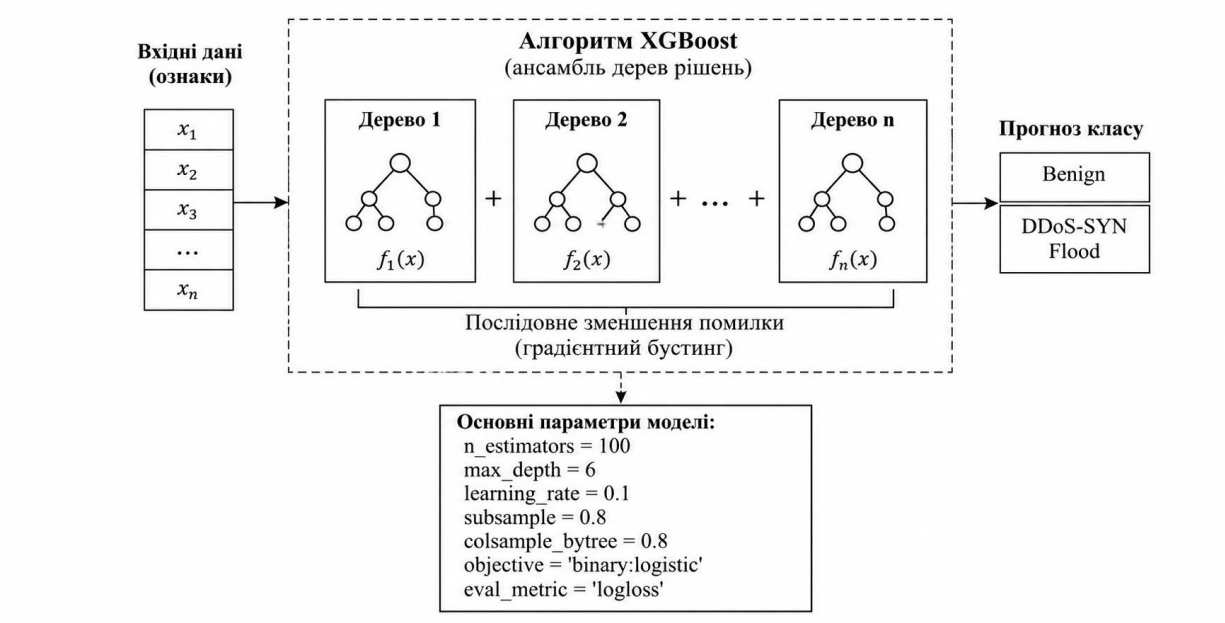


Рис. 1. Схема роботи алгоритму XGBoost на основі ансамблю дерев рішень

Для дослідження підходу до виявлення DDoS SYN Flood-атак у мережах Інтернету речей використано набір даних CICIoT2023, розроблений Canadian Institute for Cybersecurity при Університеті Нью-Брансвіка [3]. На його основі сформовано вибірку мережевого трафіку для бінарної класифікації двох класів: Benign, що відповідає нормальній мережевій активності, та DDoS-SYN Flood, що характеризує трафік досліджуваного типу атаки.

Розглянутий підхід ґрунтується на аналізі характеристик мережевого трафіку із застосуванням алгоритму XGBoost. Його реалізація передбачає формування вибірки, попередню обробку даних, розподіл на навчальну й тестову частини, навчання класифікаційної моделі та оцінювання результатів виявлення атак. Структурну схему досліджуваного підходу наведено на рис. 2.



Рис.2. Структурна схема підходу до виявлення DDoS SYN Flood-атак на основі алгоритму XGBoost

У початковому експерименті використано 628 385 записів, із них 362 361 — клас Benign і 266 024 — клас DDoS-SYN Flood. Програмну реалізацію здійснено мовою Python у середовищі Google Colab. Під час підготовки даних нескінченні значення замінено на пропуски з подальшим заповненням нулями. Вибірку розділено у співвідношенні 80 % для навчання та 20 % для тестування. Для моделі XGBoost задано 100 дерев рішень, максимальну глибину 6 і швидкість навчання 0,1.

За результатами початкового експерименту модель XGBoost продемонструвала точність класифікації 100 % на тестовій вибірці, що містила 125 677 записів. Згідно з матрицею помилок, правильно класифіковано 72 472 записи нормального трафіку (Benign) та 53 205 записів DDoS-SYN Flood-атак. Хибнопозитивних і хибнонегативних результатів не зафіксовано.

Для перевірки впливу окремих характеристик мережевого трафіку проведено додаткові експерименти з послідовним виключенням ознак Number та syn_count. Після повторного навчання точність класифікації залишилася на рівні 100 %. Найбільш важливою для моделі виявилася ознака syn_flag_number, пов’язана з активністю TCP SYN-пакетів.

Отримані результати свідчать про здатність алгоритму XGBoost розрізняти нормальний трафік і DDoS SYN Flood-атаки в межах сформованої вибірки. Водночас дослідження обмежується бінарною класифікацією на наборі даних CICIoT2023. Подальша робота передбачає перевірку моделі на незалежних даних, розширення кількості класів атак і порівняння з іншими алгоритмами машинного навчання.

Перелік посилань:

1. Fagan M., Megas K., Scarfone K., Smith M. IoT Device Cybersecurity Capability Core Baseline. NIST IR 8259A. 2020. DOI: https://doi.org/10.6028/NIST.IR.8259A

2. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of KDD 2016. P. 785–794. DOI: https://doi.org/10.1145/2939672.2939785

3. Canadian Institute for Cybersecurity. CICIoT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment. University of New Brunswick. URL: https://www.unb.ca/cic/datasets/iotdataset-2023.html

Хоменко Микита Сергійович
студент групи ІКБм-61, ННІЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАХИЩЕНОГО ОБМІНУ ДАНИМИ НА БАЗІ ПРОТОКОЛУ SSL/TLS

Розглянуто, як протокол TLS забезпечує конфіденційність, цілісність і автентичність даних під час передавання мережею. Пояснено, чим TLS 1.3 відрізняється від попередніх версій і які вразливості та помилки конфігурації найчастіше знижують захищеність з'єднань. Окрему увагу приділено практиці: навіть сучасний протокол не захищає, якщо сервер налаштовано неправильно. Сформульовано мету дослідження, яке можна виконати на безкоштовному програмному забезпеченні в умовах ізольованого стенду.

Дані в інформаційних системах постійно передаються мережею, де їх можна перехопити або змінити. Тому для захисту каналу між клієнтом і сервером використовують протокол TLS. Він шифрує трафік, перевіряє його цілісність і дає змогу переконатися, що клієнт спілкується саме з тим сервером, який йому потрібен. Остання версія протоколу, TLS 1.3, визначена в RFC 8446[1].

Настанови NIST SP 800-52 Rev. 2 вимагають від державних серверів і клієнтів підтримки TLS 1.2 з шифронаборами на основі FIPS-алгоритмів, а також вимагають підтримки TLS 1.3 з 1 січня 2024 року[2]. Водночас NIST переглядає цей документ і у 2026 році збирав коментарі щодо його оновлення, зокрема щодо того, чи й надалі рекомендувати серверам підтримку TLS 1.2. Однак на практиці багато систем досі працюють зі старими версіями та слабкими шифронаборами заради сумісності.

Історія протоколу показує, що проблеми виникають і в самому протоколі, і в його реалізаціях. Огляд відомих атак на TLS і DTLS, зокрема SSL stripping, атак на padding oracle, RC4 та стиснення (CRIME, BREACH), наведено в RFC 7457[3]. Атака POODLE використала те, що браузери повторюють невдалі з'єднання зі старішими версіями протоколу, аж до SSL 3.0, тож мережевий зловмисник може примусити клієнта знизити версію й відновити відкритий текст. Вразливість Heartbleed у бібліотеці OpenSSL дозволяла зчитувати пам'ять сервера, а її наслідки та реакцію адміністраторів дослідники вивчали за допомогою масових сканувань інтернету. Атака Sweet32 показала, що шифри з 64-бітним блоком, зокрема 3DES, вразливі до колізійних атак, бо межа безпечного обсягу даних для них становить лише близько 32 ГБ. Тому застарілі протоколи й шифри потрібно вимикати, а не лишати «про запас».

Окрема проблема перехоплення трафіку. За оцінкою дослідників, перехоплюється приблизно 5-10% HTTPS з'єднань, а проміжні пристрої й антивірусне забезпечення, які це роблять, майже завжди знижують захищеність з'єднання[4]. Особливо небезпечно, коли користувачі звикають ігнорувати

попередження браузера про недовірений або самопідписаний сертифікат, адже саме це відкриває шлях до атаки «людина посередині» (MITM).

Для оцінювання захищеності TLS-з'єднання перевіряють версії протоколу, набір шифрів, коректність ланцюга сертифікатів[2] і наявність механізмів на кшталт HSTS. Ці перевірки можна виконувати безкоштовними засобами (testssl.sh, sslyze, Nmap, Wireshark), що робить дослідження доступним без дорогого обладнання та ліцензій.

Мета роботи - дослідити вплив конфігурації сервера на захищеність TLS-з'єднання та розробити рекомендації щодо його посилення. Складність у тому, що вразливості потрібно відтворювати лише на власному ізольованому стенді, а результати перевіряти порівнянням стану «до» і «після» посилення.

Перелік посилань:

1. RFC 8446: The Transport Layer Security (TLS) Protocol Version 1.3 | RFC Editor. RFC Editor. URL: [1] (date of access: 28.9.2026).
2. SP 800-52 Rev. 2, Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations | CSRC. NIST Computer Security Resource Center | CSRC. URL: [2] (date of access: 30.9.2026).
3. RFC 7457: Summarizing Known Attacks on Transport Layer Security (TLS) and Datagram TLS (DTLS) | RFC Editor. RFC Editor. URL: [4] (date of access: 4.10.2026).
4. The Security Impact of HTTPS Interception - NDSS Symposium. NDSS Symposium. URL: [8] (date of access: 19.10.2026).

Власюк Ігор Дмитрович
аспірант групи КБа-31
Національний університет «Львівська політехніка»
Львів, Україна

КРИТЕРІЇ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ СИСТЕМ ЖУРНАЛІЗАЦІЇ КІБЕРІНЦИДЕНТІВ

Системи журналізації кіберінцидентів забезпечують збирання та обробку даних, що необхідні для виявлення й розслідування цих інцидентів. Але лише висока швидкість роботи цих систем не дає гарантії збереження усіх важливих записів або своєчасного виявлення загроз. В роботі розглянуто критерії оцінювання ефективності систем журналізації та запропоновано підхід до порівняння конфігурацій з урахуванням втрат даних, затримок оброблення, повноти виявлення кіберінцидентів й кількості хибних спрацювань. Використовуючи умовний приклад, продемонстровано, що вибір конфігурації потребує одночасного врахування набору показників відповідно до вимог операційного центру безпеки.

Ключові слова: журналізація, кіберінцидент, SOC, критерії оцінювання, відтворюваність.

Журнали подій є інформаційною основою виявлення та розслідування кіберінцидентів в операційному центрі безпеки (Security Operations Center, SOC). Проте висока швидкість надходження записів сама по собі не характеризує їх придатності для аналізу, а правильне виділення полів не гарантує виявлення атаки. Порівняння систем за одним показником може приховувати втрати даних або неприйнятне навантаження на аналітика. Тому надання об’єктивної оцінки потребує узгодження показників із функціями системи та умовами її застосування.

Стан і напрями розвитку технологій журналізації розглянуто в огляді [1]. Дослідження впливу парсингу показує, що якість структурного аналізу журналів не є універсальним показником результативності детектора аномалій [2]. У препринті [3] досліджено збереження ознак атак під час зведення записів до стандартних схем. Порівняння методів виявлення аномалій також засвідчує необхідність врахування обчислювальних витрат поряд із якістю класифікації [4]. Отже, оцінювання має охоплювати як окремі компоненти, так і кінцевий результат оброблення журналів.

Метою роботи є визначення критеріїв і порядку порівняння систем журналізації для конкретного сценарію роботи SOC. Запропонований підхід використовує три складові: відповідність показника функції системи, незалежний еталон для його розрахунку та відповідність конфігурації обов’язковим вимогам. Об’єктивність розглядається як перевірюваність оцінки за явно встановлених умов. Універсальність порогових значень чи рейтингу програмних продуктів не передбачається.

Функціональні рівні оцінювання наведено в таблиці 1. Для кожного показника необхідно визначити одиницю обліку, контрольну точку та часовий інтервал. Еталон формують незалежно від результатів оцінюваного компонента. За відсутності таких даних відповідну характеристику системи не можна вважати перевіреною.

У межах роботи система журналізації охоплює збирання, структурне оброблення та передавання записів для подальшого виявлення ознак кіберінцидентів. Запис журналу не ототожнюється з підтвердженим кіберінцидентом. Вимоги до строків зберігання, цілісності архіву й контролю доступу розглядаються окремо й потребують відповідних сценаріїв перевірки.

Таблиця 1. Матриця функціонального оцінювання

Рівень	Показники	Еталон та інтерпретація
Збирання	Частка доставлених записів; кількість дублікатів; затримка доставлення	Реєстр початкових записів. Успішне доставлення не гарантує збереження даних після збою.
Парсинг	Частка правильно розпізнаних шаблонів і виділених полів	Незалежна ручна розмітка. Коректність парсингу не гарантує виявлення атак.
Виявлення	Точність спрацювань; повнота виявлення; кількість хибних тривог за годину	Мітки цільових подій. Результати стосуються лише перевіреного сценарію.

Частка втрат показує, яка частина початкових унікальних записів не надійшла до контрольної точки за встановлений час. Дублікати враховують окремо. Затримку доставлення описують медіаною та 95-м процентилем, зазначаючи спосіб синхронізації годинників. Її оцінюють разом із втратами: виключення записів із великою затримкою штучно покращує результат.

На рівні парсингу еталонні дані містять очікувані шаблони та значення полів, необхідних для аналізу. Після нормалізації перевіряють збереження цих значень. На рівні виявлення точність позитивних спрацювань (precision) визначають як частку правильних спрацювань серед усіх позитивних спрацювань, а повноту (recall) — як частку виявлених позитивних випадків серед усіх позитивних випадків за еталонною розміткою [4]. Одиницею оцінювання може бути окремий запис, сеанс або часове вікно. Під час порівняння систем вона має залишатись незмінною.

Спочатку визначають сценарій роботи системи й допустимі межі втрат, затримки, кількості хибних тривог і витрат ресурсів. Окремо задають мінімальну повноту виявлення, щоб зменшення кількості тривог не досягалося шляхом

пропуску загроз. Системи порівнюють за однакового вхідного потоку та апаратного середовища. У протоколі фіксують версії програм, налаштування, розміри записів, інтенсивність потоку та тривалість вимірювання.

Навчальну, валідаційну й тестову вибірки розділяють до налаштування моделі; поріг спрацювання обирають за валідаційними даними. Вимірювання охоплюють прогрівання, стабільне й пікове навантаження та спорожнення черг. Фіксують кількість повторів, розкид результатів, навантаження на процесор і використання пам’яті.

Для інтегрованої системи додатково оцінюють повний шлях від початкового запису до формування тривоги. Позитивні випадки, не виявлені через втрату записів на попередніх етапах оброблення, враховують як пропущені. Еталонна розмітка має відповідати меті перевірки: виявлення системних збоїв не підтверджує здатності розпізнавати кібератаки. Кількість хибних тривог за годину є одним із показників навантаження на аналітиків SOC, проте не тотожна частці хибнопозитивних рішень серед усіх негативних прикладів.

Умовний розрахунковий приклад. Нехай конфігурації А і Б обробляють однаковий потік із 100000 записів протягом години за однакових обмежень на обчислювальні ресурси. Еталонна розмітка охоплює 3600 неперекривних часових вікон, із яких 100 позначені як позитивні. Для кожного вікна допускається не більш як одна тривога. У прикладі встановлено такі критерії прийнятності: частка втрат не перевищує 0,1 %, 95-й процентиль затримки становить не більш як 2 с, повнота виявлення становить щонайменше 90 %, а кількість хибних тривог не перевищує 20 за годину. Наведені граничні значення використовують лише для ілюстрації методики, а не як нормативні вимоги.

Для конфігурації А прийнято втрату 50 записів, 95-й процентиль затримки 0,8 с, виявлення 95 позитивних вікон і 30 хибних тривог. Для конфігурації Б відповідні значення становлять 80 записів, 1,4 с, 93 вікна та 12 тривог. Частки втрат становлять 0,05 % і 0,08 %, а повнота — 95 % і 93 %. Обидві конфігурації відповідають вимогам щодо втрат, затримки та повноти виявлення, проте конфігурація А перевищує допустиму кількість хибних тривог. За заданих умов прийнятною є лише конфігурація Б, попри більші втрати та затримку. Наведені значення є авторськими припущеннями, а не результатами випробування програмного забезпечення.

Серед конфігурацій, що виконують усі обов’язкові вимоги, обирають ту, яка відповідає пріоритетам SOC. Якщо таких конфігурацій немає, це також фіксують. Для відтворення оцінки зберігають версії даних і програм, розмітку, налаштування та результати запусків. Зміна навантаження або складу цільових подій потребує повторної перевірки.

Запропонований підхід визначає умови порівняння систем журналізації через функціональні критерії, незалежні еталони та вимоги конкретного сценарію. Умовний приклад ілюструє процедуру, але не підтверджує її

результативності на практиці. Подальші дослідження мають перевірити відтворюваність оцінок і вплив граничних значень на результати порівняння.

Перелік посилань:

1. Vasylyshyn S., Vlasiuk I., Susukailo V. Cyber Incident Logging Technologies: Current State and Future Directions. Social Development and Security. 2025. Vol. 15, No. 3. P. 201–214. URL: https://doi.org/10.33445/sds.2025.15.3.18.

2. Khan Z. A., Shin D., Bianculli D., Briand L. C. Impact of log parsing on deep learning-based anomaly detection. Empirical Software Engineering. 2024. Vol. 29. Article 139. URL: https://doi.org/10.1007/s10664-024-10533-w.

3. Holeman R., Hastings J., Vaidyan V. M. Beyond Collection: Measuring the Detection Efficacy of Modern Security Logging Standards. Препринт. 2026. arXiv:2605.05531. URL: https://doi.org/10.48550/arXiv.2605.05531.

4. A comprehensive study of machine learning techniques for log-based anomaly detection / S. Ali et al. Empirical Software Engineering. 2025. Vol. 30. Article 129. URL: https://doi.org/10.1007/s10664-025-10669-3.

Яковенко Я.І.
студент групи ІКБм-61, ННКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАХИСТУ ІНФОРМАЦІЙНОЇ СИСТЕМИ ОРГАНІЗАЦІЇ НА ОСНОВІ КОНЦЕПЦІЇ ZERO TRUST ІЗ ВИКОРИСТАННЯМ MICROSOFT ENTRA ID ТА MICROSOFT INTUNE

Розглянуто, як концепцію Zero Trust можна реалізувати в корпоративній інформаційній системі за допомогою хмарних сервісів Microsoft Entra ID та Microsoft Intune. Пояснено, чому периметрова модель захисту вже не відповідає сучасним умовам, на яких принципах будується Zero Trust і які засоби дають змогу перевіряти користувача та його пристрій під час кожного запиту доступу. Окрему увагу приділено умовному доступу й політикам відповідності пристроїв. Сформульовано мету дослідження, практичну частину якого передбачено виконати в окремому тестовому тенанті.

Ключові слова: Zero Trust, Microsoft Entra ID, Microsoft Intune, умовний доступ, відповідність пристроїв.

Традиційний підхід до захисту ґрунтується на периметрі: усе, що перебуває всередині корпоративної мережі, вважається довіреним. Така модель втрачає ефективність, коли працівники підключаються з будь-яких мереж, а дані та застосунки переходять у хмару. Зловмиснику вже не потрібно «пробивати» мережевий кордон: достатньо отримати облікові дані користувача або скомпрометувати його пристрій. Саме тому концепція Zero Trust, яка з’явилася на початку 2010-х років і нині закріплена в документах NIST [1], виходить із протилежного принципу: жодному запиту не можна довіряти лише через його розташування в мережі.

Zero Trust — не окремий продукт, а підхід до побудови захисту. Згідно з NIST SP 800-207, доступ до кожного ресурсу надається лише після перевірки, а рішення ухвалюється з урахуванням стану користувача, пристрою та контексту запиту [1]. Microsoft описує цей підхід трьома принципами: перевіряти явно, надавати мінімально необхідні права та виходити з припущення, що компрометація вже відбулася [3]. Для оцінювання того, наскільки організація просунулася на цьому шляху, використовують моделі зрілості, наприклад модель CISA, яка охоплює п’ять напрямів: ідентичності, пристрої, мережі, застосунки та дані [2].

Microsoft Entra ID відповідає за ідентичності: він автентифікує користувачів, підтримує багатофакторну автентифікацію, оцінює ризик входу й дає змогу керувати привілейованим доступом. Microsoft Intune відповідає за пристрої: він налаштовує їх, перевіряє відповідність вимогам безпеки (шифрування диска, актуальність системи, наявність захисту) і передає результат перевірки до Entra ID. Завдяки цьому політика умовного доступу може ухвалити рішення на основі кількох сигналів одночасно: хто входить, з якого пристрою, з

якого місця та наскільки такий вхід ризикований [4]. Саме цей зв’язок між двома сервісами є основою практичної реалізації Zero Trust у середовищі Microsoft 365.

Водночас впровадження Zero Trust має власні складнощі. Надто суворі політики можуть блокувати легітимних користувачів і ускладнювати роботу, а надто м’які залишають прогалини в захисті. Тому важливо не лише налаштувати політики, а й перевірити їх дію на конкретних сценаріях: вхід із нового пристрою, із ризикової локації, з пристрою, що не відповідає вимогам, або з викраденими обліковими даними. До того ж такі перевірки не можна виконувати в робочому середовищі організації, де помилкова політика здатна заблокувати доступ працівникам, тож їх доцільно проводити в окремому тестовому тенанті.

Мета роботи — розробити технологію захисту інформаційної системи організації на основі концепції Zero Trust із використанням Microsoft Entra ID та Microsoft Intune і перевірити її дію на лабораторному стенді. Складність полягає в тому, що вимоги Zero Trust сформульовано на рівні принципів, тому їх потрібно перетворити на конкретні політики, які поєднують ідентичність користувача, стан пристрою та контекст доступу, і водночас не порушити зручність роботи.

Для оцінювання ефективності такої технології передбачено використати кілька груп показників: частку запитів доступу, що перевіряються політиками; частку пристроїв, які відповідають вимогам безпеки; кількість заблокованих небезпечних сценаріїв входу; а також кількість хибних блокувань легітимних користувачів. Останній показник важливий так само, як і перші, бо надмірні обмеження знижують продуктивність і підштовхують працівників до обхідних шляхів. Додатково результати зіставлятимуться з рівнями зрілості моделі CISA [2].

Перелік використаної літератури

1. Rose S., Borchert O., Mitchell S., Connelly S. Zero Trust Architecture. NIST Special Publication 800-207. Gaithersburg: National Institute of Standards and Technology, 2020. 59 p. DOI: 10.6028/NIST.SP.800-207
2. Zero Trust Maturity Model. Version 2.0. Cybersecurity and Infrastructure Security Agency (CISA), 2023. URL: https://www.cisa.gov/zero-trust-maturity-model
3. Microsoft Learn. Zero Trust Guidance Center. Microsoft, 2026. URL: https://learn.microsoft.com/security/zero-trust/
4. Microsoft Learn. Conditional Access overview. Microsoft Entra documentation. Microsoft, 2026. URL: https://learn.microsoft.com/entra/identity/conditional-access/overview

Хідзик Я.Ю.
студентка групи ІКБм-51, ННІКЗІ ДУІКТ
Київ, Україна

АКТУАЛЬНІ ПРОБЛЕМИ КІБЕРБЕЗПЕКИ. РОЗСЛІДУВАННЯ КІБЕРІНЦИДЕНТІВ

Розслідування кіберінцидентів – це процес встановлення причин, масштабу та джерела кібератаки, а також збирання цифрових доказів для їх подальшого аналізу. В умовах війни кібератаки стали частиною гібридної агресії росії проти України, тому кожен інцидент у державних органах, енергетиці чи зв'язку потребує не лише швидкого відновлення роботи, а й ретельного дослідження. Цифрові докази легко втратити, а саме вони дають змогу встановити винних. Від того, наскільки правильно діє команда в перші години інциденту, залежить і безпека систем, і результат розслідування.

Ключові слова: розслідування кіберінцидентів, воєнний стан, цифрова криміналістика, цифрові докази, критична інфраструктура.

Від початку повномасштабного вторгнення кібератаки на українські державні органи, оператори зв'язку, енергетичні об'єкти та фінансові установи стали складовою збройного конфлікту й часто координуються з іншими формами тиску. У таких умовах розслідування кіберінциденту допомагає відновити критичні послуги та формує докази, які можуть бути використані в національних і міжнародних провадженнях.

Закон України «Про основні засади забезпечення кібербезпеки України» визначає кіберінцидент як подію або ряд несприятливих подій, зокрема з ознаками можливої кібератаки, що загрожують безпеці систем електронних комунікацій та електронних інформаційних ресурсів. Реагування на нього закон описує як структуровану сукупність дій із підготовки, виявлення й аналізу, мінімізації шкоди та запобігання повторенню [1, ст. 1]. Такий підхід доповнює життєвий цикл NIST SP 800-61, що охоплює підготовку, виявлення та аналіз, стримування, ліквідацію та відновлення, а також післяінцидентну діяльність [2, с. 21–41]. Ці фази не є суворо послідовними: під час ліквідації діяльність нерідко повертається до виявлення й аналізу [2, с. 21], а отже, помилки перших годин реагування важко виправити пізніше.

У воєнний час тиск на швидкість реагування зростає, адже недоступність зв'язку, енергопостачання чи державних сервісів безпосередньо впливає на населення й оборонну спроможність. Проте стримування інциденту, зокрема вимкнення чи ізоляція системи, може знищити тимчасові дані. Настанова NIST SP 800-86 зазначає, що такі дані у багатьох випадках слід збирати першочергово, хоча пріоритет залежить також від ймовірної цінності даних і витрат зусиль [3, с. 3-3]. Стратегію стримування слід обирати так, щоб за можливості не втратити докази і водночас оцінити, як простій системи вплине на роботу організації [3, с. 3-5]. У воєнних умовах цей компроміс загострюється, бо обладнання може бути зруйноване внаслідок обстрілів, тому журнали подій, що зберігаються лише

локально, ризикують зникнути разом із системою. Централізоване пересилання журналів на захищені сервери, яке NIST пропонує насамперед для захисту від їх епідробки [3, с. 3-3], за таких умов виконує ще й функцію резервування.

Збереження доказів важливе і з юридичного погляду. Щоб дані могли бути використані в провадженні, необхідно перевіряти їх цілісність (порівнюючи контрольні суми оригіналу й копії), документувати кожну дію та фіксувати, хто і коли мав доступ до доказів (англ. chain of custody) [3, с. 3-4]. Водночас NIST наголошує, що ця настанова не може бути основою розслідування кримінальних правопорушень [3, с. 1-1], тому допустимість доказів визначається національним процесуальним законодавством. Атрибуцію атак ускладнює те, що зловмисники можуть застосовувати засоби приховування або знищення даних, і кожен збережений артефакт підвищує шанси на достовірні висновки. Для співпраці з іншими державами важливі механізми термінового збереження комп'ютерних даних. Конвенція про кіберзлочинність передбачає таке збереження як на національному рівні, так і за запитом іншої сторони [4, ст. 16, 29].е

Щоб не обирати між швидкістю та доказами під тиском, організації слід підготуватися заздалегідь. Необхідно розробити плани реагування, що визначають, які дані фіксуються до вжиття заходів стримування; погодити з керівництвом наперед рішення про виведення критичних систем із роботи для збору даних [3, с. 2-5]; організувати централізоване зберігання журналів подій, зокрема на резервних майданчиках поза зоною ризику; підготувати персонал до збереження доказів у стислі терміни; налагодити взаємодію з Держспецзв'язку (зокрема з урядовою командою CERT-UA), кіберполіцією та СБУ ще до виникнення інциденту.

Отже, в умовах війни ефективність розслідування кіберінциденту залежить від рішень, ухвалених до атаки та в перші години реагування. Поєднання швидкого відновлення з фіксацією тимчасових даних і веденням обліку доступу до доказів дає змогу зберегти стійкість критичної інфраструктури та створити доказову базу для притягнення винних до відповідальності.

Використані джерела

1. Про основні засади забезпечення кібербезпеки України: Закон України від 05.10.2017 № 2163-VIII. URL: https://zakon.rada.gov.ua/laws/show/2163-19
2. Cichonski P., Millar T., Grance T., Scarfone K. Computer Security Incident Handling Guide: NIST Special Publication 800-61 Rev. 2. Gaithersburg: NIST, 2012. 79 p.
3. Kent K., Chevalier S., Grance T., Dang H. Guide to Integrating Forensic Techniques into Incident Response: NIST Special Publication 800-86. Gaithersburg: NIST, 2006. 121 p. URL: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-86.pdf
4. Конвенція про кіберзлочинність: Рада Європи, 23.11.2001. URL: https://zakon.rada.gov.ua/laws/show/994_575

Попелюк Софія
студентка групи ІКБМ-51, ННІКБЗІ ДУІКТ
Київ, Україна

АКТУАЛЬНІ ПРОБЛЕМИ КІБЕРБЕЗПЕКИ. АТАКИ З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

Розглянуто еволюцію кібератак із застосуванням штучного інтелекту (ШІ): від генерації фішингових повідомлень до шкідливого програмного забезпечення з вбудованими великими мовними моделями та агентних кібероперацій із мінімальною участю людини. Запропоновано класифікацію ШІ-загроз за ступенем автономності, проаналізовано дані CERT-UA за 2026 рік, кейси LameHug (APT28), GTG-1002 та перевантаження вебресурсу українського ритейлера ШІ-системами. Окреслено напрями адаптації систем захисту: стійку до фішингу автентифікацію, контроль звернень до LLM-сервісів, поведінкову детекцію та керування ботами.

Постановка проблеми. Штучний інтелект (ШІ) перетворився з допоміжного інструмента на повноцінний чинник кіберзагроз. За даними CERT-UA, у першому півріччі 2026 року в Україні зафіксовано 3137 кіберінцидентів, що на 8 % більше, ніж за попередні півроку; ознаки використання ШІ виявлялися на різних етапах атак — від тексту листа й дизайну вебсторінки до скриптів, завантажувачів та окремих функцій шкідливого програмного забезпечення [1]. Водночас саме українські органи виконавчої влади стали ціллю одного з перших публічно задокументованих зразків шкідливого ПЗ з інтегрованою великою мовною моделлю (LLM) [2]. Метою роботи є систематизація сучасних загроз, пов’язаних зі ШІ, та визначення того, як вони змінюють вимоги до систем захисту.

Класифікація. За ступенем автономності доцільно виокремити чотири рівні використання ШІ зловмисниками. Перший — ШІ-підсилений соціальний інжиніринг: генерація безпомилкових і персоналізованих повідомлень будь-якою мовою, клонування голосу та відеодипфейки для вішингу й шахрайства від імені керівника. Другий — ШІ як засіб підвищення продуктивності атакувальника: пришвидшення написання й адаптації коду та аналізу вразливостей. Третій — шкідливе програмне забезпечення з вбудованою LLM, яке отримує завдання у вигляді текстового опису й під час виконання динамічно генерує команди під конкретне середовище. Четвертий — агентні кібероперації, в яких ШІ автономно виконує більшість тактичних дій, а людина лише схвалює ключові рішення. Окремо слід розглядати два суміжні явища: атаки на самі ШІ-системи (prompt injection, отруєння даних), що виходить за межі цієї роботи, та ненавмисне навантаження на вебресурси з боку легітимних ШІ-систем, яке за наслідками подібне до відмови в обслуговуванні.

Дані CERT-UA. За оцінкою CERT-UA, у першому півріччі 2026 року застосування ШІ набуло системного характеру: один сценарій швидко адаптується під іншу установу, професію чи регіон, так само як код програми й

спосіб її запуску [1]. Показовими стали три кампанії березня–травня 2026 року зі шкідливим ПЗ AGEWHEEZE, у яких зловмисники імітували повідомлення CERT-UA, оновлення системи захищеного документообігу Верховної Ради України та застосунок Brave1 Security Hub; одна з підроблених сторінок мала технічні ознаки створення за допомогою ШІ [1]. Водночас грамотний текст, коментарі в коді чи характерна структура файлів самі по собі не доводять застосування ШІ й не можуть бути підставою для атрибуції [1].

Кейс LameHug. 10 липня 2025 року CERT-UA отримала повідомлення про розсилання листів зі скомпрометованих облікових записів від імені посадовців міністерств на адреси органів виконавчої влади; вкладений архів містив завантажувач на Python, названий LameHug [2]. Кампанію із середнім рівнем впевненості пов'язано з групою APT28 (UAC-0001). Принципова відмінність зразка в тому, що він не містить жорстко закодованих команд: текстові описи цілей, закодовані у форматі Base64, надсилаються через API сервісу Hugging Face до відкритої моделі Qwen2.5-Coder-32B-Instruct, а у відповідь повертається послідовність команд для збору відомостей про систему, пошуку документів у типових каталогах і передавання їх зловмисникам через SFTP або HTTP POST [2]. Відсутність складних засобів маскування свідчить, на думку дослідників Cato Networks, скоріше про апробацію підходу на українських цілях, ніж про повноцінну операцію [2].

Кейс GTG-1002. У листопаді 2025 року компанія Anthropic оприлюднила звіт про припинення кампанії кібершпигунства, яку, за її оцінкою з високим ступенем впевненості, здійснювала група GTG-1002, підтримувана китайською державою [3]. За даними звіту, оператори доручили агентному ШІ-інструменту, інтегрованому із зовнішніми засобами, виконання 80–90 % тактичної роботи кампанії: людське втручання було потрібне лише у 4–6 критичних точках ухвалення рішень [3]. Водночас дані походять від самого розробника платформи, а частина фахівців вказує на відсутність публічних індикаторів компрометації, тож висновки потребують незалежного підтвердження.

Кейс «Текстиль-Контакт». У серпні 2026 року власник групи компаній повідомив, що сайт інтернет-магазину tk.ua різко сповільнився: спершу припускали DDoS-атаку, збій або дії хакерів, проте причиною виявилося навантаження від запитів ШІ-систем [4]. За його словами, ШІ-системи звертаються до сайту понад 10 млн разів на добу (близько 115 запитів на секунду), а відновлення стабільної роботи без вимкнення магазину зайняло два тижні, зокрема через обмеження доступу для ШІ-систем із регіонів, що створювали найбільше навантаження [4]. Цей інцидент не є кібератакою у вузькому розумінні, адже наміру завдати шкоди не було, проте за наслідками він подібний до DoS і засвідчує, що ШІ-системи стають самостійним чинником ризику для доступності інформаційних ресурсів. Наведені цифри — зі слів власника компанії; незалежного аналізу не оприлюднено.

Тенденції. Наведені кейси засвідчують три тенденції. По-перше, ШІ знижує вартість і поріг входу: якість соціальної інженерії та швидкість підготовки атак зростають без пропорційного зростання кваліфікації виконавця. По-друге, зростає адаптивність: команди формуються під конкретну систему безпосередньо під час виконання, тому статичні сигнатури втрачають ефективність. По-третє, змінюється співвідношення темпів: автономні агенти діють швидше за оператора-людину, тож реагування лише силами аналітиків SOC стає недостатнім.

Напрями протидії. Для адаптації систем захисту доцільно: 1) впровадити автентифікацію, стійку до фішингу (FIDO2, passkeys), та верифікацію чутливих запитів незалежним каналом як відповідь на дипфейки й ШІ-фішинг; 2) блокувати запуск небезпечних типів вкладень (.pif, .scr, скрипти) та обмежувати використання інтерпретаторів на робочих станціях; 3) контролювати вихідний трафік до публічних API LLM-сервісів із нетипових процесів — це практичний індикатор компрометації, як у випадку LameHug; 4) перейти від сигнатурного виявлення до поведінкової детекції (EDR, NDR, UEBA), автоматизувати реагування (SOAR) і застосовувати ШІ на боці захисту; 5) для вебресурсів запровадити керування ботами: обмеження частоти запитів, WAF/CDN, політики доступу для ШІ-краулерів; 6) регулярно навчати персонал із застосуванням ШІ-симуляцій фішингу.

Висновки. Загрози, пов’язані зі ШІ, за короткий час пройшли шлях від допоміжного інструмента для створення фішингових повідомлень до компонента шкідливого програмного забезпечення, керованих агентів і самостійного чинника ризику для доступності ресурсів. Дані CERT-UA та кейси LameHug, GTG-1002, «Текстиль-Контакт» свідчать про початок нової фази, хоча їхня реальна ефективність і масштаби потребують подальшого вивчення. Перспективним є розроблення моделей загроз для ШІ-атак та методів виявлення поведінки шкідливого ПЗ, керованого LLM.

Література:

1. Кіберзагрози: Україна. Звіт CERT-UA за І півріччя 2026 р. (за оглядом Fintech Insider, 01.10.2026). URL: https://fintechinsider.com.ua/shtuchnyj-intelekt-u-kiberatakah-ta-zrostannya-kilkosti-inczydentiv-na-8-klyuchovi-trendy-kiberzagroz-v-ukrayini/

2. CERT-UA Discovers LAMEHUG Malware Linked to APT28, Using LLM for Phishing Campaign. The Hacker News, July 2025. URL: https://thehackernews.com/2025/07/cert-ua-discovers-lamehug-malware.html

3. Disrupting the first reported AI-orchestrated cyber espionage campaign. Anthropic, November 2025. URL: https://www.anthropic.com/news/disrupting-AI-espionage

4. Textile-Contact’s online store hit by AI systems, not hackers. The New Voice of Ukraine, August 2026. URL: https://english.nv.ua/nation/textile-contact-online-store-owned-by-ukrainian-group-of-companies-attacked-by-artificial-intel-50630109.html

Науменко Д.М.
студент групи ІКБм-62, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ВРАЗЛИВОСТЕЙ ВЕБ-ДОДАТКІВ ОРГАНІЗАЦІЇ

Розглянуто технологію виявлення вразливостей веб-додатків організації на основі поєднання статичного, динамічного та ручного тестування безпеки. Описано типові класи вразливостей відповідно до OWASP Top 10, принципи роботи засобів SAST, SCA і DAST, а також місце тестування на проникнення. Запропоновано схему процесу виявлення вразливостей із пріоритизацією за шкалою CVSS та інтеграцією перевірок у процес розроблення (CI/CD).

Ключові слова: веб-додаток, вразливість, OWASP Top 10, SAST, DAST, тестування на проникнення, CVSS.

Веб-додатки є одним з основних каналів взаємодії організації з клієнтами та партнерами і водночас однією з найпоширеніших цілей кібератак. Корпоративні портали, інтернет-магазини, особисті кабінети та API доступні з мережі Інтернет, обробляють персональні й фінансові дані та часто мають доступ до внутрішніх систем. Тому своєчасне виявлення вразливостей веб-додатків є необхідною умовою захисту інформаційних ресурсів організації [1].

Найбільш поширені класи вразливостей систематизовано в переліку OWASP Top 10. До них належать порушення контролю доступу, помилки конфігурації безпеки, ін’єкції (SQL, OS command, XSS), криптографічні помилки, недоліки автентифікації, використання вразливих і застарілих компонентів, недоліки журналювання та моніторингу [1]. Значна частина цих вразливостей виникає через помилки під час розроблення та супроводу, тому їх виявлення має здійснюватися на всіх етапах життєвого циклу додатка.

Першим етапом технології є інвентаризація: формування реєстру веб-додатків організації із зазначенням власника, технологічного стеку, точок входу (URL, API), класу оброблюваних даних і критичності. Без такого реєстру частина ресурсів, зокрема тестові та застарілі сервіси, залишається поза увагою.

Статичний аналіз безпеки (SAST) виконується над вихідним кодом без його запуску і дозволяє виявляти небезпечні конструкції: відсутність фільтрації вхідних даних, формування SQL-запитів конкатенацією рядків, жорстко закодовані паролі та ключі. Аналіз складу програмного забезпечення (SCA) перевіряє сторонні бібліотеки та залежності на наявність відомих вразливостей із бази CVE. Перевагою цих методів є раннє виявлення проблем, недоліком — значна кількість хибнопозитивних спрацювань.

Динамічне тестування (DAST) проводиться над працюючим додатком за принципом «чорної скриньки»: сканер обходить сторінки та API, надсилає модифіковані запити й аналізує відповіді сервера. Поширеними засобами є OWASP ZAP і Burp Suite [3]. DAST виявляє вразливості, які проявляються лише

під час виконання: помилки конфігурації, недоліки керування сесіями, ін’єкції. Автоматизовані засоби доповнюються ручним тестуванням на проникнення за методологіями OWASP Web Security Testing Guide [2] та NIST SP 800-115 [4], що дозволяє знайти логічні вразливості та помилки контролю доступу, недоступні для сканерів. Структуру процесу наведено на рис. 1.



Рис. 1. Схема технології виявлення вразливостей веб-додатків організації

Результати всіх етапів об’єднуються в єдиному реєстрі вразливостей, де виконуються дедуплікація та верифікація знахідок. Для пріоритизації використовується система оцінювання CVSS, яка враховує складність експлуатації, необхідні привілеї та вплив на конфіденційність, цілісність і доступність [5]. Остаточний пріоритет визначається з урахуванням критичності самого додатка: вразливість середнього рівня в публічному платіжному сервісі може бути небезпечнішою за критичну у внутрішньому тестовому середовищі. Для кожного рівня встановлюються терміни усунення, після чого виконується повторна перевірка.

Ефективність технології суттєво зростає при інтеграції перевірок у процес розроблення за підходом DevSecOps. SAST і SCA запускаються автоматично під час кожного злиття коду в конвеєрі CI/CD, DAST — після розгортання в тестовому середовищі, а ручне тестування на проникнення проводиться перед випуском значущих оновлень і періодично, наприклад раз на рік. Критичні вразливості блокують випуск нової версії до їх усунення.

Практичне застосування описаної технології дозволяє виявляти вразливості на ранніх етапах, коли їх усунення найдешевше, скоротити «вікно вразливості» між появою помилки та її виправленням і підвищити загальний рівень захищеності веб-ресурсів організації. Для ефективного впровадження рекомендується: вести актуальний реєстр веб-додатків, поєднувати автоматизовані та ручні методи тестування, пріоритизувати вразливості з урахуванням бізнес-критичності, контролювати терміни їх усунення та навчати розробників принципам безпечного програмування.

Література

1. OWASP Top Ten. URL: https://owasp.org/www-project-top-ten/ (дата звернення: 23.10.2026)
2. OWASP Web Security Testing Guide. URL: https://owasp.org/www-project-web-security-testing-guide/ (дата звернення: 23.10.2026)
3. ZAP (Zed Attack Proxy) Documentation. URL: https://www.zaproxy.org/docs/ (дата звернення: 23.10.2026)
4. NIST SP 800-115. Technical Guide to Information Security Testing and Assessment. URL: https://csrc.nist.gov/pubs/sp/800/115/final (дата звернення: 23.10.2026)
5. Common Vulnerability Scoring System v4.0: Specification Document. FIRST. URL: https://www.first.org/cvss/v4-0/specification-document (дата звернення: 23.10.2026)

Скидан Данило Романович
аспірант факультету інформаційних технологій,
Київський національний університет імені Тараса Шевченка,Київ,Україна

Щебланін Юрій Миколайович
канд. техн. наук, ст. наук. співроб.,
Київський національний університет імені Тараса Шевченка, Київ, Україна

ЗАХИСТ КОРПОРАТИВНИХ WI-FI МЕРЕЖ МЕТОДАМИ КОМБІНОВАНОГО ШИФРУВАННЯ ТА КОНТРОЛЮ СЕСІЙ

Анотація. Розглянуто підхід до підвищення кіберстійкості корпоративних інформаційних систем за рахунок поєднання контролю стану мережевих сесій, раннього відсіювання аномальних пакетів і подальшої криптографічної перевірки. Акцент зроблено на зменшенні впливу втрат і перестановки пакетів, повторних запитів та інтенсивного аномального трафіку на доступність корпоративних сервісів. Запропоновано багаторівневу логіку оброблення трафіку, яку можна інтегрувати в шлюзи, міжмережеві екрани або вузли захищеного доступу.
Ключові слова: кібербезпека, корпоративні бездротові мережі, Wi-Fi, кіберстійкість, мережеві сесії, аномальний трафік, ентропійний аналіз, комбіноване шифрування, протокол HCEP.

Корпоративні бездротові мережі (Wi-Fi) об’єднують внутрішні вебзастосунки, системи документообігу, бази даних, засоби віддаленого доступу, хмарні сервіси та інфраструктуру обміну інформацією між підрозділами. Порушення доступності хоча б одного критичного компонента може впливати на бізнес-процеси, спричиняти затримки в обробленні даних і перешкоджати роботі працівників. Тому забезпечення кібербезпеки таких систем має охоплювати не лише конфіденційність інформації, а й здатність інфраструктури зберігати працездатність за нестабільного мережевого середовища та підвищеного навантаження.

Одним із практичних викликів є розмежування легітимних мережевих відхилень і трафіку, який може бути ознакою зловмисної активності. Втрата або затримка пакетів не обов’язково свідчить про атаку: причиною можуть бути перевантаження каналу, перешкоди в бездротовому сегменті чи зміна маршруту [1, с. 15]. Водночас велика кількість повторних, застарілих або некоректних пакетів здатна витрачати обчислювальні ресурси захисного вузла та погіршувати доступність сервісів. Отже, жорстка перевірка послідовності пакетів може бути надто чутливою до нормальних мережевих коливань, тоді як відсутність ранньої фільтрації створює додаткове навантаження на інфраструктуру.

Метою роботи є обґрунтування багаторівневого підходу до захисту корпоративних бездротових мереж, у якому первинний контроль стану

мережевої сесії виконується до ресурсоємних криптографічних операцій, а остаточне рішення приймається з урахуванням результатів автентифікації та перевірки цілісності. Такий підхід не замінює міжмережеві екрани, системи виявлення вторгнень або механізми автентифікації, а доповнює їх на рівні оброблення мережевих потоків [2, с. 24].

Запропонований підхід передбачає три послідовні рівні. На першому рівні виконується перевірка базових характеристик пакета: належності до очікуваної сесії, коректності службових полів, відповідності дозволеному напрямку обміну та наявності ознак повторного передавання. Для реалізації другого рівня контролю застосовується механізм ковзного вікна станів, який фіксує номери послідовності прийнятих пакетів. Замість ресурсоємного зберігання кожного номера використовується бітова маска, де кожен біт відповідає статусу пакета в межах визначеного діапазону. Якщо номер пакета виходить за нижню межу вікна, він автоматично відхиляється як застарілий або потенційний Replay-запит. Паралельно з цим модуль ентропійного аналізу обчислює міру хаосу в заголовках пакетів на невеликих часових інтервалах. Раптові стрибки ентропії сигналізують про спроби сканування мережі або початок DDoS-атаки, що дає змогу динамічно звужувати ширину ковзного вікна для посилення контролю без зупинки легітимних сесій. [3, с. 4]. На третьому рівні проводяться криптографічна перевірка, автентифікація та перевірка цілісності, для чого використовується адаптивний протокол комбінованого шифрування (HCEP).(Рис.1)

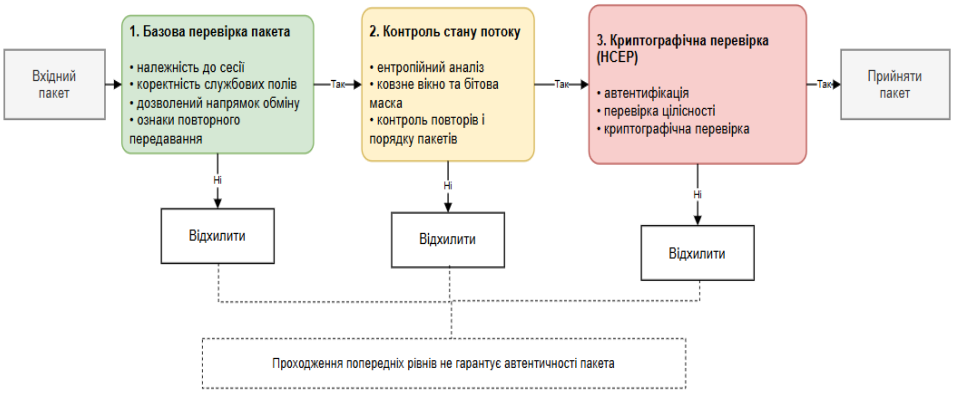


Рис.1. Послідовні рівні оброблення мережевого трафіку

Ключовим елементом є розмежування попереднього фільтрування та остаточного підтвердження легітимності. Попередній контроль може відхиляти очевидні дублікати, пакети поза допустимим вікном і повідомлення, що не відповідають контексту сесії. Проте проходження цього етапу не повинно вважатися доказом автентичності: пакет усе одно має пройти криптографічну перевірку. Це зменшує ризик того, що оптимізація продуктивності послабить модель безпеки. Для корпоративного середовища також доцільно реєструвати

причини відхилення, кількість повторів, частоту помилок автентифікації та рівень завантаження захисного вузла, щоб корелювати ці події із системами моніторингу безпеки.

Перевага контролю стану полягає в тому, що він дає змогу відрізняти допустиму перестановку пакетів від повторного надходження вже оброблених даних. Ширина вікна визначає компроміс між толерантністю до затримок і витратами на підтримання стану: надто мале вікно може спричинити відхилення легітимних пакетів, а надто велике – збільшити накладні витрати та діапазон номерів, які потрібно контролювати. Тому параметри мають визначатися з урахуванням архітектури корпоративної мережі, характеру трафіку, допустимої затримки й ресурсів шлюзу. Для критичних систем доцільно передбачити контрольоване оновлення параметрів і тестування змін у середовищі, яке не впливає на виробничі сервіси.

З погляду протидії аномальному навантаженню рання перевірка допомагає зменшити кількість пакетів, які проходять повний цикл оброблення. Це особливо важливо для вузлів віддаленого доступу, VPN-шлюзів і сервісів, що обробляють багато коротких запитів. Водночас сам контроль стану не може гарантувати захист від розподілених атак на відмову в обслуговуванні: він має застосовуватися разом з обмеженням частоти запитів, фільтрацією на периметрі, моніторингом аномалій, резервуванням критичних компонентів і планом реагування на інциденти. Подібна багаторівнева архітектура дає змогу розподіляти захисні функції між мережевим периметром, шлюзом доступу та засобами централізованого моніторингу.

Для оцінювання ефективності рішення доцільно використовувати набір показників, що відображає як безпекові, так і експлуатаційні властивості: частку помилково відхиленого легітимного трафіку, кількість повторних пакетів, які були відсіяні до криптографічної перевірки, пропускну здатність, затримку оброблення, використання CPU та пам’яті, а також час відновлення сервісу після пікового навантаження. Окремо слід оцінювати хибнопозитивні спрацювання, оскільки надмірно агресивна фільтрація може блокувати легітимну роботу корпоративних застосунків. Перевірка має охоплювати нормальне навантаження, втрати та перестановку пакетів, повторні запити й контрольовані сценарії аномальної інтенсивності.

Таким чином, кіберстійкість корпоративної бездротової мережі варто розглядати як результат узгодженої роботи адаптивного комбінованого шифрування, контролю стану мережевих сесій і механізмів ентропійного аналізу. Раннє відсіювання очевидно некоректного або повторного трафіку може зменшити непотрібні обчислення, тоді як збереження обов’язкової криптографічної перевірки протоколом HCEP забезпечує цілісність моделі безпеки. Подальший етап дослідження має передбачати перевірку такого підходу в корпоративному тестовому середовищі з різними типами сервісів і профілями

навантаження, без таких вимірювань кількісний ефект для конкретної організації не слід вважати встановленим.

Перелік використаної літератури

[1] Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). Zero Trust Architecture (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207

[2] IEEE. (2021). IEEE Standard for Information Technology—Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks—Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11-2020. https://doi.org/10.1109/IEEESTD.2021.9363693

[3] Скидан Д. Р., Щебланін Ю. Захист Wi-Fi мереж за допомогою адаптивного методу шифрування. *Кібербезпека: освіта, наука, техніка*. 2026. Т. 2, № 34. С. 63–69. DOI: https://doi.org/10.28925/2663-4023.2026.34.1261.

Азіз АБІБУЛАЄВ,
аспірант кафедри захисту інформації, ІКТА
Андріан ПІСКОЗУБ,
к. т. н., доцент кафедри захисту інформації, ІКТА
Національний університет «Львівська політехніка»

ХИБНОПОЗИТИВНІ СПРАЦЮВАННЯ ДЕТЕКТОРІВ ВЕБ-АТАК У РЕАЛЬНОМУ ТРАФІКУ

Анотація. Навчені детектори атак на веб-застосунки та API показують майже ідеальні лабораторні метрики, проте в експлуатації перевантажують аналітиків SOC хибнопозитивними спрацюваннями. У тезі розглянуто гіпотезу, за якою навчений на стендових атаках детектор у першому наближенні реагує на відмінність запиту від нормального трафіку навчання. Умова покриття за явного припущення обмежує частку хибних тривог. Детектор, навчений на URI та заголовках власного нормального трафіку, дав на двох реальних доменах 0,42 % і 0,61 % тривог порівняно з 99,5 % після навчання на інших джерелах. На мережевій пастці (honeypot) підхід виявився непридатним. Для SOC рекомендовано оцінювати детектор на власному трафіку і застосовувати для попереднього відбору подій.

Для центрів моніторингу безпеки (SOC) навчений детектор атак на веб-застосунки та API є привабливою заміною сигнатурних правил WAF (Web Application Firewall). Огляд публікацій показує, що детектори мають майже ідеальні метрики на еталонних наборах, проте наш досвід свідчить, що у виробничому середовищі вони часто непридатні через хибні тривоги на звичайних запитах. Ця проблема відома як помилка базової частоти (base-rate fallacy). У роботі [1, с. 193–194] показано, що за рідкісних атак навіть невелика частка хибних тривог робить хибними більшість сповіщень, а її автор вважає, що через це оператор схильний їх ігнорувати. Розрив приховують і відомі методологічні пастки [2, с. 3974, 3976], де модель може спиратися на сторонні артефакти даних, наприклад, на їхнє джерело. Іншою пасткою є оцінювання лише в лабораторних або стендових умовах. У тезі на чотирьох наборах реального трафіку досліджено, від чого залежить частка хибних тривог навченого детектора, а повні результати й цей огляд опубліковано в [3, с. 2].

Коли атаки навчальної вибірки породжено кількома інструментами на стенді, вони утворюють вузькі кластери, і вивчене правило може в першому наближенні реагувати на відмінність запиту від нормального трафіку навчання, а не на зловмисний вміст. Ми називаємо це гіпотезою виродження в доповнення нормальної області й наводимо для неї евристичне обґрунтування та часткову емпіричну підтримку. Якщо гіпотеза справджується, частка хибних тривог в

експлуатації залежить передусім від зсуву нормального трафіку відносно навчального.

Покриття означено без участі класифікатора, а кожен запит e подано TF-IDF-вектором $\varphi(e)$ символьних n-грам його декодованого вмісту. Позначимо через $N_r(\mathcal{D}^{tr})$ множину векторів, косинусна відстань яких до найближчого вектора нормального трафіку навчання не перевищує r. Нормальний трафік навчання покриває нормальний трафік цільового домену з радіусом r і похибкою ε, якщо виконується умова (1):

$$\mathcal{D}^{\mathrm{tar}}\left(\varphi(e) \notin N_r(\mathcal{D}^{\mathrm{tr}})\right) \leq \varepsilon \quad (1)$$

де $\mathcal{D}^{tr}$ і $\mathcal{D}^{tar}$ позначають розподіли нормального трафіку навчання і цільового домену. Величина ε обмежує частку цільових нормальних запитів поза множиною $N_r(\mathcal{D}^{tr})$, і за умови, що область прийняття детектора її містить, виконується межа (2):

$$FPR^{\mathrm{tar}} \leq \varepsilon + \rho \quad (2)$$

де FPR^{tar} позначає частку хибних тривог на нормальному трафіку цільового домену, а ρ є її складовою, спричиненою легітимними запитами, які за змістом неможливо відрізнити від атаки. Включення множини $N_r(\mathcal{D}^{tr})$ в область прийняття є припущенням, оскільки навчання саме по собі його не гарантує. Покриття виміряно для 13 пар навчальних і цільових даних. Якщо вважати, що ρ не перевищує одного відсоткового пункту, межа (2) виконувалася в 10 із них, проте змістовно обмежувала частку тривог лише у трьох внутрішньодоменних випадках. В інших трьох парах межа тривіальна, бо $\varepsilon \approx 1$, у чотирьох завищена, а в решті трьох не виконувалася, бо область прийняття не містила цю множину повністю. Тому межа не гарантує результату для інших доменів і не дає кількісно передбачити частку тривог.

Межа (2) робить зменшення ε метою проєктування, і найнадійніше мети досягають, беручи нормальний трафік навчання з того домену, де детектор працюватиме. Так побудовано логістичну регресію на векторах $\varphi(e)$, яку навчали відрізняти нормальний трафік навчального періоду цільового домену від атак зі стенда і з мережевої пастки (honeypot) SR-BH. Частку тривог визначено як частку нормальних за мітками запитів, які детектор позначив як атаку. На сайті електронної комерції Zanbil і в архіві SEC EDGAR вона становила 0,42 % і 0,61 % відповідно, а після навчання на трафіку інших джерел на Zanbil сягала 99,5 % (рис. 1). Натомість навчений на стендовому трафіку детектор із текстовими та числовими ознаками й лабораторною площею під ROC-кривою понад 0,98 позначав як атаку 83 % нормальних запитів Zanbil. Повнота (recall) детекторів Zanbil і EDGAR на наборах атак, які не залежать від цільового домену, була не нижчою за 93,0 %, і цей мінімум припадає на набір ModSecurity-2025 із заблокованих правилами WAF-запитів. Лише цей набір та атаки іншого інструмента не брали участі в навчанні.

Частка тривог не дорівнює частці хибних тривог, бо серед нормальних за мітками запитів трапляються приховані атаки. Ручна перевірка 100 тривог Zanbil з найвищою оцінкою показала, що 61 з них є саме такими атаками. Якщо серед тривог із нижчою оцінкою частка атак не вища, ніж серед перевірених, то атаками є щонайбільше 61 % тривог, а оцінка частки хибних тривог становить 0,16–0,42 %. На єдиному оціненому джерелі з кількома піддоменами частка тривог знизилася з 20,0 % до 1,42 %, коли часовий поділ виконали окремо в кожному піддомені. Проте повнота на ModSecurity-2025 впала з 93,4 % до 77,8 %. Для таких доменів результат вказує на потребу збирати нормальний трафік навчання з кожного піддомену і перенавчати детектор, коли їхній набір змінюється. На мережевій пастці SR-BH, де нормальний за мітками трафік здебільшого був автоматизованим скануванням, цей підхід дав частку тривог 53,5 %.

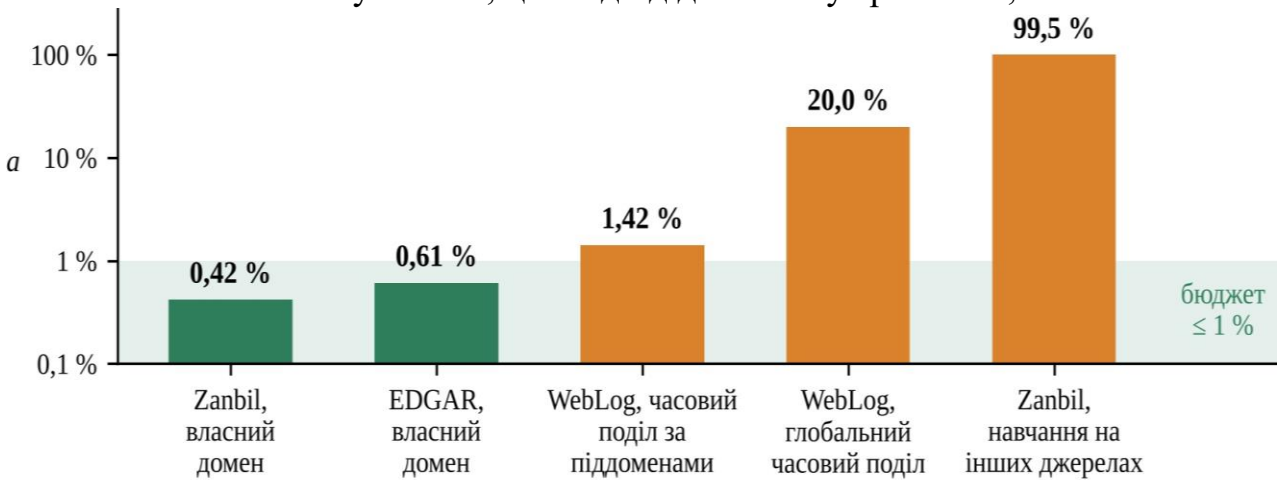


Рис. 1. Частка тривог *a* на нормальному трафіку за різних режимів навчання, де зеленим позначено значення в межах бюджету 1 %, а помаранчевим понад нього [3, с. 3, 26, 36]

Детектор також порівняно з фіксованим словником регулярних виразів із восьми груп шаблонів, а не з робочим набором правил WAF. На SR-BH детектор виявив усі атаки, а правила 45,8 %, проте повноту детектора виміряно всередині джерела навчання. Частка тривог у правил на Zanbil нижча й становить 0,09 %.

Для SOC критерієм приймання детектора в експлуатацію має бути частка тривог на нормальному трафіку власного домену разом із повнотою на незалежних наборах атак, а не лабораторна метрика на чужих даних. Поріг слід підбирати під бюджет тривог на окремій калібрувальній вибірці. Якщо на атаки припадає 0,1 % потоку і детектор знаходить їх усі, то за частки тривог 0,42 % його точність (precision) становить близько 19 %. Проте кількість подій, які має переглянути аналітик, зменшується приблизно в 190 разів. Тому детектор доцільно використовувати для попереднього відбору подій, а не для автоматичного блокування.

Висновки стосуються лише URI і заголовків, бо дані цільових доменів не містять тіл запитів. Нормальний трафік і атаки взято з різних джерел, тож модель

могла частково розрізняти саме джерело даних. Поведінкові та контекстні атаки детектор не охоплює, бо аналізує лише окремий запит.

Отже, у проведених експериментах частка тривог залежала передусім від того, на якому нормальному трафіку детектор навчали, і в якому домені його застосовували, тоді як висока лабораторна метрика її не передбачала. На нашу думку, навчений на власному трафіку детектор варто поєднувати з правилами, а покриття планується підтримувати неперервним навчанням за участю аналітика.

Список використаних джерел

1. Axelsson S. The Base-Rate Fallacy and the Difficulty of Intrusion Detection. *ACM Transactions on Information and System Security*. 2000. Vol. 3, No. 3. P. 186–205. DOI: https://doi.org/10.1145/357830.357849.

2. Dos and Don'ts of Machine Learning in Computer Security / D. Arp et al. *Proceedings of the 31st USENIX Security Symposium* (Boston, 10–12 August 2022). Boston : USENIX Association, 2022. P. 3971–3988. URL: https://www.usenix.org/conference/usenixsecurity22/presentation/arp (дата звернення: 26.10.2026).

3. Abibulaiev A., Pukach P., Vovk M. Limits of Web Attack Detection in Real-World Production Environments. *Applied Sciences*. 2026. Vol. 16, No. 20. Article 10340. DOI: https://doi.org/10.3390/app162010340.

Пашенько С. П.,
студент групи ІКБМ-51, ННІКБЗІ ДУІКТ,
Київ, Україна

РЕКОНСТРУКЦІЯ КІБЕРІНЦИДЕНТІВ ШЛЯХОМ КОРЕЛЯЦІЇ ЦИФРОВИХ СЛІДІВ

У тезах розглянуто проблему відновлення послідовності кіберінциденту за неповними та різнорідними цифровими слідами. Запропоновано доказово орієнтовану послідовність розслідування, що поєднує збереження даних, нормалізацію часових позначок, кореляцію подій та перевірку альтернативних пояснень. Окрему увагу приділено матриці доказових зв'язків, яка дозволяє відокремити достовірно зафіксовані факти від аналітичних припущень. На умовному прикладі показано, як події з журналів кінцевої точки, мережі й автентифікації утворюють перевірювану хронологію. Окреслено обмеження підходу, пов'язані з неповнотою журналювання та похибками часу.

Ключові слова: кіберінцидент, цифрові докази, кореляція подій, журналювання, цифрова криміналістика, хронологія подій.

Актуальність. У корпоративній інфраструктурі відомості про одну атаку можуть бути розподілені між кінцевими пристроями, мережевим обладнанням, системами ідентифікації, хмарними сервісами та засобами моніторингу. Окреме сповіщення про підозрілий процес або нетипове мережеве з'єднання лише вказує на можливу проблему, але не пояснює початковий вектор проникнення, послідовність дій або масштаб інциденту. Метою роботи є запропонувати перевірювану реконструкцію подій, що відокремлює факти від гіпотез. Відповідно до NIST SP 800-61 Rev. 3, діяльність із реагування доцільно розглядати в контексті управління кіберризиками та безперервного вдосконалення організаційних процесів [1].

Збереження цифрових слідів. Перед інтенсивним аналізом необхідно визначити, які джерела можуть бути втрачені через перезапуск пристрою, ротацію журналів або зміни в хмарному середовищі. Визначаються відповідальні особи, порядок доступу та перелік отриманих матеріалів. Для кожної копії доцільно фіксувати час отримання, походження, застосований інструмент і контрольну суму, а всі подальші операції виконувати з робочою копією. Така дисципліна відповідає загальним принципам цифрової криміналістики, описаним у NIST SP 800-86, та документування ланцюга збереження доказів у RFC 3227 [2; 4]. Хеш підтверджує незмінність файлу, але не достовірність його вмісту.

Нормалізація подій. Далі записи з різних джерел переводяться до спільного формату: час із зазначенням часового поясу, ідентифікатор вузла або облікового запису, тип події, первинне джерело та посилання на збережений доказ. Потрібно враховувати різницю між часом створення, реєстрації та отримання події колектором, а також можливе відхилення годинників. Події впорядковуються за часом, проте близькість двох записів ще не доводить причинного зв'язку між ними. Додаткові поля, наприклад ідентифікатор процесу, сеансу входу або мережевого з'єднання, допомагають перевірити, чи описують журнали одну й ту саму дію.

Кореляція та формування гіпотез. Пропонується пов’язувати події не тільки за часовою близькістю, а й за спільними сутностями: хостом, користувачем, процесом, IP-адресою, доменом і типом активності. Для опису можливої мети дій можна застосовувати тактики й техніки MITRE ATT&CK, які слугують аналітичним словником, а не автоматичним доказом конкретної атаки [3]. Коли виявлено підозріле виконання сценарію, слід перевірити, чи пов’язане воно з входом користувача, створенням дочірнього процесу та вихідним мережевим з’єднанням. Альтернативою може бути легітимна адміністративна дія.

Матриця доказових зв’язків. Як практичний результат аналітичного узагальнення запропоновано оформлювати кожен ключовий висновок у вигляді запису з п’яти полів: твердження; первинне джерело; ознака зв’язку з іншими подіями; альтернативне пояснення; статус перевірки. Статус «підтверджено» надається лише тоді, коли запис перевірено за доступними первинними даними; статус «ймовірно» використовується для пояснення, яке має неповне підтвердження; статус «невідомо» позначає прогалину, а не відсутність атаки. На відміну від простого списку індикаторів компрометації, така матриця допомагає побачити, на яких фактах тримається кожна частина реконструкції та яку інформацію ще потрібно зібрати.

Умовний приклад застосування. Припустімо, що о 08:41 зареєстровано перехід за посиланням із підозрілого листа, о 08:43 - запуск командного інтерпретатора з офісного застосунку, а о 08:46 - DNS-запит до невідомого домену. Ці три спостереження утворюють підставу для гіпотези про початкове зараження, проте не дозволяють одразу визнати її доведеною. Для перевірки необхідно зіставити обліковий запис, ідентифікатори процесів, походження листа, часові позначки, мережеві потоки та журнали захисних засобів. Якщо частина даних відсутня, це прямо фіксується у звіті. Приклад є синтетичним і використовується виключно для демонстрації логіки методики.

Висновки. Узгоджена хронологія кіберінциденту повинна спиратися на простежувані цифрові докази, а не лише на одночасність сповіщень. Запропонована схема створює основу для оформлення результатів розслідування. Її очікувана користь полягає у зменшенні необґрунтованих висновків та кращій передачі справи між аналітиками. Подальший напрям дослідження - практична перевірка на навчальних наборах журналів.

Список використаних джерел

1. 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: 24.10.2026).
2. Kent K., Chevalier S., Grance T., Dang H. Guide to Integrating Forensic Techniques into Incident Response. NIST SP 800-86. 2006. URL: https://doi.org/10.6028/NIST.SP.800-86 (дата звернення: 24.10.2026).
3. MITRE ATT&CK. Enterprise Tactics. URL: https://attack.mitre.org/tactics/enterprise/ (дата звернення: 25.10.2026).
4. Brezinski D., Killalea T. Guidelines for Evidence Collection and Archiving. RFC 3227. 2002. URL: https://www.rfc-editor.org/rfc/rfc3227 (дата звернення: 25.10.2026).

Levterov G.A., bachelor student
Kharkiv's National University of Radio Electronics (KhNURE)
Levterov A. A., professor
National University of Civil Protection of Ukraine
Kharkiv, Ukraine

ANALYSIS CYBER INCIDENT PREVENTION IN THE MOBILE SECTOR OF THE UKRAINIAN ENERGY AND DEFENSE COMPLEXES UNDER CONDITIONS POWER SUPPLY DEGRADATION

The comprehensive scientific and technical analysis of vulnerabilities within the mobile segment of Ukraine's critical information infrastructure and defense sector for the 2024–2026 period has been described. The study synergistic effects of combined kinetic and cyber operations amidst hybrid warfare and persistent power supply disruptions have been researched. Indicators of Compromise were identified, and operational security protocols were reviewed. The comparative analysis of mobile device management systems has been performed with a focus on the resilience of distributed electric power networks.

Keywords: Mobile segment, Critical information infrastructure, Defense sector.

The contemporary doctrine of hybrid warfare involves the deep integration of destructive cyber operations and conventional kinetic strike capabilities. In the context of the protracted conflict in Ukraine, the critical information infrastructure (CII), specifically the energy sector and operational military command systems, is subjected to coordinated destructive impacts. The characteristic feature of the current stage (2024–2026) is the shifting focus of attacking subcontractors of the russian intelligence services (specifically, the APT groups Sandworm/UAC-0165 and Gamaredon/UAC-0010) toward the mobile segment of the CII [1].

The implementation of scheduled stabilization and emergency power outages, alongside total blackouts, leads to the forced decentralization of operational management. Engineering and technical personnel of distribution system operators (oblenenergos) and tactical-level operators of the Armed Forces of Ukraine are compelled to massively diversify their communication channels, transferring operational planning and the coordination of restoration works to mobile terminals. This de facto erodes the classical concept of a protected network perimeter and shifts the critical vulnerability node to the end-user endpoints [2]. The primary risk factors under conditions of rolling degradations of centralized power supply include: Erosion of the Trusted Transport Environment. The overloading and subsequent shutdown of cellular base stations force personnel to utilize unauthorized public or volunteer Wi-Fi access points without implementing cryptographic tunnels. This opens a vector for executing MitM attacks and intercepting telemetric data [3]. Compromise of the BYOD Concept. The forced utilization of personal endpoints under a BYOD framework within critical industrial control loops due to a logistical deficit of specialized secure hardware platforms. Degradation of Situational Awareness and Vigilance. Heightened

psychological pressure on personnel under emergency conditions leads to a sharp decline in the cognitive filtering of incoming information, exponentially increasing the efficacy of social engineering techniques.

Analysis of operational data from the government computer emergency response team CERT-UA [4] allows for the classification of profiled mobile threats based on the targeted platform type and the complexity of the deployed mechanisms.

iOS Platform (DarkSword Exploit Kit). Vector: Target — command nodes and senior engineers. Uses Watering Hole attacks via compromised news/government portals to deploy a 0-day WebKit (Safari) chain. Execution: One-Click/Zero-Click web page access triggers a browser sandbox escape, forces Kernel Privilege Escalation, and injects a fileless spyware module into RAM. Network IoCs: Outbound HTTPS to C2 sqwas.shapelie[.]com; traffic to compromised nodes novosti.dn[.]ua and 7aac.gov[.]ua. System IoCs: Anomalous WebContent RAM spikes; missing crash logs due to automated purging.

Android Platform (CamelSpy Trojan). Vector: Group UAC-0244. Weaponized packages disguise as defense software (tactical maps, logistics, fuel vouchers). Execution: Exploits Android Accessibility Services to execute screen scraping, log keystrokes, and directly extract local Signal/WhatsApp SQLite databases. Network IoCs: Background WebSockets to C2 for exfiltrating /sdcard/DCIM/data. System IoCs: Rogue apps with randomized names requesting unauthorized READ_SMS, RECORD_AUDIO, and ACCESS_FINE_LOCATION privileges.

Minimizing the risk of compromise to the mobile segment of CII requires the implementation of strict preventive protocols. To minimize the risk of compromise to the mobile segment of CII, a two-level matrix of preventive protocols is being implemented, delineating personnel procedures for both normal operations and system failure scenarios: operational-dispatch and engineering personnel at thermal power plants (TPP) and regional energy companies: required to conduct audits of active sessions (connected devices) in the «Signal» messenger every 12 hours; route traffic exclusively through encrypted corporate VPN gateways when connecting to Starlink nodes; and immediately activate radio isolation mode (airplane mode) upon detecting anomalies in the ICS/SCADA environment. Strictly prohibited are the unauthorized transmission of photos or videos of infrastructure damage to public communication channels, logging into personal cloud storage services (iCloud, Google Drive) from work terminals, and storing authentication credentials in unencrypted (cleartext) form. Tactical and officer-level personnel of the Armed Forces of Ukraine: required to maintain permanent hardware-level activation of "Lockdown Mode"; access restricted situational awareness networks strictly via dynamic tokens; and locally wipe cached map data upon completion of combat duty. Strictly prohibited are the installation of third-party software packages (*.apk) from external sources and the transmission of precise deployment coordinates outside of specialized, secure cryptographic applications.

Ensuring the integrity of the CII mobile perimeter is unattainable without the deployment of centralized MDM systems. Under conditions of degraded external communication channels, the architectural autonomy of the management platform becomes the critical selection parameter. Comparative technical analysis of MDM/EMM management architectures the comparative analysis of cloud-based (Microsoft Intune) and on-premises (MobileIron/Ivanti) MDM architectures, evaluated against resilience criteria during transport network degradation, revealed the following patterns: Microsoft Intune (Cloud-Native): Operates on a SaaS model, requiring constant integration with Azure ID and external Microsoft repositories. It exhibits critically low fault tolerance during outages: if backbone communication links are severed, endpoints lose connectivity to the management cloud, completely blocking the application and updating of security policies. Isolation of operational environments is implemented at the logical level via App Protection policies, which are strictly dependent on the device's global network status. A key advantage is its deep integration with Microsoft Defender for Endpoint, providing proactive threat telemetry analytics. MobileIron/Ivanti (Hybrid/On-Premise): Deployed locally within the secure perimeter of CII. It demonstrates high fault tolerance during total infrastructure failure, as the local MobileIron Core server continues to broadcast compliance policies via internal backup microwave links. It ensures strict environment containerization based on AppConnect technology, physically isolating the work workspace from the personal one at the file system level. It integrates with local MTD modules for passive anomaly analysis within industrial automation protocols.Scientifically justified choice for CII assets (e.g., thermal power plants) under the permanent threat of destructive kinetic impacts leans heavily toward the On-Premise architecture of MobileIron/Ivanti. Local sovereignty over the management server guarantees the continuity of endpoint compliance monitoring even during a complete network isolation of the national energy segment from the global Internet.

Traditional approaches based on passive perimeter defense exhibit absolute failure against specialized exploit kits like DarkSword and trojans of the CamelSpy class under blackout conditions. Transitioning to a paradigm of Cyber Resilience based on ZTA, strict MDM containerization, and automated behavioral MTD analysis represents the only viable scenario for sustaining the operability of the state's energy and defense complexes. To mitigate and prevent such incidents, the following systematic solutions must be enforced: 1) For management and dispatch personnel at CII facilities, a mandatory protocol (a "Lockdown" mode) must be implemented to ensure continuous device operation under maximum functional restrictions. The MDM policies should restrict the execution of dynamic scripts, disable media file previews in messaging apps, and block unauthorized connections via USB interfaces; 2) Energy companies should transition from cloud-based MDM systems to hybrid or fully on-premise infrastructure setups. Local authentication servers must be integrated with backup satellite communication channels (e.g., Starlink) to create a closed, autonomous

control loop that is resilient to physical network failures; 3) Information security units at CII facilities should integrate local monitoring systems with state-run cyber-threat data aggregation platforms to enable the automated, real-time receipt of IoCs from CERT-UA.

References

1. CERT-UA. Аналітичний звіт щодо кібероперацій російської федерації проти КІІ України за І квартал 2026 року.URL: https://cert.gov.ua/
2. Національна академія наук України Інститут проблем моделювання в енергетиці Вплив деградації енергомереж на стійкість ІТ-інфраструктури диспетчеризації. Київ, 2025. — С. 45-52.
3. NIST Special Publication 800-124 Revision 2. Guidelines for Managing the Security of Mobile Devices in the Enterprise. National Institute of Standards and Technology, 2023.
4. Держспецзв'язку України. Специфікація експлойт-кита DarkSword та заходи протидії атакам, що таргетуються, на iOS пристрої. Технічний бюлетень CERT-UA #1142, 2026.

Ахмедов А.Ф.
студент групи ІКБ-32, ННІКБЗІ ДУІКТ,
Київ, Україна

ФОРМУВАННЯ КРИТЕРІЇВ ДЕТЕРМІНОВАНОГО БІОМЕТРИЧНОГО ПРОТОКОЛУ В IoT-ІНФРАСТРУКТУРІ

Розглянуто концепцію детермінованої біометричної автентифікації в IoT-інфраструктурі без збереження специфічних для суб'єкта допоміжних даних. Проаналізовано обмеження відомих підходів до захисту біометричних даних, а також класичного векторного квантування та локально-чутливого хешування. За результатами аналізу сформовано комплекс технічних та криптографічних критеріїв детермінованого біометричного протоколу, дотримання яких є передумовою реалізації концептуальної моделі цифрової ідентифікації людини identity-as-existence («ідентичність як факт існування») та підвищення стійкості IoT-інфраструктури з автономними периферійними вузлами.

Ключові слова: IoT-інфраструктура, біометрична автентифікація, детермінований протокол, допоміжні дані, шумостійкість, дискретизація біометричних ознак, identity-as-existence.

Домінування серверних моделей автентифікації зумовлює вразливість IoT-інфраструктури через обмежену інтероперабельність, наявність єдиної точки відмови та концентрацію біометричних даних у централізованих сховищах. Модель identity-as-existence та гіпотеза про можливість створення детермінованого біометричного протоколу генерації стабільного коду високої ентропії без збереження специфічних для суб'єкта допоміжних даних [1] окреслюють напрямок підвищення стійкості IoT-інфраструктури, проте залишають невирішеним питання формалізації відповідних вимог.

Відомі способи перетворення неперервних біометричних ознак у дискретне подання мають обмеження щодо детермінованої відтворюваності в умовах шуму вимірювання та вимагають формування, використання та зберігання специфічних для суб'єкта допоміжних даних, тоді як класичне векторне квантування та локально-чутливе хешування є вразливими до внутрішньокласових варіацій. Відсутність узгодженого набору технічних і криптографічних критеріїв для детермінованого біометричного протоколу в автономних периферійних вузлах ускладнює оцінювання наявних технологій та стримує цілеспрямований пошук перспективних рішень.

Метою роботи є аналіз методів дискретизації біометричних ознак та обґрунтування комплексу критеріїв детермінованого біометричного протоколу автентифікації в IoT-інфраструктурі.

Способи захисту біометричних даних Fuzzy Vault, Fuzzy Commitment та Fuzzy Extractors [2, 3] використовують коди з виправленням помилок або

поліноміальні конструкції з метою стійкості до шуму вимірювання, однак потребують формування та зберігання специфічних для суб'єкта публічних допоміжних даних, що збільшує поверхню атаки, зумовлює ризики витоку статистичної інформації про вихідний біометричний вектор та ускладнює досягнення технологічної автономності периферійних IoT-вузлів. Крім того, обмежена виправна здатність застосованих кодів накладає суворі обмеження на максимально допустимий рівень шуму, що знижує гнучкість системи.

Класичне векторне квантування та локально-чутливе хешування [4] не потребують зберігання індивідуальних допоміжних даних, проте виявляються чутливими до положення вхідного вектора відносно меж дискретизуючих структур. Навіть незначний внутрішньокласовий шум може змістити вектор у сусідню комірку або хеш-кошик, внаслідок чого формується інший дискретний ідентифікатор, що порушує критично важливу для автентифікації технологічну вимогу детермінованої відтворюваності.

Отже, наявні науково-практичні методи не забезпечують одночасної детермінованої відтворюваності стабільного коду високої ентропії в умовах внутрішньокласового шуму за відмови від збереження специфічних для суб'єкта допоміжних даних, що вимагає розроблення системи критеріїв детермінованого біометричного протоколу автентифікації.

Результати дослідження дозволяють формалізувати комплекс критеріїв детермінованого біометричного протоколу генерації стабільного коду високої ентропії без збереження специфічних для суб'єкта допоміжних даних.

1. Детермінованість. Повторні вимірювання біометричних ознак того самого суб'єкта в межах властивого сенсорам внутрішньокласового шуму мають приводити до формування ідентичного дискретного подання.
2. Відсутність специфічних для суб'єкта допоміжних даних. Біометричний протокол виключає формування, зберігання чи передачу будь-яких особистих даних, оскільки всі операційні параметри мають бути суто системними.
3. Універсальність системних параметрів. Параметри перетворення вхідного біометричного вектора в дискретне подання визначаються виключно класом сенсора і є єдиними для всіх суб'єктів, що унеможливлює індивідуальне навчання або калібрування системи.
4. Шумостійкість. Протокол гарантує тотожність дискретного подання, якщо збурення вхідного вектора не перевищують межі внутрішньокласових варіацій IoT-сенсорів.

5. Обчислювальна придатність. Алгоритм забезпечує низьку обчислювальну складність для функціонування на низькоресурсних автономних периферійних IoT-пристроях без залучення додаткової пам'яті чи апаратних ресурсів.
6. Компактність дискретного подання. Структура сформованого ідентифікатора є фіксованою та компактною, що оптимізує операції зберігання й порівняння даних на вузлах IoT з обмеженим обсягом обчислювальних ресурсів.

Комплексне дотримання зазначених критеріїв є передумовою реалізації моделі цифрової ідентифікації людини identity-as-existence та забезпечення автономної детермінованої автентифікації в IoT-інфраструктурі.

Проаналізовано критичні обмеження ключових підходів до перетворення неперервних біометричних ознак у стабільне дискретне подання. Показано, що жодне з відомих рішень не забезпечує одночасно відсутності специфічних для суб'єкта допоміжних даних, детермінованості, шумостійкості та прийнятної обчислювальної складності для малопотужних IoT-пристроїв. Сформовано комплекс із шести критеріїв детермінованого біометричного протоколу в IoT-інфраструктурі, які становлять теоретичну та практичну основу для розробки й об'єктивного оцінювання нових технічних рішень.

Подальші дослідження доцільно спрямувати на розроблення методу перетворення біометричних ознак, який повністю задовольнятиме зазначений комплекс критеріїв, а також на його експериментальну верифікацію в умовах функціонування автономних периферійних вузлів IoT.

Перелік посилань:

1. Ахмедов А.Ф. Децентралізована біометрична ідентифікація як механізм підвищення стійкості IoT-інфраструктури // Всеукраїнська науково-практична конференція «Цифрова трансформація кібербезпеки» від 29 квітня 2026 р. Київ: ДУІКТ, 2026.
URL: https://duikt.edu.ua/uploads/p_3086_30075970.pdf
2. Juels A., Sudan M. A Fuzzy Vault Scheme // Designs, Codes and Cryptography, 2006. Vol. 38, P. 237-257.
URL: https://doi.org/10.1007/s10623-005-6343-z
3. Dodis Y., Ostrovsky R., Reyzin L., Smith A. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data // SIAM Journal on Computing, 2008. Vol. 38, No. 1. P. 97-139.
URL: https://doi.org/10.1137/060651380
4. Indyk P., Motwani R. Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality // Proceedings of the 30th Annual ACM Symposium on Theory of Computing, New York: ACM, 1998. P. 604-613.
URL: https://doi.org/10.1145/276698.276876

Мельник В.І.
студент групи ІКБМ-61, ННКБЗІ ДУІКТ,
Київ, Україна

ТЕХНОЛОГІЯ АВТОМАТИЗОВАНОГО РЕАГУВАННЯ НА ФІШИНГОВІ АТАКИ З ВИКОРИСТАННЯМ SOAR

Фішинг – це різновид кібератаки, що використовує шахрайські електронні листи, текстові повідомлення, веб-ресурси або телефонні дзвінки, щоб шляхом обману користувачів отримати конфіденційну інформацію. Фішингові атаки використовують методи соціальної інженерії, чим суттєво відрізняються від інших типів кібератак.Традиційне реагування на фішингові атаки передбачає ручний аналіз підозрілих повідомлень, перевірку індикаторів компрометації та виконання заходів з нейтралізації загроз, що потребує значних часових і людських ресурсів. Автоматизація цього процесу за допомогою SOAR-систем дає змогу прискорити оброблення інцидентів, зменшити вплив людського фактору та підвищити ефективність використання.

Ключові слова: фішинг, SOAR-системи, соціальна інженерія, плейбук.

За результатами щорічного аналітичного звіту Verizon DBIR фішинг залишається одним з найпоширеніших початкових векторів компрометації організацій [1]. Використання генеративних моделей дозволяє зловмисникам масово створювати переконливі листи збільшуючи кількість звернень користувачів до SOC аналітиків. Традиційний процес обробки одного звернення включає наступні дії: отримання листа від користувача, вивантаження службових заголовків, перевірку результатів автентифікації відправника, аналіз посилань та вкладень в ізольованому середовищі, звірення з базами індикаторів компрометації, пошук такого ж листа в інших поштових скриньках та його видалення, блокування шкідливих доменів і складання звіту. Кожна така дія може виконуватись у різних інтерфейсах, що вимагає від працівника від десятьох хвилин до кількох годин уваги. Протягом цього часу лист може залишатись доступним іншим адресантам, що збільшує ймовірність переходу за шкідливим посиланням і збільшує показник середнього часу реагування.

Для оптимізації цього процесу можна використовувати SOAR-системи, які за допомогою програмних інтерфейсів об’єднують засоби захисту та інформаційні системи організації, виконують повторювані дії без участі людини й ведуть облік інцидентів. Логіка реагування на фішингові інциденти описується у вигляді плейбуків – формалізованих сценаріїв, що складаються з послідовних дій, умовних переходів і точок підтвердження аналітиком.

Переважно технологія автоматизованого реагування на фішингову атаку складається з п’яти етапів. На першому етапі плейбук запускається автоматично: від клавіші «Повідомити про фішинг» у поштовому клієнті, від поштового шлюзу або від правила кореляції в SIEM-системі. На другому етапі з повідомлення вилучаються артефакти: адреси відправника та шляху доставки, результати

перевірок SPF, DKIM та DMARC, усі посилання, імена та хеш-суми вкладень. На третьому етапі виконується запит до сервісів репутації, перевірка вкладень у пісочниці, перевірка в платформі обміну індикаторами MISP та аналіз віку домену. На четвертому етапі за результатами перевірка обчислюється інтегральна оцінка ризику повідомлення:

$$R = \sum_{i=1}^{n} w_i \cdot s_i, \sum_{i=1}^{n} w_i = 1 \quad (1)$$

де s_i – нормалізована оцінка і-го джерела, w_i – ваговий коефіцієнт цього джерела, n – кількість джерел. Порогові значення T_1 і T_2 визначають гілку плейбука: при R < T_1 повідомлення вважається безпечним, інцидент закривається; при $T_1 \leq R < T_2$ випадок передається аналітику зі зібраним контекстом; при $R \geq T_2$ автоматично виконуються дії стримування – пошук та видалення листа в усіх пошукових скриньках, додавання посилання і домену до списків блокування на проксі-сервері та міжмережевому екрані, а за наявності введення облікових даних – примусове скидання пароля й відкликання активних сесій.

На п’ятому етапі система автоматично повідомляє користувача про результат, формує тікет із повним журналом дій та передає зібрані індикатори в базу знань для подальшого використання.

Разом із тим автоматизація має обмеження, які необхідно враховувати при проєктуванні. Сервесні облікові записи SOAR мають широкі повноваження щодо видалення пошти, зміни правил мережевого захисту, тому для них застосовується принцип мінімальних привілеїв, а їхні дії повністю журналюються. Для критичний дій при середніх значенням ризику зберігається підтвердження дій людиною. Плейбуки слід тестувати на навчальних кампаніях, оскільки помилка в логіці може призвести до видалення легітимних листів. Також варто зазначити, що процес реагування має відповідати рекомендаціям NIST SP 800-61 щодо обробки інцидентів інформаційної безпеки [2].

Отже, впровадження SOAR-систем перетворює реагування на фішинг на керований та відтворюваний процес. Автоматичне вилучення артефактів, оцінка ризику та стримування загрози в масштабі всієї організації скорочують час реакції, розвантажують аналітиків від рутинних завдань і дозволяють спрямувати їхні зусилля на розслідування складних атак.

Перелік посилань:

1. Verizon: Wireless, Internet, TV and Phone Services | Official Site. URL: https://www.verizon.com/business/resources/executivebriefs/2026-dbir-executive-summary.pdf (дата звернення: 22.10.2026).

2. *NIST Technical Series Publications*. URL: https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-61r2.pdf (дата звернення: 22.10.2026).

Кіх М. В.,
д. ф., асистент каф. БІТ ННІКТА
НУ «Львівська політехніка»,
Львів, Україна

ВИКОРИСТАННЯ АУДІОЕНТРОПІЇ ДЛЯ ФОРМУВАННЯ ПОЧАТКОВОГО СТАНУ ГІБРИДНОГО ГЕНЕРАТОРА ПСЕВДОВИПАДКОВИХ ПОСЛІДОВНОСТЕЙ

Розглянуто проблему недостатньої або неконтрольованої ентропії початкового стану генераторів псевдовипадкових послідовностей, що може створювати передумови для передбачення вихідних даних і компрометації криптографічних механізмів. Систематизовано вимоги до придатності аудіосигналу як джерела початкового стану гібридного генератора: вимірювальні критерії вихідного сигналу, а також оцінювання мінімальної ентропії та безперервні перевірки працездатності джерела за NIST SP 800-90B. Визначено чинники та показники запланованої експериментальної перевірки впливу акустичного середовища й пристрою запису на частку придатних записів.

Ключові слова: кібербезпека, аудіоентропія, генератор псевдовипадкових послідовностей, початковий стан, джерело ентропії, повторне відтворення сигналу.

Проблема кібербезпеки полягає в тому, що стійкість криптографічних механізмів залежить не лише від алгоритму, а й від непередбачуваності початкового стану генератора псевдовипадкових послідовностей (ГПВП). Якщо початкове значення має низьку ентропію, повторюється або може бути відтворене зловмисником, вихід генератора потенційно стає передбачуваним. Це створює ризики для криптографічних ключів, одноразових значень, токенів та інших даних, що застосовуються в автентифікації й захищеному обміні. Отже, конкретна проблема, яку розглядає робота, — зменшення ризику компрометації ГПВП через слабке, повторно використане або контрольоване джерело початкової ентропії.

Аудіосигнал, отриманий мікрофоном, містить фізичні варіації — квантувальний і тепловий шум тракту, — придатні для формування початкових даних. Водночас аудіо не можна автоматично вважати випадковим: тиша, циклічний звук, автоматичне регулювання підсилення, шумозаглушення та фільтрація операційної системи або навмисно поданий запис можуть знижувати непередбачуваність. Тому завдання полягає не в перетворенні відліків на біти, а в перевірці придатності входу.

Мета роботи — визначити вимоги до придатності аудіосигналу як джерела початкового стану гібридного ГПВП і сформулювати вимірювані критерії контролю. Аудіозапис (48 кГц, 16 біт) перетворюється на біти додаванням за модулем 2 молодших бітів трьох послідовних відліків. З отриманих бітів формуються початкові значення N генераторів Джиффі, кожен із яких має три регістри зсуву зі зворотним зв’язком довжиною 59, 53 і 47 біт, а виходи об’єднуються нелінійною булевою функцією [3]. Повний початковий стан

становить N × 159 біт (954–1431 біт для N = 6–9). Ці біти можна вважати ентропією лише тією мірою, якою її підтверджує оцінювання: послідовність із 1431 біта з мінімальною ентропією 0,5 біт/біт містить приблизно 715 біт ентропії, що значно менше за розмір стану.

Для аналізу безпеки необхідно розділяти три рівні перевірки: (1) придатність і непередбачуваність джерела ентропії; (2) коректність формування початкового стану та захист від повторного використання; (3) статистичні властивості вихідної послідовності. Набори NIST SP 800-22, Diehard і TestU01 виявляють статистичні відхилення у виході, проте проходження тестів не доводить криптографічної стійкості [2], а достатність ентропії на вході оцінюють окремо. Для джерела ентропії NIST SP 800-90B передбачає оцінювання мінімальної ентропії та безперервні перевірки працездатності, зокрема тест повторень (RCT) і адаптивний тест частки (APT) [1].

Таблиця 1. Загрози аудіоджерелу та вимірювані критерії контролю

Загроза	**Ознака в сигналі або бітах**	**Критерій контролю**
Тиша або майже сталий сигнал	Низький рівень RMS, висока частка нульових відліків, малий набір різних значень	Відхилення запису за порогами RMS і частки нулів; мінімальна ентропія нижча за задану (SP 800-90B)
Дублювання каналів, обробка сигналу в ОС	Збіг відліків лівого й правого каналів; нетипова статистика молодших бітів	Контроль збігу каналів; вимкнення автоматичного регулювання підсилення та шумозаглушення; порівняння статистики молодших бітів
Повторне відтворення запису	Збіг фрагментів сигналу в різних сеансах	Облік використаних сегментів (хеші); унікальність сеансу; захист тракту захоплення
Навмисний акустичний вплив	Знижена непередбачуваність за контрольованого шуму	Модель загроз; комбінування з незалежним джерелом ентропії

Порогові значення критеріїв залежать від пристрою та середовища, тому їх слід визначати експериментально, а не задавати універсально. У поточній реалізації запис вважається придатним, якщо виконано п’ять статистичних перевірок: ентропія Шеннона $H \geq 0{,}98$, χ^2-тест і тест серій ($p > 0{,}01$), відносна різниця кількості нулів і одиниць та середня автокореляція за лагами 1–100 не перевищують заданих порогів; у разі невдачі запис повторюють (за замовчуванням до п’яти спроб) [3]. Частку записів, що проходять верифікацію з першої спроби, у типових умовах (офіс, житлове приміщення) попередньо оцінено як понад 95 %, що потребує систематичної перевірки.

Ці перевірки є статистичними й не вимірюють ентропію джерела: ентропія Шеннона відображає лише частоти нулів і одиниць у вихідній послідовності. Так, для шести тестових записів середня автокореляція становила 0,0023–0,0028, тобто значно нижче порога, тому цей критерій відсікає лише грубі дефекти. Тому

запропоновано доповнити їх оцінкою мінімальної ентропії (оцінювач найчастішого значення за SP 800-90B) та діагностикою вихідного сигналу. Для перевірки впливу умов заплановано серії записів в акустичних середовищах (тиха кімната, офіс, вулиця) та на різних пристроях за незмінних частоти дискретизації й підсилення; показники — мінімальна ентропія, частка записів, що пройшли верифікацію, і середня кількість спроб.

Ці заходи є вимогами до реалізації та подальшої перевірки, а не гарантією безпеки. Якщо аудіо використовується як єдине джерело, безпека залежить від акустичного середовища та можливості атакувальника впливати на нього, тому аудіо доцільно розглядати як компонент формування початкового стану й за потреби поєднувати з незалежним криптографічно придатним джерелом ентропії [3].

Висновки

Виокремлено проблему кібербезпеки — ризик передбачуваності криптографічних значень через недостатню або повторно використану ентропію початкового стану ГПВП. Придатність аудіосигналу як джерела слід підтверджувати вимірювальними критеріями: показниками вихідного сигналу, оцінкою мінімальної ентропії та безперервними перевірками за NIST SP 800-90B; статистичних тестів виходу для цього недостатньо. Подальша робота — експериментальне оцінювання впливу акустичного середовища й пристрою запису на ентропію джерела та перевірка сценаріїв контрольованого аудіовпливу.

Список використаних джерел:

1. Turan M. S. et al. *Recommendation for the Entropy Sources Used for Random Bit Generation*. NIST Special Publication 800-90B. 2018. DOI: https://doi.org/10.6028/NIST.SP.800-90B.

2. Bassham L. E. et al. *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. NIST Special Publication 800-22, Revision 1a. 2010. DOI: https://doi.org/10.6028/NIST.SP.800-22r1a.

3. Кіх М., Нємкова О. Практична схема гібридного генератора псевдовипадкових послідовностей на основі аудіоентропії та нелінійних булевих функцій. *Кібербезпека: освіта, наука, техніка*. 2025. № 2 (30). С. 180–194. DOI: https://doi.org/10.28925/2663-4023.2025.30.959.

Молнар Віталій Васильович,
аспірант кафедри захисту інформації,
Національний університет «Львівська політехніка»

ПОРІВНЯННЯ СПОСОБІВ АВТОМАТИЗОВАНОГО РЕАГУВАННЯ НА НЕБЕЗПЕЧНІ ЗМІНИ КОНФІГУРАЦІЇ AWS

Анотація. Розглянуто автоматизоване реагування на небезпечні зміни конфігурації ресурсів AWS. Порівняно періодичну перевірку стану ресурсів і запуск коригувальної дії за подією зміни. Описано взаємодію хмарних сервісів, умови запуску обробників та перевірку результату. На прикладах публічного доступу до сховища та відкритого віддаленого доступу пояснено відмінності між цими підходами. Обґрунтовано їх застосування для планового контролю конфігурації та усунення відомих порушень без очікування чергової періодичної перевірки.

Публічний доступ до сховища або надмірно відкрите правило мережевого доступу можуть з’явитися після звичайної зміни конфігурації. Щоб усунути порушення автоматично, потрібно визначити, як його виявити, яку дію виконати та як перевірити результат. Від способу виявлення залежить момент запуску цієї дії.

Рекомендації NIST пов’язують виявлення, реагування та відновлення із загальним управлінням кіберризиками [1]. Для хмарної системи це означає потребу узгодити автоматизовані дії з її політикою безпеки.

Періодична перевірка. У цьому варіанті конфігурацію ресурсу перевіряють за розкладом. AWS Config зіставляє її із заданим правилом. Якщо виявлено порушення, його усунення виконує сценарій AWS Systems Manager Automation. Важливо, що AWS Config підтримує також запуск перевірок за змінами конфігурації [2]. Далі розглядається саме періодичний варіант.

Реагування за подією. AWS CloudTrail реєструє виклики програмних інтерфейсів хмарних сервісів. Події таких викликів надходять до Amazon EventBridge, де правила визначають їх подальше передавання [3]. Подія зміни конфігурації може запустити функцію AWS Lambda з наперед визначеною коригувальною дією. Обидва маршрути реагування показано на рис. 1.

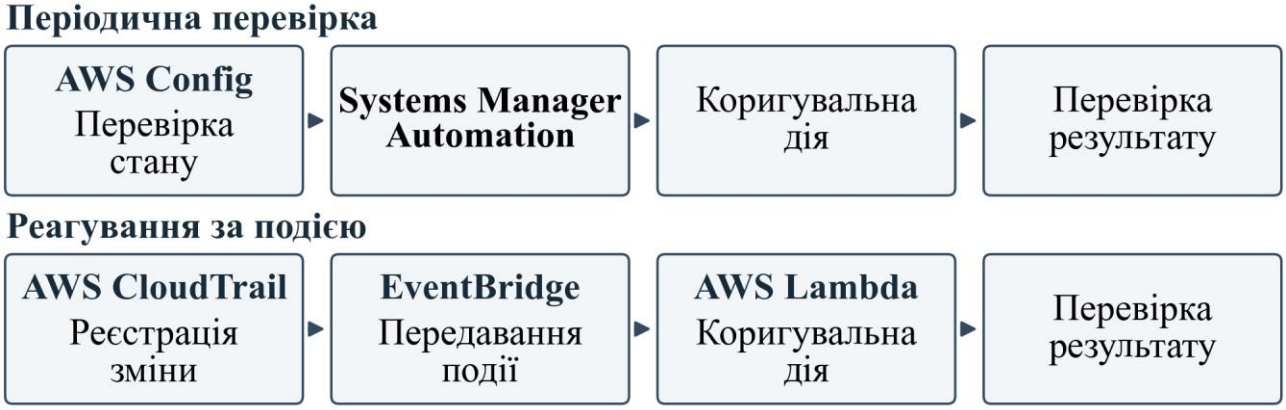


Рис. 1. Маршрути автоматизованого реагування в AWS

Періодична перевірка зручна для планового контролю відповідності. Вона оцінює стан ресурсу на момент перевірки та не потребує окремого правила для кожного способу, яким виникло порушення. Водночас зміна, внесена між перевірками, може залишатися невиявленою до наступного оцінювання. Частоту перевірок потрібно вибирати з урахуванням призначення ресурсу та наслідків відкритого доступу.

Для подійного маршруту необхідно налаштувати реєстрацію потрібних операцій, їх доставлення до EventBridge та умови виклику обробника. Затримка тут залежить від передавання події та виконання обробника. Очікування чергової періодичної перевірки до цього маршруту не входить.

Приклади застосування. Описані архітектури використано в авторському дослідженні, опублікованому у співавторстві [4]. Для сховища Amazon S3 розглядалося усунення дозволу публічного доступу з політики бакета. Для групи безпеки вилучалося правило, що відкривало SSH-доступ із будь-якої адреси. Залежно від архітектури ці зміни виконував сценарій Systems Manager Automation або функція Lambda. Кінцевий стан ресурсу в обох випадках мав відповідати тій самій політиці безпеки.

Правило реагування має пов’язувати конкретне порушення з дозволеною коригувальною дією. У ньому визначають ресурс, умову запуску та очікуваний результат. Відмінності маршрутів наведено в табл. 1.

Таблиця 1. Порівняння маршрутів реагування

Ознака	Періодична перевірка	Реагування за подією
Запуск	Настання часу перевірки	Надходження події зміни
Дані для рішення	Поточна конфігурація ресурсу	Дані події та стан ресурсу
Виконавець	Systems Manager Automation	AWS Lambda
Затримка	Очікування перевірки та виконання сценарію	Доставлення події та виконання обробника

Для S3 потрібно розрізняти дозволений публічний доступ і помилкову зміну політики. Відкритість сховища зі статичними матеріалами може бути передбачена його призначенням. Правило контролю має враховувати такі винятки. Для сховища з обмеженим доступом сценарій коригування повинен прибирати саме небезпечний дозвіл, зберігаючи інші потрібні права. Тому правила реагування слід прив’язувати до політики конкретного ресурсу та перевіряти до ввімкнення автоматичних змін.

Для SSH-доступу завдання полягає в усуненні надмірно широкого мережевого дозволу. Це не обов’язково потребує повного вимкнення віддаленого адміністрування. Дозволені підключення з погоджених адрес мають зберігатися. Сценарій повинен визначати конкретне небезпечне правило, а перевірка результату має підтверджувати як усунення порушення, так і збереження потрібного доступу.

Подія описує операцію, яка вже відбулася. До моменту її оброблення конфігурацію могли змінити повторно або виправити вручну. Сам факт надходження події тому недостатній для безумовного коригування. Обробник має перевірити, чи зберігається порушення та чи стосується воно потрібного ресурсу. Для періодичного контролю аналогічно важлива відповідність між станом, за яким ухвалено рішення, і станом на момент виконання зміни.

Вибір способу реагування. Періодичний контроль доцільний, коли потрібно регулярно оцінювати ресурси за спільними правилами відповідності. Подійний маршрут доцільний для відомих небезпечних змін, які можна розпізнати за даними події та усунути визначеною дією. Його застосування залежить від наявності потрібних подій і повноти налаштувань їх оброблення.

Підходи можна поєднувати. Подія запускає усунення відомого порушення, а планова перевірка додатково оцінює відповідність поточного стану. Якщо обидва маршрути можуть змінювати той самий ресурс, їхні дії потрібно узгодити. Обробник має перевіряти, чи зберігається порушення, і виконувати лише необхідні зміни. Таке поєднання потребує окремої перевірки.

Завершення сценарію автоматизації не слід ототожнювати з підтвердженням безпечного стану. Після коригування потрібно перевірити конфігурацію та зберегти результат виконання. Повноваження обробника мають обмежуватися потрібними ресурсами й операціями. Це особливо важливо для систем, де частина публічного доступу дозволена: автоматизація повинна враховувати погоджені винятки.

Слід також передбачити дії на випадок невдалого коригування. Якщо обробнику бракує повноважень або ресурс недоступний для зміни, відсутність підтвердження має бути помітною відповідальному адміністратору. Для подальшого розбору потрібні відомості про ресурс, виявлене порушення, виконану операцію та причину помилки. Це дає змогу продовжити усунення вручну та скоригувати правила автоматизації без припущення, що кожний запуск завершується успішно.

Висновки. Періодична перевірка та реагування за подією відрізняються насамперед способом запуску усунення порушення. Плановий контроль оцінює відповідність ресурсів у визначені моменти, а подійний маршрут пов’язує зміну конфігурації із запуском обробника. Для обох потрібні чіткі правила коригування та перевірка кінцевого стану. Вибір архітектури слід узгоджувати з характером порушень і вимогами до захисту конкретної системи.

Список використаних джерел

1. 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. DOI: 10.6028/NIST.SP.800-61r3.
2. AWS. Components of an AWS Config Rule. AWS Config User Guide. URL: https://docs.aws.amazon.com/config/latest/developerguide/evaluate-config_components.html.
3. AWS. AWS service events delivered via AWS CloudTrail. Amazon EventBridge User Guide. URL: https://docs.aws.amazon.com/eventbridge/latest/userguide/eb-service-event-cloudtrail.html.
4. Molnar V., Sabodashko D., Opirskyy I. A comparative performance analysis of automated cloud security remediation architectures for AWS cloud environments. International Journal of Information and Computer Security. 2026. Vol. 30, No. 3. P. 373–393. DOI: 10.1504/IJICS.2026.154858.

Костенко Яків Олександрович
Здобувач вищої освіти групи УБДМ-61, ДУІКТ

СИСТЕМА МОНІТОРИНГУ ТА АНАЛІЗУ АНОМАЛЬНОЇ ПОВЕДІНКИ КОРИСТУВАЧІВ У КОРПОРАТИВНИХ МЕРЕЖАХ НА ОСНОВІ SYSTECA UAM

У сучасних умовах стрімкої цифровізації бізнес-процесів та переходу багатьох компаній на гібридний формат роботи забезпечення безперервності роботи компаній стало одним із головних завдань кібербезпеки. Традиційно фокус уваги фахівців з інформаційної безпеки був зміщений на захист від зовнішніх атак, таких як DDoS-атаки, фішингові кампанії, використання вразливостей нульового дня та експлойтів. Проте реальна практика та світова статистика інцидентів свідчать, що значна частина критичних порушень конфіденційності, цілісності та доступності даних пов'язана саме з внутрішніми загрозами. Скомпрометовані облікові записи легітимних співробітників, зловмисні дії інсайдерів або елементарна халатність персоналу можуть завдати бізнесу непоправної фінансової та репутаційної шкоди. Саме тому класичного периметрального захисту мережі вже недостатньо — сучасним компаніям потрібен глибинний, безперервний контроль за тим, що відбувається безпосередньо всередині корпоративної мережі на кінцевих точках (endpoints).

Використання спеціалізованих рішень класу UAM (User Activity Monitoring — моніторинг активності користувачів) є логічним та критично необхідним кроком для ефективного управління вищезазначеними ризиками. У цій доповіді розглядається сучасний підхід до побудови системи моніторингу та аналізу поведінки користувачів на базі комплексної платформи Systeca UAM (раніше відомої як Ekran System).

Головна перевага впровадження платформи Systeca полягає в тому, що система не просто збирає розрізнені текстові логи подій, а фіксує дії користувачів у зручному для аналізу індексованому відеоформаті з глибокою прив'язкою до метаданих. До таких метаданих належать натискання клавіш (кейлоггінг), запущені процеси, відкриті файли, відвідані веб-ресурси та введені команди в терміналі. Це дозволяє офіцерам служби безпеки не лише констатувати факт інциденту, але й бачити повний візуальний контекст дій співробітника або зовнішнього підрядника (Third-party vendor), що є критично важливим для швидкого та об'єктивного розслідування інцидентів інформаційної безпеки.

Проте для забезпечення справжньої кіберстійкості інфраструктури недостатньо лише пасивно фіксувати події — потрібно вміти діяти на випередження та виявляти загрози на ранніх етапах. Для вирішення цього завдання в Systeca UAM інтегровано потужний модуль UEBA (User and Entity

Behavior Analytics — аналітика поведінки користувачів та сутностей). За допомогою алгоритмів машинного навчання (Machine Learning) система протягом певного часу вивчає типову повсякденну поведінку кожного працівника і формує його унікальний базовий «нормальний» профіль.

Після завершення періоду навчання будь-яке суттєве відхилення від цього профілю розцінюється системою як аномалія. Наприклад, якщо бухгалтер, який зазвичай працює в обліковій системі з 9:00 до 18:00, раптом о другій годині ночі починає масово копіювати клієнтські бази даних на несанкціонований USB-накопичувач або намагається отримати доступ до нетипових для його ролі серверів та баз даних, система миттєво генерує алерт високого рівня критичності.

З точки зору забезпечення безперервності бізнесу (Business Continuity), ключову роль у таких рішеннях відіграє функція автоматизованого реагування на інциденти. Платформа Systeca дозволяє гнучко налаштувати правила, за якими у разі виявлення критичної аномалії система може вжити автоматичних превентивних заходів: миттєво заблокувати поточну сесію користувача, вивести на екран попереджувальне повідомлення про порушення політики безпеки або заблокувати підключений периферійний пристрій ще до того, як конфіденційні дані покинуть межі корпоративної мережі.

Такий комплексний підхід кардинально змінює загальну стратегію управління ризиками в організації. По-перше, знижується час реакції (MTTR): замість того, щоб тижнями розслідувати наслідки витоку даних, служба безпеки реагує на загрозу в режимі реального часу, маючи під рукою відеодокази. По-друге, мінімізуються простої: завдяки детальному та прозорому аудиту розслідування інцидентів проходить набагато швидше, що дозволяє оперативніше відновлювати нормальну роботу IT-інфраструктури. По-третє, забезпечується надійний контроль підрядників та повна прозорість дій зовнішніх IT-спеціалістів.

Висновок. Впровадження сучасної системи моніторингу та аналізу аномальної поведінки на основі платформи Systeca UAM дозволяє трансформувати корпоративну інформаційну безпеку від застарілого реактивного до ефективного проактивного формату. Це не просто інструмент тотального контролю, а невід'ємна складова комплексної стратегії кіберстійкості, яка гарантує, що людський фактор, незалежно від того, чи є він навмисним, чи випадковим, не стане причиною фатальної зупинки бізнес-процесів компанії.

Література:

1. Офіційна документація та технічні специфікації платформи Syteca (Ekran System).

2. Корченко О.Г. Системи захисту інформації: принципи побудови та управління ризиками. – К.: ДУІКТ, 2023. – 256 с.

3. NIST Special Publication 800-137. Information Security Continuous Monitoring (ISCM) for Federal Information Systems and Organizations.

Ялбуган Ф. Д.
студент групи ІКБм-11, ННІКЗІ ДУІКТ, Київ, Україна

ГІБРИДНЕ ОЦІНЮВАННЯ КІБЕРРИЗИКІВ ОРГАНІЗАЦІЙ-КОНТРАГЕНТІВ ЗА OSINT-ДАНИМИ З ВИКОРИСТАННЯМ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ

Розглянуто задачу попереднього оцінювання кіберризиків організацій-контрагентів за відкритими джерелами. Запропоновано гібридний підхід, у якому велика мовна модель лише виділяє ризикові події з доказовими цитатами, а підсумкову оцінку обчислює детермінована модель з урахуванням основної категорії ризику, давності подій і монотонної агрегації. Серії повторних запусків показали, що RAG-відбір контексту усуває розкид оцінки, а аналіз чутливості на 63 збережених результатах для 8 організацій – що ранжування організацій стійке до зміни калібрувальних коефіцієнтів моделі.

Ключові слова: OSINT, кіберризик, велика мовна модель, RAG, контрагент, відтворюваність, аналіз чутливості.

Вступ

Атаки через постачальників і підрядників залишаються одним із поширених векторів компрометації організацій, тому перед початком співпраці доцільно оцінити кібербезпековий профіль контрагента. Типова перевірка контрагентів охоплює переважно юридичні та фінансові відомості, а комерційні платформи зовнішнього рейтингу кібербезпеки недоступні багатьом організаціям через вартість. Водночас відкриті джерела містять сигнали про кібератаки, витоки даних і шахрайські кампанії, пов’язані з організацією, проте їх ручний аналіз трудомісткий.

Великі мовні моделі (LLM) автоматизують аналіз таких текстів, але їхні числові оцінки нестабільні: різні моделі виставляють різні бали для однакових вхідних даних, а зниження температури генерації не гарантує стабільності повторних оцінок [1]. Підхід, у якому LLM інтерпретує зміст, а оцінку розраховує детермінований механізм, запропоновано для сценаріїв уразливостей [2], а поєднання LLM-виділення сутностей із виведенням за правилами – для побудови графів знань зі звітів CERT-UA [3]. Для оцінювання кіберризику окремої організації за неструктурованими OSINT-публікаціями такий підхід залишається маловивченим.

Метою роботи є розроблення та перевірка гібридного методу оцінювання кіберризиків організацій-контрагентів за OSINT-даними, що забезпечує відтворюваність і пояснюваність оцінки. Робота розвиває підхід, запропонований автором у дипломному дослідженні [4].

1. Гібридний метод оцінювання

Метод реалізовано у програмному прототипі. За назвою організації формуються пошукові запити з тематичними ознаками кіберризику, знайдені

сторінки очищуються, оцінюються за релевантністю й довірою до джерела та перевіряються на дублювання. Далі RAG-відбір у векторному сховищі ChromaDB залишає фрагменти, найближчі до тем кіберризику, і лише їх передає мовній моделі Gemini. Модель виділяє ризикові події – структуровані записи з категорією, датою, статусом, доказовими цитатами та ознакою зв'язку з організацією, – після чого події нормалізуються правилами (рис. 1). Числового бала LLM не виставляє: вона лише класифікує ознаки подій.



Рис. 1. Конвеєр гібридного оцінювання кіберризиків

Бал події s_i ($0 \le s_i \le 1$) обчислюється як базовий вплив (тяжкість з урахуванням упевненості, вага категорії ризику, довіра до джерела, корекція зв'язку з організацією, бонус за наявність цитати), помножений на модифікатори напряму впливу, пом'якшення, кібербезпекового контексту та актуальності. Порівняно з базовою версією моделі [4] внесено три зміни. По-перше, враховується вага лише основної категорії: раніше ваги всіх категорій події сумувалися, тож подія з кількома мітками отримувала завищений бал. По-друге, актуальність події визначається за її датою:

$$m_t = max\left(m_{min}, 0{,}5^{\frac{a}{T}}\right) \quad (1)$$

де a – вік події в місяцях на дату збору даних, T = 12 міс. – період напіврозпаду, m_{min} = 0,5 – мінімальний множник. Для подій без дати застосовується модифікатор статусу (активна, історична, усунена). По-третє, бали подій агрегуються монотонно:

$$S = 100\left(1 - \left(1 - s_{(1)}\right)\prod_{k=2}^{n}\left(1 - \lambda s_{(k)}\right)\right) \quad (2)$$

де $s_{(1)} \ge s_{(2)} \ge \ldots \ge s_{(n)}$ – упорядковані бали подій, λ = 0,5 – коефіцієнт внеску неосновних подій. На відміну від попередньої формули 0,6·max + 0,4·avg, нова подія не може знизити підсумковий бал S. Рівень ризику визначається за порогами 15, 35, 60 і 80 балів (низький, середній, високий, критичний).

2. Експериментальна перевірка

Використано 63 збережені результати LLM-аналізу для 8 організацій банківського, телекомунікаційного й енергетичного секторів та міжнародних IT-

компаній, отримані у квітні – червні 2026 р. Перерахунок оцінок виконано без повторних звернень до LLM, назви організацій анонімізовано.

Відтворюваність перевірено серіями з п’яти повторних запусків на однакових вхідних даних (табл. 1). За однієї й тієї самої моделі без RAG-відбору бал змінювався в межах 9,5 бала, тоді як з RAG-відбором він був однаковим в усіх запусках. Для іншої моделі без RAG розкид сягав 11,2 бала зі зміною рівня ризику, а в додатковій серії один із п’яти запусків завершився невдало. Усі доказові цитати (294 з 294) підтверджено в текстах джерел.

Таблиця 1 – Розкид підсумкового бала в серіях повторних запусків

Організація	**Модель**	**RAG**	**Бал: мін.–макс.**	**Рівень ризику**
А (банк)	Gemini 3.1 Flash-Lite	так	72,4–72,4	високий
Б (телеком)	Gemini 3.1 Flash-Lite	так	43,5–43,5	середній
Б (телеком)	Gemini 3.1 Flash-Lite	ні	25,2–34,6	низький
Б (телеком)	Gemini 2.5 Flash	ні	58,4–69,7	середній / високий

Примітка: бали обчислено формулою, чинною на момент експерименту [4].

Чутливість моделі до калібрування оцінено методом Монте-Карло: у 1000 вибірках усі калібрувальні коефіцієнти (ваги тяжкості, категорій і довіри до джерел, модифікатори статусу та впливу) випадково змінювалися в межах ±30 %, а ранжування 8 організацій порівнювалося з базовим за коефіцієнтом τ Кендалла. Для запропонованої моделі середнє $\tau = 0{,}94$, $\tau \geq 0{,}7$ досягнуто в 100 % вибірок, рівень ризику організації – у 82 % випадків; для базової формули [4] ці показники становлять 0,79, 83 % і 72 %. Додаткове варіювання параметрів T (6–24 міс.), m_{min} (0,25–0,75) і λ (0,25–1) знижує частку збережених рівнів до 61 % за $\tau = 0{,}89$.

На історичних прогонах (різні версії конвеєра) стандартне відхилення бала для двох організацій зменшилося (19,2 → 6,7; 13,6 → 11,4), а для третьої зросло (8,7 → 11,7) через непослідовне виділення дат подій мовною моделлю. Так, для організації Б юридична подія грудня 2023 р. (арешт корпоративних прав) за базовою формулою давала високий рівень ризику (74,4), а з урахуванням основної категорії та давності на дату збору даних – середній (41,5).

Висновки

Запропоновано гібридний метод оцінювання кіберризиків організацій-контрагентів за OSINT-даними, у якому мовна модель лише виділяє та класифікує події, а бал обчислюється детермінованою моделлю з урахуванням основної категорії ризику, давності подій і монотонної агрегації. Експерименти показали, що в серіях повторних запусків RAG-відбір контексту усунув розкид оцінки, а відносне ранжування організацій стійке до зміни калібрувальних коефіцієнтів і стійкіше, ніж за базовою формулою. Водночас абсолютні рівні ризику залежать від параметрів затухання та агрегації, а нестабільне виділення дат подій збільшує розкид оцінки. Подальші дослідження передбачають

експертне калібрування коефіцієнтів методом аналізу ієрархій, перевірку на ширшій вибірці та порівняння з оцінюванням безпосередньо мовною моделлю.

Список використаної літератури

1. Lau F. Same Input, Different Scores: A Multi Model Study on the Inconsistency of LLM Judge. arXiv:2603.04417. 2026. URL: https://arxiv.org/abs/2603.04417.
2. Ünal N. M., Çeliktaş B. Automating cyber risk assessment with public LLMs: an expert-validated framework and comparative analysis. IEEE Access. 2026. Vol. 14. P. 47754–47778. DOI: 10.1109/ACCESS.2026.3678044.
3. Turskyi V. A Formal Model for Constructing Sensitive Data Graphs from Cyber Reports using Large Language Models. Theoretical and Applied Cybersecurity. 2025. Vol. 7, No. 2. P. 98–107. DOI: 10.20535/tacs.2664-29132025.2.338785.
4. Ялбуган Ф. Д. Методи та засоби автоматизованого оцінювання кіберризиків організацій на основі OSINT-даних, RAG та LLM : дипломна робота на здобуття ступеня бакалавра : 125 Кібербезпека. Київ : КПІ ім. Ігоря Сікорського, 2026. 131 с.

Алексанян Андрій Ельдарович
студен групи КНм-21, ННІТ, ДУШКТ
Київ, Україна

ВИЯВЛЕННЯ АНОМАЛЬНОЇ ПОВЕДІНКИ КОРИСТУВАЧІВ МЕТОДАМИ МАШИННОГО НАВЧАННЯ ДЛЯ ЗАБЕЗПЕЧЕННЯ КІБЕРБЕЗПЕКИ КОРПОРАТИВНИХ СИСТЕМ

Розглянуто роль аналізу поведінки користувачів у забезпеченні кібербезпеки корпоративних інформаційних систем. Показано, як методи машинного навчання можуть застосовуватися для виявлення відхилень у типових сценаріях використання ІКТ-послуг, зокрема під час аналізу входів до системи, частоти запитів і характеристик мережевої активності. Описано основні етапи такого підходу: підготовку журналів подій, формування ознак, вибір моделі та оцінювання її результатів. Окрему увагу приділено обмеженням, хибним спрацюванням і захисту приватності користувачів.

Ключові слова: кібербезпека корпоративних інформаційних систем, машинне навчання, поведінка користувачів, ІКТ-послуги, виявлення аномалій.

Корпоративні інформаційні системи забезпечують роботу з даними, внутрішніми сервісами та бізнес-процесами організацій. Їхня безпека залежить не лише від технічних засобів захисту, а й від того, як користувачі отримують доступ до ресурсів і взаємодіють із ними. Компрометація облікового запису, зловживання наданими правами або нетипова активність можуть залишатися непоміченими, якщо контроль ґрунтується лише на фіксованих правилах. Тому аналіз поведінки користувачів ІКТ-послуг може доповнювати традиційні засоби захисту та допомагати своєчасно звертати увагу на підозрілі події.

У межах корпоративної системи поведінка користувача може описуватися послідовністю подій: часом входу, частотою звернень до сервісів, кількістю невдалих спроб автентифікації, типами запитів і змінами пристрою або мережевого середовища. Окремий показник не є достатньою підставою вважати дію небезпечною, однак сукупність ознак у контексті попередньої активності може допомогти визначити відхилення, що потребують додаткової перевірки. Саме тут перетинаються завдання кібербезпеки корпоративних інформаційних систем і тема прогнозування поведінки користувачів ІКТ-послуг: оцінювання очікуваної активності може стати одним з елементів виявлення потенційних загроз.

Для аналізу таких даних можна розглядати два загальні підходи. За наявності прикладів подій, позначених як нормальні або підозрілі, застосовують алгоритми класифікації, наприклад дерева рішень або випадковий ліс. Якщо розмічених прикладів недостатньо, можуть використовуватися методи виявлення аномалій, які шукають суттєві відхилення від типових шаблонів. Реалізацію експериментального прототипу мовою Python можна побудувати з використанням pandas для підготовки даних і scikit-learn для навчання та

оцінювання моделей. На етапі оцінювання доцільно аналізувати precision, recall і F1-score поряд із загальною точністю, оскільки хибні спрацювання створюють додаткове навантаження на фахівців, а пропущені інциденти можуть становити ризик для системи [1; 2].

Використання машинного навчання не означає, що кожне відхилення є кіберінцидентом. Результати моделі доцільно застосовувати як сигнал для додаткової перевірки, поєднуючи їх із правилами безпеки, журналами подій та оцінкою фахівця. На надійність підходу впливають неповнота даних, зміни робочих звичок, дисбаланс між звичайними та підозрілими подіями, а також поява нових сценаріїв атак. Не менш важливо дотримуватися принципів мінімізації даних, обмежувати доступ до журналів активності та визначати обґрунтовані строки їх зберігання. Управління такими ризиками має бути частиною загальної програми кібербезпеки організації [3].

Кібербезпека корпоративних інформаційних систем може бути посилена за допомогою аналізу та прогнозування поведінки користувачів ІКТ-послуг. Методи машинного навчання дають змогу виявляти закономірності й відхилення в журналах подій, доповнюючи правила контролю доступу та моніторингу. Практичне застосування такого підходу потребує належної підготовки даних, вибору відповідних метрик, перевірки результатів і захисту приватності. Отже, прогнозування поведінки користувачів є доречним напрямом використання машинного навчання в межах забезпечення кібербезпеки корпоративних інформаційних систем.

Список використаних джерел

1. Scikit-learn developers. Model evaluation: quantifying the quality of predictions. URL: https://scikit-learn.org/stable/modules/model_evaluation.html.

2. Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey. ACM Computing Surveys. 2009. Vol. 41, No. 3. Article 15. DOI: https://doi.org/10.1145/1541880.1541882.

3. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. 2024. DOI: https://doi.org/10.6028/NIST.CSWP.29.

Москаленко Н. В.
асистент, КНУ імені Тараса Шевченка, Україна
Сокульський О. Є.
доцент, к.т.н., ВІТІ імені Героїв Крут, Київ, Україна
Топольськов Є.О.
доцент, к.т.н., КНУ імені Тараса Шевченка, Україна

ШТУЧНИЙ ІНТЕЛЕКТ ЯК СКЛАДОВА ЯКОСТІ ПЗ

З появою штучного інтелекту кардинально змінюються підходи до розробки програмного забезпечення. Використання методів та засобів штучного інтелекту на всіх етапах створення дозволяє не тільки скоротити витрати ресурсів, але й значно підвищити ефективність та якість отриманих результатів. Однією з складових розробки є написання технічної документації, від якості та повноти якої в цілому залежить якість використання та функціонування програмного забезпечення. Це доволі трудомісткий процес, в якому з успіхом використовується штучний інтелект.

Ключові слова. Програмне забезпечення, штучний інтелект, генерація, код, документація.

Розробка програмного забезпечення включає декілька етапів та базується на дотриманні концепції SDLC(Software Development Life Cycle). З появою штучного інтелекту досягнення в цій галузі все більше з успіхом використовуються в розробці програмного забезпечення. Це генерація та автодоповнення коду, тестування, пошук помилок, рефакторинг, оптимізація та багато іншого. Ваговою складовою створення є написання технічної документації. Засоби з використанням штучного інтелекту роблять цей рутинний, але дуже важливий процес, більш простим та ефективним. Документи, які можна створити, використовуючи сучасні засоби:

- ✓ PRD(Product Requirements Document) – вимоги до продукту;
- ✓ BRD(Business Requirements Document) – бізнес-вимоги;
- ✓ схеми хмарної інфраструктури та карти залежностей мікрослужб;
- ✓ REST API(Representational State Transfer Application Programming Interface), які визначають архітектурний стиль взаємодії між програмами через інтернет;
- ✓ SDK(Software Development Kit) – перелік інструментів для створення ПЗ;
- ✓ SOP(Standard Operating Procedure) – покрокові інструкції для виконання рутинних, виробничих або системних завдань;
- ✓ Runbook– інструкції для серверів на основі стандартів(GDPR, ISO 27001);
- ✓ Стуктури та опис бази знань;
- ✓ Troubleshooting – пошук та виправлення збоїв;
- ✓ Release Notes – оновлення ПЗ та плани тестування (UAT).

Слід зауважити, що все більшою популярністю користується використання при розробці технічної документації інтелектуального помічника Microsoft Copilot, який орієнтований на роботу з текстовими документами.

Перелік посилань:

1. О. Г. Руденко, О. О. Безсонов, С. П. Євсеєв, О. Б. Ахієзер, Ю. І. Зайцев ; за заг. ред. С. П. Євсеєва. Штучний інтелект. Нейромережева обробка інформації : архітектури, навчання, застосування. Львів «Новий Світ-2000», 2026. – 426 с. ISBN 978-966-418-517-9

2. Створення шаблону технічної документації за допомогою ШІ. [Електронний ресурс]. Режим доступу:https://word.cloud.microsoft/create/uk/blog/technical-documentation-template-ai/

Круть Ростислав Андрійович
магістрант групи ІКБм-62, ДУІКТ, Київ, Україна

ТЕХНОЛОГІЇ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ КОМП'ЮТЕРНИХ МЕРЕЖ ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ VPN

У тезах розглянуто використання віртуальних приватних мереж для захисту корпоративного мережевого трафіку під час передавання через недовірені канали. Визначено функції VPN-тунелю, роль протоколів IPsec та WireGuard і вимоги до автентифікації, керування ключами, розмежування доступу та моніторингу. Запропоновано критерії оцінювання VPN-рішення за показниками безпеки й продуктивності. Показано, що VPN захищає канал зв'язку, але має застосовуватися разом із засобами контролю доступу, захистом кінцевих пристроїв і журналюванням подій.

Ключові слова: VPN, комп'ютерна мережа, кібербезпека, IPsec, WireGuard, захищений канал.

Постановка проблеми. Віддалена робота, хмарні сервіси та взаємодія між територіально розподіленими підрозділами потребують передавання корпоративних даних через мережі, яким організація не може повністю довіряти. Відкритий канал створює ризик перехоплення або зміни трафіку. Віртуальна приватна мережа (VPN) формує логічний захищений канал поверх наявної мережевої інфраструктури. Проте VPN не усуває ризики, пов'язані зі скомпрометованим пристроєм, викраденими обліковими даними чи надмірними правами користувача. Безпека залежить від конфігурації VPN і його інтеграції в загальну систему захисту мережі.

Мета роботи. Метою є визначення принципів і вимог до застосування VPN для захисту комп'ютерних мереж, а також критеріїв оцінювання безпеки та практичної придатності такого рішення.

Принцип роботи VPN. VPN створює тунель між користувацьким пристроєм і шлюзом або між мережевими шлюзами. Пакети передаються тунелем у захищеному вигляді. Залежно від протоколу та налаштувань криптографічні механізми забезпечують конфіденційність, цілісність даних і автентифікацію вузлів. У корпоративних мережах поширені сценарії віддаленого доступу користувача до внутрішніх ресурсів і з'єднання мереж різних підрозділів. Політика доступу має визначати, до яких систем дозволено підключатися через тунель.

Технології та налаштування. IPsec є набором протоколів мережевого рівня. IKEv2 виконує автентифікацію вузлів і встановлює асоціації безпеки, а ESP забезпечує захист IP-трафіку [1, с. 5; 2, с. 3]. WireGuard автентифікує вузли за відкритими ключами [3, с. 4]. Вибір рішення має враховувати сумісність із мережевим обладнанням і клієнтськими пристроями, вимоги до масштабування, спосіб розповсюдження ключів і можливості централізованого адміністрування. Порівняння слід проводити для конкретної інфраструктури, адже результат залежить від конфігурації, навантаження та характеристик кінцевих вузлів.

Режим маршрутизації впливає на захист: повний тунель спрямовує весь трафік через шлюз, а роздільний — лише вибрані адреси або підмережі. Роздільний режим може зменшити затримку, але залишає частину трафіку поза VPN. Тому маршрути, DNS і поведінку при відмові тунелю потрібно перевірити відповідно до моделі загроз. Оскільки VPN захищає з'єднання лише між кінцевими точками, він не замінює захист пристроїв, сегментацію мережі та контроль привілеїв.

Вимоги до впровадження. Для зменшення ризику несанкціонованого доступу потрібні надійна автентифікація, а за можливості — багатофакторна перевірка користувачів. У разі зміни ролі працівника слід переглядати його дозволи; після звільнення, компрометації чи втрати пристрою потрібно відкликати облікові дані або ключі, які більше не мають використовуватися. Доступ через VPN варто надавати лише до ресурсів, потрібних для робочих завдань, із сегментацією мережі та окремими правилами для адміністративного доступу. Кінцеві пристрої мають отримувати оновлення та перевірятися засобами захисту. Для конфігурації IPsec доцільно спиратися на настанови NIST щодо планування, розгортання й експлуатації IPsec VPN [4, с. 59–60].

Моніторинг та оцінювання. Роботу VPN необхідно супроводжувати журналюванням підключень, відмов автентифікації, змін конфігурації та подій керування ключами. Моніторинг дає змогу виявляти незвичну активність і перевіряти дотримання політик доступу. Ефективність рішення оцінюють за двома групами критеріїв. До критеріїв безпеки належать належна автентифікація, захист ключів, цілісність конфігурації, обмеження доступу та повнота журналів. До експлуатаційних критеріїв належать пропускна здатність, затримка, стабільність з'єднання, масштабованість і час відновлення після збою. Перевіряти слід також поведінку системи у випадках втрати з'єднання, відкликання ключа та блокування облікового запису. Для порівняльного тестування слід фіксувати однакові умови: пропускну здатність каналу, обсяг переданих даних і конфігурацію вузлів. Окремо визначають час встановлення тунелю та накладні витрати шифрування. Це дає змогу відокремити вплив VPN від відмінностей тестового середовища.

Висновки. VPN дає змогу захистити передавання даних через недовірені мережі й організувати керований віддалений доступ до корпоративних ресурсів. Для належного рівня захисту потрібні правильно налаштований протокол, надійна автентифікація, контроль ключів, мінімально необхідні права, сегментація мережі та аналіз журналів подій. Отже, VPN є одним із компонентів комплексного захисту комп'ютерної мережі. Подальше дослідження передбачає порівняння VPN-рішень у визначеній мережевій топології за показниками безпеки, продуктивності та зручності адміністрування.

Список використаних джерел

1. Kaufman C., Hoffman P., Nir Y., Eronen P., Kivinen T. Internet Key Exchange Protocol Version 2 (IKEv2). RFC 7296. IETF, 2014.
2. Kent S. IP Encapsulating Security Payload (ESP). RFC 4303. IETF, 2005.
3. Donenfeld J. A. WireGuard: Next Generation Kernel Network Tunnel. NDSS, 2017.
4. Barker E., Dang Q., Frankel S., Scarfone K., Wouters P. Guide to IPsec VPNs. NIST SP 800-77 Rev. 1, 2020.

Власенко Г. В.
студентка групи ІКБм-11, ННІКЗІ ДУІКТ, Київ, Україна

КРИТЕРІЇ ФОРМУВАННЯ ОБМЕЖЕНОГО НАБОРУ СЦЕНАРІЇВ ВАЛІДАЦІЇ МЕХАНІЗМІВ КІБЕРЗАХИСТУ НА ОСНОВІ MITRE ATT&CK

Розглянуто формування обмеженого набору сценаріїв валідації кіберзахисту на основі MITRE ATT&CK. Запропоновано враховувати релевантність сценаріїв до профілю загроз і приріст покриття пов’язаних із ними захисних заходів. Для умовної корпоративної інформаційної системи зіставлено відбір за релевантністю і послідовний відбір за приростом покриття. Використано відкриті зіставлення ATT&CK із CIS Safeguards. Порівняння показало різницю в розрахунковому покритті та компроміс із середньою релевантністю.

Ключові слова: MITRE ATT&CK, CIS Controls, кіберзахист, покриття, релевантність.

Вступ

Валідація механізмів кіберзахисту потребує відбору сценаріїв, релевантних профілю загроз інформаційної системи. Основою для їх формування може слугувати MITRE ATT&CK — база знань, що систематизує тактики й техніки зловмисників [1]. Оскільки ATT&CK містить понад 200 технік і 400 підтехнік, їх повне охоплення в умовах обмежених ресурсів практично недоцільне, що зумовлює необхідність пріоритизації сценаріїв.

Існуючі підходи пріоритизують ATT&CK-техніки за поширеністю, актуальністю загроз, впливом та іншими характеристиками. Зокрема, Top ATT&CK Techniques враховує поширеність, критичні точки атак і можливість захисних дій [2], а Abo-alian et al. застосовують багатокритеріальний аналіз для відбору технік емуляції зловмисника [3].

Водночас висока індивідуальна пріоритетність сценаріїв не гарантує доцільності їх спільного включення до обмеженого набору, оскільки вони можуть бути пов'язані з однаковими захисними можливостями. Метою роботи є обґрунтування критеріїв формування такого набору з урахуванням релевантності кіберзагроз і перекриття захисних можливостей.

1. Обґрунтування критеріїв формування набору сценаріїв

Введемо поняття сценарію валідації механізму кіберзахисту — це опис перевірки, що передбачає відтворення визначеної поведінки зловмисника та оцінювання реакції відповідних захисних механізмів за попередньо встановленими очікуваними результатами. Із метою спрощення моделі, кожному сценарію ставиться у відповідність одна основна техніка або підтехніка MITRE ATT&CK.

Формування набору сценаріїв передбачає попереднє визначення множини кандидатів відповідно до характеристик досліджуваної інформаційної системи. До цієї множини включаються сценарії, передумови реалізації яких відповідають її технологічному середовищу, архітектурі та визначеним умовам перевірки.

Для подальшого відбору сценаріїв пропонується використовувати два критерії: релевантність кіберзагрозам та додаткове покриття захисних можливостей. Перший характеризує відповідність сценарію профілю загроз досліджуваної системи, тоді як другий дозволяє врахувати перекриття захисних можливостей між уже відібраними сценаріями та кандидатами на включення до набору.

Релевантність R_i визначається з урахуванням актуальності загрози T_i та потенційного впливу її реалізації I_i. Обидва чинники оцінюються за трирівневою шкалою (1 — низький, 2 — помірний, 3 — високий), а підсумкова оцінка визначається як $R_i = T_i + I_i - 1, R_{\text{і}} \in \{1, 2, 3, 4, 5\}$. Актуальність встановлюється за відомостями про використання відповідної техніки зловмисниками та її відповідність профілю загроз ІС, а вплив — за можливим порушенням конфіденційності, цілісності й доступності інформаційних активів.

Критерій додаткового покриття відображає кількість захисних можливостей, які може охопити сценарій порівняно з уже сформованим набором. Для встановлення зв'язків між поведінкою зловмисників і захисними можливостями можуть використовуватися відкриті зіставлення ATT&CK із відповідними технічними контролями, зокрема матеріали Center for Threat-Informed Defense [4].

2. Методичний підхід до формування набору технік

Нехай S — відібраний набір кандидатів, M_i — множина унікальних ідентифікаторів CIS Safeguards, пов’язаних із технікою сценарію i, а $M(S)$ — об’єднання цих множин для всіх сценаріїв із S. Позначення $|X|$ означає потужність множини X. Тоді додаткове покриття $\Delta_i(S)$ та покриття набору $F(S)$ визначаються так:

$$\Delta_i(S) = \left|{}^{M_i}\!/_{M(S)}\right|; \quad F(S) = |M(S)|.$$

Встановлюється бюджет $|S| \leq B$ та поріг $R_i \geq R_{min}$, де R_{min} — порогове значення релевантності, нижче якого сценарії не розглядаються. Базовий підхід обирає кандидатів за спаданням, а за рівності оцінок — за кількістю пов’язаних Safeguards. У запропонованій процедурі на кожному кроці додається кандидат із найбільшим додатковим покриттям $\Delta_i(S)$; за рівності приросту — із більшим R_i. Відбір припиняється після досягнення або за нульового приросту для решти кандидатів.

3. Демонстраційне застосування

Розглянуто умовну корпоративну ІС із Windows 11, Windows Server, доменними обліковими записами, електронною поштою, RDP та мережевими сервісами. Було ілюстративно відібрані 10 кандидатів (табл. 1) як застосовні до модельного Windows-орієнтованого корпоративного середовища та представляють різні етапи/типи поведінки зловмисника. Зіставлення M_i отримано

з наданого набору CTID для CIS Controls 8.1.2 та Enterprise ATT&CK 19.1. Оцінки T_i та I_i задано як ілюстративні припущення для цієї моделі експертним методом.

Таблиця 1 — Вихідні сценарії та пов'язані CIS Safeguards.

ID	ATT&CK ID	Сценарій	T	I	R	$\lvert M \rvert$
S1	T1595.002	Сканування вразливостей	1	2	2	0
S2	T1566.002	Фішингове посилання	2	2	3	6
S3	T1059.001	Виконання PowerShell	3	3	5	5
S4	T1136.002	Створення доменного облікового запису	2	2	3	2
S5	T1053.005	Заплановане завдання	2	3	4	1
S6	T1055	Впровадження коду в процес	2	2	3	3
S7	T1686.003	Зміна брандмауера Windows	2	2	3	1
S8	T1021.001	Переміщення через RDP	3	2	4	11
S9	T1041	Ексфільтрація каналом C2	3	3	5	3
S10	T1499.002	Виснаження ресурсу сервісу	2	3	4	1

Встановивши обмеження у $B = 3$ та $R_{min} = 3$, до розгляду допущено 9 сценаріїв (усі, окрім S1). За базовим підходом обрання трьох найбільш релевантних сценаріїв обраними є S3, S8 та S9 ($F = 18, \bar{R} = 4{,}67$). За підходом послідовного відбору для найбільшого додаткового покриття обраними є S2, S3, S8 ($F = 21, \bar{R} = 4{,}00$). В обох наборах є по одному повторному зіставленню (враховано лише унікальні Safeguards). Приріст кількості пов'язаних Safeguards становить 3, або 16,7%, але супроводжується зниженням середньої релевантності. Перебором усіх 84 можливих трійок з 9 кандидатів підтверджено, що $F = 21$ є максимальним для цієї вибірки. Значення порогу та бюджету обрано умовно для компактної демонстрації процедури; у практичному застосуванні бюджет і мінімально допустима релевантність задаються організацією відповідно до ресурсів та її профілю ризику.

Висновки

Обґрунтовано два критерії формування обмеженого набору сценаріїв — релевантність загрозам та додаткове покриття пов'язаних захисних заходів — за умови попередньої відповідності кандидатів досліджуваному середовищу. У демонстраційному прикладі запропонована процедура збільшила розрахункове покриття з 18 до 21 CIS Safeguards порівняно з відбором за релевантністю, проте знизила її середнє значення. Подальше дослідження має враховувати нерівнозначність захисних заходів, трудомісткість перевірок і результати фактичного виконання сценаріїв у контрольованому середовищі.

Список використаної літератури

1. MITRE. (n.d.). MITRE ATT&CK. Retrieved October 14, 2026, from https://attack.mitre.org/

2. Center for Threat-Informed Defense. (2024, August 14). Top ATT&CK Techniques. https://ctid.mitre.org/projects/top-attack-techniques

3. Abo-alian, A., Youssef, M., & Badr, N. L. (2025). A data-driven approach to prioritize MITRE ATT&CK techniques for active directory adversary emulation. Scientific Reports, 15, Article 27776. https://doi.org/10.1038/s41598-025-12948-x

4. Center for Threat-Informed Defense. (2026, October 5). CIS Controls for Threat Mitigations. https://ctid.mitre.org/projects/cis-controls-for-threat-mitigations/

Коршиков О.В.
студент групи УБДм-61, ННІКБЗІ ДУІКТ
Мужанова Т.М.
доцент кафедри УКБЗІ, ННІКБЗІ ДУІКТ

ЕВОЛЮЦІЯ ВИМОГ ДО ФАХІВЦІВ З КІБЕРБЕЗПЕКИ В УМОВАХ ЦИФРОВІЗАЦІЇ

У роботі проаналізовано ключові зміни в галузі ІТ та кібербезпеки за останні десятиріччя, зокрема перехід до архітектури Zero Trust, хмарних обчислень, поглиблення автоматизації і розвиток штучного інтелекту. На основі цих трансформацій окреслено нові вимоги до кваліфікації сучасних кіберфахівців. Наголошено, що для успішної діяльності професіонал з кібербезпеки має володіти не лише глибокими Hard Skills (хмарна безпека, програмування, ML-моделі), а й розвиненими Soft Skills, які охоплюють, зокрема, командну взаємодію, професійне спілкування, антикризове управління і вміння формувати підтримку керівництва для рішень з кібербезпеки.

Ключові слова: цифровізація, вимоги до фахівця з кібербезпеки, Hard&Soft Skills кіберфахівця.

Упродовж останніх 15-20 років відбувся фундаментальний зсув у підходах до ІТ й забезпечення інформаційної та кібербезпеки (ІКБ). Основними передумовами такої еволюції були перехід на гібридні форми роботи, переміщення інформаційних ресурсів у хмари, розширення масштабів автоматизацію усіх можливих процесів організації [1].

Розвиток сфери ІТ докорінно змінив вимоги до кваліфікації спеціалістів з ІКБ. Сучасний фахівець повинен володіти принципово новим набором як технічних (Hard Skills), так і міжособистісних (Soft Skills) якостей [2].

Ці зміни на рівні технічних умінь і навичок можна чітко простежити за такими ключовими технологічними напрямами як архітектура безпеки, життєвий цикл розробки ПЗ, розгортання інфраструктури, моніторинг і реагування на події кібербезпеки.

З огляду на те, що *архітектура безпеки* пережила поступовий перехід від захисту периметра архітектури нульової довіри (Zero Trust), фахівець з кібербезпеки має вміти впроваджувати мікросегментацію мережі, доступ з урахуванням контексту (місця, часу тощо), і принцип найменших привілеїв.

В управлінні *інфраструктурою захисту* класичне налаштування фізичного обладнання і локальних мереж замінив менеджмент динамічних віртуальних середовищ, що вимагає від спеціаліста глибокої експертизи у сфері хмарної безпеки.

Затребуваний кіберфахівець має володіти навичками *програмування* на рівні, що дозволяє створювати індивідуальний захисний інструментарій і реалізувати концепцію Security as Code (SaC), за якої перевірки й політики безпеки автоматично виконуються як програмний код.

Оскільки зараз обробку масштабних потоків хмарних даних замість людей тепер виконує штучний інтелект, аналітикам з *моніторингу загроз* критично

важливо вміти тренувати ML-моделі та володіти навичками промпт-інжинірингу для роботи з ШІ-асистентами [3].

Поряд із цим галузь кібербезпеки інтегрувалася в загальні бізнес-процеси організації, через що Soft Skills стали не менш важливими за технічну кваліфікацію. Ефективна діяльність у цій сфері тепер базується на комунікації та взаємодії з людьми.

До базових нетехнічних компетенцій сучасного кіберфахівця належать:

- *командна взаємодія* - як здатність координувати спільні зусилля колективу, оперативно розподіляти завдання й обмінюватися інформацією для узгодження зусиль із кіберзахисту;
- *професійна дипломатія*, яка характеризується вмінням вести переговори, проявляти емпатію і знаходити компроміси у комунікації зі стейкхолдерами;
- *ведення дискусії й формування підтримки керівництва* для ІКБ шляхом здатності передавати складні технічні питання, аналізувати загрози мовою фінансів і бізнес-показників, обґрунтовуючи доцільність інвестицій у захист чи ШІ-рішення;
- забезпечення *кібергігієни*, що охоплює ефективне навчання й наставництво персоналу нетехнічних відділів з питань безпечної поведінки у цифровому просторі і запобігання загрозам фішингу й соціальної інженерії;
- *управління кризами* через уміння ухвалювати виважені й точні рішення під часовим тиском в умовах реальних інцидентів, демонструючи стресостійкість і розвинений емоційний інтелект [4].

Схема основних навичок фахівця з кібербезпеки показана на рис. 1.

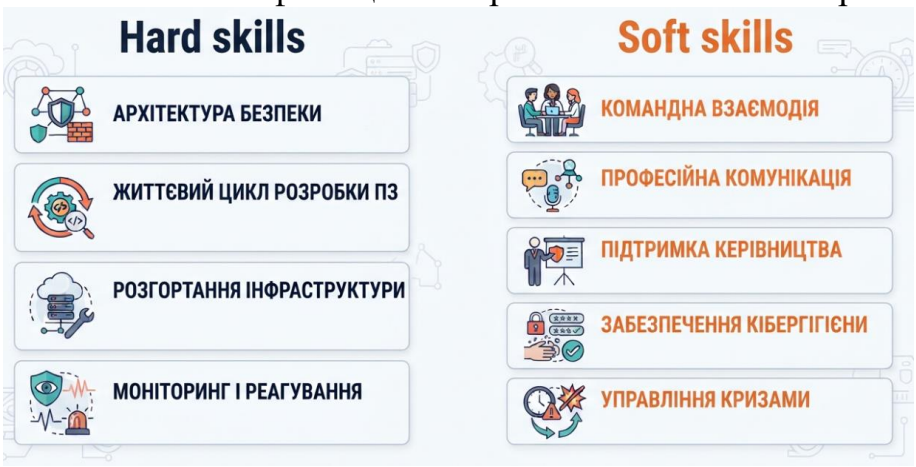


Рис. 1. Основні навички фахівця з кібербезпеки

Отже, стрімкий розвиток ІТ і хмарних рішень докорінно змінив парадигму забезпечення кібербезпеки. Сучасна кібербезпека трансформувалася з ізольованої технічної функції на інтегрований бізнес-процес. Відповідно, успіх у цій сфері сьогодні визначається синергією високої технічної експертизи (зокрема в автоматизації та застосуванні ШІ) й розвинених міжособистісних навичок,

необхідних для ефективного управління кризами, навчання персоналу та комунікації зі стейкхолдерами.

Список використаних джерел

1. Ветлицька О.С., Треньов М.Г. Проблеми кіберстійкості ІКТ-систем в умовах цифрової трансформації. *Телекомунікаційні та інформаційні технології*. 2024. № 1 (82). С. 64-72. DOI: 10.31673/2412-4338.2024.016472

2. Daniel Bendler, Michael Felderer. Competency Models for Information Security and Cybersecurity Professionals: Analysis of Existing Work and a New Model. *ACM Transactions on Computing Education*. 2023. 23. P.1-33. DOI: 10.1145/3573205.

3. Business and Technology Insights and Trends. Gartner Top 10 Strategic Technology Trends for 2027. *Gartner*. URL: https://www.gartner.com/en/insights

4. Duncan Whitley. 10 Soft Skills that will Benefit your Cybersecurity Career. *Centri*. URL: https://www.centri.org/blog/posts/10-soft-skills-for-cybersecurity-careers

Миловидов В. Ю.
асистент кафедри СТКБ,
ННІКБЗІ, ДУІКТ,
Київ, Україна

ОЦІНЮВАННЯ ВНЕСКУ КОНТЕКСТНИХ ОЗНАК У НЕЙРОМЕРЕЖЕВЕ ВИЯВЛЕННЯ ФІШИНГОВИХ ВЕБРЕСУРСІВ

Анотація. Досліджено вплив контекстних характеристик вебсторінок на нейромережеве виявлення фішингу. На корпусі з 79 315 пар URL–HTML зіставлено базові ознаки, п’ять груп контексту та їх комбінації. Встановлено, що контекст підвищує повноту виявлення, проте його користь залежить від складу ознак і супроводжується зміною частки хибних спрацювань. Найпослідовніший внесок у дослідах додавання груп показали характеристики посилань і ресурсів. Обґрунтовано необхідність перевіряти перенесення результатів на доменно відокремлену тестову вибірку.

Ключові слова: фішинг, виявлення фішингових вебресурсів, контекстні ознаки, нейронні мережі, аналіз URL та HTML, відбір ознак, хибні спрацювання, інформаційна безпека.

Виявлення фішингових вебресурсів ускладнюється тим, що форми авторизації, зовнішні посилання та вбудовані сценарії властиві як шахрайським, так і легітимним сторінкам. Їх наявність сама по собі недостатня для висновку про загрозу. Значення має узгодженість елементів: чи відповідає заявлений бренд домену, куди спрямована парольна форма та як пов’язані посилання зі структурою сторінки. Водночас розширення набору характеристик потребує перевірки: додаткові ознаки можуть відображати особливості навчального корпусу, які не зберігаються на нових даних.

Цей напрям має наукове обґрунтування. У дослідженні PHILTER, опублікованому на USENIX Security 2026, проаналізовано 55 підходів і показано, що високої загальної точності недостатньо для оцінювання практичної придатності детектора. Автори звертають увагу на обмежене тестування різноманітних легітимних сторінок і пов’язаний із цим ризик хибних тривог [1]. У PhishIntention спільно аналізуються імітація бренду та намір отримати облікові дані, причому останній уточнюється взаємодією зі сторінкою [2]. Це підкреслює доцільність дослідження зв’язків між ознаками.

Метою роботи є оцінювання внеску груп контекстних ознак та їх комбінацій у якість нейромережевої класифікації вебресурсів із одночасним урахуванням повноти виявлення фішингу й хибних спрацювань на легітимних сторінках.

Розроблено програмний застосунок мовою Python для статичного аналізу URL і збереженого HTML. Базовий набір містить 69 ознак: 48 характеристик адреси та 21 характеристику HTML. Додатково сформовано 70 контекстних ознак у п’яти групах: форми та введення даних; відповідність бренду й домену; посилання та зовнішні ресурси; статичні індикатори сценаріїв і навігації; призначення та текстове наповнення сторінки. Сформовано також 24 явні комбінації, зокрема поєднання парольної форми з невідповідністю бренду, зовнішнім адресатом або прихованим iframe. Повна схема охоплює 163 ознаки.

Експериментальну базу становили 79 315 пар URL–HTML із 12 заздалегідь визначених фрагментів PhreshPhish. Автори цього корпусу окремо розглядають витоки між вибірками та нереалістичні співвідношення класів як причини завищених оцінок детекторів [3]. Після очищення й доменних виключень сформовано навчальну вибірку з 29 873 сторінок, вибірку відбору ознак із 9 653, калібрувальну з 9 916 і фінальну тестову з 13 104 сторінок. Остання містить 7 315 легітимних і 5 789 фішингових сторінок із 12 434 доменів. Між частинами немає збігів доменів, канонізованих URL і нормалізованого HTML.

Основним класифікатором обрано багатошаровий персептрон з одним прихованим шаром із 10 нейронів, активацією ReLU та оптимізатором Adam. Навчання тривало 80 епох із розміром пакета 64 та швидкістю навчання 0,001; параметри нормалізації визначалися лише на навчальних даних. Порівняно додавання кожної групи, усі десять пар груп, вилучення груп із повної схеми та варіант без явних комбінацій. Основні конфігурації навчено з п’ятьма, групові порівняння — з трьома значеннями початкового стану генератора випадкових чисел. Разом із контрольними моделями виконано 84 окремі навчання.

Оцінювали повноту виявлення, частку хибних спрацювань і стандартизовану часткову площу під ROC-кривою в діапазоні цих спрацювань до 1% (pAUC). Серед наперед визначених кандидатів обирали найменший набір, середня pAUC якого на вибірці відбору відставала від найкращої не більш як на 0,005. Цій умові відповідав набір із 139 ознак: базові характеристики й усі п’ять груп контексту без явних комбінацій. Такий відбір не означає глобальної оптимальності серед усіх можливих підмножин. Поріг кожної моделі визначали на калібрувальній вибірці за умови частки хибних спрацювань не вище 1%, після чого переносили без змін на тестові дані.

За середніми результатами п’яти запусків базова модель виявила 65,46% фішингових сторінок за 1,66% хибних спрацювань. Обраний контекстний варіант зі 139 ознаками підвищив повноту до 72,28%, а частка хибних спрацювань становила 1,84%. Отже, приріст повноти на 6,82 процентного пункту супроводжувався зростанням частки хибних тривог на 0,18 процентного пункту. Середня pAUC змінилася з 0,7383 до 0,7437, причому її покращення спостерігалося у трьох із п’яти запусків. Калібрувальне обмеження 1% не збереглося на тесті, що вказує на необхідність перевірки порогів за зміни даних.

Найпослідовніший внесок у дослідах додавання окремих груп показали характеристики посилань і ресурсів: середня pAUC зросла на 0,0312 на вибірці відбору та на 0,0166 на фінальному тесті. До цієї групи належать частки внутрішніх і зовнішніх посилань, кількість зовнішніх доменів, розміщення сценаріїв, зображень та iframe. Натомість переваги пар «форми — посилання» та «форми — бренд», виявлені під час відбору, на тесті не підтвердилися. Додавання 24 явних комбінацій до всіх контекстних груп знизило середню тестову pAUC з 0,7422 до 0,7293 за однакових трьох початкових станів. Отриманий результат не

дає підстав вважати ускладнення схеми ознак самостійним чинником покращення.

Практичне значення частки хибних тривог залежить від поширеності загрози. За припущення, що фішинг становить 1% потоку з 10 000 сторінок, показники контекстної моделі одного зафіксованого запуску відповідають приблизно 71 правильному виявленню та 156 хибним тривогам. Це розрахунковий сценарій за незмінних характеристик помилок, а не результат експлуатації. Він обґрунтовує використання класифікатора для пріоритезації перевірок, а не автоматичного блокування всіх позначених ресурсів.

Додатково перевірено перенесення URL-моделі на корпус PhiUSIIL, описаний у журналі Computers & Security [4]. Після очищення та виключення доменних перетинів оцінено 215 891 запис: повнота становила 41,39%, частка хибних спрацювань — 0,62%. Ця перевірка стосується лише URL-ознак і не є зовнішнім підтвердженням переваги HTML-контексту.

Науковий внесок роботи полягає в порівняльному оцінюванні груп контексту та перевірці перенесення ефектів їх поєднання на доменно відокремлені дані. Практичне значення має встановлений компроміс між додатковими виявленнями й хибними тривогами, а також відсутність виграшу від механічного розширення набору комбінацій. Результати обмежені відібраною частиною корпусу, його мітками та статичним аналізом HTML. Перспективою є перевірка на незалежному корпусі URL–HTML і дослідження стійкості порогів у часі. Запропонований застосунок доцільно розглядати як засіб первинного відбору підозрілих ресурсів для подальшої перевірки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Alam M., Rahman M. L., Paul S. K., Hays A. W., Hussain A., Huq M. I., Saxena N. SoK: PHILTER: Uncovering Security and Functional Gaps in AI-based Phishing Website Detection Literature via an LLM-based Reasoning Framework. 35th USENIX Security Symposium. 2026. P. 4981–5000. URL: https://www.usenix.org/conference/usenixsecurity26/presentation/alam.
2. Liu R., Lin Y., Yang X., Ng S. H., Divakaran D. M., Dong J. S. Inferring Phishing Intention via Webpage Appearance and Dynamics: A Deep Vision Based Approach. 31st USENIX Security Symposium. 2022. P. 1633–1650. URL: https://www.usenix.org/conference/usenixsecurity22/presentation/liu-ruofan.
3. Dalton T., Gowda H., Rao G., Pargi S., Hadj Khodabakhshi A., Rombs J., Jou S., Marwah M. PhreshPhish: A Real-World, High-Quality, Large-Scale Phishing Website Dataset and Benchmark. arXiv:2507.10854v2. 2026. Препринт. URL: https://arxiv.org/abs/2507.10854v2.
4. Prasad A., Chandra S. PhiUSIIL: A diverse security profile empowered phishing URL detection framework based on similarity index and incremental learning. Computers & Security. 2024. Vol. 136. Article 103545. DOI: https://doi.org/10.1016/j.cose.2023.103545.

Нижник А.О.
аспірант кафедри захисту інформації, Національний університет «Львівська політехніка», Львів, Україна

МЕТОД ВИЯВЛЕННЯ АНОМАЛЬНИХ ПОСЛІДОВНОСТЕЙ КОМАНД MAVLINK У СИСТЕМАХ КЕРУВАННЯ БПЛА

Запропоновано детектор, який оцінює кожну польотну команду MAVLink у контексті попередніх команд, відповідей автопілота та стану апарата і повертає одне з трьох рішень: PASS, ALARM або UNKNOWN, коли даних бракує. Аномалією вважається порушення заданої політики. Детектор перевірено на симуляторі PX4 v1.17.0 зі штатними наземними станціями QGroundControl і MAVSDK у 100 сеансах. Він не дав жодної тривоги на 35 штатних командах, а втручання, що порушують правила, і однакові з ними дозволені дії оператора отримали однакові рішення. Головне обмеження задає період телеметрії стану 1,028 с: зліт під час посадки лишається непоміченим, якщо випереджає чергове повідомлення телеметрії.

Ключові слова: MAVLink, БПЛА, PX4, контекстний аналіз команд, виявлення аномалій.

Постановка проблеми. MAVLink, протокол зв'язку між наземною станцією та автопілотом БПЛА, задає формат кожного повідомлення окремо і не визначає, які послідовності команд безпечні. Чи доречна команда, залежить від попередніх команд, відповідей автопілота і стану апарата. Звідси дві труднощі для детектора. Повторна команда або незвичний порядок дій можуть бути нормою: протокол вимагає повторити команду, якщо підтвердження не надійшло, а відповідь ACCEPTED означає лише, що автопілот спробує її виконати [1]; QGroundControl штатно надсилає зліт раніше за ARM. І навпаки, коректна команда від довіреної наземної станції може бути небезпечною в поточному стані апарата, як вимкнення двигунів у польоті.

Системи виявлення вторгнень для MAVLink [2] навчають нейромережу на послідовностях ідентифікаторів повідомлень і виявляють flooding-атаки; зміст команд і стан апарата вони не враховують. Системи DATUM [3] і Platum [4] перевіряють послідовності повідомлень за формальною специфікацією протоколу, але оцінені лише за накладними витратами. Як такий детектор поводиться на польотних командах реальних наземних станцій, зокрема на повторах і дозволених діях, схожих на порушення, не показано. У цій роботі таку перевірку виконано для детектора з чотирма правилами, де аномалія означає порушення заданої політики, а не статистичне відхилення.

Метод. Детектор підключено до проксі між наземною станцією та автопілотом: він отримує копії всіх повідомлень і нічого в них не змінює. За телеметрією автопілота детектор відстежує стан апарата: чи увімкнені двигуни, який режим активний, на землі апарат чи в повітрі; дані, старші за 3 с, вважаються невідомими. Кожну польотну команду (ARM і DISARM, тобто увімкнення та вимкнення двигунів, зліт, посадка, повернення, утримання позиції) детектор звіряє з чотирма правилами.

R1: примусове вимкнення двигунів (force DISARM), коли апарат у повітрі.

R2: команда, відправник або адресат поза дозволеним профілем.

R3: повторний зліт з іншими параметрами протягом 2 с після першого, який ще не відхилено і не виконано.

R4: зліт під час посадки.

На відміну від двійкового класифікатора детектор має третю відповідь. ALARM означає порушення хоча б одного правила, PASS, що порушень не знайдено, а коли правилу бракує свіжих даних, детектор повертає UNKNOWN і не вгадує. Правила навмисно спрацьовують і на дозволені дії оператора з таким самим виглядом (аварійний DISARM, виправлення висоти, переривання посадки), бо детектор не знає наміру. Цю ціну політики вимірюють окремо.

Експеримент. Стенд: симулятор PX4 v1.17.0 (SITL, модель SIH quadx, MAVLink 2 без підписування), штатні наземні станції QGroundControl v5.1.4 і MAVSDK, один канал керування через проксі; стенд реалізовано мовою Rust (rustc 1.97.0, Ubuntu 26.04). Проведено 100 сеансів трьох видів: 43 штатні польоти, 20 сеансів із повторним зльотом через задану затримку після команди посадки і 37 сеансів із втручаннями та контрольними діями. У кожному сеансі наперед вибрано одну цільову команду, за якою оцінюють рішення; загальну частку виявлених «атак» не обчислюють: спрацювання правила на події, створеній для нього, перевіряє лише реалізацію. Фактичну реакцію автопілота встановлюють за його бортовим журналом; джерело кожної команди та її дозволеність відомі лише оцінювачу, а не детектору. Ті самі записи обробляють п'ять варіантів детектора: D0 перевіряє лише профіль відправників і команд; D1 додатково перевіряє порядок команд, але без параметрів і стану апарата; D2 застосовує всі чотири правила; ще два варіанти D2 позбавлені відповідно підтверджень автопілота (D2−ACK) і телеметрії стану (D2−STATE).

Таблиця 1. Рішення детекторів на цільових командах (A/U/P: ALARM/UNKNOWN/PASS)

Сценарій (цільова команда)	Тип, правило	n	D0	D1	D2	D2−ACK	D2−STATE
B1: зліт MAVSDK	штатна	3	0/0/3	0/0/3	0/0/3	0/0/3	0/3/0
B2n: зліт QGC	штатна	5	0/0/5	0/0/5	0/0/5	0/0/5	0/5/0
B2r: зліт QGC після відмови ARM	штатна	5	0/0/5	0/0/5	0/0/5	0/0/5	0/5/0
B3a: повтор ARM без підтвердження	штатна	5	0/0/5	0/0/5	0/0/5	0/0/5	0/0/5
B3t: повтор TAKEOFF без підтвердження	штатна	5	0/0/5	0/0/5	0/0/5	0/0/5	0/5/0
B4m: зміна режиму HOLD → RTL	штатна	2	0/0/2	0/0/2	0/0/2	0/0/2	0/0/2
B5a: ARM після автоматичного DISARM	штатна	5	0/0/5	0/0/5	0/0/5	0/0/5	0/0/5
B5f: HOLD після відновлення зв'язку	штатна	5	0/0/5	0/0/5	0/0/5	0/0/5	0/0/5
B4k: аварійний DISARM у польоті	дозволена, R1	2	0/0/2	2/0/0	2/0/0	2/0/0	0/2/0
B4v: виправлення висоти зльоту	дозволена, R3	2	0/0/2	0/0/2	2/0/0	2/0/0	0/2/0
B4p-s: зліт під час посадки, L видно	дозволена, R4	2	0/0/2	0/0/2	2/0/0	2/0/0	0/2/0
B4p-e: зліт одразу після LAND	дозволена, R4	2	0/0/2	0/0/2	1/0/1	1/0/1	0/2/0
T01: зліт через 0,1 с після LAND	дозволена, R4	4	0/0/4	0/0/4	1/0/3	1/0/3	0/4/0
T03: через 0,3 с	дозволена, R4	4	0/0/4	0/0/4	2/0/2	2/0/2	0/4/0
T06: через 0,6 с	дозволена, R4	4	0/0/4	0/0/4	2/0/2	2/0/2	0/4/0
T10: через 1,0 с	дозволена, R4	4	0/0/4	0/0/4	4/0/0	4/0/0	0/4/0
T15: через 1,5 с	дозволена, R4	4	0/0/4	0/0/4	4/0/0	4/0/0	0/4/0
A1: RTL від стороннього джерела	втручання, R2	3	3/0/0	3/0/0	3/0/0	3/0/0	3/0/0

Сценарій (цільова команда)	Тип, правило	n	D0	D1	D2	D2−ACK	D2−STATE
A2f: внесений аварійний DISARM	втручання, R1	5	0/0/5	5/0/0	5/0/0	5/0/0	0/5/0
A3: підмінена висота в повторі TAKEOFF	втручання, R3	4	0/0/4	0/0/4	4/0/0	4/0/0	0/4/0
A4-s: затриманий повтор, L видно	втручання, R4	5	0/0/5	0/0/5	5/0/0	5/0/0	0/5/0
A4-e: затриманий повтор після LAND	втручання, R4	5	0/0/5	0/0/5	1/0/4	1/0/4	0/5/0
A5l: внесений LAND	втручання	5	0/0/5	0/0/5	0/0/5	0/0/5	0/0/5
A5r: внесений RTL	втручання	5	0/0/5	0/0/5	0/0/5	0/0/5	0/0/5
A2c: звичайний DISARM у польоті	контроль	5	0/0/5	5/0/0	0/0/5	0/0/5	0/0/5

Результати. Детектор не заважає штатній роботі. На 35 штатних командах, серед них повтори після втраченого підтвердження і зліт через QGroundControl, D0, D1, D2 і D2−ACK не дали ні тривоги, ні UNKNOWN (табл. 1). Вибірки малі: результат 0 з 5 узгоджується з часткою хибних тривог до 0,52 (95 % довірчий інтервал), тому нульову частку він не доводить.

Правила виявляють заплановані втручання, але не відрізняють їх від дозволених дій. D2 дав 38 тривог на 100 командах і жодного UNKNOWN. Внесене аварійне вимкнення двигунів, підмінену висоту зльоту і затриманий зліт після зафіксованої посадки він позначив у 5/5, 4/4 і 5/5, а такі самі дозволені дії оператора у 2/2, 2/2 і 2/2. Команду повернення від стороннього джерела виявили всі п'ять варіантів за чужим ідентифікатором. Внесені посадку та повернення у штатному форматі від довіреної наземної станції пропущено в 10 з 10 випадків, як і передбачалося. На контрольному звичайному вимкненні двигунів у польоті, відхиленому автопілотом, D1 дав п'ять хибних тривог, а D2 жодної.

Головне обмеження задає телеметрія, а не правила. В усіх 34 сеансах, де зліт надходив під час посадки, автопілот уже за 16–24 мс після команди посадки перейшов у режим посадки, але приймав зліт і переривав посадку: за 10 с апарат набирав 9,8–13,9 м. Детектор же дізнається про посадку лише з чергового повідомлення телеметрії стану, а вони надходять раз на 1,028 с. Тому рішення залежало від затримки між командами посадки і зльоту: за 0,1; 0,3; 0,6; 1,0 і 1,5 с посадку було видно в 1/4, 2/4, 2/4, 4/4 і 4/4 сеансів, а затриманий повтор зльоту одразу після команди посадки позначено лише в 1 з 5. В одному з сеансів із затримкою 0,6 с повідомлення стану надійшло на 43 мс пізніше за команду зльоту: детектор ще бачив режим утримання позиції в повітрі й дав PASS, а автопілот перервав посадку й набрав 11,7 м. Так 12 із 34 таких зльотів отримали PASS через запізнення телеметрії. Урізані варіанти це підтверджують: без телеметрії стану 65 зі 100 рішень стають UNKNOWN, серед них усі 35 тривог за R1, R3 і R4, а без підтверджень автопілота жодне рішення не змінилося. Усі 500 рішень збіглися з очікуваннями, записаними до запуску.

Висновки. Детектор на чотирьох правилах сумісний зі штатною роботою двох наземних станцій, ухвалює визначене рішення на кожній команді, поки є свіжа телеметрія стану, і однаково позначає втручання та дозволені дії, які виглядають однаково. Його можливості обмежує телеметрія: зліт під час посадки лишається непоміченим, якщо випереджає повідомлення про стан. Метод не

перевіряє справжності відправника і не знає наміру оператора; команду в штатному форматі від довіреної наземної станції він не помічає, а одноразова зміна висоти без попереднього зльоту обходить R3. Висновки стосуються одного симуляційного стенда, двох наземних станцій і заданих сценаріїв; для реального апарата і радіоканалу потрібна окрема перевірка.

Література

1. MAVLink Guide. Command Protocol. URL: https://mavlink.io/en/services/command.html (звернення 18.10.2026).
2. Liu R., Zhao X., Liu N., Li J., Wang G. UAV Intrusion Detection Based on iTransformer for MAVLink Message ID Sequence. Proc. of the 15th Int. Conf. on Communication and Network Security (ICCNS 2025). ACM, 2025. P. 61–68. DOI: 10.1145/3789456.3789458.
3. Amorim A., Taylor M., Kann T., Harrison W. L., Leavens G. T., Joneckis L. Enforcing MAVLink Safety & Security Properties via Refined Multiparty Session Types. NASA Formal Methods (NFM 2025). LNCS, vol. 15682. Springer, 2025. P. 1–10. DOI: 10.1007/978-3-031-93706-4_1.
4. Amorim A., Gazzillo P., Taylor M., Joneckis L. From High-Level Types to Low-Level Monitors: Synthesizing Verified Runtime Checkers for MAVLink. NASA Formal Methods (NFM 2026). LNCS, vol. 16622. Springer, 2026. P. 99–119. DOI: 10.1007/978-3-032-28079-4_5.

Чалий О.В.
КНУ імені Тараса Шевченка,
Київ, Україна
Толюпа С.В. д-р техн. наук, проф.
КНУ імені Тараса Шевченка,
Київ, Україна

ГРАФ ЗНАНЬ ДЛЯ ВИЯВЛЕННЯ НЕВІДПОВІДНОСТЕЙ В АУДИТІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Розглянуто концепцію використання графа знань для встановлення невідповідностей у доказовому забезпеченні аудиту інформаційної безпеки. Запропоновано представляти критерії аудиту, об'єкти перевірки, аудиторські процедури, джерела даних, докази та знахідки як взаємопов'язані сутності. Визначено, що графове представлення дає змогу виявляти відсутні або некоректні зв'язки, часові й контекстні невідповідності, неповноту походження доказів і суперечності між джерелами. Наголошено на необхідності підтвердження виявлених невідповідностей аудитором.

Ключові слова: граф знань, аудит інформаційної безпеки, аудиторський доказ, невідповідність, доказова прогалина.

Вступ. Сучасні інформаційні системи характеризуються розподіленою інфраструктурою та використанням різнорідних засобів захисту. Під час аудиту інформаційної безпеки опрацьовуються нормативні документи, внутрішні політики, описи активів, журнали подій, конфігураційні файли та результати сканування вразливостей. Відомості з цих джерел можуть бути неповними, застарілими або суперечливими, що ускладнює перевірку виконання вимог системи управління інформаційною безпекою [2, 3].

Висновок щодо відповідності об'єкта встановленим вимогам має ґрунтуватися на об'єктивних і придатних до перевірки доказах [1]. Для цього необхідно простежити зв'язок між нормативною вимогою, об'єктом перевірки, аудиторською процедурою, отриманим доказом і знахідкою. Втрата або неправильне визначення одного з таких зв'язків може призвести до необґрунтованого висновку.

Засоби автоматизації аудиту переважно виконують окремі операції: інвентаризацію активів, сканування вразливостей, накопичення журналів або ведення вимог і контролів у GRC-платформах. Їх результати часто залишаються у відокремлених системах та не утворюють єдиного доказового представлення аудиту [4]. Метою роботи є визначення можливостей графа знань для виявлення невідповідностей і доказових прогалин без автономного формування аудиторського висновку.

Методичний підхід. Запропоновано концептуальне графове представлення аудиторської інформації. Вершинами графа є нормативне джерело, критерій аудиту, захід контролю, об'єкт перевірки, інформаційний актив, аудиторська процедура, вимога до доказу, джерело даних, фактичний

доказ і аудиторська знахідка. Семантичні зв’язки між ними можуть мати значення «визначає», «застосовується до», «перевіряється», «потребує», «отримано з», «підтверджує», «спростовує» та «обґрунтовує».

Аналіз полягає у зіставленні очікуваного доказового шляху з фактично представленими сутностями та зв’язками. Перевіряються наявність обов’язкового доказу, його належність до визначеного об’єкта й періоду, повнота відомостей про походження та узгодженість із матеріалами інших джерел. Такий підхід є концептуальним: у роботі не пропонується окрема математична модель або програмна реалізація.

Результати та обговорення. Граф знань дає змогу виявляти кілька типів неточностей. До них належать відсутність обов’язкової сутності; відсутній або помилковий зв’язок між доказом, критерієм і об’єктом; невідповідність сфери застосування; часові невідповідності; неповнота відомостей про походження; суперечності між джерелами. Наприклад, сервер може бути позначений IP-адресою у звіті сканера, іменем вузла в журналі подій та обліковим номером у реєстрі активів. Семантичні зв’язки дають змогу розглядати ці записи як відомості про один об’єкт і не дублювати результати перевірки.

Інший приклад стосується щоквартального перегляду привілейованих облікових записів. Критерій аудиту пов’язується із системою керування доступом, процедурою перевірки та очікуваними звітами. Якщо у графі наявні звіти лише за окремі квартали, це вказує на неповноту доказового обґрунтування. Водночас відсутність звіту не доводить, що перегляд не проводився. Тому результатом аналізу має бути повідомлення про доказову прогалину, яка потребує перевірки аудитором.

Суперечність може виникнути, коли політика вимагає журналювання адміністративних дій, технічна конфігурація свідчить про вимкнення відповідного механізму, а звіт підрозділу декларує повне виконання вимоги. Поєднання цих матеріалів навколо одного критерію робить суперечність видимою та забезпечує простежуваний доказовий шлях від вимоги до знахідки.

Обмеженням підходу є те, що структурна повнота графа не гарантує достатності й достовірності доказів. Автоматично встановлений зв’язок також може не враховувати контекст діяльності організації. Тому сутності, зв’язки та виявлені прогалини повинні отримувати підтверджений статус лише після професійного оцінювання аудитором.

Висновки. Граф знань створює основу для структурованого подання аудиторської інформації та виявлення відсутніх, неправильно пов’язаних, застарілих або суперечливих доказів. Практичним результатом є пояснюване повідомлення про елемент доказового обґрунтування, який потребує уточнення. Це підвищує простежуваність аудиту та знижує ризик формування висновків на основі неповних матеріалів, зберігаючи остаточне рішення за аудитором.

Список використаних джерел:

1. ISO 19011:2018. Guidelines for auditing management systems. International Organization for Standardization, 2018.
2. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection. Information security management systems. Requirements. ISO/IEC, 2022.
3. ISO/IEC 27007:2020. Information security, cybersecurity and privacy protection. Guidelines for information security management systems auditing. ISO/IEC, 2020.
4. Chalyi O., Toliupa S. Comparison of International Audit Standards for Information Security with Automation Perspectives. Cybersecurity: Education, Science, Technique. 2025. No. 3(31). P. 579–588. DOI: 10.28925/2663-4023.2025.31.1046.

Козяр Михайло Олександрович
студент групи ІКБмз-11, ННІКБЗІ ДУІКТ, Київ, Україна

ІНТЕГРАЦІЯ DEVSECOPS-ПРАКТИК У GITOPS-КОНВЕЄР ДЛЯ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ХМАРНОЇ ІНФРАСТРУКТУРИ

Розглянуто підхід до підвищення безпеки хмарної інфраструктури шляхом інтеграції практик DevSecOps у GitOps-конвеєр. Проаналізовано типові загрози для хмарних середовищ, пов'язані з помилками конфігурації та компрометацією CI/CD, і запропоновано послідовність контрольних точок безпеки на етапах коду, збирання, розгортання та експлуатації. Підхід співвіднесено з функціями NIST CSF 2.0. Запропоновано показники оцінювання ефективності контрольних точок і шляхи зменшення хибних спрацювань.

Ключові слова: кібербезпека, хмарні технології, DevSecOps, GitOps, CI/CD, NIST CSF.

Міграція корпоративних сервісів у хмарні середовища змінила модель захисту інформації: мережевий периметр поступився місцем ідентифікаторам, API та конфігураціям як основним об'єктам атаки. Значна частка інцидентів у хмарі виникає не через вразливості провайдера, а через помилки на боці клієнта в межах моделі спільної відповідальності [1]: надмірні права доступу, відкриті сховища даних, секрети у відкритому вигляді в репозиторіях. Водночас практики інфраструктури як коду (IaC) та GitOps, за яких бажаний стан інфраструктури описується в Git-репозиторії й автоматично застосовується конвеєром, роблять сам конвеєр доставки критичним активом: його компрометація дає зловмиснику контроль над усім середовищем [2].

Метою роботи є визначення набору контрольних точок безпеки, які можна вбудувати в GitOps-конвеєр без втрати швидкості доставки змін. Для цього розглянуто типові загрози для хмарних середовищ і сформовано відповідні до них перевірки на кожному етапі конвеєра.

Аналіз інцидентів і методик оцінювання ризиків [2] дозволяє виділити чотири групи загроз. По-перше, компрометація облікових даних: ключі доступу, токени та паролі, випадково збережені в коді або змінних середовища конвеєра. По-друге, помилки конфігурації: публічні сховища даних, надто широкі правила мережевого доступу, відсутність шифрування. По-третє, атаки на ланцюг постачання: підміна залежностей, зловмисні образи контейнерів, зміна збірного скрипту. По-четверте, зловживання самим конвеєром: надмірні права його облікового запису, виконання непідтвердженого коду в збірних агентах.

Пропонується розглядати конвеєр як послідовність чотирьох етапів, на кожному з яких виконуються автоматизовані перевірки. На етапі коду застосовуються пошук секретів (pre-commit-хуки та перевірка на стороні сервера), захищені гілки, обов'язкове рев'ю змін і підписані коміти. Це зменшує ймовірність потрапляння облікових даних і шкідливого коду в основну гілку, а

підписані коміти дозволяють встановити автора кожної зміни інфраструктури.

На етапі збирання виконуються статичний аналіз коду (SAST), аналіз залежностей (SCA), генерація переліку компонентів (SBOM), сканування образів контейнерів та підписання артефактів. Ці заходи відповідають рекомендаціям щодо захисту ланцюга постачання програмного забезпечення в CI/CD-конвеєрах [3]. Підписаний образ разом із SBOM стає перевіреним артефактом, який на подальших етапах можна автентифікувати перед розгортанням, а не лише довіряти назві й тегу образу.

На етапі розгортання IaC-описи (Terraform, маніфести Kubernetes) перевіряються на небезпечні конфігурації, а політики як код (policy as code) блокують зміни, що порушують вимоги безпеки: образи без підпису, контейнери з привілейованим режимом, публічні сховища, відсутність шифрування. Сервісний обліковий запис конвеєра отримує мінімально необхідні права, а статичні ключі замінюються короткоживучими обліковими даними, які видаються на час виконання завдання через федерацію ідентичності.

На етапі експлуатації забезпечується безперервний контроль відхилення фактичної конфігурації від стану, зафіксованого в Git, централізоване журналювання дій та виявлення аномалій. У моделі GitOps будь-яка ручна зміна розглядається як відхилення й автоматично скасовується, що водночас ускладнює непомітне закріплення зловмисника в середовищі.

Запропоновану послідовність доцільно співвіднести з NIST CSF 2.0 [4]. Етапи коду та збирання підтримують функції Identify і Protect (керування активами, доступом і ланцюгом постачання), етап розгортання - функцію Protect, етап експлуатації - функції Detect, Respond і Recover. Єдині політики, ролі та відповідальність за конвеєр належать до функції Govern, яка з’явилася у версії 2.0 і підкреслює роль управління в кібербезпеці.

Ефективність запропонованого підходу доцільно оцінювати за невеликою кількістю показників: часом від появи вразливості або помилки конфігурації до її виправлення, часткою змін, що пройшли перевірки без ручного втручання, кількістю виявлених секретів у репозиторіях і часткою хибних спрацювань. Скорочення першого показника та зростання другого за стабільних значень решти свідчить про те, що безпека інтегрована в процес доставки й не гальмує його.

Основним ризиком такого підходу є надмірна кількість перевірок, що спричиняє хибні спрацювання й знижує довіру команд до інструментів безпеки. Тому впроваджувати контрольні точки доцільно поетапно: спочатку блокувальні перевірки лише для критичних ризиків (секрети в коді, публічні сховища, адміністративні права), решту - у режимі попередження з поступовим переведенням у блокувальний режим. Окремої уваги потребує захист самого конвеєра: розмежування прав, ізоляція збірних агентів і журналювання їхніх дій.

Отже, вбудовування контрольних точок безпеки у GitOps-конвеєр

перетворює його з потенційної точки компрометації на основний механізм забезпечення та підтвердження безпеки хмарної інфраструктури. Подальші дослідження пов’язані з кількісною оцінкою впливу таких перевірок на час доставки змін і частку хибних спрацювань.

Література:

1. AWS Shared Responsibility Model. Amazon Web Services. URL: https://aws.amazon.com/compliance/shared-responsibility-model/

2. OWASP Top 10 CI/CD Security Risks. OWASP Foundation. URL: https://owasp.org/www-project-top-10-ci-cd-security-risks/

3. NIST SP 800-204D. Strategies for the Integration of Software Supply Chain Security in DevSecOps CI/CD Pipelines. NIST, 2024.

4. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. NIST, 2024.

Колесникова А. Я.
студентка групи ІКБМ-51, ННІКБЗІ ДУІКТ,
Київ, Україна

ОЦІНЮВАННЯ КІБЕРЗАХИЩЕНОСТІ ХМАРНИХ СЕРЕДОВИЩ НА БАЗІ МАТЕМАТИЧНИХ МОДЕЛЕЙ ТА СТАНДАРТІВ ISO

У тезах розглядається актуальна проблема забезпечення кібербезпеки під час міграції корпоративних бізнес-процесів до хмарних провайдерів. Проаналізовано ключові загрози, зокрема неправильну конфігурацію систем безпеки, несанкціонований доступ та недостатній захист API-інтерфейсів. Досліджено комплексний метод математичного оцінювання кіберзахищеності постачальників хмарних послуг (IaaS, PaaS, SaaS) для проведення аудиту інфраструктури, що складається з 11 етапів. Обґрунтовано необхідність імплементації міжнародних стандартів ISO/IEC 27017 та 27018 для захисту конфіденційної інформації та мінімізації ризиків витоку даних.

Ключові слова: кібербезпека, хмарні провайдери, міграція сервісів, аудит інфраструктури, інформаційна безпека, ISO/IEC 27017, захист даних.

У сучасному світі спостерігається активна міграція бізнес-процесів компаній до хмарних сервісів, таких як Amazon Web Services, Microsoft Azure, Google Cloud Platform тощо, з метою зменшення витрат на підтримку власної інфраструктури та підвищення гнучкості [2, с. 110]. Незважаючи на очевидні переваги, проблема кібербезпеки залишається відкритою, оскільки неправильне налаштування орендованих ресурсів створює високий ризик витоку конфіденційної інформації [2, с. 111–112].

Як зазначають провідні компанії у сфері інформаційної безпеки (Proofpoint, CrowdStrike, Check Point), основними загрозами для організацій є неправильна конфігурація систем захисту, можливість несанкціонованого доступу, недостатній захист API-інтерфейсів та відсутність ешелонованих систем протидії кібератакам [1, с. 75]. На сьогодні більшість компаній спираються на базовий стандарт ISO/IEC 27001, проте він орієнтований переважно на приватні (Private) мережі і не надає чіткого плану дій для систем, опублікованих у загальнодоступній мережі Інтернет [1, с. 75–76]. Саме тому виникає об’єктивна потреба у застосуванні спеціалізованих галузевих стандартів, таких як ISO/IEC 27017 (засоби контролю інформаційної безпеки для хмарних послуг) та ISO/IEC 27018 (захист персональних даних у публічних хмарах) [3, 4], а також у розробці дієвих математичних методів аудиту провайдерів. Метою роботи є опис методу оцінювання кіберзахищеності хмарних сервісів на базі математичної моделі [1, с. 76].

Для комплексного аудиту постачальників хмарних послуг (за моделями IaaS, PaaS, SaaS тощо) пропонується використовувати метод, який базується на 11 ключових етапах [1, с. 76]. Оцінювання здійснюється за допомогою системи балів від 0 (мінімальний бал) до 5 (максимальний бал) для кожного параметра [1, с. 76].

Етапи 1–10 передбачають послідовну оцінку конкретних модулів хмарного середовища:

1. **GPCSP** (General Points) – опис загальних положень постачальника послуг [1, с. 76].
2. **NCSP** (Network) – рівень захищеності мережі, її пропускна здатність та протидія DDoS-атакам [1, с. 77].
3. **SCSP** (Storage) – кіберзахищеність накопичувачів та можливість відновлення масивів даних [1, с. 79].
4. **SRCSP** (Server) – здатність серверного обладнання протидіяти фізичному впливу [1, с. 79].
5. **VCSP** (Virtualization) – ступінь захищеності середовища віртуалізації провайдера [1, с. 79–80].
6. **OSCSP** (Operation System) – оцінка рівня безпеки обраних операційних систем [1, с. 80].
7. **CTCSP** (Container Technology) – захищеність розгорнутих контейнерів [1, с. 81].
8. **RCSP** (Runtime) – оцінка систем моніторингу та забезпечення безперервної роботи [1, с. 81].
9. **ACSP** (Application) – спроможність пропонованого застосунку протидіяти атакам [1, с. 82].
10. **DCSP** (Data) – виявлення потенційних джерел витоку даних компанії [1, с. 82].

Варто зазначити, що система оцінювання адаптується під тип сервісу: наприклад, модуль операційної системи (OSCSP) не оцінюється для моделі IaaS, але є обов'язковим для моделей SaaS та PaaS [1, с. 77]. Імплементація настанов зі стандартів ISO/IEC 27017 та ISO/IEC 27018 [3, 4] на етапах аудиту мережі та обробки даних дозволяє забезпечити повну відповідність інфраструктури міжнародним вимогам.

На фінальному 11-му етапі (RECSP) обчислюється загальний коефіцієнт кіберзахищеності (CC), що становить відсоткове відношення суми набраних балів до максимально можливої їх кількості для конкретного типу хмарного сервісу (наприклад, 615 балів для SaaS або 330 балів для IaaS) [1, с. 82-83]. На основі отриманого коефіцієнта формується кінцева рекомендація для компанії-замовника: хмарний сервіс може отримати статус «Рекомендується до використання (Recommended)», «Можливе до використання (Maybe)» або «Не рекомендується до використання (No Recommended)» [1, с. 83].

Список літератури

1. Педченко Є. М., Іванченко І. С. Метод оцінювання кіберзахищеності хмарних сервісів об'єктів інформаційної інфраструктури. *Сучасний захист інформації.* 2024. № 3(59). С. 75–84. DOI: 10.31673/2409-7292.2024.030008

2. Pedchenko Y., Ivanchenko Y., Ivanchenko I., Lozova I., Jancarczyk D., Sawicki P. Analysis of modern cloud services to ensure cybersecurity. *Procedia Computer Science*. 2022. Vol. 207. P. 110–117. DOI: 10.1016/j.procs.2022.09.043

3. ISO/IEC 27017:2015. Information technology — Security techniques — Code of practice for information security controls based on ISO/IEC 27002 for cloud services.

4. ISO/IEC 27018:2019. Information technology — Security techniques — Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors.

Алексанян Андрій Ельдарович
Студент групи КНм-21, ННІТ, ДУІКТ
Київ, Україна

ІНТЕГРАЦІЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ В SOC ДЛЯ ВИЯВЛЕННЯ АНОМАЛЬНОЇ ПОВЕДІНКИ КОРИСТУВАЧІВ

Розглянуто підхід до інтеграції моделей машинного навчання в центр операцій безпеки (SOC) для аналізу поведінки користувачів ІКТ-послуг. Запропоновано концептуальну послідовність оброблення подій: збір журналів, формування поведінкових ознак, оцінювання відхилень, кореляція сигналів і передавання пріоритетних сповіщень аналітику. Визначено критерії оцінювання такого рішення та основні ризики його впровадження, зокрема хибні спрацювання, зміни звичайної поведінки й захист персональних даних.
Ключові слова: SOC, машинне навчання, поведінкова аналітика, виявлення аномалій, моніторинг кібербезпеки, ІКТ-послуги.

Центр операцій безпеки (Security Operations Center, SOC) забезпечує централізований моніторинг подій, виявлення інцидентів і координацію реагування. На практиці SOC може отримувати значну кількість сповіщень від засобів захисту, журналів автентифікації, мережевих сервісів і корпоративних застосунків. Якщо аналіз спирається лише на фіксовані правила, нетипова активність облікового запису може залишитися непоміченою або, навпаки, породити надмірну кількість хибних сигналів. Тому актуальним є підхід, за якого поведінкова аналітика на основі машинного навчання доповнює наявні механізми моніторингу та допомагає визначати події, що потребують першочергової перевірки.

Доцільно розглядати таку інтеграцію як послідовність взаємопов’язаних етапів. На першому етапі збираються журнали входів, результати автентифікації, час і частота звернень до сервісів, типи пристроїв та інші доступні події. На другому етапі дані очищуються, узгоджуються за часом і перетворюються на ознаки, що описують активність користувача у визначеному часовому вікні. Далі модель оцінює відхилення від індивідуального або групового профілю поведінки. Отриманий сигнал не слід трактувати як доказ атаки: його потрібно зіставити з іншими подіями, правилами безпеки та контекстом роботи користувача.

Для практичного використання SOC важливо не лише виявити відхилення, а й визначити порядок їх опрацювання. Концептуально пріоритет сповіщення може формуватися з урахуванням оцінки аномальності, критичності ресурсу, рівня привілеїв облікового запису та наявності пов’язаних подій. Така схема дає змогу відокремлювати одиничні нетипові дії від сукупностей ознак, які разом становлять більший ризик. Вагові коефіцієнти та порогові значення мають визначатися під час перевірки на даних конкретної організації, а не задаватися

універсально. У SOC результат аналізу може бути переданий аналітику як пояснюване сповіщення із зазначенням ознак, що вплинули на його пріоритет.

За наявності розмічених прикладів нормальної та підозрілої активності можна розглядати класифікаційні алгоритми. Якщо підтверджених інцидентів мало, доцільним напрямом є методи виявлення аномалій, які навчаються на типових сценаріях поведінки. Для реалізації прототипу мовою Python можуть застосовуватися засоби підготовки даних і бібліотеки машинного навчання, зокрема pandas та scikit-learn. Оцінювання має враховувати precision, recall і F1-score, а також кількість хибних спрацювань та час, потрібний аналітику для опрацювання сповіщення. Самої загальної точності недостатньо, оскільки частка підозрілих подій зазвичай значно менша за частку звичайної активності [1; 2].

Поведінка працівників змінюється залежно від завдань, графіків і робочих обставин, тому модель потребує періодичної перевірки та контролю дрейфу даних. Для зменшення ризику необґрунтованих висновків сповіщення машинного навчання мають підтримувати, а не замінювати рішення фахівця. Важливими є також мінімізація зібраних даних, обмеження доступу до журналів, контроль строків зберігання та документування правил використання результатів. Ці заходи слід узгоджувати з процесами управління ризиками й реагування на інциденти, що є складовими системного підходу до кібербезпеки [3].

Інтеграція поведінкової аналітики на основі машинного навчання в SOC може допомогти впорядкувати потік подій і зосередити увагу аналітиків на найбільш значущих відхиленнях. Запропонована концепція охоплює підготовку журналів, формування ознак, оцінювання аномальності, кореляцію подій і визначення пріоритету сповіщень. Подальша перевірка підходу має передбачати тестування на репрезентативних даних, порівняння алгоритмів і вимірювання не лише якості виявлення, а й операційного навантаження на SOC.

Список використаних джерел

1. Scikit-learn developers. Model evaluation: quantifying the quality of predictions. URL: https://scikit-learn.org/stable/modules/model_evaluation.html.

2. Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey. ACM Computing Surveys. 2009. Vol. 41, No. 3. Article 15. DOI: https://doi.org/10.1145/1541880.1541882.

3. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. 2024. DOI: https://doi.org/10.6028/NIST.CSWP.29.

Вострих Д.О.
студент групи ІКБм-61, ННІКБЗІ, ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАХИСТУ СУЧАСНИХ МЕРЕЖ НА ОСНОВІ PALO ALTO NGFW

Нові виклили кібербезпеці мереж з'являються ледь не щодня, а з появою ШІ інструментів конвенціональні методи захисту стають все менш надійними. У роботі проаналізовано недоліки стандартних систем захисту мереж і розглянуто технологію захисту на основі Palo Alto Networks NGFW. Зокрема механізми ідентифікації застосунків, користувачів та пристроїв та їх заємодія для попердження вторгнення чи несанкціонованих дій.

Ключові слова: кібербезпека, Palo Alto Networks, міжмережевий екран нового покоління, мережевий трафік.

Розвиток цифрової інфраструктури, поширення хмарних сервісів і збільшення кількості кіберзагроз виманабит удосконалення технологій захисту комп'ютерних мереж. Традиційні міжмережеві екрани, які переважно контролюють трафік на основі IP-адрес, портів і мережевих протоколів, вже не є додастатнім. Зловмисники можуть використовують зашифровані канали зв'язку, взламані легітимні програми та наявні у системі вразливості. З появою ШІ інструментів у них також з'явилося значно більше можливостей для найрізноманітніших спроб прокнинення, часто через зовсім неочікуванні вектори, які закрити традиційними системами є майже неможливим. Звіт Palo Alto Networks[1] виявив ШІ інструменти як значний фактор успіху та знаходженню нових шляхів атаки на цифрову інфраструктуру. У зв'язку з цим актуальним є застосування міжмережевих екранів нового покоління (Next-Generation Firewall, NGFW), які можуть працювати зі значно більшим контекстом, ніж традиційні системи.

Palo Alto Networks Next-Generation Firewall[2] — це міжмережевий екран наступного покоління, що поєднує традиційні методи фільтрації трафіку з функціями контролю додатків, пристроїв, моніторингу мережевої активності та виявлення та попередження інцидентів. Для цього використовуються наступні технології:

- App-ID – забезпечує ідентифікацію додатків незалежно від використовуваних порітв та протоколів, що дозволяє обмежувати використання небажаних сервісів, контролювати доступ до дозволених застосунків та виявляти аномальні поведінки додактів чи користувачів.
- User-ID – пов'язує активність з користувачем з можиливістю інтеграції з Active Directory або інших постачальників ідентифікації.
- Content-ID – дозволяє перевіряти вміст трафіку за допомогою Deep Packet Inspection та контролювати передачу даних. Також ця технологія може

співпрацювати зі вбудованою системою запобіганню вторгнення(IPS) та хмарною пісочнецею (Wildfire) для перевірки файлів [2].

- Device-ID – забезпечує контроль та індентифікацію IoT пристроїв.

Практичне застосування системи передбачає створення коплексних політик, що регулюють взаємодію між сегментами мережі, користувачами, застосунками, пристроями та зовнішніми ресурсами. Наприклад, можна відокремити робочі станції користувачів від серверів, дозволити доступ до служб відповідно до потреб і повноваженнь користувачів і заблокувати несанкціоновані з'єднання. Для підвищення ефективності системи необхідно використовувати принцип мінімальних привілеїв. Для великих мереж також є можливість використання централізованої панелі керування або хмарна панель, що поєднують багато інших сервісів. Перевагою такого підходу є значний контроль того, що може відбудватися в мережі, що значно звужує потенційні вектори атак.

Palo Alto Networks NGFW забезпечує копмлексний підхід до захисту сучасних комп'ютерних мереж завдяки поєднанню ідентифікації застосунків і користувачів, контролю вмісту трафіку, фільтрації ресурсів, файлів та аналізу поведінки користувачів. Однак ватор зазначити, що система сама по собі не гарантує абсолютного захисту, адже неправильно налаштовані політики, застарілі сигнатури і відстуність переіодичного їх перегляду можуть значно вплинути на її ефективність.

Перелік посилань:

1. Palo Alto Networks Unit 42 Global Incident Response Report 2026. Режим доступу: https://www.paloaltonetworks.com/content/dam/pan/en_US/assets/pdf/unit42/Unit42-Global-Incident-Response-Report.pdf
2. What Is a Next-Generation Firewall (NGFW)? A Complete Guide. Режим доступу: https://www.paloaltonetworks.com/cyberpedia/what-is-a-next-generation-firewall-ngfw.
3. App-ID Режим доступу: https://docs.paloaltonetworks.com/ngfw/administration/app-id.

Куценко Олексій Сергійович
студент гр. УБДМ-61
ННІЗІ, ДУІКТ, м. Київ

ЗАХИСТ IoT-ПРИСТРОЇВ ТЕЛЕМАТИЧНИХ СИСТЕМ НА ЕТАПАХ ВПРОВАДЖЕННЯ ТА ЕКСПЛУАТАЦІЇ

Анотація. *У тезах розглянуто особливості захисту IoT-пристроїв у складі телематичних систем моніторингу транспорту та стаціонарних об'єктів. Проаналізовано типову архітектуру телематичної системи та визначено основні вектори атак на GPS-трекери, канали передачі даних і серверну платформу. Обґрунтовано доцільність розподілу заходів захисту відповідно до етапів життєвого циклу системи – впровадження та експлуатації. Запропоновано перелік базових організаційно-технічних заходів для кожного етапу, а також підхід до виявлення аномалій у навігаційних даних, зокрема ознак глушіння та спуфінгу сигналів GNSS.*

Ключові слова: *Інтернет речей, телематична система, GPS-трекер, GNSS-спуфінг, життєвий цикл, кібербезпека.*

Стрімке поширення технологій Інтернету речей (IoT) призвело до того, що телематичні системи стали невід'ємною частиною інфраструктури логістичних, аграрних, комунальних підприємств, а також організацій оборонного сектору. Основою таких систем є GPS-трекери – спеціалізовані IoT-пристрої, що визначають місцеположення об'єкта, збирають дані з датчиків і передають їх на сервер через мобільну мережу. На відміну від класичних вузлів корпоративної мережі, ці пристрої працюють автономно, мають обмежені обчислювальні ресурси, перебувають поза контрольованим периметром і фізично доступні стороннім особам.

В умовах воєнного стану інформація про місцеположення та маршрути руху транспорту набуває ознак чутливої, а її витік або підміна можуть мати наслідки, що виходять далеко за межі економічних збитків. Водночас, як свідчить практика впровадження телематичних рішень, забезпечення безпеки часто зводиться до вибору надійного виробника обладнання, тоді як заводські налаштування пристроїв і серверної платформи залишаються незмінними. Це суперечить вимогам національного законодавства щодо захисту інформаційних систем та об'єктів критичної інфраструктури [4].

Типова телематична система має чотири рівні: кінцевий пристрій (трекер із GNSS-приймачем, GSM/LTE-модемом, SIM-карткою та інтерфейсами RS-232, RS-485, CAN, 1-Wire); канал передачі даних (мережа мобільного оператора, протоколи TCP/UDP, службові SMS-команди); серверна платформа (сервер прийому даних, база даних, веб-інтерфейс, прикладний програмний інтерфейс); клієнтські застосунки та інтеграції з обліковими системами. Кожен рівень формує власну поверхню атаки. Більшість категорій вразливостей OWASP IoT Top 10 [3] – слабкі паролі, незахищені мережеві сервіси, відсутність безпечного механізму оновлення, незахищена передача даних, небезпечні налаштування за

замовчуванням, недостатній фізичний захист – безпосередньо застосовні до телематичного обладнання.

До характерних для телематичних систем загроз доцільно віднести такі. По-перше, віддалене переналаштування трекера за допомогою SMS або мережевих команд у разі використання пароля за замовчуванням чи відсутності обмеження номерів-відправників. Це дозволяє перенаправити потік даних на сторонній сервер або керувати вихідними лініями пристрою (наприклад, реле блокування двигуна). По-друге, передача даних у відкритому вигляді за протоколами поверх TCP/UDP, що уможливлює їх перехоплення та модифікацію. По-третє, фізичний доступ до сервісних інтерфейсів пристрою та вилучення SIM-картки для несанкціонованого використання. По-четверте, глушіння та спуфінг сигналів GNSS, які дають змогу отримати недостовірні координати без компрометації самого пристрою. По-п’яте, компрометація облікових записів платформи моніторингу (слабкі паролі, відсутність двофакторної автентифікації, спільні облікові записи диспетчерів).

Аналіз зазначених загроз показує, що значна частина вразливостей закладається саме на етапі впровадження системи, а проявляється вже під час її експлуатації. Тому запропоновано структурувати заходи захисту відповідно до етапів життєвого циклу, спираючись на базовий набір можливостей кібербезпеки IoT-пристроїв, визначений у NISTIR 8259A [1], та базові вимоги стандарту ETSI EN 303 645 [2].

На етапі впровадження пропонується: обирати обладнання з урахуванням безпекових можливостей (шифрування каналу, захищене оновлення вбудованого програмного забезпечення, парольний захист команд); проводити інвентаризацію з фіксацією IMEI, ICCID SIM-картки, версії прошивки та прив’язки до об’єкта. Замінювати паролі за замовчуванням і вимикати невикористовувані інтерфейси та сервіси. Обмежувати керування через SMS «білим списком» номерів. Використовувати приватний APN з маршрутизацією лише до сервера телематики, а за можливості – VPN-тунель між мережею оператора та серверною частиною. На серверній платформі запроваджувати розмежування прав доступу, двофакторну автентифікацію та мінімізацію відкритих мережевих портів.

На етапі експлуатації ключовими є: централізоване керування версіями прошивок і контрольоване оновлення. Моніторинг стану пристроїв (втрата зв’язку, нештатні перезавантаження, зміна конфігурації, заміна SIM-картки), журналювання дій користувачів платформи та періодичний аудит конфігурацій, наявність процедури реагування, що передбачає виведення скомпрометованого пристрою з експлуатації, блокування SIM-картки та повторне конфігурування.

Окремим напрямом є виявлення аномалій у навігаційних даних, яке дозволяє протидіяти атакам на GNSS без модифікації апаратної частини. Ознаками глушіння або спуфінгу можуть слугувати: різкі «стрибки» координат, фізично неможлива швидкість переміщення між двома послідовними точками,

розбіжність між GNSS-координатами та орієнтовним місцеположенням за даними базових станцій, раптове зменшення кількості видимих супутників або зростання геометричного фактора точності. Поєднання цих ознак у правилах кореляції на серверній платформі дозволяє автоматично позначати недостовірні дані та сповіщати оператора.

Отже, захист IoT-пристроїв у складі телематичних систем не може обмежуватися вибором обладнання чи окремими технічними засобами. Розподіл заходів за етапами життєвого циклу дає змогу усунути типові вразливості ще під час впровадження та підтримувати належний рівень захищеності під час експлуатації, а аналіз аномалій навігаційних даних підвищує стійкість системи до атак на GNSS, актуальних для сучасних умов. Подальші дослідження доцільно спрямувати на формалізацію правил виявлення аномалій та оцінювання їх ефективності на експериментальному прототипі.

Список використаних джерел

1. Fagan M., Megas K. N., Scarfone K., Smith M. IoT Device Cybersecurity Capability Core Baseline : NISTIR 8259A. Gaithersburg : National Institute of Standards and Technology, 2020. DOI: 10.6028/NIST.IR.8259A.

2. ETSI EN 303 645 V2.1.1 (2020-06). Cyber Security for Consumer Internet of Things: Baseline Requirements. Sophia Antipolis : ETSI, 2020.

3. OWASP Internet of Things Top 10 – 2018. OWASP Foundation. URL: https://owasp.org/www-project-internet-of-things/

4. Про основні засади забезпечення кібербезпеки України : Закон України від 05.10.2017 № 2163-VIII. URL: https://zakon.rada.gov.ua/laws/show/2163-19

Піскунов Костянтин Валерійович
студент групи ІКБм-61, ННІЗІ ДУІКТ,
Київ, Україна

ТЕХНОЛОГІЯ ЗАПОБІГАННЯ ФІШИНГОВИМ АТАКАМ, СПРЯМОВАНИМ НА ВИКРАДЕННЯ ОБЛІКОВИХ ДАНИХ КОРИСТУВАЧІВ

Запропоновано технологію захисту корпоративних облікових даних, яка поєднує фільтрацію повідомлень, контроль вебресурсів, фішингостійку автентифікацію та реагування на компрометацію. Визначено порядок застосування захисних механізмів, обмеження підходу й показники оцінювання його ефективності. Очікуваний результат полягає у зниженні ризику компрометації облікових записів завдяки перевірці повідомлень і сторінок до введення даних та прив'язці автентифікації до легітимного сервісу.

Ключові слова: фішинг, облікові дані, багатофакторна автентифікація, FIDO2, WebAuthn, контроль доступу.

Фішингові повідомлення спонукають користувачів переходити на підроблені сторінки та розкривати паролі й одноразові коди. У дослідженні D. Lain, K. Kostiainen та S. Čapkun за участю понад 14 тисяч працівників упродовж 15 місяців показано, що поширений формат навчання під час симуляцій не гарантує підвищення стійкості до фішингу [1]. Додатковою проблемою є передавання викраденого коду справжньому сервісу в реальному часі. За NIST SP 800-63B-4, автентифікація з ручним введенням одноразового коду не вважається фішингостійкою [2]. Це обґрунтовує необхідність технічного захисту, який зменшує залежність від уважності користувача.

Метою роботи є обґрунтування технології запобігання викраденню та зловмисному використанню облікових даних шляхом узгодженого застосування засобів виявлення фішингу, контролю автентифікації та припинення підозрілих сеансів. Технологію запропоновано для організації з керованими робочими пристроями та централізованим постачальником ідентичності. Модель загроз охоплює підроблені сторінки входу, посилання в листах і QR-кодах, ретрансляцію одноразових кодів та використання викрадених сеансових токенів.

Перший етап захисту передбачає перевірку повідомлень. Поштовий шлюз аналізує відправника, вкладення й посилання, зокрема адреси, закодовані в QR-кодах. Перевірки SPF, DKIM і DMARC застосовують для протидії підміні домену відправника. DMARC узгоджує домен у полі From з доменом, підтвердженим SPF або DKIM [3]. Водночас успішна перевірка не підтверджує безпечність змісту листа: повідомлення може надійти зі скомпрометованого облікового запису або схожого домену.

На наступному етапі здійснюється контроль вебресурсу. У керованому браузері перевіряють повне доменне ім'я, перенаправлення, репутацію адреси та

відповідність форми входу переліку погоджених доменів. Підтверджені фішингові ресурси блокують; невідомі сторінки з ознаками імітації корпоративного входу передають на додаткову перевірку. Наявність HTTPS сама по собі не є підставою дозволяти введення даних. Контроль має спрацьовувати до їх введення, оскільки шкідливий сценарій може зчитувати символи ще до натискання кнопки надсилання.

Для корпоративного входу застосовують FIDO2/WebAuthn з локальною перевіркою користувача за PIN-кодом або біометричною ознакою. Автентифікатор створює криптографічну відповідь, прив'язану до домену легітимного сервісу, без передавання секретного ключа вебсайту. Сервер перевіряє одноразовий виклик, походження запиту, ідентифікатор сервісу та цифровий підпис [2]. Для відновлення доступу передбачають резервний автентифікатор та перевірену процедуру, яка запобігає обходу захисту через слабший спосіб входу.

Взаємодію компонентів запропонованої технології наведено на рис. 1. Перевірки повідомлень і вебресурсів доповнюються фішингостійкою автентифікацією, а їхні події використовуються для виявлення інцидентів та оновлення індикаторів загроз.

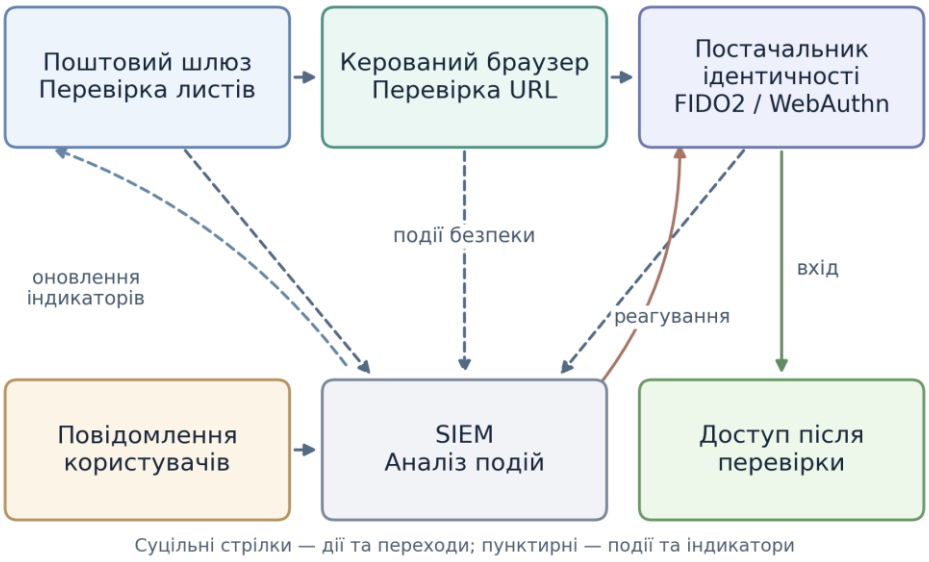


Рис. 1. Взаємодія компонентів технології запобігання фішингу

Події поштового шлюзу, браузера й постачальника ідентичності передають до системи керування інформацією та подіями безпеки (SIEM). За поєднання переходу на підтверджений фішинговий ресурс і підозрілого входу запускають перевірку інциденту. У разі підтвердження компрометації припиняють сеанси постачальника ідентичності та окремих застосунків, відкликають підтримувані системою токени й відновлюють захищений доступ [2]. Дії виконують за узгодженими процедурами реагування.

У поштовому клієнті передбачають кнопку повідомлення про підозрілий лист. Після перевірки такого повідомлення нові індикатори передають засобам блокування. Дослідження [1] підтверджує практичну придатність залучення працівників до колективного виявлення фішингу. Навчання зосереджують на розпізнаванні підозрілих запитів і швидкому повідомленні про них.

Правило ухвалення рішення розділяє підтверджені загрози та евристичні ознаки. Збіг із перевіреним індикатором зумовлює блокування; лише підозріла структура адреси потребує додаткової перевірки. Сам факт нового пристрою не повинен автоматично вважатися компрометацією. Особливість запропонованого порядку полягає в поєднанні блокування фішингової сторінки до введення даних із перевіркою права на вхід незалежно від того, чи була сторінка вчасно розпізнана.

Для оцінювання ефективності технології передбачено порівняння базової фільтрації, фільтрації з одноразовими кодами та запропонованого комплексу на однакових сценаріях і тестових облікових записах. Окремо враховують частку виявлених фішингових ресурсів серед усіх фішингових ресурсів, частку помилково заблокованих легітимних ресурсів, затримку перевірки та частку успішних захоплень облікового запису. Додатково вимірюють час від підтвердження інциденту до припинення доступу. Числові значення цих показників мають бути отримані під час випробувань.

Впровадження технології пропонується починати з пілотної групи користувачів. Для неї реєструють основний і резервний автентифікатори, визначають дозволені домени входу та перевіряють надходження подій до SIEM. На тестових облікових записах відтворюють перехід за QR-кодом, імітацію сторінки входу, ретрансляцію коду та втрату автентифікатора. Для кожного сценарію фіксують етап зупинення атаки, спрацювання контролів і доступність легітимних сервісів. За результатами пілотування уточнюють правила блокування та порядок відновлення доступу. Перед розгортанням у всій організації перевіряють сумісність сервісів, визначають відповідальних за обробку повідомлень користувачів і документують порядок припинення скомпрометованих сеансів.

Межі застосування визначаються станом кінцевого пристрою та повнотою інтеграції. Фішингостійка автентифікація не усуває загрозу викрадення вже створеного сеансу шкідливим програмним забезпеченням. Тому її доповнюють захистом пристроїв, моніторингом сеансів і контрольованим відновленням доступу.

Отже, запропонована технологія охоплює доставлення фішингового повідомлення, відкриття сторінки, автентифікацію та подальше використання сеансу. Її практична цінність полягає в узгодженні захисних дій на цих етапах. Очікуване зменшення кількості компрометацій потребує підтвердження випробуваннями з одночасним контролем помилкових блокувань.

Перелік посилань:

1. Phishing in Organizations: Findings from a Large-Scale and Long-Term Study [Електронний ресурс] / D. Lain, K. Kostiainen, S. Čapkun // 2022 IEEE Symposium on Security and Privacy (SP). 2022. — Режим доступу: https://doi.org/10.1109/SP46214.2022.9833766 (дата звернення: 23.10.2026).
2. Digital Identity Guidelines: Authentication and Authenticator Management. SP 800-63B-4 [Електронний ресурс] // National Institute of Standards and Technology. 2025. — Режим доступу: https://doi.org/10.6028/NIST.SP.800-63B-4 (дата звернення: 23.10.2026).
3. Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489 [Електронний ресурс] / M. Kucherawy, E. Zwicky. 2015. — Режим доступу: https://doi.org/10.17487/RFC7489 (дата звернення: 23.10.2026).

Черненко Д. А.
студентка групи ІКБм-62, ННІКБЗІ ДУІКТ,
Київ, Україна

ТЕХНОЛОГІЇ ЗАХИСТУ КОРПОРАТИВНИХ ДАНИХ НА БАЗІ MICROSOFT PURVIEW DLP

У сучасних умовах масового переходу організацій на хмарні сервіси та гібридні формати роботи зростає ризик випадкових або навмисних витоків конфіденційної інформації. Обсяги даних, що циркулюють через електронну пошту, месенджери та робочі станції працівників, роблять традиційні засоби захисту недостатньо ефективними.

Впровадження технологій запобігання витокам даних (Data Loss Prevention, DLP) стає необхідною умовою забезпечення кіберстійкості бізнесу. У цьому контексті рішення Microsoft Purview DLP забезпечує наскрізний моніторинг, класифікацію та автоматизований захист корпоративної інформації в єдиній хмарній екосистемі.

Ключові слова: к інформаційна безпека, витік даних, Microsoft Purview DLP, класифікація даних, політики безпеки.

Платформа Microsoft Purview DLP пропонує комплексний підхід до виявлення та запобігання втраті інформації на основі уніфікованого аналізу контексту та контенту. Основу технологічного стеку складає багаторівнева система класифікації, яка поєднує декілька фундаментальних механізмів виявлення критичної інформації:

- типи чутливої інформації - використання зумовлених шаблонів (наприклад, регулярні вирази для пошуку номерів банківських карток, паспортних даних чи системних секретів);
- навчені класифікатори - використання моделей машинного навчання для ідентифікації складного та неструктурованого контенту, такого як вихідний код, фінансова звітність, юридичні договори чи медичні картки;
- точне співставлення даних - створення криптографічних цифрових відбитків структурних баз даних для виявлення точних збігів персональних записів із мінімізацією рівня хибнопозитивних спрацювань;
- мітки чутливості - інтеграція з Microsoft Information Protection для прямого маркування та захисного шифрування файлів на рівні самого документу.

Захист даних реалізується шляхом побудови гнучких політик DLP, які охоплюють усі ключові канали можливої компрометації інформації:

- хмарні сервіси та корпоративні комунікації - безперервний моніторинг та аналіз вмісту в Exchange Online, SharePoint Online, OneDrive for Business та каналах і чатах Microsoft Teams;
- кінцеві пристрої - контроль дій користувачів на робочих станціях під управлінням операційних систем Windows та macOS (включаючи заборону копіювання файлів на зовнішні USB-накопичувачі, друку, створення скріншотів, вивантаження у неавторизовані веб-браузери чи копіювання в буфер обміну);

- хмарні застосунки сторонніх розробників - розширення дії політик безпеки на зовнішні SaaS-платформи (Salesforce, Box, Google Drive) через інтеграцію з Microsoft Defender for Cloud Apps;
- локальні сховища - сканування та класифікація даних на файлових серверах та локальних ресурсах SharePoint.

Важливою перевагою Microsoft Purview DLP є гнучка система реагування на події та підтримка концепції мінімізації впливу на бізнес-процеси. Замість суто жорсткого блокування дій система підтримує адаптивні сценарії: від відображення контекстних підказок із можливістю надання бізнес-обґрунтування дії до автоматичного шифрування та повного блокування несанкціонованого передавання контенту.

Процес впровадження систем DLP на базі Microsoft Purview рекомендується здійснювати поетапно:

1. Класифікація - визначення критичних активів, вибір відповідних SIT та створення міток чутливості.
2. Аудит та тестування - розгортання політик без застосування блокувань для оцінки обсягів потенційних витоків, аналізу хибних спрацювань та корегування правил.
3. Активна нейтралізація - перехід до реального блокування несанкціонованих дій та підключення сповіщень для адміністраторів SOC.
4. Інтеграція з Insider Risk Management - поєднання аналізу контенту з аналізом поведінкових чинників ризику (наприклад, аномальна активність працівника перед звільненням).

Ефективне функціонування політики DLP у Microsoft Purview забезпечується за рахунок її інтеграції з консоллю *Microsoft Defender XDR* та засобами *SIEM/SOAR* (зокрема Microsoft Sentinel). Це дозволяє службі безпеки отримувати деталізовані сповіщення про спроби витоку даних у режимі реального часу, аналізувати контекст і автоматизувати сценарії реагування. Такий підхід суттєво скорочує час виявлення та нейтралізації інцидентів (MTTD/MTTR) без залучення значних ручних ресурсів.

Окрім контролю за збереженими даними та їх використанням на кінцевих точках, платформа надає гнучкі інструменти для захисту даних у русі (Data in Motion). Механізми *Network DLP* та інтеграція з сервісами поштової маршрутизації дозволяють автоматично аналізувати вихідний трафік, застосовувати примусове TLS-шифрування або блокувати відправку листів і файлів стороннім реципієнтам, якщо контент містить конфіденційні маркери чи не відповідає внутрішнім регламентам безпеки компанії.

Таким чином, використання Microsoft Purview DLP дозволяє побудувати цілісну та масштабну систему захисту корпоративних даних відповідно до концепції Zero Trust. Це дозволяє організаціям мінімізувати ризики репутаційних і фінансових втрат від витоків інформації, підвищити обізнаність співробітників

щодо цифрової гігієни та забезпечити повну відповідність вимогам міжнародних і національних стандартів безпеки (GDPR, ISO/IEC 27001, PCI DSS).

Перелік посилань:

1. Про захист інформації в інформаційно-комунікаційних системах: Закон України від 05.07.1994 № 80/94-ВР.
URL: https://zakon.rada.gov.ua/laws/show/80/94-%D0%B2%D1%80 (дата звернення: 26.10.2026).
2. Microsoft Purview Data Loss Prevention overview. *Microsoft Learn*. 2024.
URL: https://learn.microsoft.com/en-us/purview/dlp-learn-about-dlp (дата звернення: 26.10.2026).

Райченко В. С.
Студент групи ІКБм-51, ННІКБЗІ ДУІКТ
Київ, Україна

ІНТЕГРОВАНИЙ ПІДХІД ДО ЗАХИСТУ КОРПОРАТИВНИХ ІНФОРМАЦІЙНИХ СИСТЕМ: МОНІТОРИНГ, УПРАВЛІННЯ ВРАЗЛИВОСТЯМИ ТА АВТОМАТИЗАЦІЯ РЕАГУВАННЯ

Анотація. Розглянуто інтегрований підхід до захисту корпоративних інформаційних систем, що поєднує три контури: профілактику (управління активами та вразливостями), виявлення (SIEM) і реагування (SOAR). Обґрунтовано необхідність спільної моделі активів, ризик-орієнтованої пріоритизації та автоматизації за принципом «людина в контурі». Запропоновано метрики оцінки ефективності: MTTD, MTTR та покриття активів моніторингом.

Корпоративні інформаційні системи (КІС) є середовищем, у якому зосереджені ключові бізнес-процеси, персональні та комерційні дані організації. Розширення хмарних сервісів, віддалена робота, інтеграція з партнерами через API та зростання кількості кінцевих точок збільшують поверхню атаки, а в умовах воєнного стану інтенсивність і складність атак на українські організації зростають. Тому завдання захисту КІС перестає зводитися до встановлення окремих засобів і потребує системного підходу [1, 2].

Типовою проблемою корпоративних ландшафтів інформаційної безпеки є фрагментація. Системи моніторингу, сканери вразливостей, засоби захисту даних і кінцевих точок працюють автономно, мають окремі моделі активів та власні інтерфейси, а аналітик центру моніторингу змушений вручну зіставляти дані з кількох консолей. Наслідками є дублювання сповіщень, втома аналітиків, затримка між виявленням вразливості та її усуненням, неповна картина інциденту й нестабільна якість реагування, що залежить від досвіду конкретного фахівця.

Пропонується розглядати захист КІС як три взаємопов’язані контури, що відповідають функціям Identify/Protect, Detect та Respond/Recover у структурі NIST Cybersecurity Framework 2.0 [1]. Перший – контур профілактики: інвентаризація активів, безперервне управління вразливостями й оновленнями. Другий – контур виявлення: централізований збір і кореляція подій у системі SIEM, моніторинг доступу до даних та стану кінцевих точок. Третій – контур реагування: оркестрація та автоматизація дій у системі SOAR за формалізованими плейбуками відповідно до процесу обробки інцидентів, описаного в NIST SP 800-61 [3].

Контур профілактики спирається на актуальний реєстр активів. Управління вразливостями має бути ризик-орієнтованим: пріоритет визначається не лише оцінкою критичності за CVSS, а й наявністю відомого експлойту, значущістю активу для бізнесу та його експозицією. Це узгоджується з рекомендаціями щодо

планування корпоративного керування оновленнями [4] і дає змогу спрямувати обмежені ресурси на усунення тих недоліків, які найімовірніше будуть використані зловмисниками.

У контурі виявлення SIEM виконує роль єдиної точки кореляції: події із серверів, мережевого обладнання, хмарних сервісів, систем автентифікації та засобів захисту даних нормалізуються й аналізуються за правилами, що відображають тактики й техніки реальних атак. Якість виявлення визначається повнотою джерел журналів, коректністю їх розбору та регулярним донастроюванням правил для зменшення кількості хибних спрацювань. Окремої уваги потребує моніторинг активності в базах даних та привілейованих облікових записів, адже саме дані є кінцевою метою більшості атак.

Контур реагування вирішує проблему часу. Для типових сценаріїв (фішинг, ураження кінцевої точки, компрометація облікового запису, витік даних) складаються плейбуки, у яких автоматизовано збагачення даних про інцидент, перевірку індикаторів компрометації, створення заявки, сповіщення відповідальних осіб та заходи локалізації, наприклад ізоляцію вузла чи блокування облікового запису. Принципово, що дії з потенційно руйнівним впливом на бізнес-процеси виконуються за принципом «людина в контурі», тобто потребують підтвердження аналітика. Такий підхід скорочує час обробки та робить реагування відтворюваним.

Інтеграція контурів забезпечується спільною моделлю активів і двосторонніми потоками даних: SIEM отримує контекст про вразливості та критичність активів для підвищення пріоритету сповіщень, а результати розслідувань повертаються до процесу управління вразливостями й до уточнення правил виявлення. Ефективність оцінюється метриками: середній час виявлення (MTTD) і реагування (MTTR), частка критичних вразливостей, усунутих у регламентований строк, покриття активів моніторингом, частка хибних спрацювань. Організаційну основу становить система управління інформаційною безпекою за ISO/IEC 27001 [2], яка закріплює ролі, процедури та цикл постійного вдосконалення.

Окремий напрям інтеграції стосується захисту самих даних. Моніторинг доступу до баз даних і файлових сховищ, контроль привілейованих користувачів та виявлення аномальної активності постачають SIEM найцінніші події: масові вибірки записів, зміни прав доступу, звернення в неробочий час. Поєднання цих подій із даними про вразливості та про поведінку облікових записів дає змогу виявляти ланцюжки атаки на ранніх етапах, коли локалізація ще не потребує зупинки бізнес-процесів.

Практичне впровадження підходу потребує поетапності. Доцільно починати з найкритичніших активів і невеликої кількості пілотних сценаріїв, щоб перевірити якість джерел подій, коректність кореляційних правил і безпечність автоматичних дій. Серед ризиків слід враховувати неповноту журналів, помилки

в плейбуках, надмірну залежність від автоматизації та перевантаження аналітиків кількістю сповіщень. Для їх зниження необхідні тестування плейбуків перед запуском у промислову експлуатацію, управління змінами, регулярне навчання персоналу та періодичний перегляд правил виявлення за результатами розслідуваних інцидентів.

Отже, інтегрований підхід, що об’єднує профілактику, виявлення та автоматизоване реагування навколо спільної моделі активів і вимірюваних метрик, дає змогу перейти від набору розрізнених засобів до керованої системи кіберзахисту КІС. Подальші дослідження пов’язані із застосуванням методів машинного навчання для пріоритизації сповіщень і оцінкою ефективності автоматизації за показниками MTTD та MTTR.

Література

1. NIST. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29. Gaithersburg, 2024.
2. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection – Information security management systems – Requirements.
3. Cichonski P., Millar T., Grance T., Scarfone K. Computer Security Incident Handling Guide. NIST SP 800-61 Rev. 2. 2012.
4. Souppaya M., Scarfone K. Guide to Enterprise Patch Management Planning: Preventive Maintenance for Technology. NIST SP 800-40 Rev. 4. 2022.

Рядий Д.О.
студент групи ІКБм-22, ННІКБЗІ ДУІКТ
Київ, Україна

КОНТЕКСТНИЙ АНАЛІЗ HTTP-ЗАПИТІВ У СИСТЕМІ ЗАХИСТУ ВЕБ-ДОДАТКІВ OPEN-APPSEC

Контекстний аналіз HTTP-запитів є одним із підходів до захисту веб-додатків, що передбачає врахування вмісту запитів та особливостей їх використання в конкретному середовищі. Застосування машинного навчання в open-appsec дає змогу поєднувати виявлення ознак атак з аналізом характерного для веб-додатка трафіку. Такий підхід спрямований на розпізнавання шкідливих запитів і зменшення кількості хибних спрацювань зі збереженням доступності легітимних функцій веб-додатка.

Ключові слова: захист веб-додатків, open-appsec, контекстний аналіз, HTTP-запити, машинне навчання, WAF.

Захист веб-додатків потребує виявлення шкідливих запитів за умов різноманітності користувацьких даних і способів їх передавання. Однакові послідовності символів можуть бути як складовою атаки, так і допустимим вмістом запиту, зокрема під час роботи з програмним кодом, пошуковими виразами або текстовими редакторами. Через це перевірка без урахування призначення окремих параметрів може призводити до хибних спрацювань і блокування легітимних операцій. Актуальним є застосування контекстного аналізу, який враховує особливості використання конкретного веб-додатка.

Контекстний аналіз HTTP-запитів передбачає оцінювання їхнього вмісту у взаємозв’язку з призначенням ресурсу, структурою параметрів і характером взаємодії користувачів із системою. Це дає змогу уточнювати значення підозрілих конструкцій залежно від того, де та за яких умов вони передаються. Одним із засобів, що реалізує такий підхід для захисту веб-додатків і прикладних програмних інтерфейсів, є open-appsec.

Для аналізу запитів open-appsec використовує машинне навчання, поєднуючи попередньо навчену модель із моделлю, що формується безпосередньо в захищуваному середовищі. Попередньо навчена модель забезпечує розпізнавання загальних ознак шкідливого вмісту, а модель середовища враховує характерний для конкретного веб-додатка трафік. Завдяки цьому оцінювання запитів спирається як на ознаки можливих атак, так і на особливості використання захищуваного ресурсу [1].

Оброблення HTTP-запиту охоплює декодування й нормалізацію його вмісту, пошук індикаторів атаки та контекстне оцінювання. На початковому етапі аналізуються адреса ресурсу, заголовки та дані, передані в тілі повідомлення. Нормалізація приводить різні представлення даних до вигляду, придатного для подальшої перевірки, зменшуючи вплив кодування на розпізнавання підозрілих конструкцій. Після виявлення індикаторів атаки система уточнює оцінку запиту

з урахуванням його джерела, цільової URL-адреси, конкретного поля введення та відомостей про захищуване середовище [2].

Практичне значення контекстного аналізу можна пояснити відмінностями у призначенні полів веб-додатка. Фрагмент HTML-коду в редакторі публікацій може бути допустимим вмістом, тоді як аналогічні дані в параметрі числового ідентифікатора не відповідають очікуваному формату. Отже, наявність окремих спеціальних символів або конструкцій потребує оцінювання разом з умовами їх використання. Водночас відповідність запиту типовій поведінці не гарантує його безпечності, тому врахування контексту має доповнювати аналіз ознак атаки.

Особливістю такого підходу є залежність якості контекстного оцінювання від повноти відомостей про легітимну роботу веб-додатка. Різноманітність користувацьких операцій, зміни функціональності та поява нових форматів даних впливають на характер трафіку. Формування моделі середовища потребує достатнього обсягу репрезентативних запитів, які відображають звичайне використання ресурсу. Тому в open-appsec передбачено режим навчання та виявлення, що дає змогу накопичувати такі відомості перед переходом до блокування відповідно до стану навчання моделі й рекомендацій системи [3].

У системі захисту веб-додатка open-appsec виконує роль додаткового рівня перевірки вхідних запитів. Його застосування має поєднуватися із серверною валідацією даних, належним контролем доступу та усуненням вразливостей програмного коду. Таке поєднання важливе, оскільки аналіз HTTP-трафіку не охоплює всіх можливих порушень логіки роботи веб-додатка та правил доступу до його ресурсів.

Отже, контекстний аналіз HTTP-запитів засобами open-appsec поєднує виявлення ознак атак з урахуванням особливостей захищуваного середовища. Використання машинного навчання створює передумови для точнішого розмежування легітимного та шкідливого трафіку й зменшення кількості хибних спрацювань. Результативність такого захисту залежить від якості сформованої моделі, коректності налаштувань і його узгодження з іншими заходами безпеки веб-додатка.

Перелік посилань:

1. open-appsec. Machine Learning Engine — Technology. URL: https://www.openappsec.io/tech.
2. open-appsec Documentation. Contextual Machine Learning. URL: https://docs.openappsec.io/concepts/contextual-machine-learning.
3. open-appsec Documentation. Track Learning and Move From Learn/Detect to Prevent. URL: https://docs.openappsec.io/how-to/configuration-and-learning/track-learning-and-move-from-learn-detect-to-prevent.

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

ЗАЛЕЖНІСТЬ ІНФОРМАТИВНОСТІ ОЗНАК МЕРЕЖЕВОГО ТРАФІКУ ВІД ТИПУ АТАК

Інформативність ознак мережевого трафіку визначає їхню придатність для розмежування легітимної та шкідливої активності. Різні типи атак характеризуються відмінностями в часових, кількісних та об'ємних параметрах мережевих потоків, тому значущість окремих характеристик може змінюватися залежно від сценарію. Урахування цих відмінностей є основою обґрунтованого відбору ознак, зменшення надлишковості даних і перевірки стійкості моделей виявлення атак.

Ключові слова: мережевий трафік, інформативність ознак, виявлення атак, відбір ознак, машинне навчання, класифікація.

Виявлення атак у мережевому трафіку ґрунтується на аналізі характеристик, що відображають особливості обміну даними між вузлами. Мережеві потоки описуються тривалістю, кількістю пакетів, їхніми розмірами, інтервалами надходження та інтенсивністю передавання. Зокрема, такі характеристики формуються засобом CICFlowMeter і представлені в наборі даних CSE-CIC-IDS2018 [1]. Вибір ознак визначає, які властивості трафіку будуть доступні моделі під час навчання та класифікації.

Інформативність ознаки характеризує її здатність допомагати розрізняти класи об'єктів у межах конкретного завдання. Для виявлення атак важливими є характеристики, значення або поєднання яких відрізняються для нормальної та шкідливої активності. Ознака може бути малокорисною окремо, проте набувати значення у взаємодії з іншими параметрами. Тому відбір потребує врахування як індивідуальних властивостей характеристик, так і залежностей між ними [2].

Відмінності між сценаріями атак зумовлюють потребу в аналізі різних груп ознак. Для інтенсивних атак на доступність доцільно досліджувати частоту надходження пакетів, кількість потоків та обсяг переданих даних. Для автоматизованих звернень — повторюваність і часові характеристики активності. Водночас статистичний опис потоку не завжди відображає зміст прикладного запиту. Вибір характеристик має спиратися на механізм атаки та доступні спостереження, оскільки саме по собі відхилення параметрів трафіку ще не доводить наявності шкідливої поведінки [3].

Застосування незмінного набору ознак передбачає, що виявлені під час навчання закономірності залишатимуться корисними надалі. Проте зміни служб, навантаження та складу мережевої активності можуть обмежувати перенесення моделі в інше середовище [3]. Характеристики, значущі для кількох сценаріїв, можуть утворювати спільну підмножину, але їхню універсальність необхідно перевіряти. Збіг відібраних ознак може відображати не лише властивості атак, а й особливості збору та розмітки даних [4].

Для оцінювання інформативності застосовуються статистичні методи, аналіз взаємної інформації та методи відбору, вбудовані в алгоритми машинного навчання. Вони дають змогу визначати характеристики, корисні для класифікації, та формувати скорочені набори ознак. Водночас індивідуальне ранжування не завжди враховує взаємодію між параметрами, тому якість відібраної підмножини потребує окремої перевірки [2].

Відбір характеристик дозволяє зменшувати розмірність даних і витрати на їх оброблення. Виключення надлишкових ознак може спростити модель без погіршення її прогностичних властивостей, однак цей результат потребує перевірки [2]. Набір, достатній для поширених атак, може виявитися недостатнім для рідкісних сценаріїв. Тому оцінювання має охоплювати повноту виявлення шкідливої активності та частоту хибних спрацювань, а не обмежуватися загальною точністю [3; 4].

Важливе значення має розмежування даних для відбору ознак і перевірки моделі. Використання тестової вибірки під час визначення значущих характеристик створює витік інформації та може призводити до завищених оцінок якості. Також необхідно враховувати репрезентативність даних і відповідність умов тестування передбачуваному застосуванню системи [4]. Перевірка на інших часових періодах або сценаріях допомагає з’ясувати, наскільки відібрані характеристики зберігають свою корисність.

Залежність інформативності від умов спостереження створює передумови для адаптивного відбору ознак. Перспективним напрямом є поєднання характеристик, корисність яких підтверджується в різних сценаріях, із додатковими показниками поточної активності. Такий підхід потребує узгодженого оновлення набору ознак і моделі, а також перевірки того, чи покращує перегляд характеристик виявлення атак.

Отже, інформативність ознак мережевого трафіку залежить від характеру атак, властивостей середовища та складу навчальних даних. Урахування цих чинників дає змогу обґрунтовано формувати набори характеристик і зменшувати їхню надлишковість. Перевірка значущості ознак у різних сценаріях є основою оцінювання здатності моделі зберігати якість виявлення за зміни мережевої активності.

Перелік посилань:

1. Canadian Institute for Cybersecurity. CSE-CIC-IDS2018 on AWS. University of New Brunswick. URL: https://www.unb.ca/cic/datasets/ids-2018.html.

2. Guyon I., Elisseeff A. An Introduction to Variable and Feature Selection. Journal of Machine Learning Research. 2003. Vol. 3. P. 1157–1182. URL: https://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf.

3. Sommer R., Paxson V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. 2010 IEEE Symposium on Security and Privacy. 2010. P. 305–316. DOI: 10.1109/SP.2010.25.

4. Arp D., Quiring E., Pendlebury F., Warnecke A., Pierazzi F., Wressnegger C., Cavallaro L., Rieck K. Dos and Don’ts of Machine Learning in Computer Security. 31st USENIX Security Symposium. 2022. P. 3971–3988. URL: https://www.usenix.org/conference/usenixsecurity22/presentation/arp.

Medynskyi O. R.
PhD Student, Group Kba-21, ICTA, NULP
Susukailo V. A.
PhD, Senior Lecturer, ICTA, NULP

A DIGITAL IMAGE VERIFICATION METHODOLOGY FOR CYBER INCIDENT INVESTIGATION IN THE ERA OF GENERATIVE ARTIFICIAL INTELLIGENCE

Abstract. The paper addresses the problem of establishing the authenticity of digital images that appear as artifacts in cyber incident investigations, given the widespread availability of generative models (GANs, VAEs, diffusion models). A multi-level image verification methodology integrated into the incident response lifecycle is proposed. It combines metadata analysis, frequency analysis in the DCT domain, neural network detectors and explainable artificial intelligence (XAI) methods. A decision rule with an explicit uncertainty zone is defined, together with requirements for preserving the chain of custody and ensuring the evidentiary admissibility of the results.

Keywords: *cyber incident investigation, digital forensics, synthetic images, diffusion models, DCT analysis, XAI.*

Cyber incident investigation has traditionally relied on event logs, network traffic, and memory and disk images. At the same time, a growing share of incidents – phishing campaigns, business email compromise (BEC), social engineering fraud, bypassing of know-your-customer (KYC) procedures, and information operations – is built around multimedia content. The spread of diffusion models has made photorealistic image generation available to anyone without special skills. As a result, the analyst faces a new task: before an image can be used as evidence or as an indicator of compromise, its origin and authenticity must be established. An error at this stage leads either to false attribution of the incident or to the evidence being ruled inadmissible.

Current incident response recommendations [1] and the international standard for handling digital evidence [2] regulate the identification, collection, acquisition and preservation of artifacts in detail, but they do not include procedures for verifying whether content is synthetic. Individual detectors, in turn, generalize poorly: models trained on images produced by generative adversarial networks lose much of their accuracy on images produced by diffusion models [3]. Moreover, repeated JPEG compression, rescaling and screenshots, which are typical of messengers and corporate e-mail, destroy the most informative features. Consequently, the output of a single detector cannot be a sufficient basis for a conclusion in an incident report.

BEC incidents are a typical example: to support a pretext, the attacker sends the finance department generated images of identity documents, screenshots of payment orders or “photographs” of signed contracts. In such cases the image itself defines the attack vector, and the conclusion about its origin determines the incident classification, the damage assessment and any subsequent legal action.

The aim of this work is to develop an image verification methodology that is embedded in the cyber incident investigation process and ensures the reproducibility

and evidentiary admissibility of the results.

The proposed methodology consists of four levels of analysis. Level 1 is acquisition and preservation: the original file is obtained from the primary source (mailbox, cloud storage, endpoint), its SHA-256 hash is computed before any processing, and a chain of custody log is maintained [2]; all further actions are performed only on a working copy. Level 2 is metadata and context analysis: examination of EXIF/XMP data, checking for a C2PA manifest (Content Credentials), JPEG quantization tables, consistency between the claimed capture device and the file parameters, and consistency of timestamps with the incident timeline. The absence of metadata does not by itself indicate synthesis, whereas markers left by generative services are a strong indicator. Level 3 is frequency analysis: computation of the two-dimensional discrete cosine transform (DCT) and the averaged image spectrum. Upsampling operations in generators leave periodic artifacts that appear as anomalous peaks in the high-frequency region of the spectrum [4]. Level 4 is an ensemble of neural network detectors (CNN, ViT) trained on different generator families, with decisions explained by XAI methods (Grad-CAM, SHAP), which makes it possible to localize the image regions that influenced the result.

Robustness to anti-forensic techniques is an important aspect. An attacker may deliberately strip or forge metadata, recompress the image or add noise to mask frequency artifacts, and apply adversarial perturbations against a specific detector. For this reason, the methodology does not rely on any single feature: an inconsistency between levels (for example, “clean” camera metadata combined with pronounced spectral anomalies) is itself treated as an indicator of manipulation. The regulatory context is also taken into account: Article 50 of the EU Artificial Intelligence Act requires providers of generative systems to ensure machine-readable marking of synthetic content. This will make Level 2 more informative over time, but it does not remove the need for independent methods, since such marking can be removed. Images containing personal data must be processed in line with the GDPR principles of data minimization and purpose limitation.

The results of Levels 2–4 are aggregated into an integral score:

$$S = \sum_{i=1}^{n} w_i \cdot p_i, \quad (1)$$

where $p_i \in [0; 1]$ is the probability of synthetic origin estimated by the i-th method; w_i is the weight of the method ($\Sigma w_i = 1$), determined on a validation set with regard to the method’s robustness to the transformations detected in the file; n is the number of methods applied. The decision is made using three zones: $S \geq \tau_h$ – “signs of synthetic origin detected”; $S \leq \tau_l$ – “no signs of synthesis detected”; $\tau_l < S < \tau_h$ – “inconclusive”, which requires expert involvement. The uncertainty zone is essential, as it prevents categorical conclusions where the data do not support them.

The methodology is integrated into the incident response lifecycle [1] as follows. During the detection and analysis phase, a SOC analyst performs initial triage (Levels 1–2), and a full analysis is initiated if there is any suspicion. The results are documented

in a separate section of the incident report, specifying the methods, model versions, S values and explanation maps. Confirmed synthetic images are converted into indicators of compromise (cryptographic and perceptual hashes, distribution domains and accounts), which are used in SIEM rules and shared as threat intelligence through platforms such as MISP. During the post-incident activity phase, the weights and thresholds are refined using the new samples.

The methodology will be evaluated on a test set that combines authentic photographs with images produced by GAN- and diffusion-based generators, each subjected to the transformations typical of real distribution channels (messenger recompression, rescaling, screenshots). Since the cost of errors in an investigation is asymmetric, the evaluation focuses not only on overall accuracy but also on the false positive rate at a fixed true positive rate, the share of cases assigned to the uncertainty zone, and the stability of the explanation maps under these transformations.

Thus, the proposed methodology closes the gap between standardized incident response procedures and the task of establishing the authenticity of multimedia artifacts. Its multi-level structure reduces dependence on any single detector, while the use of XAI and a chain of custody log ensures the evidentiary admissibility of the results. Further research will focus on determining the weights on datasets that reproduce real-world image distribution channels and on evaluating the robustness of the methodology against adversarial attacks.

References

1. Nelson A., Rekhi S., Souppaya M., Scarfone K. Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile : NIST Special Publication 800-61 Rev. 3. Gaithersburg : National Institute of Standards and Technology, 2025. DOI: 10.6028/NIST.SP.800-61r3.

2. ISO/IEC 27037:2012. Information technology – Security techniques – Guidelines for identification, collection, acquisition and preservation of digital evidence. Geneva : ISO, 2012.

3. Corvi R., Cozzolino D., Zingarini G., Poggi G., Nagano K., Verdoliva L. On the detection of synthetic images generated by diffusion models. ICASSP 2023 – IEEE International Conference on Acoustics, Speech and Signal Processing. Rhodes Island, 2023. P. 1–5.

4. Frank J., Eisenhofer T., Schönherr L., Fischer A., Kolossa D., Holz T. Leveraging frequency analysis for deep fake image recognition. Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020. Vol. 119. P. 3247–3258.

Сушанов Павло Вікторович
студент групи ІКБм-22, ННІКБЗІ ДУІКТ, Київ, Україна

ТЕХНОЛОГІЯ ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ ВІД ФІШИНГОВИХ АТАК

Анотація. У роботі розглянуто комплексний підхід до протидії фішингу в інформаційних системах. Проаналізовано способи, за допомогою яких зловмисники намагаються отримати паролі, заволодіти обліковими записами або змусити працівників відкрити небезпечні вкладення. Описано роль поштових шлюзів, перевірки доменів і посилань, механізмів SPF, DKIM та DMARC, багатофакторної автентифікації й навчання персоналу. Запропоновано послідовність виявлення, оцінювання та обробки підозрілих повідомлень, а також показники для перевірки результативності захисту. Практична цінність підходу полягає у зниженні ймовірності компрометації облікових записів і зменшенні наслідків фішингових інцидентів.

Ключові слова: фішинг, кібербезпека, інформаційні системи, електронна пошта, SPF, DKIM, DMARC, багатофакторна автентифікація.

Актуальність дослідження. Фішинг є одним із поширених способів проникнення до корпоративного цифрового середовища. Для атаки можуть використовуватися листи, що зовні нагадують повідомлення від керівника, банку, служби підтримки чи партнера організації. Одержувача намагаються переконати перейти на копію справжнього вебсайту, відкрити файл або надати конфіденційну інформацію. Якщо користувач виконує таку дію, зловмисник може отримати доступ до пошти, робочих сервісів і внутрішніх даних.

Визначити фішингове повідомлення лише за однією ознакою не завжди можливо: легітимні листи також можуть містити термінові прохання чи вкладення, а шахрайські повідомлення часто добре імітують справжні. Тому необхідно перевіряти відправника, домен, фактичну адресу посилання, вкладення та відповідність повідомлення звичному робочому процесу.

Мета роботи — сформувати послідовний підхід до захисту інформаційних систем, який об’єднує технічні інструменти, правила організації та підготовку користувачів.

Рівні захисту та обробка інформації. Основою технології є багаторівнева перевірка електронної пошти. Поштовий шлюз може оцінювати репутацію відправника, результати доменної автентифікації, текст повідомлення, посилання та вкладені файли. Додаткові відомості надходять із журналів поштової системи, засобів захисту робочих станцій, браузерів і повідомлень самих працівників. Зіставлення цих даних допомагає з’ясувати, кому надійшов лист і чи були після його отримання підозрілі дії.

Для зменшення ризику підміни домену варто налаштувати SPF, DKIM і DMARC. SPF перевіряє, чи має сервер право надсилати листи від імені певного

домену. DKIM використовує цифровий підпис для перевірки цілісності повідомлення та його зв’язку з доменом підписанта. DMARC визначає, як обробляти повідомлення, що не пройшли потрібні перевірки, і дає змогу отримувати звіти. Водночас ці технології не гарантують виявлення всіх атак: шахрай може використати схожий домен або зловмисно скористатися справжнім скомпрометованим обліковим записом.

Для подальшого аналізу доцільно зберігати час надходження листа, адресу й домен відправника, результати перевірок SPF/DKIM/DMARC, виявлені URL-адреси, відомості про вкладення та список одержувачів.

Дані слід захищати від несанкціонованого доступу, а порядок і строк їх зберігання визначити внутрішніми правилами. Наявність журналів полегшує пошук інших копій шкідливого повідомлення та аналіз можливого масштабу інциденту.

Виявлення фішингу та реагування на інциденти. На початковому етапі повідомлення перевіряється за низкою індикаторів: домен відправника не відповідає заявленій організації, посилання веде на інший сайт, у листі міститься нетипова вимога терміново переказати кошти або передати пароль, чи додано файл, який не очікувався одержувачем. За потреби URL-адреси перевіряються за репутаційними джерелами або в ізольованому середовищі. Окрема підозріла ознака є приводом для додаткового аналізу, але не завжди достатньою підставою для остаточного висновку.

Після перевірки повідомленню присвоюється рівень ризику. Лист із низьким ризиком може бути доставлений, повідомлення із сумнівними ознаками — позначене попередженням або передане на ручний розгляд, а лист із високим ризиком — заблоковане чи поміщене в карантин. Правила слід перевіряти на прикладах легітимної та шкідливої пошти, щоб не створювати надмірної кількості хибних спрацювань. Працівникам потрібно надати простий спосіб пересилати підозрілі повідомлення команді безпеки.

Якщо з’ясовано, що користувач відкрив небезпечне посилання або ввів пароль на підробленій сторінці, слід оперативно оцінити масштаб події. Залежно від обставин потрібно завершити активні сеанси, змінити пароль, перевірити журнали входів, перевірити пристрій засобами захисту й обмежити його мережеву взаємодію, якщо є ознаки зараження. Для критичних сервісів варто застосовувати багатофакторну автентифікацію, бажано стійку до фішингу, наприклад passkeys або апаратні ключі безпеки. Дії персоналу мають відповідати затвердженому плану реагування.

Організаційні заходи та обізнаність персоналу. Ефективність технологічного захисту значною мірою залежить від поведінки користувачів. Працівників потрібно навчати перевіряти адресу відправника й домен сайту, не відкривати несподівані вкладення та не передавати паролі чи одноразові коди через посилання з листів. Якщо повідомлення містить фінансове доручення або

незвичний запит від керівника, його варто підтвердити іншим відомим каналом зв’язку. Навчальні симуляції можуть допомогти закріпити навички, якщо після них надаються зрозумілі пояснення.

Організації доцільно встановити порядок повідомлення про фішинг, визначити відповідальних за аналіз листів і встановити строки первинного реагування. Додаткове зниження ризику забезпечують мінімально необхідні права доступу, окремі адміністративні облікові записи, регулярне оновлення програмного забезпечення та перевірені резервні копії. Такі заходи не замінюють фільтрацію пошти, але обмежують можливі наслідки, якщо небезпечне повідомлення все ж дійде до користувача.

Методика оцінювання ефективності. Перевіряти технологію варто на контрольованому наборі тестових листів, який містить як фішингові приклади, так і звичайну кореспонденцію. У вибірку можна включити повідомлення з підробленими доменами, небезпечними посиланнями, вкладеннями та цільовими запитами, а також легітимні листи зі схожими ознаками. Налаштування виконують на одній частині набору, а підсумкову оцінку — на окремій, щоб уникнути завищення результатів.

Для оцінювання застосовують точність виявлення (Precision), повноту (Recall), F1, частку хибних спрацювань, кількість легітимних листів, помилково поміщених у карантин, і час від отримання повідомлення до його блокування. Також можна відстежувати, як часто користувачі повідомляють про підозрілі листи та як змінюються результати навчальних симуляцій. Показники слід розглядати разом: висока повнота за надмірної кількості хибних спрацювань може ускладнювати роботу, а успіх на тестовій вибірці не гарантує аналогічного результату в реальних умовах.

Висновки. У роботі обґрунтовано необхідність багаторівневого захисту інформаційних систем від фішингу. Запропонований підхід поєднує аналіз електронної пошти, перевірку доменів і посилань, застосування SPF, DKIM та DMARC, багатофакторну автентифікацію, моніторинг подій і навчання персоналу. Для результативної роботи системи важливо не лише виявити підозрілий лист, а й забезпечити зрозуміле реагування та можливість перевірити причини ухваленого рішення.

Список використаних джерел

1. Cybersecurity and Infrastructure Security Agency (CISA). Recognize and Report Phishing. URL: https://www.cisa.gov/secure-our-world/recognize-and-report-phishing (дата звернення: 28.10.2026).
2. National Institute of Standards and Technology (NIST). Digital Identity Guidelines: Authentication and Authenticator Management. SP 800-63B-4. URL: https://csrc.nist.gov/pubs/sp/800/63/B/4/final (дата звернення: 28.10.2026).
3. Kitterman S. Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1. RFC 7208. IETF, 2014. URL: https://www.rfc-editor.org/info/rfc7208/ (дата звернення: 28.10.2026).
4. Herr T., Levine J. Domain-Based Message Authentication, Reporting, and Conformance (DMARC). RFC 9989. IETF, 2026. URL: https://www.rfc-editor.org/info/rfc9989/ (дата звернення: 28.10.2026).

Тишик Іван Ярославович

ВИЯВЛЕННЯ КІБЕРЗАГРОЗ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНОЇ КОРЕЛЯЦІЇ ПОДІЙ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Анотація. Розглянуто підхід до виявлення кіберзагроз на основі інтегрованого аналізу подій інформаційної безпеки із застосуванням методів штучного інтелекту. Запропоновано здійснювати централізований збір і нормалізацію подій із різнорідних джерел, виявлення аномальної активності та кореляцію взаємопов'язаних подій. Для встановлення відповідності виявленої активності характерним сценаріям кібератак запропоновано використовувати базу знань MITRE ATT&CK. Розглянуто можливість реалізації підходу засобами SIEM, UEBA та SOC для підвищення ефективності виявлення складних багатоступеневих кіберзагроз.

Сучасні інформаційно-комунікаційні системи характеризуються значною кількістю взаємопов'язаних компонентів, розподіленістю інформаційних ресурсів та постійним зростанням інтенсивності мережевої взаємодії. За таких умов традиційні засоби виявлення кіберзагроз, орієнтовані переважно на аналіз окремих подій або використання заздалегідь визначених сигнатур, не завжди забезпечують своєчасну ідентифікацію складних багатоступеневих атак. Особливо актуальною ця проблема є для атак, під час яких зловмисники використовують легітимні облікові записи, штатні засоби операційної системи та здійснюють горизонтальне переміщення між вузлами корпоративної мережі.

Перспективним напрямом підвищення ефективності моніторингу кібербезпеки є інтегрований аналіз подій інформаційної безпеки із застосуванням методів штучного інтелекту. Методи машинного навчання дають змогу здійснювати аналіз значних масивів подій та виявляти закономірності й аномалії, які складно встановити традиційними сигнатурними методами [1, с. 1158–1162]. На відміну від аналізу ізольованих спрацювань окремих засобів захисту, такий підхід передбачає формування комплексного уявлення про розвиток потенційної атаки на основі кореляції подій, отриманих із різнорідних джерел.

Метою роботи є підвищення ефективності виявлення кіберзагроз шляхом інтегрованого аналізу подій інформаційної безпеки, визначення аномальної активності засобами штучного інтелекту та її кореляції з відомими техніками реалізації кібератак

Наукова новизна запропонованого підходу полягає у поєднанні поведінкового аналізу аномальної активності з контекстною кореляцією різнорідних подій інформаційної безпеки та їх зіставленням із закономірностями реалізації кібератак, представленими у MITRE ATT&CK. На відміну від підходів, які ґрунтуються на аналізі окремих спрацювань або наперед визначених сигнатур, передбачається аналіз взаємопов'язаної послідовності подій з урахуванням їх часових, поведінкових та логічних зв'язків. Це створює можливість виявлення ознак розвитку багатоступеневої атаки навіть у випадках, коли окремі події не перевищують встановлених порогів спрацювання засобів захисту.

Джерелами даних для такого аналізу можуть бути журнали операційних систем і серверів, мережеве обладнання, системи автентифікації, контролери домену Active Directory, міжмережеві екрани, IDS/IPS, засоби захисту кінцевих точок та SIEM-системи. Централізований збір, попередня обробка і нормалізація цих даних створюють основу для встановлення часових, логічних та поведінкових взаємозв'язків між подіями. Централізоване керування журналами подій є важливою складовою процесів моніторингу та аналізу стану інформаційної безпеки [3, с. 2-1–2-4].

Запропонований підхід передбачає послідовне виконання таких етапів: збір подій інформаційної безпеки з різних джерел, їх попередню обробку та нормалізацію, виявлення аномальної активності із застосуванням методів штучного інтелекту, кореляцію взаємопов'язаних подій та формування сповіщення про потенційну кіберзагрозу. Спрощену структуру запропонованого процесу наведено на рис. 1.

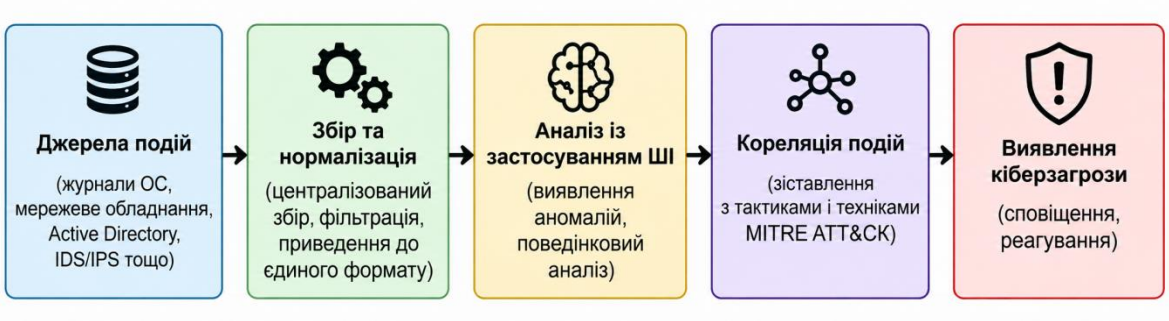


Рис. 1. Спрощена схема процесу виявлення кіберзагроз на основі інтелектуальної кореляції подій інформаційної безпеки

Окрема подія інформаційної безпеки не завжди містить достатньо інформації для однозначної ідентифікації атаки. Наприклад, невдала спроба автентифікації, запуск PowerShell або звернення до мережевого ресурсу можуть бути легітимними діями. Водночас їх поєднання з успішною автентифікацією з нетипового вузла, використанням адміністративних привілеїв, спробами отримання облікових даних та подальшим горизонтальним переміщенням може свідчити про розвиток кібератаки. Тому важливим завданням є не лише виявлення окремих аномалій, а й встановлення взаємозв'язків між ними.

Для визначення таких залежностей доцільно застосовувати методи машинного навчання та поведінкового аналізу. Виявлення аномалій передбачає встановлення шаблонів поведінки, які не відповідають очікуваній або типовій активності системи [4, с. 2–4]. Це дає змогу формувати профілі типової активності користувачів, вузлів і процесів та виявляти відхилення від сформованого базового рівня. Ознаками аномальної поведінки можуть бути нехарактерний час автентифікації, значна кількості невдалих входів, використання підвищених привілеїв, запуск нетипових процесів, зміна системних параметрів та незвична мережева активність.

Важливим елементом запропонованого підходу є кореляція виявленої активності з відомими закономірностями реалізації кібератак. Для цього доцільно використовувати базу знань MITRE ATT&CK, у якій систематизовано тактики та техніки, які застосовуються зловмисниками [2]. Виявлена послідовність подій може бути зіставлена з такими етапами атаки, як початковий доступ (Initial Access), виконання (Execution), закріплення (Persistence), отримання облікових даних (Credential Access), підвищення привілеїв (Privilege Escalation) та горизонтальне переміщення (Lateral Movement).

Особливого значення такий підхід набуває для моніторингу середовища Active Directory, де компрометація одного облікового запису може стати початковим етапом подальшого розвитку атаки. Кореляція подій автентифікації, змін привілеїв, запуску системних утиліт, звернень до контролерів домену та аномальної мережевої активності дає можливість виявляти ознаки компрометації, які складно встановити шляхом аналізу окремого журналу подій.

Практична реалізація запропонованого підходу може здійснюватися в межах центру моніторингу безпеки (SOC) із використанням SIEM як базової платформи централізованого збору та первинної кореляції подій. Поєднання SIEM із технологіями UEBA та засобами машинного навчання дає змогу доповнити традиційні правила кореляції механізмами поведінкового аналізу та автоматизованого виявлення аномальної активності.

Важливим аспектом практичного застосування запропонованого підходу є можливість його адаптації до особливостей конкретної інформаційної системи при зміні характеру кіберзагроз. Накопичення даних про події безпеки, результати їх аналізу та підтверджені інциденти створює основу для подальшого уточнення моделей виявлення аномальної активності.

Таким чином, інтегрований аналіз подій інформаційної безпеки із застосуванням штучного інтелекту є перспективним підходом до виявлення сучасних кіберзагроз. Поєднання даних із різнорідних джерел, виявлення аномальної поведінки та кореляція отриманих результатів із відомими техніками атак дають змогу підвищити точність ідентифікації потенційних інцидентів. Подальші дослідження доцільно спрямувати на вдосконалення механізмів автоматизованої кореляції аномальної активності з послідовностями технік MITRE ATT&CK та експериментальну перевірку запропонованого підходу в корпоративних мережах.

Список використаних джерел

1. Buczak A. L., Guven E. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys & Tutorials. 2016. Vol. 18, No. 2. P. 1153–1176.
2. MITRE. *MITRE ATT&CK: Adversarial Tactics, Techniques, and Common Knowledge*. The MITRE Corporation.
3. Scarfone K., Souppaya M., Cody A., Orebaugh A. *Guide to Computer Security Log Management*. NIST Special Publication 800-92. Gaithersburg: National Institute of Standards and Technology, 2006.
4. Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey. *ACM Computing Surveys*. 2009. Vol. 41, No. 3. Article 15.

УДК 004.056:316.77

Верголяс О. О., к.ю.н.
Харківський національний університет Повітряних Сил імені Івана Кожедуба
ORCID: 0000-0002-9780-1298

КОГНІТИВНА БЕЗПЕКА КРИТИЧНИХ АУДИТОРІЙ В УМОВАХ ІНФОРМАЦІЙНО-ПСИХОЛОГІЧНИХ ОПЕРАЦІЙ, ПІДСИЛЕНИХ ГЕНЕРАТИВНИМ ШТУЧНИМ ІНТЕЛЕКТОМ

Анотація. Розглянуто захист критичних аудиторій від інформаційно-психологічних операцій, підсилених генеративним штучним інтелектом. Показано, що поєднання масштабної генерації синтетичного контенту з адресним мікротаргетингом зміщує вплив на визначені групи, тоді як наявні засоби зосереджено на виявленні артефактів, а не на стійкості самої аудиторії. Запропоновано контур когнітивної безпеки аудиторії, у якому технічне виявлення синтетики є вхідним сигналом, а захист будується на інокуляційному та організаційному рівнях. Визначено критерій критичності аудиторії.

Генеративний штучний інтелект здешевив і пришвидшив виробництво переконливого синтетичного контенту: відеофейків, синтетичних персон, масового персуазивного тексту. За третьою доповіддю Європейської служби зовнішніх справ про загрози іноземного інформаційного маніпулювання та втручання (FIMI) у 2024 році застосування ШІ в таких кампаніях помітно зросло, а найбільшим за обсягом виявився не гучний deepfake, а масова генерація «наповнювального» тексту, що імітує суспільний консенсус і обходить ручну модерацію [1]. Коли така генерація поєднується з адресним психографічним мікротаргетингом, вплив зміщується від масового мовлення до прицільного тиску на визначені аудиторії. В Україні їх коло стале: військовослужбовці, їхні родини, населення прифронтових громад, оператори критичної інфраструктури, чиї рішення мають безпосередню безпекову ціну; когнітивний вплив на цей сегмент має вимірювану воєнно-економічну ціну [2]. Об'єкт захисту тому зміщується з периметра інформаційних систем на когнітивну стійкість самої аудиторії.

Наявні технічні засоби зосереджено на виявленні артефактів синтетики: для відеофейків запропоновано детектори на основі трансферного навчання (модель Xception) з високими значеннями ROC-AUC на еталонному корпусі FaceForensics++ [3]. Такі детектори відповідають на питання, чи є контент підробленим, але не на питання, як захистити аудиторію, на яку його розраховано; зв'язок між технічним виявленням і когнітивною стійкістю груп-мішеней лишається слабко опрацьованим [2].

Пропонується розглядати когнітивну безпеку критичних аудиторій як окремий контур захисту. Він описує вектор впливу за стадіями каналу (виробництво синтетичного матеріалу, адресна доставка, засвоєння аудиторією)

і вводить критерій критичності аудиторії: пріоритетність групи визначає не її чисельність, а функція її рішень і рівень доступу до безпеково значущих процесів у поєднанні з вразливістю до конкретного наративу [2]. Технічне виявлення синтетики [3] тут є входом, а не самоціллю: його результати живлять два подальші рівні. Інокуляційний рівень передбачає, що попереднє ознайомлення аудиторії з прийомами маніпуляції (префбанкінг) підвищує опірність до дезінформації, що підтверджено польовим експериментом із короткими відеороликами на платформі YouTube [4]. Організаційний рівень охоплює адресне попередження критичних груп і системні заходи з підвищення когнітивної стійкості.

Запропонований контур переносить акцент із виявлення підробленого контенту на стійкість тих, кому він адресований, і прив'язує технічний вхідний сигнал до інокуляційних та організаційних заходів. Він доповнює технічні детектори, а не замінює їх. Подальші роботи доцільно спрямувати на емпіричне вимірювання ефекту інокуляції для конкретних критичних груп [4].

Список літератури

1. Third EEAS Report on Foreign Information Manipulation and Interference Threats. European External Action Service. March 2025. URL: https://www.eeas.europa.eu/eeas/3rd-eeas-report-foreign-information-manipulation-and-interference-threats-0_en (дата звернення: 30.06.2026).
2. Воєнно-економічні наслідки когнітивного впливу на систему воєнної та національної безпеки України / О. Семененко та ін. Social Development and Security. 2025. Т. 15, № 3. С. 308–321. DOI: 10.33445/sds.2025.15.3.27.
3. Deepfake Detection Using Transfer Learning-Based Xception Model / V. Rajakumareswaran та ін. Advanced Information Systems. 2024. Vol. 8, No. 2. P. 89–98. DOI: 10.20998/2522-9052.2024.2.10.
4. Psychological inoculation improves resilience against misinformation on social media / J. Roozenbeek та ін. Science Advances. 2022. Vol. 8, No. 34. Article eabo6254. DOI: 10.1126/sciadv.abo6254.

Svedeniuk O.V.
PhD Student, Group Kba-21, ICTA, NULP
Kurii Y. O.
PhD, Senior Lecturer, ICTA, NULP

APPLICATION OF AI MODELS FOR AUTOMATING IOT INFRASTRUCTURE COMPLIANCE WITH ETSI EN 303 645 AND IEC 62443 STANDARDS

Abstract. This paper investigates methods for automated assessment and formalization of baseline security requirements for consumer and industrial (IIoT) devices based on ETSI EN 303 645 and IEC 62443 standards. We propose the application of artificial intelligence models for continuous monitoring of IoT infrastructure configurations and verifying their compliance with regulatory requirements. A hybrid methodology combining deterministic software control tools with intelligent NLP and ML algorithms is considered. This approach significantly reduces auditing time and increases the overall security level of heterogeneous networks.

Keywords: IoT, IIoT, artificial intelligence, ETSI EN 303 645, IEC 62443, cybersecurity, compliance, audit automation.

The rapid expansion of Internet of Things (IoT) and Industrial Internet of Things (IIoT) technologies across critical infrastructure is accompanied by a proportional rise in cybersecurity threats. The convergence of Information Technology (IT) and Operational Technology (OT) networks, combined with extreme device heterogeneity, resource-constrained hardware microcontrollers, and legacy industrial communication protocols such as Modbus TCP, MQTT, and OPC UA frequently lacking native encryption and authentication mechanisms, creates severe systemic vulnerabilities. To mitigate these operational risks, international regulatory standards such as ETSI EN 303 645 [1] for consumer IoT baseline security and the IEC 62443 [2] series for industrial automation and control systems have been developed.

However, implementing and verifying these security standards in complex operational technology environments encounters substantial technical difficulties. Traditional compliance assessment relies heavily on manual configuration auditing, static credential checks, and periodic vulnerability scanning across hundreds of heterogeneous network nodes. This process is time-consuming, resource-intensive, highly expensive, and inherently prone to human error [3]. Furthermore, modern IIoT environments are highly dynamic: continuous onboarding of unmanaged edge devices, transient network topologies, dynamic IP assignments, and frequent firmware updates necessitate real-time auditing, which cannot be achieved using conventional, periodic manual methods.

To overcome these structural limitations and eliminate audit visibility gaps, a hybrid methodology is proposed, integrating deterministic software tools (vulnerability scanners, configuration management frameworks) with adaptive artificial intelligence (AI) models. This dual-engine architecture establishes a unified compliance framework where deterministic tools evaluate known static configuration baselines while AI

models continuously monitor non-linear behavioral telemetry. The automated compliance verification pipeline comprises two interdependent core stages: intelligent requirements formalization and continuous automated infrastructure state assessment.

At the requirement formalization stage, domain-adapted Natural Language Processing (NLP) models, including fine-tuned Transformer architectures, parse unstructured security policies, technical specifications, and regulatory directives. By extracting semantic entities, operational constraints, and access control rules, these NLP models construct formal ontologies and automatically map extracted requirements onto specific device profiles. They transform complex natural language text into machine-readable logic, structured Open Policy Agent (OPA) specifications, and executable verification scripts. This process creates a dynamic, continuously updated benchmark knowledge base without relying on manual rule engineering.

At the continuous assessment stage, a low-latency stream processing pipeline ingests heterogeneous telemetry data, including PCAP network traffic dumps, Syslog events, eBPF kernel tracepoints, and JSON-formatted configuration states from edge IoT/IIoT nodes. High-dimensional feature vectors extracted from this data (such as packet inter-arrival times, payload entropy, protocol header fields, and system API call sequences) are processed through supervised classification algorithms (e.g., XGBoost) and unsupervised deep anomaly detection models (e.g., Convolutional Autoencoders or Graph Neural Networks). By evaluating real-time operational behavior against formal rule-sets derived from IEC 62443-4-2 (specifically system integrity and zone-conduit segmentation controls), the system dynamically detects policy drift, privilege escalation, and unauthorized inter-zone interactions, instantly flagging them as compliance violations [4, p. 114070].

The proposed hybrid model achieves operational synergy: deterministic software tools execute high-speed baseline validation, while AI models perform deep behavioral analysis to identify complex, multi-vector compliance violations. Furthermore, integrating automated risk scoring, alert prioritization, and closed-loop policy enforcement—such as dynamic network isolation—significantly reduces Mean Time to Detect (MTTD) compliance drift within Zero Trust security architectures. Consequently, transitioning from periodic manual audits to Continuous Compliance substantially mitigates human risk, minimizes operational overhead, and enhances critical infrastructure resilience against modern cyber threats.

References:

1. ETSI EN 303 645 V2.1.1 (2020-06). CYBER; Cyber Security for Consumer Internet of Things: Baseline Requirements.

2. IEC 62443-4-2:2019. Security for industrial automation and control systems – Part 4-2: Technical security requirements for IACS components. IEC, 2019.

3. NIST IR 8259A. IoT Device Cybersecurity Capability Core Baseline. National Institute of Standards and Technology, 2020.

4. Bagaa M., Taleb T., Bernabe J. B., Skarmeta A. A Machine Learning Security Framework for IoT Systems // IEEE Access. 2020. Vol. 8. P. 114066–114077.

Ковалівська К. В.
студентка групи 1-8м, ФЕМП, ДТЕУ
Київ, Україна
Науковий керівник: Дивак В. В.
канд. пед. наук, доцент кафедри КНтІС, ДТЕУ
Київ, Україна

ДОСЛІДЖЕННЯ РІВНЯ КІБЕРБЕЗПЕКИ УКРАЇНИ В УМОВАХ ВОЄННОГО СТАНУ ТА ПОСИЛЕННЯ АГРЕСІЇ РФ

У дослідженні проаналізовано сучасний стан кібербезпеки України в умовах воєнного стану та посилення агресії рф. Розглянуто позиції України у Глобальному індексі кібербезпеки та динаміку його основних складових. Досліджено структуру кібератак за ключовими секторами у 2025 р. Визначено найбільш уразливі напрями цифрової інфраструктури. Обґрунтовано необхідність подальшого розвитку технічних засобів захисту, кадрового потенціалу, систем реагування на кіберінциденти та зміцнення кіберстійкості держави.

Ключові слова: кібербезпека, кіберзагрози, кібератаки, цифрова безпека, воєнний стан, критична інфраструктура, інформаційний захист

Стрімкий розвиток цифрових технологій, зокрема електронної комерції, блокчейну, Інтернету речей, комп’ютерного інжинірингу та бездротового зв’язку, суттєво трансформує економічну діяльність. Водночас ускладнення процесів у цифровому середовищі супроводжується зростанням кіберзагроз, що особливо актуалізує питання кібербезпеки України в умовах воєнного стану [1].Кібербезпека є важливою складовою захисту цифрового простору держави та передумовою його стабільного розвитку. У зв’язку з цим на рис. 1 представлено динаміку індексу кібербезпеки України.

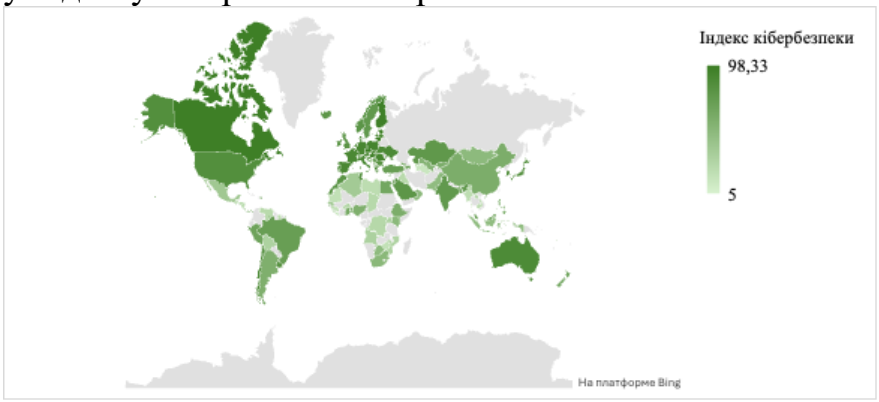


Рис. 2. Індекс кібербезпеки Україні в рейтингу Глобального індексу кібербезпеки в 2024 р.

Джерело: [2]

Як свідчать дані рис. 2, у 2024 р. Україна посіла 59-те місце у світовому рейтингу Глобального індексу кібербезпеки, що характеризує сформований рівень спроможності держави забезпечувати захист цифрового середовища. Такий результат набуває особливого значення в умовах воєнного стану, коли Україна систематично стикається з кіберзагрозами та атаками, пов'язаними зі збройною агресією рф.

Для поглиблення оцінки впливу воєнних умов на кібербезпеку України доцільно проаналізувати динаміку ключових складових відповідного індексу, представлену на рис. 2.

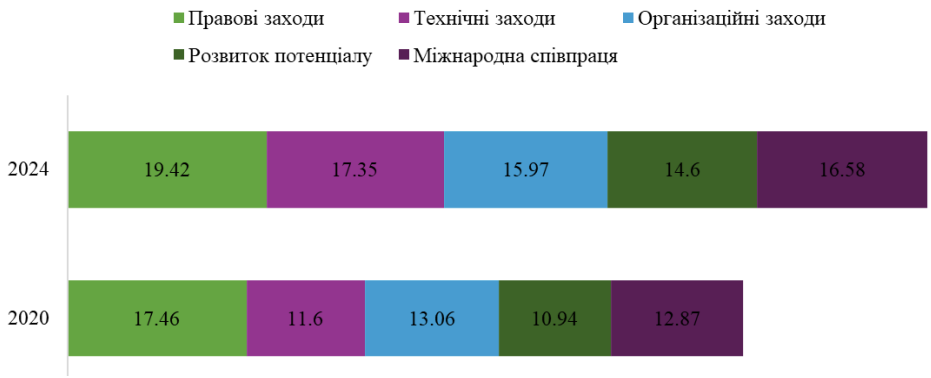


Рис. 2. Динаміка складових Глобального індексу кібербезпеки України у 2020 та 2024 рр.

Джерело: [3-4]

Аналіз рис. 2 свідчить, що за досліджуваний період найбільш помітне покращення відбулося у сфері технічного забезпечення кібербезпеки. Цьому сприяли модернізація цифрової інфраструктури, удосконалення реагування на кіберінциденти та впровадження сучасних стандартів захисту інформації. У результаті Україна поступово скоротила розрив із європейськими країнами за рівнем розвитку кібербезпеки.

Особливо значний вплив на сферу кібербезпеки мали події 2022–2024 рр., коли в умовах повномасштабної агресії рф суттєво зросла інтенсивність кібератак. Власне це посилило потребу в сучасних засобах захисту інформації, цифрових систем і мережевої інфраструктури. Структуру кібератак за основними напрямами в Україні наведено на рис. 3.

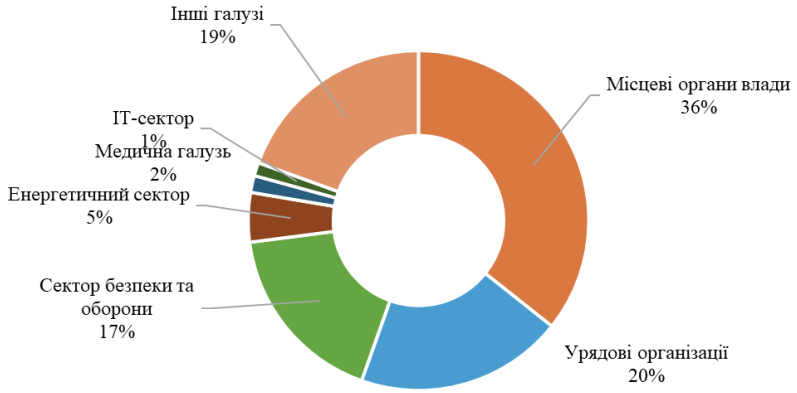


Рис. 3. Структура кібератак за основними напрямами в Україні у 2025 р., %
Джерело: [5]

Аналіз рис. 3 свідчить про збереження значного рівня кіберзагроз для України в умовах воєнного стану. Найбільша частка атак припадала на місцеві органи влади, що відображає підвищену увагу до адміністративних ресурсів та інформаційних систем регіонального рівня. Значна кількість інцидентів також була спрямована на урядові установи та сектор безпеки й оборони, які залишаються одними з основних цілей кібератак.

Водночас кіберзагрози для енергетичного, IT- та медичного секторів створюють ризики для стабільного функціонування критичної інфраструктури. Зростання кількості зафіксованих кіберінцидентів у 2025 р. може бути пов'язане як із посиленням інтенсивності атак, так і з удосконаленням систем їх виявлення та реагування. Саме це актуалізує подальший розвиток технічних засобів захисту та зміцнення кіберстійкості державних і приватних установ. Водночас підвищений рівень цифрових загроз стимулює розвиток українського ринку кібербезпеки та впровадження сучасних технологій інформаційного захисту.

Отже, проведене дослідження засвідчило, що в умовах воєнного стану та посилення агресії рф кібербезпека набуває стратегічного значення для забезпечення стійкості держави. Попри зростання інтенсивності кібератак, Україна демонструє поступове зміцнення системи кіберзахисту, насамперед завдяки розвитку технічної інфраструктури та механізмів реагування на кіберінциденти. Найбільш уразливими залишаються органи державної і місцевої влади, сектор безпеки й оборони, енергетика, IT та медицина. Подальше підвищення кіберстійкості потребує модернізації засобів захисту, удосконалення систем моніторингу загроз та посилення захисту критичної інформаційної інфраструктури.

Перелік використаної літератури

1. Краус К. М., Краус Н. М., Штепа О. В. Цифрова трансформація кібербезпеки на мікрорівні в умовах воєнного стану. Innovation and Sustainability. 2022. № 3. С. 26–37.

2. Global Cybersecurity Index 2024. 2024. URL: https://www.itu.int/epublications/publication/global-cybersecurity-index-2024 (дата звернення 26.09.2026)
3. Global Cybersecurity Index 2020: Measuring Commitment to Cybersecurity. International Telecommunication Union. Geneva, 2021. URL: https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E (дата звернення 26.09.2026)
4. Global Cybersecurity Index 2024: 5th Edition. International Telecommunication Union. Geneva, 2024. URL: https://www.itu.int/epublications/publication/global-cybersecurity-index-2024 (дата звернення 26.09.2026)
5. CERT-UA у 2025 році опрацювала майже 6000 кіберінцидентів: кількість ворожих атак зросла на 37%. Державна служба спеціального зв'язку та захисту інформації України. URL: https://cip.gov.ua/ua/news/cert-ua-u-2025-roci-opracyuvala-maizhe-6000-kiberincidentiv-kilkist-vorozhikh-atak-zrosla-na-37 (дата звернення 26.09.2026)

Ровда Володимир Володимирович, аспірант
Державний університет
інформаційно-комунікаційних технологій,
м. Київ

ІНТЕЛЕКТУАЛЬНІ МЕТОДИ ОПТИМІЗАЦІЇ КАНАЛЬНОГО ТА МЕРЕЖЕВОГО РІВНІВ У ВИСОКОНАВАНТАЖЕНИХ ЛОКАЛЬНИХ РАДІОМЕРЕЖАХ

Розглянуто проблеми забезпечення ефективності локальних радіомереж в умовах високого навантаження та динамічної деградації ресурсів вузлів. Запропоновано два взаємопов'язані методи: адаптивний інтелектуальний MAC-протокол для мінімізації колізій на основі нечіткої логіки та предиктивного аналізу.

Вступ та постановка проблеми

Функціонування сучасних локальних радіомереж, зокрема в умовах автономних сенсорних інфраструктур та промислового Інтернету речей (IoT), характеризується постійною деградацією ресурсів вузлів та непередбачуваними відмовами. На канальному рівні класичні протоколи множинного доступу (CSMA/CA) базуються на механізмі двійкового експоненційного відступу. У високонавантажених сценаріях цей підхід є надто грубим: він не враховує поточну кількість активних конкурентів та динаміку завантаження каналу, що призводить до нелінійного зростання ймовірності колізій та катастрофічного падіння пропускної здатності.

На мережевому рівні традиційний підхід до управління топологією є переважно реактивним: переконфігурування відбувається лише після фізичної відмови вузла. Це спричиняє періоди втрати зв'язку та лавиноподібне перевантаження альтернативних шляхів. Фундаментальне протиріччя полягає у конфлікті між необхідністю проактивного переконфігурування топології для запобігання збоям та обмеженими обчислювальними й енергетичними ресурсами вузлів. Метою роботи є розробка інтелектуальних методів адаптивного доступу до каналу та проактивного переконфігурування топології для забезпечення синергетичного ефекту підвищення надійності та часу життя мережі.

Удосконалений метод оптимізації MAC-протоколу (AI-MAC)

Запропонований метод AI-MAC замінює класичний механізм фіксованого відступу на інтелектуальний контролер, який динамічно оптимізує параметри вікна конкуренції для кожного вузла. Стан кожного вузла описується сукупністю локальних змінних: оцінкою кількості активних конкурентів, довжиною черги пакетів, класом пріоритету трафіку, історією колізій за ковзним вікном та середньою затримкою обслуговування. Ймовірність колізії для конкретного вузла залежить від добутку ймовірностей утримання від передачі всіх інших активних вузлів у поточному часовому слоті.

Цільова функція методу полягає у максимізації сукупної пропускної здатності сегмента мережі при одночасній мінімізації ймовірності колізій та дотриманні обмежень затримки для кожного класу трафіку:

$$\max_{\{CWi(t)\}} J^{MAC} = \sum_{i=1}^{N} \left[\omega_1 S_1(t) - \omega_2 P_{col,i} - \omega_3 \frac{di(t)}{D_{\max}^{ci}} \right] \quad (1)$$

де

Si (t) — миттєва пропускна здатність *i*-го вузла;

Pcol,i(t)— ймовірність колізії;

di(t) — середня затримка;

$D^{ci}{}_{max}$ — максимально допустима затримка для класу пріоритету

$\omega_1, \omega_2, \omega_3$ — вагові коефіцієнти, що балансують цілі.

Архітектура AI-MAC.

Метод реалізується через чотири модулі:

Оцінка конкурентів: кожен вузол безперервно оцінює кількість активних суперників на основі аналізу зайнятості каналу. Для згладжування шумових флуктуацій у нестаціонарних умовах використовується модифікований фільтр Калмана.

Нечіткий контролер вікна конкуренції: на основі вхідних змінних (кількість конкурентів, довжина черги, історія колізій) нечітка система виведення з трикутними функціями належності визначає оптимальний розмір вікна конкуренції. Дефазифікація виконується методом центру ваги, що забезпечує обчислювальну легкість.

Пріоритетне диференціювання: на відміну від стандарту, де параметри фіксовані, AI-MAC динамічно адаптує різницю між класами трафіку залежно від глобального рівня завантаження сегмента, гарантуючи обслуговування критичного трафіку навіть у періоди пікового навантаження.

Предиктивне уникнення колізій: використовуючи історію колізій, модуль прогнозує ймовірність колізії на наступний слот за допомогою експоненційного згладжування. Якщо прогнозована ймовірність перевищує порогове значення, вузол превентивно збільшує вікно конкуренції у визначену кількість разів або відкладає передачу.

Комплексний метод превентивного переконфігурування топології (ProTop)

Для запобігання розривам зв'язку через деградацію ресурсів запропоновано метод ProTop, який моделює мережу як динамічний зважений граф. Кожен вузол відстежує свій стан, включаючи поточну енергію, швидкість її розряду, відношення сигнал/шум, навантаження та прогнозований залишковий час життя.

Прогнозування залишкового часу життя вузла здійснюється на основі нелінійної екстраполяції, яка враховує не лише поточну швидкість розряду, але й її прискорення (другу похідну).

Детекція критичних подій. Критична подія, що потребує переконфігурування, визначається як ситуація, за якої прогнозований стан мережі через заданий горизонт прогнозування порушує обмеження функціонування. Це відбувається, якщо прогнозований час життя будь-якого вузла-ретранслятора менший за горизонт прогнозування, якщо очікуване відношення сигнал/шум на критичному ребрі падає нижче мінімального порогу, або якщо число обумовленості матриці зв'язності графу перевищує допустиму межу, що свідчить про наближення мережі до стану розпаду.

Оптимізація топології. При виявленні критичної події формується задача пошуку нової конфігурації графу, яка мінімізує цільову функцію штрафу

Алгоритм ProTop. Метод реалізує трирівневе прогнозування (коротко-, середньо- та довгострокове) та подієвий тригер ініціації замість періодичного. При спрацюванні тригера виконується ідентифікація локальної зони впливу для уникнення глобального перерахунку, генерується набір кандидатних топологій (шляхом перепризначення ролей кластерних голів та активації резервних ребер), обирається найкраща згідно з функцією (3), і здійснюється плавна міграція (спочатку активуються нові ребра, потім деактивуються старі) з подальшою верифікацією та можливістю відкату.

Результати імітаційного моделювання

Для верифікації методів було проведено дискретно-подієве імітаційне моделювання в середовищі Python.

Результати для AI-MAC (сегмент мережі з 50 вузлів):

При зростанні кількості вузлів від 10 до 50 класичний протокол DCF показує стрімке зростання ймовірності колізій (з 0.12 до 0.45). Запропонований AI-MAC утримує цей показник на рівні 0.08–0.18, що відповідає зниженню колізій на 60%. Сукупна пропускна здатність сегмента зростає на 25–35% порівняно з DCF. Індекс справедливості Джайна залишається на рівні не нижче 0.88, а обмеження затримки для високопріоритетного трафіку (менше 10 мс) гарантовано дотримуються навіть за умов пікового навантаження, тоді як класичний підхід порушує їх вже при 30 вузлах.

Результати для ProTop (мережа зі 100 вузлів):

Реактивний підхід генерує в середньому 12–18 розривів зв'язку на 1000 слотів. Метод ProTop зменшує цей показник до 2–4 розривів (зниження в 4–5 разів), оскільки критичні ситуації нейтралізуються до моменту фізичної відмови. Мережа з ProTop зберігає 80% активних вузлів на 40% довше порівняно з реактивним підходом завдяки превентивному розвантаженню вузлів. Подієвий механізм ініціації переконфігурування генерує на 55% менше службового трафіку порівняно з періодичним переконфігуруванням.

Висновки

Розроблено інтелектуальний MAC-протокол AI-MAC, який інтегрує нечіткий контролер, предиктивний модуль уникнення колізій та динамічне пріоритетне диференціювання. Це дозволило знизити ймовірність колізій на 60% та збільшити пропускну здатність на 25–35% у високонавантажених сегментах.

Запропоновано комплексний метод ProTop превентивного переконфігурування топології. Використання нелінійного прогнозування залишкового часу життя (з урахуванням прискорення деградації) та подієвого механізму ініціації забезпечує зменшення розривів зв'язку в 4–5 разів та подовження часу життя мережі на 40%.

Доведено, що інтеграція методів AI-MAC та ProTop створює синергетичний ефект: дані про зміну топології автоматично передаються на MAC-рівень для коригування нечіткого контролера, що забезпечує комплексну адаптацію мережі до динамічних умов при мінімальному службовому навантаженні.

Література

1. *IEEE 802.11-2020 - IEEE Standard for Information technology--Telecommunications and information exchange between systems.*
2. *Zhan C., Dong Y. A Survey on Machine Learning for MAC Protocols. // IEEE Communications Surveys & Tutorials, 2023.*

3. *Wang M. et al. Proactive Topology Management in IoT Networks: A Comprehensive Survey. // IEEE Internet of Things Journal, 2022.*
4. *Bianchi G., Tinnirello I. Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network. IEEE INFOCOM, 2003.*
5. *Naparstek O., Cohen K. Deep multi-user reinforcement learning for distributed dynamic spectrum access. IEEE Trans. Wireless Commun., 2019. Vol. 18, No. 1. P. 310–323.*
6. *Menache I., Ozdaglar A. Network Games: Theory, Models, and Dynamics. Morgan & Claypool, 2011.*
7. *Banchs A., Vollero L. A dynamic scheduling approach to flexible QoS support in IEEE 802.11e. IEEE Wireless Commun., 2005.*

УДК 004.056.53:519.217

Лукінов Я.В., здобувач ОС «бакалавр»
Нечитайло Ю.А., доцент
Державний біотехнологічний університет (ДБТУ)
Харків, Україна

ВИКОРИСТАННЯ СТОХАСТИЧНИХ МОДЕЛЕЙ ТА СТАТИСТИЧНОГО АНАЛІЗУ ДЛЯ ВИЯВЛЕННЯ Й ПРОТИДІЇ КІБЕРАТАКАМ У КІБЕРФІЗИЧНИХ СИСТЕМАХ

У роботі досліджується проблематика виявлення та протидії кібератакам у кіберфізичних системах за допомогою стохастичних моделей і методів статистичного аналізу. Обґрунтовано неефективність традиційних детермінованих засобів захисту проти прихованих втручань. Проаналізовано застосування алгоритмів на базі адаптивних кумулятивних сум (CUSUM), методів зниження розмірності (t-SNE), байєсівської фільтрації та моделей простору станів. Доведено, що статистичні підходи забезпечують високу інтерпретованість результатів, гарантують низький рівень хибних спрацювань і потребують значно менших обчислювальних ресурсів. Окреслено перспективи створення гібридних систем безпеки для критичної інфраструктури.

Ключові слова: кіберфізичні системи, виявлення аномалій, кібербезпека, стохастичні моделі, статистичний аналіз, приховані атаки, машинне навчання.

Стрімке впровадження кіберфізичних систем (КФС), які безперервно інтегрують обчислювальні, мережеві та фізичні процеси (зокрема комплекси автоматизованого управління SCADA, програмовані логічні контролери PLC та промислові пристрої інтернету речей IIoT), у критичну інфраструктуру (енергетичний сектор, системи водопостачання, транспортні мережі, промислове виробництво) супроводжується експоненційним зростанням кількості та масштабу кібератак. Ці атаки все частіше спрямовуються безпосередньо на дестабілізацію фізичного рівня функціонування систем [2, с. 2]. Традиційні сигнатурні методи виявлення вторгнень та жорстко детерміновані правила безпеки демонструють критично низьку ефективність проти новітніх, таргетованих і прихованих (stealthy) атак. Це зумовлено тим, що такі підходи не враховують стохастичну динаміку фізичного процесу та імовірнісні закономірності його поведінки в штатному режимі [1, с. 3]. Зловмисник, який діє повільно, маніпулюючи даними в межах статичних порогових значень детекторів, здатний тривалий час залишатися непоміченим, що зрештою може призвести до руйнування обладнання чи масштабних екологічних катастроф [1, с. 4]. За таких умов безальтернативним стає застосування математичного апарату стохастичних моделей (марковські процеси, моделі простору станів) та методів багатокритеріального статистичного аналізу для синтезу адаптивних систем виявлення аномалій. Такі системи здатні предиктивно розпізнавати найменші відхилення від імовірнісного профілю нормальної роботи КФС [3, с. 2; 4, с. 11].

Метою дослідження є комплексний аналіз та наукова систематизація сучасних підходів до застосування стохастичних моделей і методів статистичного аналізу для раннього виявлення й ефективної протидії кібератакам у кіберфізичних системах. Окрім того, завданням роботи є визначення концептуальних переваг і технологічних обмежень цих підходів порівняно з класичними детермінованими парадигмами та методами глибинного машинного навчання.

Одним із найбільш показових та емпірично підтверджених прикладів застосування статистичного підходу є метод адаптивного до поточного стану системи виявлення аномалій на основі кумулятивної суми залишків (state-aware CUSUM). Цей алгоритм був успішно апробований на базі промислового водоочисного комплексу SWaT [1, с. 7]. Специфіка таких об'єктів критичної інфраструктури полягає у наявності значної фізичної інерції (наприклад, зміна рівня води в резервуарах або концентрації хімічних реагентів відбувається поступово). На відміну від традиційних методів, CUSUM здатний накопичувати незначні відхилення в часі, що робить його надзвичайно чутливим до stealthy-атак. Такі атаки реалізуються шляхом мікроманіпуляцій, коли зловмисник намагається непомітно змінити стан системи, обережно маневруючи в межах допустимих статичних порогових значень сигналізації. У запропонованій математичній моделі похибка (residual) між фактично виміряним сенсорами і теоретично оціненим значенням нормалізується відносно історичного середнього показника для конкретного етапу технологічного процесу. Це дозволяє динамічно адаптувати пороги чутливості залежно від поточних режимів роботи (наприклад, заповнення, фільтрація, скидання), суттєво звужуючи «простір маневру» для зловмисника, та мінімізувати частку хибних спрацювань до 99,2 % порівняно з традиційними стан-незалежними детекторами, які генерують надмірну кількість помилкових тривог під час перехідних процесів [1, с. 9–10].

Іншим фундаментальним науковим вектором є формування репрезентативних еталонних вибірок (датасетів) для об'єктивного оцінювання алгоритмів детекції у промислових мережах, де наявність якісних маркованих даних є критичною проблемою. Зокрема, інноваційний датасет ICS-Flow комплексно поєднує мережеві потокові ознаки (тривалість сесій, обсяги переданих пакетів, міжпакетні інтервали) з логами станів фізичного процесу під впливом чотирьох класів кіберзагроз: розвідувальні операції, DDoS-атаки, семантична підміна даних через механізм Man-in-the-Middle (MitM) та повторне відтворення мережевих пакетів [2, с. 5–6]. У випадку MitM-атак зловмисник не просто порушує доступність системи, а здійснює інтелектуальну модифікацію команд управління (payload), що вимагає глибокого аналізу кореляцій між мережевим і фізичним рівнями. Для глибинного аналізу внутрішньої структури цих багатовимірних даних успішно застосовується нелінійний статистичний

алгоритм зниження розмірності t-SNE (t-Distributed Stochastic Neighbor Embedding). Його використання дозволяє математично обґрунтовано трансформувати складний гіперпростір ознак у дво- або тривимірну проєкцію шляхом мінімізації дивергенції Кульбака — Лейблера між спільними розподілами ймовірностей. Це зберігає локальні зв'язки між об'єктами, допомагає візуально й аналітично виокремити просторові кластери, що відповідають різним векторам атак, гарантуючи високу сепарабельність класів та суттєво полегшуючи роботу подальших алгоритмів класифікації і кластеризації [2, с. 11].

Окремої уваги заслуговує стохастичне моделювання динаміки фізичного процесу через математичний апарат простору станів (state-space model). У цій парадигмі еволюція кіберфізичної системи описується системою рівнянь, де шум технологічного процесу (неконтрольовані збурення) та похибки вимірювання сенсорів розглядаються як випадкові величини із заздалегідь відомими законами розподілу (як правило, гауссівськими). Це максимально наближає математичну модель до реальних умов експлуатації КФС з їхньою притаманною невизначеністю. На цій базі розроблено гібридний метод нейромережевої ідентифікації системи з інтегрованою байєсівською фільтрацією (NSIBF). У межах цього підходу непараметризований фільтр Калмана (unscented Kalman filter) безперервно апроксимує розподіл ймовірностей прихованого стану КФС. На відміну від розширеного фільтра Калмана, він використовує метод сигма-точок (unscented transform), що дозволяє уникнути складного обчислення матриць Якобі та ефективніше працювати із сильно нелінійними динамічними системами. Водночас наявність аномалії ідентифікується через розрахунок відстані Махаланобіса між фактичним спостереженням і багатовимірним прогнозом. Ця метрика об'єктивно враховує коваріаційну матрицю нормальних станів, тобто дисперсію та взаємозв'язки між різними фізичними змінними (температурою, тиском, витратою) [3, с. 4–5]. Апробація цього підходу підтвердила зростання інтегральної метрики точності F1-score на 2,9–7,6 % порівняно з базовими методами прогнозування, які ігнорують стохастичну природу процесу [3, с. 8].

Систематизація зазначених підходів ґрунтовно розкрита в сучасному науковому огляді архітектур безпеки КФС, де модельно-орієнтовані статистичні алгоритми виокремлюються як фундаментальний клас безпекових рішень, що здатен функціонувати як самостійно, так і в тандемі з іншими технологіями [4, с. 15]. Хоча надскладні моделі глибинного навчання (Deep Learning) часто демонструють високу абсолютну точність детекції, вони функціонують за принципом «чорного ящика», приховуючи логіку прийняття рішень. Статистичні методи, натомість, забезпечують прозорість обчислень та високий рівень інтерпретованості (explainability) результатів. Це є критичною вимогою в умовах жорстких обмежень обчислювальних потужностей промислових мікроконтролерів (PLC), терміналів віддаленого зв'язку (RTU) та периферійних

вузлів, які просто не здатні підтримувати ресурсомісткі нейромережеві обчислення в режимі реального часу [4, с. 18–19].

Список використаних джерел

1. Ghaeini H. State-Aware Anomaly Detection for Industrial Control Systems / H. Ghaeini, A. Nasrullah, M. H. H. Nejad [et al.] // Proceedings of the ACM Conference. — 2018. — С. 1–12. — URL: https://tippenhauer.de/publication/ghaeini-18-stateaware/ghaeini-18-stateaware.pdf (дата звернення: 15.09.2026).
2. Dehlaghi-Ghadim A. Anomaly Detection Dataset for Industrial Control Systems / A. Dehlaghi-Ghadim, M. H. Moghadam, A. Balador, H. Hansson // arXiv:2305.09678 [cs.CR]. — 2023. — 15 с. — URL: https://arxiv.org/pdf/2305.09678.pdf (дата звернення: 15.09.2026).
3. Feng C. Time Series Anomaly Detection for Cyber-Physical Systems via Neural System Identification and Bayesian Filtering / C. Feng, P. Tian // arXiv:2106.07992 [cs.LG]. — 2021. — URL: https://arxiv.org/pdf/2106.07992 (дата звернення: 15.09.2026).
4. Abshari D. Cyber-Physical Systems Security: A Comprehensive Review of Anomaly Detection Techniques / D. Abshari, M. Sridhar // Internet of Things (Elsevier) / arXiv:2502.13256 [cs.CR]. — 2025. — URL: https://arxiv.org/pdf/2502.13256 (дата звернення: 15.09.2026).

Нестор Наталія Володимирівна,
доктор юридичних наук, професор, заслужений юрист *України, адвокат, заступниця генерального директора ДП «Національні інформаційні системи»*

ЗАБЕЗПЕЧЕННЯ КІБЕРСТІЙКОСТІ ДЕРЖАВНИХ РЕЄСТРІВ УКРАЇНИ В УМОВАХ ГІБРИДНИХ ЗАГРОЗ: ЄВРОПЕЙСЬКІ СТАНДАРТИ ТА МІЖНАРОДНЕ СПІВРОБІТНИЦТВО

Анотація. Досліджено кіберстійкість державних реєстрів як складову безперервності цифрової юстиції, захисту прав людини та міжнародної довіри до державних цифрових сервісів. На основі аналізу європейських актів NIS2, Cybersecurity Act та Interoperable Europe Act визначено взаємозв'язок управління кіберризиками, інтероперабельності й доказової надійності реєстрових даних. Запропоновано авторську модель міжнародного співробітництва, що поєднує координацію реагування на інциденти, взаємне навчання, перевірку цілісності даних та інституційне забезпечення відновлення послуг. Обґрунтовано необхідність адаптації європейських підходів з урахуванням правового статусу України та ризиків воєнного часу.

Цифровізація державного управління змінила значення державних реєстрів: вони стали не лише сховищами юридично значущих відомостей, а й інфраструктурною основою реалізації прав, адміністративних процедур та доступу до правосуддя. У період збройної агресії й гібридних загроз втрата доступності реєстру, несанкціонована зміна записів або порушення простежуваності операцій можуть спричинити одночасно технологічні, правові та міжнародно-репутаційні наслідки. Тому кіберстійкість доцільно визначати як здатність реєстрової системи запобігати порушенням, зберігати достовірність відомостей, підтримувати критичні функції, відновлюватися після інцидентів та забезпечувати перевірюваність виконаних дій.

Метою дослідження є обґрунтування міжнародно-інституційних механізмів підвищення кіберстійкості державних реєстрів України з урахуванням європейських стандартів. Методологічну основу становлять системний, порівняльно-правовий та інституційний підходи. Аналіз стосується нормативних принципів і організаційних моделей, а не непублічних відомостей про архітектуру чи вразливості конкретних державних систем.

Директива (ЄС) 2022/2555 (NIS2) передбачає ризик-орієнтований підхід до кібербезпеки, управління інцидентами, безперервності діяльності, безпеки ланцюгів постачання та відповідальності управлінських органів [1]. Її положення становлять орієнтир для гармонізації українських практик, однак не слід механічно ототожнювати сферу прямого застосування права ЄС із правовим режимом українських державних підприємств та реєстрів. Відповідні рішення потребують національного нормативного закріплення, визначення суб'єктів відповідальності та співвіднесення з режимами захисту інформації.

Регламент (ЄС) 2019/881 (Cybersecurity Act) закріплює інституційну роль ENISA та європейську рамку сертифікації кібербезпеки [2]. Для українського реєстрового сектору значення має не формальне копіювання сертифікаційних механізмів, а формування перевірюваних вимог до постачальників, програмних компонентів, управління вразливостями та незалежного оцінювання захищеності. Важливо поєднати такі вимоги з процедурами закупівель і контролю виконання договорів, щоб міжнародні стандарти перетворювалися на практичні критерії надійності.

Окремий міжнародний вимір пов’язаний з інтероперабельністю. Регламент (ЄС) 2024/903 (Interoperable Europe Act) формує підхід до транскордонної взаємодії цифрових публічних послуг через правову, організаційну, семантичну та технічну сумісність [3]. Для України це важливий методологічний орієнтир у процесі європейської інтеграції. Проте розширення обміну даними без належної ідентифікації, розмежування доступу та журналювання може збільшувати поверхню атаки. Отже, інтероперабельність і кіберстійкість слід проєктувати спільно, а не як послідовні, незалежні етапи.

Міжнародне співробітництво у цій сфері виходить за межі технічної допомоги. Воно охоплює дипломатичне узгодження пріоритетів цифрової інтеграції, вироблення сумісних процедур реагування на транскордонні інциденти, обмін методиками оцінювання ризиків, підготовку персоналу та залучення незалежної експертизи. Партнерства з інституціями ЄС і профільними мережами мають будуватися на визначених мандатах, законних підставах обміну інформацією та захисті персональних даних. Такий підхід перетворює кіберстійкість із внутрішнього ІТ-завдання на предмет міжнародної інституційної взаємодії.

Особливого значення набуває судово-експертний вимір. Реєстрові записи, журнали подій, резервні копії та метадані можуть використовуватися для встановлення обставин інциденту або перевірки юридично значущих дій. Їхня доказова придатність залежить не лише від захищеності системи, а й від належної фіксації походження, часу, цілісності та ланцюга збереження цифрових матеріалів. Міжнародна співпраця у сфері цифрової криміналістики має сприяти сумісності процедур документування, зберігаючи процесуальні гарантії та компетенцію національних органів.

Пропонується авторська чотирикомпонентна модель співробітництва: (1) нормативна сумісність — зіставлення українських вимог із ризик-орієнтованими принципами NIS2; (2) операційна координація — погоджені канали повідомлення про інциденти та спільні навчання; (3) доказова надійність — уніфіковані вимоги до журналювання, фіксації змін і збереження цифрових слідів; (4) інституційне відновлення — сценарії резервування, перевірки відновлених даних і безперервного надання критичних послуг. Модель є

дослідницькою пропозицією, а не описом уже запровадженого державного механізму.

У практиці міжнародних відносин така кооперація може реалізовуватися через тематичні міжурядові консультації, експертні місії, спільні протоколи навчань та проєкти технічної допомоги. Водночас участь у міжнародних механізмах не повинна підміняти внутрішню відповідальність за правомірність доступу до реєстрових відомостей. Узгодження умов міжнародного обміну інформацією потребує окремого визначення цілей, категорій даних, строків їх зберігання та відповідальних установ.

Практична реалізація моделі передбачає визначення власників ризиків, класифікацію критичних функцій, регулярні міжвідомчі навчання, документоване тестування відновлення та незалежний перегляд процедур роботи з цифровими доказами. Показниками результативності можуть бути час відновлення критичних сервісів, частка перевірених резервних копій, повнота аудиторських журналів та частота успішних навчальних сценаріїв. Такі індикатори варто застосовувати без розкриття інформації, що може створити додаткові загрози безпеці.

Отже, кіберстійкість державних реєстрів є одночасно питанням цифрового врядування, забезпечення прав людини та міжнародної співпраці. Європейські акти надають нормативні орієнтири, проте їх ефективне використання в Україні залежить від адаптації до національних компетенцій і реальних загроз. Наукова новизна запропонованого підходу полягає в інтеграції нормативного, дипломатично-інституційного, операційного та судово-експертного вимірів в єдину модель забезпечення стійкості державних реєстрів.

Список використаних джерел

1. Directive (EU) 2022/2555 of 14 December 2022 (NIS2). URL: https://eur-lex.europa.eu/eli/dir/2022/2555/oj
2. Regulation (EU) 2019/881 of 17 April 2019 (Cybersecurity Act). URL: https://eur-lex.europa.eu/eli/reg/2019/881/oj
3. Regulation (EU) 2024/903 of 13 March 2024 (Interoperable Europe Act). URL: https://eur-lex.europa.eu/eli/reg/2024/903/oj

Леляк А. В., Астраханцев А. А.

ЗАХИСТ КАНАЛІВ В СИСТЕМАХ ЕЛЕКТРОННОЇ ІДЕНТИФІКАЦІЇ ЗА ДОПОМОГОЮ ВДОСКОНАЛЕНОЇ БІБЛІОТЕКИ OPENSCP

Системи електронної ідентифікації стали фундаментальним компонентом сучасної цифрової інфраструктури. Документи електронної ідентифікації містять високочутливі персональні дані та криптографічні ключі, тому вони зазвичай зберігаються в захищеному елементі, а зв'язок з такими пристроями захищений протоколами Secure Channel Protocol (SCP) від GlobalPlatform, зокрема SCP03 та SCP11. Для вирішення ряду проблем існуючих рішень з відкритим кодом було розроблено бібліотеки OpenSCP-Java та OpenSCP-Python для забезпечення підтримки останніх версій протоколу SCP та з покращеннями архітектури. Аналіз підвищення рівня безпеки показує кількісний приріст рівня криптографічної стійкості для ентропії криптографічних викликів та стійкості MAC кодів. Якісний аналіз показує приріст рівня захищеності на 6% за критеріями НД ТЗІ 2.5-004-99.

Ключові слова: електронна ідентифікація, безпека, Secure Channel Protocol

Вступ

Електронне посвідчення особи – це набір цифрових атрибутів та ідентифікаторів, що унікально ідентифікує людину як цифрову сутність в системі. Електронне посвідчення може бути використане для ідентифікації особи, аутентифікації та цифрового підпису. Для безпечного зберігання електронних документів на пристрої користувача з максимальним рівнем захисту використовується чіп вбудованого захищеного елемента (embedded Secure Element, eSE).

Для комунікації з захищеними елементами (Secure Element, SE) використовуються команди APDU (Application Protocol Data Unit) зі специфікації ISO/IEC 7816-4 [1], що передаються через канал, захищений безпековим протоколом організації Global Platform. Безпека забезпечується протоколами SCP (Secure Channel Protocol) версій 3 (SCP03) та 11 (SCP11), що використовуються як для комунікації між захищеним елементом і сервером видачі документів для персоналізації чипа, так і для комунікації між захищеним елементом і додатком гаманця для передачі електронного документа за запитом зчитувача.

Бібліотеки OpenSCP

Імплементація протоколів SCP на захищеному елементі постачається виробником чипа та не може бути замінена. Існуючі імплементації протоколів SCP для ОСЕ з відкритим кодом мають наступні недоліки:

- Неповна імплементація, що не підтримує останні версії протоколів – SCP03 v1.2 та SCP11 v1.4.
- Жорстка політика ліцензування (здебільшого LGPL-3.0), що не дозволяє використання в комерційному закритому коді.
- Імплементація протоколу тісно зв’язана з специфікою конкретної апаратної частини, що не дозволяє легко змінити канал передачі даних чи апаратну частину.

Для подолання цих проблем, були створені бібліотеки OpenSCP-Java [2] та OpenSCP-Python [3]. Запропоноване технічне рішення має наступні переваги:

- Рішення постачається в форматі бібліотек, що дозволяє інтеграцію з будь-якою вже існуючою клієнтською логікою.
- В рішенні чітко відокремлена імплементація протоколів SCP від специфіки апаратної частини, що дозволяє клієнту використовувати рішення з будь-яким варіантом з'єднання до захищеного елемента.
- Рішення підтримує останні версії протоколів SCP та випущено під ліцензією Apache 2.0, що не забороняє клієнтам будувати власні рішення.

Оцінка покращення рівня захищеності

Рівень захищеності оцінювався для двох систем зберігання електронних посвідчень особи, де документ зберігається на захищеному елементі. З захищеним елементом взаємодіють сервер видачі електронних посвідчень та додаток гаманця. В якості існуючої системи розглядається типове рішення, де реалізована імплементація SCP протоколів для OCE (Off-Card Entity). Для аналізу рівня захищеності в роботі оцінювалась криптографічна стійкість (кількісний аналіз) та критерії захищеності інформації згідно з НД ТЗІ 2.5-004-99 [4] (якісний аналіз).

Кількісний аналіз підвищення захищеності показав, що запропонована система використовує режим S16, що збільшує розмір виклику до 16 байтів. Ентропія криптографічного виклику зростає в $F_H = \frac{H_{S16}}{H_{S8}} = \frac{2^{128}}{2^{64}} = 2^{64}$ рази, що робить атаку методом повного перебору обчислювально недоцільною. Також стійкість MAC в запропонованій системі використовує повне значення MAC розміром в 16 байт, що зменшує ймовірність підробки значення $MACF_{Pf} = \frac{P_{forge}^{S8}}{P_{forge}^{S16}} = \frac{2^{-64}}{2^{-128}} = 2^{64}$і унеможливлює практичне здійснення атак на колізії MAC. Оцінка захищеності взаємної автентифікації визначалась як $Security \approx \min(H_{challenge}, H_{MAC}, H_{key})$. Рівень безпеки існуючої системи: $Security_{S8} = \min(64,64,128) = 64$ bits. Рівень безпеки запропонованої системи: $Security_{S16} = \min(128,128,128) = 128$ bits, що збільшує рівень захищеності в два рази.

Якісний аналіз підвищення захищеності базувався на наступній оцінці існуючої системи за критеріями НД ТЗІ 2.5-004-99: КД-4, КА-4, КО-1, КК-1, КВ-4, ЦД-4, ЦА-4, ЦО-1, ЦВ-3, ДР-1, ДС-2, ДЗ-1, ДВ-2, НР-2, НИ-3, НК-2, НО-3, НЦ-3, НТ-2, НВ-2, НА-2, НП-2, Г-2 (Архітектура), Г-3 (Середовище розробки), Г-3 (Послідовність розробки), Г-3 (Середовище функціонування), Г-1 (Документація), Г-1 (Випробування КЗЗ).

В запропонованому рішенні пропрієтарний формат сертифікатів від Global Platform (має обмежену взаємодію та сумісність) замінено на розповсюджений формат сертифікатів X.509 для інфраструктури відкритих ключів. Це покращує оцінку системи за критерієм «Ідентифікація та автентифікація при обміні» з НВ-

2 до НВ-3 за рахунок виконання наступної умови: «Використовуваний протокол автентифікації повинен забезпечувати можливість однозначного підтвердження джерела об'єкта незалежною третьою стороною».

Основним архітектурним покращенням системи, що використовує бібліотеки OpenSCP, є чітке відокремлення логіки протоколу від логіки передачі даних по конкретному фізичному каналу зв’язку. Це покращує рівень гарантій за критерієм «Архітектура» з Г-2 до Г-3, оскільки починає виконуватися умова «КЗЗ повинен складатися з добре визначених і максимально незалежних компонентів. Кожний з компонентів повинен бути спроектований відповідно до принципу мінімуму повноважень».

Відокремлення імплементації протоколу SCP в запропонованій системі дозволило створити нові ізольовані тести на правильність імплементації захисту каналу передачі даних. Ці тести виявили дрібні недоліки в початковій імплементації, що були виправлені в запропонованому рішенні. Це покращило рівень гарантій за критерієм «Випробування комплексу засобів захисту» з Г-1 до Г-2 за рахунок виконання вимоги: «“Розробник повинен усунути або нейтралізувати всі знайдені “слабкі місця” і виконати повторне тестування КЗЗ для підтвердження того, що виявлені недоліки були усунені і не з'явилися нові “слабкі місця”».

Конвертація оцінки за критеріями НД ТЗІ в числові значення показало, що використання бібліотек OpenSCP в екосистемі електронних посвідчень особи дозволяє покращити рівень захищеності системи на 6%.

Висновки

Існуючі рішення з відкритим кодом мають обмежену підтримку останніх версій SCP протоколів, жорсткі політики ліцензування та тісний зв’язок логіки протоколу з логікою апаратної частини, що сильно обмежує їх використання в комерційних та крос-платформених системах електронних документів. Для подолання цих проблем були розроблені бібліотеки OpenSCP, що підтримують останні версії SCP протоколів (SCP03 v1.2, SCP11 v1.4), мають чітко розділену імплементацію протоколів SCP та апаратно-залежної частини, та були випущені ліцензією Apache 2.0, що дозволяє їх використання в закритому пропрієтарному коді.

Кількісна оцінка рівня захищеності показує збільшення ентропії криптографічного виклику та криптографічної стійкості MAC (Message Authentication Code) в 2^{64} рази, і збільшення рівня безпеки взаємної автентифікації в 2 рази. Якісна оцінка рівня захищеності за НД ТЗІ 2.5-004-99 демонструє підвищення рівня захищеності запропонованої системи на 6% за рахунок наступних критеріїв – «Ідентифікація та автентифікація при обміні», «Відкат», «Архітектура» та «Випробування комплексу засобів захисту».

Список використаних джерел

1. International Organization for Standardization. ISO/IEC 7816-4:2020. Ідентифікаційні картки. Картки з інтегральною схемою. Частина 4: Організація, безпека та команди для обміну даними. 2020. ISO Online Browsing Platform.
2. Samsung. OpenSCP-Java [Електронний ресурс] : програмне забезпечення. GitHub. Репозиторій Samsung OpenSCP-Java.
3. Samsung. OpenSCP-Python [Електронний ресурс] : програмне забезпечення. GitHub. Репозиторій Samsung OpenSCP-Python.
4. Адміністрація Державної служби спеціального зв’язку та захисту інформації України. НД ТЗІ 2.5-004-99 [Електронний ресурс] : нормативний документ. 1999. PDF.

Нетребко Є.Г.
студент групи 125-23-3,
Національний технічний університет «Дніпровська політехніка»,
Дніпро, Україна

Мєшков В.І.
доктор філософії, доцент кафедри безпеки інформації та телекомунікацій,
Національний технічний університет «Дніпровська політехніка»
Дніпро, Україна

АНАЛІЗ УМОВ ВИБОРУ ЗАСОБІВ БАГАТОФАКТОРНОЇ АВТЕНТИФІКАЦІЇ ДЛЯ ЗАХИСТУ ОБЛІКОВИХ ЗАПИСІВ

Досліджено умови вибору засобів багатофакторної автентифікації (MFA) для захисту облікових записів. Порівняно SMS-коди, TOTP, push-підтвердження та FIDO2/WebAuthn за стійкістю до фішингу й обмеженнями застосування. Визначено пріоритетні категорії облікових записів для інформаційних систем університету, запропоновано послідовність упровадження MFA з урахуванням відновлення доступу. Отримані результати можуть бути використані для вибору засобів автентифікації відповідно до ризиків і технічних можливостей сервісів.

Ключові слова: багатофакторна автентифікація, MFA, обліковий запис, фішинг, FIDO2, WebAuthn.

Використання лише паролів не забезпечує достатнього захисту облікових записів: облікові дані можуть бути викрадені внаслідок фішингу, повторного використання паролів або витоків. Особливої уваги потребують електронна пошта, хмарні й навчальні сервіси та адміністративні профілі. CISA рекомендує застосовувати багатофакторну автентифікацію (Multi-Factor Authentication, MFA) усюди, де вона доступна [1].

Багатофакторна автентифікація використовує щонайменше два фактори різних категорій: знання (пароль, PIN-код), володіння (пристрій, криптографічний ключ) або біометричну характеристику. Два паролі не утворюють MFA. Водночас різні засоби підтвердження особи мають неоднакову стійкість до атак, тому самого факту ввімкнення MFA недостатньо для обґрунтування вибору механізму [2].

Метою дослідження є визначення умов вибору засобів багатофакторної автентифікації для захисту облікових записів з урахуванням їх критичності, стійкості механізмів до фішингу та обмежень практичного застосування.

Аналіз засобів MFA. SMS-коди прості в налаштуванні, однак їх застосування пов’язане з ризиком перенесення номера на іншу SIM-картку, перехоплення повідомлень та передавання коду через підроблений сайт. Електронна пошта залежить від захищеності поштової скриньки; за NIST SP 800-63B-4 вона не допускається як позасмуговий автентифікатор [2].

Програмні автентифікатори формують одноразові паролі на основі часу (TOTP) без отримання SMS. Проте користувач може ввести такий код на фішинговому сайті. Push-підтвердження зручні, але багаторазові запити можуть спонукати до помилкового погодження. Перевірка числа зменшує цей ризик, проте не забезпечує повної стійкості до фішингу [4].

FIDO2/WebAuthn використовують асиметричну криптографію та прив'язку автентифікації до домену сервісу, що запобігає використанню відповіді на підробленому домені [2, 3]. Фізичні ключі й вбудовані автентифікатори забезпечують MFA за умови поєднання незалежних факторів, наприклад володіння пристроєм із PIN-кодом або біометричною перевіркою.

Для систематизації результатів проведено порівняння SMS-кодів, TOTP, push-підтверджень із перевіркою числа та FIDO2/WebAuthn (табл. 1). Критеріями обрано стійкість до фішингу, характерні способи обходу захисту та обмеження впровадження. Під стійкістю до фішингу розуміється неможливість використання автентифікаційної відповіді на підробленому домені навіть у разі взаємодії з ним користувача. Таке зіставлення дає змогу обґрунтувати вибір засобу залежно від наслідків компрометації облікового запису.

Таблиця 1

Порівняння засобів багатофакторної автентифікації

Засіб	**Стійкість до фішингу**	**Основні обмеження**
SMS-коди	Ні	SIM-swapping, перехоплення, передавання коду
TOTP	Ні	Передавання коду на фішинговому сайті
Push із перевіркою числа	Ні	Соціотехнічний вплив, залежність від пристрою
FIDO2/WebAuthn	Так, за належної реалізації	Сумісність, відновлення доступу

Результати порівняння свідчать, що додатковий код SMS або TOTP ускладнює вхід із викраденим паролем, проте користувач може передати його зловмиснику на фішинговій сторінці. Перевірка числа в push-запиті зменшує ризик випадкового підтвердження численних запитів, але не виключає соціотехнічного впливу. На відміну від цих засобів, FIDO2/WebAuthn застосовує криптографічну перевірку з прив'язкою до домену сервісу. Тому критичність облікового запису та ймовірність цільового фішингу мають визначати пріоритет впровадження фішингостійкого механізму; сумісність і вартість залишаються практичними обмеженнями.

Сценарій для освітнього середовища. Для облікових записів адміністраторів, працівників із розширеними повноваженнями, поштових скриньок, через які можливе відновлення інших доступів, та систем оброблення персональних даних доцільно визначити високий пріоритет застосування фішингостійкої MFA. Для звичайних навчальних профілів потрібно враховувати доступність пристроїв і сумісність сервісів, передбачаючи поетапний перехід від доступних механізмів до фішингостійких. Рішення про категорію облікового

запису доцільно ухвалювати за обсягом повноважень, цінністю доступних даних і можливими наслідками захоплення.

Запропонована послідовність упровадження охоплює інвентаризацію облікових записів, їх класифікацію, вибір засобів MFA, пілотне налаштування для групи високого ризику, перевірку відновлення доступу, розширення охоплення та аудит подій входу. Окремої уваги потребує втрата телефона або апаратного ключа: резервні коди необхідно зберігати незалежно від основного пристрою, а заміну автентифікатора проводити лише після належної перевірки особи. Інакше процедура відновлення може стати слабшою ланкою, яка дозволить обійти основний механізм MFA [4].

Для перевірки результативності впровадження запропоновано зіставляти частку облікових записів, захищених MFA, частку критичних профілів із фішингостійкою автентифікацією, середній час входу, кількість звернень щодо відновлення та результати контрольованих перевірок входу на підробленому домені. Показники потрібно порівнювати на однакових категоріях користувачів до і після пілотного впровадження. Практичні випробування в межах роботи не проводилися, тому кількісне зниження ризику не стверджується.

Висновки. Визначено умови вибору засобів MFA з урахуванням критичності облікових записів, стійкості до фішингу та обмежень використання. Порівняльний аналіз SMS, TOTP, push-підтверджень і FIDO2/WebAuthn обґрунтовує перевагу фішингостійких механізмів для облікових записів високого ризику. Запропоновано послідовність упровадження MFA в інформаційних системах університету з урахуванням безпечного відновлення доступу.

Список використаних джерел:

1. CISA. Require Multifactor Authentication. URL: https://www.cisa.gov/audiences/small-and-medium-businesses/secure-your-business/require-multifactor-authentication.

2. NIST. Digital Identity Guidelines: Authentication and Authenticator Management. NIST SP 800-63B-4. 2025. DOI: 10.6028/NIST.SP.800-63b-4.

3. CISA. Implementing Phishing-Resistant MFA. 2022. URL: https://www.cisa.gov/sites/default/files/2023-01/fact-sheet-implementing-phishing-resistant-mfa-508c.pdf.

4. OWASP Foundation. Multifactor Authentication Cheat Sheet. URL: https://cheatsheetseries.owasp.org/cheatsheets/Multifactor_Authentication_Cheat_Sheet.html.

Литвиненко Роман Геннадійович
Буняк Анастасія Володимирівна
Рябець Віктор Вікторович
Курсанти навчально-наукового інституту №4
Харківський національний університет внутрішніх справ
Науковий керівник: Лучик Світлана Дмитрівна
Доктор економічних наук, професор, професор кафедри інформаційних систем та технологій
Харківський національний університет внутрішніх справ

SIEM-СИСТЕМИ ЯК ІНСТРУМЕНТ ЦЕНТРАЛІЗОВАНОГО ЗБОРУ ТА АНАЛІЗУ ІНФОРМАЦІЇ

Розглянуто SIEM-системи як засіб централізованого збору, нормалізації та аналізу подій інформаційної безпеки з різних компонентів інфраструктури. Описано кореляцію подій, застосування поведінкової аналітики UEBA, машинного навчання та технологій SOAR для автоматизації реагування. На прикладі Microsoft Sentinel показано практичні можливості таких рішень. Визначено чинники, що впливають на ефективність SIEM: повнота джерел даних, налаштування правил виявлення, якість журналів і професійна оцінка інцидентів.

Ключові слова: SIEM-системи, кореляція подій, кібербезпека, UEBA, SOAR, журнали подій.

Сучасні організації щодня працюють із великими обсягами інформації, значна частина якої пов’язана з функціонуванням їхніх інформаційних систем. Сервери, мережеве обладнання, операційні системи та різноманітні програмні засоби фіксують дії користувачів, зміни в роботі обладнання й події, які можуть свідчити про порушення безпеки. Проте наявність таких записів сама по собі не гарантує своєчасного виявлення загроз. Якщо інформація зберігається в різних системах і аналізується окремо, встановити повну картину інциденту буває складно. Для розв’язання цієї проблеми використовують SIEM-системи (Security Information and Event Management), які об’єднують дані з різних джерел і дають змогу аналізувати їх у межах єдиного середовища.

Метою дослідження є визначення можливостей SIEM-систем у процесі збору й аналізу подій інформаційної безпеки для виявлення кіберзагроз.

Робота SIEM-систем починається з отримання інформації від різних компонентів інформаційної інфраструктури. Це можуть бути сервери, операційні системи, міжмережеві екрани, антивірусні програми, системи виявлення вторгнень і корпоративні застосунки. Оскільки кожне джерело може використовувати власний формат записів, отримані дані необхідно впорядкувати та привести до вигляду, придатного для спільного аналізу. Важливість належної роботи з журналами подій розглядається в

рекомендаціях Національного інституту стандартів і технологій США (NIST), де описано основні підходи до їх формування, передавання, зберігання та аналізу [1]. Завдяки об'єднанню таких відомостей фахівці можуть досліджувати події не окремо, а з урахуванням загальної ситуації в інформаційній системі.

Однією з найважливіших можливостей SIEM є кореляція подій. Вона пе-редбачає пошук зв'язків між записами, які надходять із різних джерел, і дає змо-гу виявляти підозрілі послідовності дій. Наприклад, спочатку система може зафіксувати кілька невдалих спроб входу до облікового запису, потім успішну автентифікацію, а згодом – звернення до інформації, до якої цей користувач зазвичай не має доступу. Розгляд цих подій у сукупності дає підстави перевірити можливість компрометації облікового запису. Якщо ж аналізувати кожен запис окремо, такий зв'язок може залишитися непоміченим. Отже, значення SIEM полягає не тільки в централізованому зберіганні журналів, а й у можливості використовувати їх для виявлення підозрілої активності та визначення подій, які потребують першочергової уваги.

Можливості аналізу розширюються завдяки використанню машинного на-вчання та поведінкової аналітики. Одним із прикладів є технологія User and Entity Behavior Analytics (UEBA), яка допомагає визначати характерні моделі поведінки користувачів і пристроїв та виявляти відхилення від них [2]. Так, пі-дозру може викликати вхід до системи в незвичний час, особливо якщо після цього відбувається спроба отримати додаткові привілеї або доступ до нетипових ресурсів. Водночас не кожне відхилення означає наявність атаки: причиною можуть бути зміна робочого графіка, виконання нових завдань або інші обставини. Тому результати автоматизованого аналізу потрібно розглядати в контексті інших подій і перевіряти перед ухваленням рішення.

Прикладом практичного застосування таких можливостей є Microsoft Sentinel – рішення для моніторингу та захисту інформаційної інфраструктури. Воно дає змогу підключати джерела даних, налаштовувати правила виявлення загроз, досліджувати інциденти й автоматизувати окремі етапи реагування [3]. Ще одним напрямом розвитку засобів кібербезпеки є SOAR (Security Orchestration, Automation and Response). Такі технології доповнюють SIEM, до-помагаючи координувати дії у відповідь на виявлені події. Наприклад, після по-яви підозрілого сповіщення система може зібрати додаткові відомості, запустити визначену процедуру перевірки та повідомити відповідального спеціаліста. За відповідних налаштувань можна також автоматизувати блокування небезпечного з'єднання або обмеження доступу. Це дає змогу скоротити час на виконання типових операцій, однак автоматичні дії мають відповідати встановленим правилам безпеки та рівню ризику.

Попри широкі можливості, ефективність SIEM залежить від того, наскіль-ки якісно організовано роботу із системою. Якщо частина джерел не передає да-ні або журнали містять неповну інформацію, це може ускладнити виявлення ін-

цидентів. Важливе значення мають також налаштування правил аналізу: надто загальні правила можуть створювати багато хибнопозитивних сповіщень, тоді як недостатньо чутливі – пропускати підозрілі події. Крім того, зберігання й оброблення значних обсягів інформації потребує ресурсів, а самі журнали необхідно захищати від несанкціонованого доступу та змін. Рекомендації CISA щодо ведення журналів і моніторингу подій також наголошують на важливості використання таких даних для контролю безпеки інформаційних систем [4]. З огляду на це, упровадження SIEM має супроводжуватися регулярним переглядом правил виявлення, визначенням пріоритетності сповіщень і контролем якості даних.

Таким чином, SIEM-системи дають змогу об’єднувати інформацію з різ-них компонентів інформаційної інфраструктури та використовувати її для своєчаснішого виявлення можливих загроз. Їхня користь особливо помітна під час аналізу взаємопов’язаних подій, які складно оцінити окремо. Застосування поведінкової аналітики, машинного навчання й автоматизації реагування додатково розширює можливості таких рішень. Водночас результат залежить від якості зібраних даних, правильності налаштувань і професійної оцінки виявлених інцидентів. Тому SIEM доцільно розглядати як один із важливих складників комплексної системи забезпечення інформаційної безпеки.

Список використаних джерел:

1. Kent K., Souppaya M. Guide to Computer Security Log Management. *National Institute of Standards and Technology*. NIST Special Publication 800-92. 2006. URL: https://csrc.nist.gov/pubs/sp/800/92/final (дата звернення: 13.10.2026)

2. Microsoft. Identify advanced threats with User and Entity Behavior Analytics (UEBA) in Microsoft Sentinel. *Microsoft Learn.* URL: https://learn.microsoft.com/en-us/azure/sentinel/identify-threats-with-entity-behavior-analytics (дата звернення: 13.10.2026)

3. Microsoft. Microsoft Sentinel documentation. *Microsoft Learn.* URL: https://learn.microsoft.com/en-us/azure/sentinel/ (дата звернення: 13.10.2026)

4. Cybersecurity and Infrastructure Security Agency. Use Logging on Business Systems. *CISA.* URL: https://www.cisa.gov/audiences/small-and-medium-businesses/secure-your-business/use-logging-on-business-systems (дата звернення: 13.10.2026)

Буханець І.С.
студент групи ІКБм-21, ННІКБЗІ, ДУІКТ,
Київ, Україна

ПІДХІД ДО ВИЯВЛЕННЯ АНОМАЛІЙ ТА ЗАГРОЗ НУЛЬОВОГО ДНЯ В IoT-МЕРЕЖАХ НА ОСНОВІ ОДНОКЛАСОВОГО МАШИННОГО НАВЧАННЯ

Сьогодні спостерігається стрімка інтенсифікація прихованих кібератак та експлуатації загроз нульового дня периферійної інфраструктури Інтернету речей (IoT). Обмежені обчислювальні ресурси кінцевих вузлів (роутерів, датчиків) унеможливлюють розгортання традиційних сигнатурних систем захисту. Це дозволяє зловмисникам тривалий час використовувати скомпрометовані headless-пристрої як мережеві плацдарми, маскуючи деструктивну діяльність під штатний режим роботи.

Для розв'язання цієї проблеми запропоновано підхід до виявлення аномалій, що базується на одноклasовому машинному навчанні. Обґрунтовано технологію побудови еталонного профілю «нормальної» поведінки пристрою IoT на основі алгоритму Isolation Forest. Навчання моделі здійснюється виключно на легітимних даних (Benign) сучасного багаторівневого крос-архітектурного датасету CIC-YNU-IoTMal 2026. Розроблено п’ятиетапну схему технологічного процесу, яка забезпечує мінімізацію простору ознак за критерієм мінімальних витрат апаратних ресурсів через рекурсивне урізання (RFE) та фільтрацію ресурсомістких системних логів (зокрема, STRACE).

Ключові слова: кібербезпека, IoT-периферія, загрози нульового дня, виявлення аномалій, одноклasове машинне навчання, Isolation Forest, CIC-YNU-IoTMal 2026, мінімізація ознак.

Стрімка інтеграція технологій Інтернету речей (IoT) у промислові, медичні та споживчі сектори супроводжується критичним відставанням рівнів їхньої захищеності від темпів розгортання інфраструктури [1; 2]. Сутність загрози безпеці в IoT-мережах полягає у виникненні аномальних поведінкових станів та експлуатації загроз нульового дня (Zero-day), які через відсутність попередньо відомих цифрових відбитків (сигнатур) не можуть бути ідентифіковані класичними засобами захисту. Пристрої периферії (маршрутизатори, датчики, контролери) функціонують у концепції headless (без інтерфейсу користувача та локальних систем адміністрування), що дозволяє зловмисникам після успішної реалізації Zero-day атак місяцями використовувати скомпрометовані вузли як приховані плацдарми для шпигунства, збору облікових даних або координації DDoS-кампаній. При цьому пристрої візуально продовжують виконувати свої штатні функції, маскуючи деструктивну діяльність.

Проблема захисту таких систем полягає в гострому обмеженні обчислювальних ресурсів IoT-периферії (низька тактова частота CPU, критичний дефіцит RAM) унеможливлюють розгортання повнофункціональних систем виявлення вторгнень (IDS) або антивірусних агентів безпосередньо на кінцевих вузлах. Ситуація ускладнюється поширенням хакерських технік типу Living off the Land, коли для закріплення в системі використовуються легітимні утиліти вбудованих ОС (на базі ядер Linux/OpenWrt), що нівелює ефективність бінарних

ML-класифікаторів через відсутність чіткої межі між класами «Норма» та «Атака».

Для вирішення цієї проблеми необхідний перехід від суцільного сигнатурного контролю до інтелектуального аналізу поведінкових метрик на основі методів однокласового машинного навчання (One-Class Machine Learning). Оскільки легітимна активність периферійного IoT-пристрою є високоциклічною та передбачуваною, побудова математичного профілю його еталонної «нормальної» поведінки за допомогою алгоритму Isolation Forest дозволяє розв'язати проблему невідомих загроз. Будь-яка спроба експлуатації уразливості нульового дня неминуче викликає мікрозміни в системних (SAR) або мережевих (PCAP) метриках. Завдяки ітераційній мінімізації простору ознак до критичного мінімуму (5–8 ключових параметрів), стає можливим розгортання легких обчислювальних моделей безпосередньо в мікропрограмах обмежених у ресурсах пристроїв, забезпечуючи детекцію аномалій у реальному часі без перевантаження апаратної частини.

Ми використали набір даних CIC-YNU-IoTMal 2026 [3], який характеризується багаторівневістю (multimodal/multilayer structure) та крос-архітектурністю. Набір даних CIC-YNU-IoTMal створено за допомогою спеціалізованого сендбоксу CIC-YNU-IoTMal-Sandbox (на базі емуляторів QEMU під управлінням OpenWrt) одночасно на трьох рівнях:

мережевий рівень (PCAP), який містить дані сирого мережевого трафіку, згенерований пристроями (понад 2.4 мільйона семплів);

рівень операційної системи (SAR), який містить дані системних звітів про завантаження апаратної частини (CPU, оперативна пам'ять, швидкість читання/запису I/O), що збираються щосекунди (близько 1.8 мільйона семплів);

рівень ядра (STRACE), який містить дані глибоких логів системних викликів (system calls), які робить бінарний файл під час виконання (понад 105 мільйонів записів).

Набір даних CIC-YNU-IoTMal охоплює 10 000 реальних зразків шкідливого програмного забезпечення, запущених на чотирьох ключових архітектурах процесорів IoT: ARM, MIPS, MIPSEL та x86. Сповіщення про нормальну поведінку (Benign) були додатково згенеровані, зокрема із залученням великих мовних моделей (LLM) для імітації легітимної активності користувачів.

CIC-YNU-IoTMal містить дані представників найвідоміших родин IoT-threats: Mirai, Bashlite (Gafgyt), DarkNexus, Rudedevil, Tsunami, а також категорії Agent та Generic.

Розглянемо зміст технології побудови профілю «нормальної» поведінки IoT-пристрою на основі Isolation Forest.

Етап 1. Еталонний збір та попереднє очищення даних (Baseline Data Ingestion). На цьому етапі формується репрезентативний набір даних, який

відображає виключно нормальну (штатну) роботу пристрою в різних експлуатаційних режимах.

Здійснюємо екстракцію еталонного профілю з CIC-YNU-IoTMal застосовуючи фільтрацію набору даних за міткою класу для вилучення записів, що марковані як Benign (нормальна активність), та повне виключення шкідливих семплів (Malware). Проводимо ресурсно-орієнтовну селекцію метрик, під час якої відбираємо виключно «легкі» ознаки із категорій PCAP (наприклад, інтервали між пакетами – IAT, обсяг переданих байтів) та системних метрик SAR (завантаження CPU, RAM), які слабкий роутер спроможний фіксувати без зниження своєї продуктивності. Потім здійснюємо синхронізацію часових рядів, агрегуючі обрані метрики у фіксовані часові вікна (наприклад, інтервали по 1–5 секунд) для згладжування випадкових мережевих сплесків.

Етап 2. Статистичне моделювання та нормалізація ознак (Feature Engineering). На цьому етапі проводиться стабілізація ознакового простору перед подачею в ізоляційний ліс для уникнення спотворень через різний масштаб метрик. Видаляємо константні ознаки шляхом виключення метрик, які в штатному режимі мають нульову дисперсію (не змінюються), оскільки вони не несуть інформативного навантаження для побудови дерев. Здійснюємо масштабування даних шляхом застосування методів нормалізації, стійких до випадкових викидів (наприклад, RobustScaler на основі медіани та інтерквартильного розмаху), щоб зберегти математичну чіткість меж «нормальності».

Етап 3. Навчання моделі еталонного профілю (Isolation Forest Core Training). На цьому етапі здійснюємо побудову математичної структури («лісу»), яка описує щільність та взаємозв'язки нормальної поведінки. Конфігуруємо гіперпараметри та навчаємо ансамбль ізоляційних дерев на комп’ютері з обмеженням параметрів під майбутнє апаратне забезпечення (наприклад, n_estimators=50, max_samples='auto'). Визначаємо поріг забруднення (Contamination Parameter) шляхом встановлення параметра contamination на мінімальний рівень (наприклад, 0.001–0.01), що враховує ймовірність незначних технічних збоїв або шумів у навчальній вибірці, які не є атаками. Обчислюємо оцінку аномальності (Anomaly Score) шляхом розрахунку еталонного розподілу оцінок для нормальних даних, де значення близькі до 0 означають повну відповідність профілю, а значення ближче до 1 – потенційне відхилення.

Етап 4. Визначення критичного порогу детекції (Threshold Optimization). На цьому етапі встановлюємо межі, за якою поведінка пристрою вважається скомпрометованою. Проводимо валідацію на змішаній вибірці шляхом тестування навченого лісу на валідаційному наборі даних CIC-YNU-IoTMal, що містить як Benign (які не застосовувалися в моделі), так і Malware (активність ботнетів, сканування). Здійснюємо фіксацію математичного порогу τ шляхом розрахунку граничного значення Anomaly Score через пошук компромісу між

мінімізацією хибних спрацьовувань (False Positive Rate) та максимізацією виявлення прихованих загроз нульового дня (Zero-day).

Етап 5. Квантування моделі та розгортання на IoT-периферії (Deployment & Local Monitoring). На цьому етапі переносимо сформований еталонний профіль безпосередньо в мікропрограму пристрою для моніторингу в реальному часі. Експортуємо у C-код шляхом конвертації структури ухвалення рішень Isolation Forest (порогові значення розщеплення ознак у вузлах дерев) у компактні статичні масиви та структури *if-else* на мові *C*. Здійснюємо локальний скринінг шляхом запуску оптимізованої моделі на роутері (OpenWrt), яка в циклі реального часу приймає поточні 5–8 метрик пристрою, пропускає їх через згенеровані дерева й обчислює Anomaly Score. Якщо рахунок перевищує встановлений поріг τ, пристрій негайно сигналізує про компрометацію.

Оскільки легітимна активність периферійного IoT-пристрою є високоциклічною та передбачуваною, побудова математичного профілю його еталонної «нормальної» поведінки за допомогою алгоритму Isolation Forest дозволяє розв'язати проблему невідомих загроз. Будь-яка спроба експлуатації уразливості нульового дня неминуче викликає мікрозміни в системних (SAR) або мережевих (PCAP) метриках. Завдяки ітераційній мінімізації простору ознак до критичного мінімуму (5–8 ключових параметрів), стає можливим розгортання ультралегких обчислювальних моделей безпосередньо в мікропрограмах обмежених у ресурсах пристроїв, забезпечуючи детекцію аномалій у реальному часі без перевантаження апаратної частини.

Перелік використаної літератури

1. The 2026 IoT Security Landscape Report [Electronic resource] / Bitdefender, NETGEAR. – October 22, 2026. – 33 p. – URL: https://blogapp.bitdefender.com/hotforsecurity/content/files/2026/10/2026-_iot_report_consumer.pdf (дата звернення: 24.10.2026).

2. OT/IoT Fields Hot Summer Cyber Trends of 2026 [Electronic resource] / Nozomi Networks Labs. – October 12, 2026. – URL: https://www.nozominetworks.com/blog/ot-iot-fields-hot-summer-cyber-trends-of-2026 (дата звернення: 24.10.2026).

3. CIC-YNU-IoTMal: A comprehensive multilayer dataset for static and dynamic analysis of IoT malware behavior / S. Dadkhah, O. D. Okey, S. A. Maret et al. // Information Systems. – 2026. – Vol. 140. – 102722. – ISSN 0306-4379. – DOI: https://doi.org/10.1016/j.is.2026.102722.

Коровкін А.В.
студент групи ІКБм-21, ННІКБЗІ, ДУІКТ,
Київ, Україна

ПІДХІД ДО ВИЯВЛЕННЯ КІБЕРАТАК НА IoT-ПЕРИФЕРІЮ НА ОСНОВІ МІНІМІЗАЦІЇ ОЗНАК СИСТЕМНОГО МОНІТОРИНГУ МЕТОДАМИ МАШИННОГО НАВЧАННЯ

Сьогодні спостерігається стрімке зростання кількості та інтенсифікації кібератак на IoT-периферію. Обмежені обчислювальні ресурси кінцевих вузлів, таких як датчики та роутери, унеможливлюють розгортання традиційних систем виявлення вторгнень (IDS), що робить їх вразливими елементами мережі. Запропоновано науково-практичний підхід до виявлення кібератак на основі інтелектуального аналізу даних безпосередньо на межі мережі. Основна ідея полягає в оптимізації та зниженні розмірності простору ознак (Feature Selection) для зменшення апаратного навантаження на процесори пристроїв IoT.

Ключові слова: кібербезпека, кібератаки, виявлення кібератак, ознаки системного моніторингу, IoT-периферія, машинне навчання.

Прогнозується, що кількість пристроїв Інтернету речей (IoT) у світі зросте більш ніж удвічі – з 19,8 мільярда у 2025 році до понад 40,6 мільярда пристроїв IoT до 2034 року (рисунок 1). У 2034 році найбільша кількість пристроїв IoT буде знаходитися в Китаї – близько 7,51 мільярда споживчих пристроїв. Пристрої IoT використовуються в усіх галузях промисловості та на споживчих ринках, причому на споживчий сегмент припадатиме близько 60 відсотків усіх пристроїв IoT або підключених пристроїв у 2025 році. Прогнозується, що ця частка залишатиметься на цьому рівні протягом наступних десяти років [1].

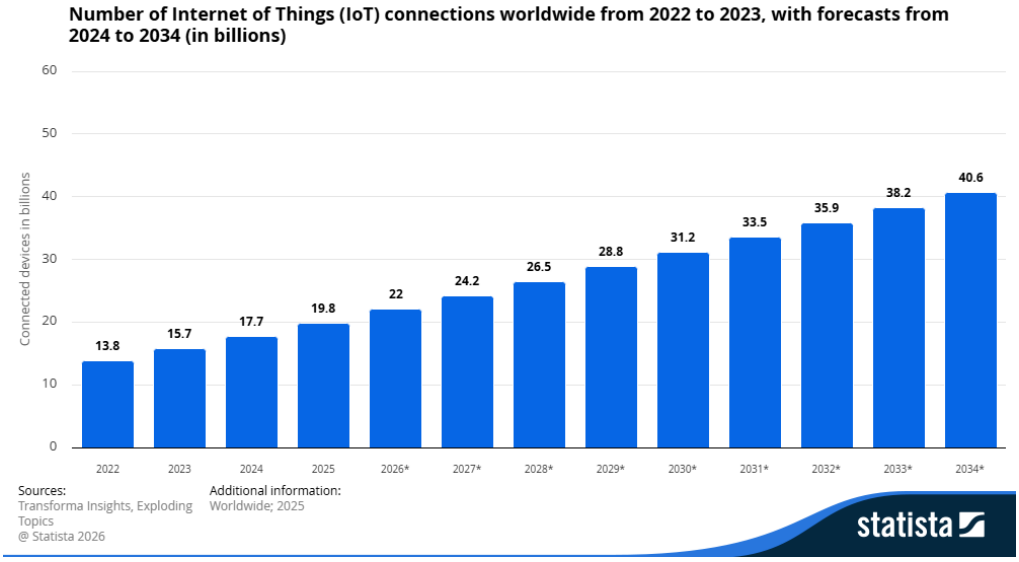


Рис. 1. Кількість під’єднаних пристроїв IoT у світі [1]

У Звіті SentinelOne [2] відмічається, що маршрутизатори будуть основними шлюзами для 75% кібератак, пов'язаних з Інтернетом речей. Згідно з даними на початок 2026 року, можна очікувати в середньому понад 820 000 атак Інтернету речей на день. Кількість пристроїв Інтернету медичних речей (IoMT) зросте і досягне 7,4 мільйона цього року. 83% пристроїв медичної візуалізації працюватимуть на непідтримуваних середовищах ОС. Глобальні атаки шкідливого програмного забезпечення для Інтернету речей вже зросли на 124%, що означає, що зараз тривають масштабні DDoS-кампанії [2].

Обмежені обчислювальні ресурси пристроїв IoT не дозволяють розгортати та застосовувати засоби активного моніторингу безпеки. Це призводить до того, що скомпрометований пристрій може функціонувати як актив зловмисника: маршрутизувати трафік, збирати облікові дані або брати участь в атаках ботнетів, на вигляд працюючи в штатному режимі. Збільшення кількості пристроїв IoT та кібератак на них, а також неможливість їх повноцінного захисту роблять ці пристрої одними з найбільш вразливих активів у будь-якому мережевому середовищі. Ефективним розв'язанням цієї проблеми є перехід від суцільного контролю до оптимізованого аналізу даних. Це обумовлює актуальність проблеми виявлення кібератак на IoT-периферію на основі мінімізації ознак системного моніторингу методами машинного навчання.

Зважаючи на такі жорсткі ресурсні обмеження та критичну вразливість інфраструктури, традиційні системи виявлення вторгнень (IDS) виявляються неефективними для захисту кінцевих вузлів. Це зумовлює гостру потребу в розробці альтернативних, легких архітектур безпеки, здатних функціонувати в умовах дефіциту обчислювальних ресурсів. Зниження розмірності простору ознак (Feature Selection) результатів системного моніторингу методами машинного навчання у цьому контексті дозволяє суттєво зменшити навантаження на процесори пристроїв IoT, зберігаючи при цьому високу точність та швидкість ідентифікації аномальної поведінки в реальному часі.

Ми використали набір даних CIC-YNU-IoTMal 2026 [3], який є комплексним багатошаровим набором даних для статичного та динамічного аналізу поведінки шкідливого програмного забезпечення Інтернету речей. CIC-YNU-IoTMal 2026 містить зібрані дані мережевого трафіку (PCAP), системних трас (STRACE) та системної активності (SAR) під час застосування різноманітних моделей поведінки шкідливого програмного забезпечення на архітектурах ARM, MIPS, MIPSEl та x86 [3].

Запропоновано технологічну схему процесу відбору критично важливого мінімуму метрик для розгортання на IoT-периферії, яку показано на рисунку 2. Розглянемо більш детально кожний етап цієї схеми.

Деталізація кожного етапу схеми

Етап 1: Фільтрація за критерієм накладних витрат. Оскільки збір кожної метрики споживає ресурси процесора (CPU) та пам'яті (RAM), на цьому етапі

відсікаються ознаки, на збирання та обробку яких витрачається більше обчислювальних ресурсів. Повністю виключаються глибокі логи STRACE (аналіз системних викликів у реальному часі на роутері занадто ресурсомісткий). Залишаються лише легкі агреговані метрики мережевих потоків (PCAP / NetFlow) та базові низькорівневі показники завантаження системи (SAR, наприклад, дельта CPU idle чи швидкість читання/запису I/O).

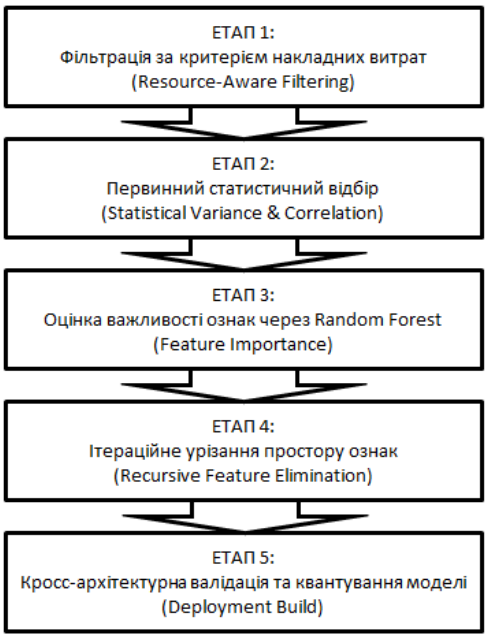


Рис. 2. Схема технологічного процесу відбору мінімуму метрик

Етап 2: Первинний статистичний відбір. На цьому етапі видаляється дані, які несуть інформаційний шуму та надлишковість:

видаляються ознаки із нульовою або майже нульовою дисперсією (метрики, які не змінюються під час атак і в режимі спокою);

виявляється мультиколінеарність даних. Якщо дві мережеві метрики сильно корелюють ($r > 0.9$), залишається лише одна – та, яку простіше та швидше обчислити процесору.

Етап 3: Оцінка важливості ознак через Random Forest. На цьому етапі застосовуємо базовий алгоритм на повному (але вже очищеному на Етапі 2) наборі даних для ранжування метрик. Навчається початкова модель Random Forest на сервері чи потужній машині. Розраховується показник Gini Importance (Mean Decrease in Impurity) або Permutation Importance для кожної ознаки. Формується рейтинг метрик від найбільш інформативних (наприклад, тривалість сесії, кількість байт у пакеті) до найменш корисних.

Етап 4: Ітераційне урізання (Recursive Feature Elimination, RFE). На цьому етапі здійснюється пошук точки балансу між точністю (Accuracy/F1-score) та кількістю ознак. Крок за кроком видаляються найслабші ознаки з рейтингу Етапу 3, і модель перенавчається. Процес зупиняється, коли розмірність простору ознак скорочується до критичного мінімуму (наприклад, від 5 до 8 ключових метрик), при якому F1-score моделі все ще залишається в межах припустимої норми (наприклад, >95%).

Етап 5: Крос-архітектурна валідація та квантування. На цьому етапі здійснюється адаптація та перевірка життєздатності моделі на реальному пристрої IoT. Оскільки датасет CIC-YNU-IoTMal кросс-архітектурний, перевіряється, чи відібрані 5-8 метрик однаково добре працюють як для ARM, так і для MIPS архітектур (запобігання перенавчанню під конкретне віртуальне середовище). Фінальна модель Random Forest експортується у вигляді масиву умовних операторів (*if-else*) мовою *C* безпосередньо у прошивку (OpenWrt/FreeRTOS). Дані з рухомою комою (*float*) конвертуються в цілі числа (*int*) для оптимізації математичних обчислень на датчиках без математичного співпроцесора (FPU).

Отже, зниження розмірності простору ознак (Feature Selection) за рахунок відсікання ресурсомістких метрик (наприклад, важких логів STRACE) є ключовою умовою для створення легких архітектур безпеки для IoT середовищ.

Використання моделі Random Forest на базі комплексного крос-архітектурного датасету CIC-YNU-IoTMal підтвердило свою доцільність. Механізм оцінки важливості ознак (Gini Importance) у поєднанні з методом рекурсивного урізання (RFE) дозволяє успішно ізолювати критичний мінімум із 5–8 найважливіших метрик.

Розроблена п'ятиетапна схема технологічного процесу забезпечує системний перехід від сирих даних до легкої моделі. Отриманий у результаті мінімум метрик дозволяє «урізати» (обмежити глибину та кількість дерев) і квантувати фінальну ML-модель у масив умовних операторів (*if-else*) мовою *C*.

Запропонований підхід уможливлює інтеграцію інтелектуальних засобів захисту безпосередньо у прошивки (OpenWrt, FreeRTOS) слабких домашніх роутерів, контролерів та датчиків. При цьому обчислювальне навантаження на процесори мінімізується, а точність ідентифікації аномалій та ботнет-активності в реальному часі зберігається на високому рівні (F1-score >95%).

Перелік використаної літератури

1. Number of Internet of Things (IoT) connections worldwide from 2022 to 2023, with forecasts from 2024 to 2034 (in billions) / Statista Research Department. Last update: May 26, 2025. URL: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/ (дата звернення: 26.10.2026).

2. Key Cyber Security Statistics for 2026 / SentinelOne. Updated Mar 12, 2026. URL: https://www.sentinelone.com/cybersecurity-101/cybersecurity/cyber-security-statistics/ (дата звернення: 26.10.2026).

3. CIC-YNU-IoTMal: A comprehensive multilayer dataset for static and dynamic analysis of IoT malware behavior / S. Dadkhah, O. D. Okey, S. A. Maret et al. // Information Systems. 2026. Vol. 140. 102722. ISSN 0306-4379. DOI: https://doi.org/10.1016/j.is.2026.102722.

Правдивий В.С.
студент групи ІКБм-21, ННІКБЗІ, ДУІКТ,
Київ, Україна

ПІДХІД ДО ВИЯВЛЕННЯ ШКІДЛИВОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ IoT НА ОСНОВІ МУЛЬТИМОДАЛЬНОГО АНАЛІЗУ ТЕЛЕМЕТРІЇ МЕТОДАМИ МАШИННОГО НАВЧАННЯ

Сьогодні спостерігається зростання кількості випадків створення та застосування шкідливого програмного забезпечення для Інтернету речей. Кібербезпека Інтернету речей є важливою, оскільки скомпрометовані підключені пристрої становлять не лише ризик для даних, вони створюють операційні та фізичні ризики. Пропонується підхід до виявлення шкідливого програмного забезпечення IoT на основі мультимодального аналізу телеметрії методами машинного навчання.

Ключові слова: кібербезпека Інтернету речей, виявлення шкідливого програмного забезпечення, ознаки системного моніторингу, машинне навчання.

Інтернет речей (IoT) – це мережа фізичних об’єктів ("речей"), обладнаних датчиками, програмним забезпеченням та іншими технологіями, необхідними для зв’язку й обміну даними з іншими пристроями в Інтернеті. Такими об’єктами є медичне обладнання, вбудовані системи, датчики, принтери або будь-які інтелектуальні побутові чи портативні пристрої. Ця мережа постійно розширюється. Так, у 2025 році у світі було активно 21,1 мільярда пристроїв Інтернету речей, що на 14% більше, ніж у попередньому році. Прогнозується, що до 2027 року це число досягне 29,7 мільярда – збільшення на 41% лише за два роки [1].

У Звіті DeXpose [2] відмічається, що у 2025 році кількість випадків застосування шкідливого програмного забезпечення для Інтернету речей зросла на 124% у порівнянні з минулим роком. Кількість пристроїв, залучених до ботнетів та кампаній зі шкідливим програмним забезпеченням, зросла більш ніж удвічі за один рік. Основною рушійною силою є вибухове зростання варіантів сімейства Mirai, на які припадало приблизно 40% спостережуваних корисних даних шкідливого програмного забезпечення для Інтернету речей [2].

На відміну від традиційного комп'ютерного злому, хакерство через Інтернет речей зазвичай використовує пристрої, які не мають екрана, інтерфейсу користувача та моніторингу безпеки, а це означає, що компрометації залишаються непоміченими протягом місяців, поки зловмисник використовує пристрій як постійний плацдарм у мережі організації [2]. Пристрої Інтернету речей дедалі частіше стають точкою входу для атак інформаційних систем та мереж організацій.

Необхідно відмітити, що пристрої Інтернету речей зазвичай розгортаються без активного моніторингу безпеки, а це означає, що скомпрометований пристрій може працювати як актив зловмисника, маршрутизуючи трафік, збираючи облікові дані або беручи участь в атаках ботнету, при цьому здаючись, що він функціонує нормально [2].

В [3] зазначається, що шкідливе програмне забезпечення продовжує

становити критичну загрозу безпеці екосистеми Інтернету речей, що зумовлено різноманітністю та динамікою мережевих середовищ. Ці умови створюють значні вразливості, що робить пристрої Інтернету речей основними цілями для складних атак із застосуванням шкідливих програм.

Отже, основними проблемами забезпечення безпеки пристроїв Інтернету речей є жорсткі обмеження щодо апаратних ресурсів (слабкі процесори, мінімальний обсяг оперативної пам'яті та енергоспоживання). Підходи, що базуються на сигнатурному аналізі (пошуку відомих фрагментів коду), виявилися неефективними в середовищі IoT через гетерогенність архітектур (пристрої Інтернету речей працюють на багатьох різних процесорних архітектурах (ARM, MIPS, PPC, x86), під кожну з яких шкідливий код компілюється окремо) та розвиток самих загроз (застосування обфускації коду, поліморфізму та атак нульового дня, які сигнатурні методи принципово не здатні розпізнати). У цих умовах виникає гостра потреба в переході від статичного аналізу коду до динамічного аналізу поведінки системи пристроїв Інтернету речей. Методи машинного навчання є найбільш перспективним напрямом розв'язанням цієї проблеми.

В роботі ми пропонуємо застосовувати підхід до виявлення шкідливого програмного забезпечення в середовищі IoT на основі мультимодального аналізу телеметрії методами машинного навчання. Ми провели дослідження на основі набору даних Канадського інституту кібербезпеки CIC-YNU-IoTMal Dataset 2026, який представляє собою комплексний багатошаровий набір даних для статичного та динамічного аналізу поведінки шкідливого програмного забезпечення Інтернету речей [3].

Набір даних CIC-YNU-IoTMal вирізняється трьома ключовими характеристиками: захоплення пакетів мережевого трафіку, що генеруються під час виконання шкідливого програмного забезпечення (PCAP); трасування системних викликів, зібрані за допомогою STRACE (STRACE); журнали системної активності (SAR), включаючи використання процесора та пам'яті під час виконання шкідливого програмного забезпечення на пристроях Інтернету речей. Динамічний аналіз зібраних даних проводився за допомогою віртуалізованої системи QEMU з прошивкою маршрутизатора OpenWRT на архітектурах ARM, MIPS, MIPSEL та x86 [3].

Сутність мультимодального машинного навчання полягає в комплексному аналізі зібраних даних щодо поведінки середовища IoT. Ми використали особливу відмінність набору даних CIC-YNU-IoTMal – комбінацію мережевих (PCAP) та системних параметрів (SAR/STRACE). Замість аналізу лише трафіку, модель вчиться бачити кореляцію. Наприклад, коли вірус починає DDoS-атаку (відбувається мережевий сплеск), модель відслідковує як одночасно змінюється навантаження на процесор та які системні виклики генеруються. Ми використали ансамблі моделей (Random Forest) із злиттям векторів ознак (Feature Fusion) на

етапі моделювання.

Розглянемо процес злиття мультимодальних даних (рисунок 1). Оскільки дані мають різну природу та часову дискретність (наприклад, мережеві пакети йдуть мілісекундами, а SAR фіксує стан системи щосекунди), головний технологічний виклик – вирівнювання гетерогенних даних за часом (Time-series Alignment).

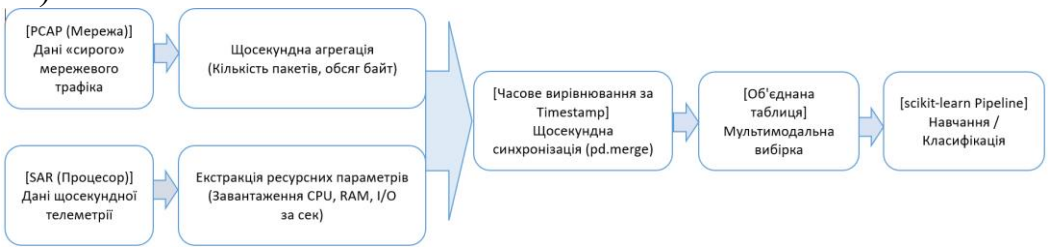


Рис. 1. Схема злиття мультимодальних даних (як способу їх подання)

Опис компонентів, які показано на рисунку 1.

1. Вхідні модальності:

джерело 1 (PCAP): потік неструктурованих асинхронних подій з високою дискретністю (мікросекунди);

джерело 2 (SAR): дискретні часові ряди стану ресурсів пристрою IoT, що знімаються з фіксованим кроком в одну секунду.

2. Блок обробки даних мережі (агрегація) здійснює віконне групування пакетів за інтервалами в одну секунду з розрахунком сумарних статистик (обсяг трафіку, інтенсивність потоку пакети/сек).

3. Блок часового вирівнювання (Time-series Alignment) виконує операцію реляційного об'єднання за допомогою індексованого ключа Timestamp. Рядки з відсутніми мережевими даними за певну секунду заповнюються нульовими значеннями.

4. Матриця ознак та Конвеєр (Pipeline): сформований мультимодальний набір даних подається на вхід scikit-learn Pipeline, де дані проходять масштабування через ColumnTransformer і передаються в ансамблевий класифікатор (RandomForest) для виявлення прихованих кореляцій між мережевими та обчислювальними аномаліями.

Таким чином, об'єднання різнорідних даних (трафіку та навантаження процесора/пам'яті) для виявлення прихованих кореляцій між мережевими та обчислювальними аномаліями в єдиний конвеєр забезпечить підвищення точності виявлення функціонування шкідливого програмного забезпечення в середовищах IoT.

Перелік використаної літератури

1. House N. IoT Security Statistics: Threats & Trends. StationX. October 2026. URL: https://app.stationx.net/articles/iot-security-statistics (дата звернення: 25.10.2026).

2. IoT Hacking Statistics 2026. Global Trends and Key Insights. DeXpose. February 10, 2026. URL: https://www.dexpose.io/iot-hacking-statistics/ (дата звернення: 25.10.2026).

3. CIC-YNU-IoTMal: A comprehensive multilayer dataset for static and dynamic analysis of IoT malware behavior / S. Dadkhah, O. D. Okey, S. A. Maret et al. Information Systems. 2026. Art. 102722.

Шевченко О.О.
студент групи ІКБм-22, ННІКБЗІ, ДУІКТ,
Київ, Україна

ТЕХНОЛОГІЯ ВИЯВЛЕННЯ НИЗЬКОІНТЕНСИВНИХ ТА ПРИХОВАНИХ КІБЕРАТАК НА IoT-ПРИСТРОЇ НА ОСНОВІ СТАТИСТИЧНОГО ПРОФІЛЮВАННЯ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ

На основі аналізу сучасного ландшафту загроз досліджено актуальну проблему виявлення низькоінтенсивних (low-rate) та прихованих кібератак на периферійні пристрої Інтернету речей (IoT). Сьогодні спостерігається тенденція до зміни стратегії зловмисників: замість масивних вторгнень пріоритет надається латентному закріпленню у мережевому периметрі (техніки Living off the Land, резидентні проксі, stealth-ботнети). Такі атаки навмисно маскуються під легітимний трафік, а дефіцит обчислювальних ресурсів кінцевих вузлів (роутерів, датчиків) унеможливлює застосування стандартних систем моніторингу, що дозволяє компрометаціям залишатися непоміченими протягом багатьох місяців.

Для розв'язання цієї проблеми запропоновано технологію автономного виявлення аномалій та загроз нульового дня (Zero-day), що базується на комбінації методів статистичного профілювання та однокласового машинного навчання (One-Class Machine Learning). Математичне моделювання еталонної «нормальної» поведінки пристрою реалізовано за допомогою алгоритму Isolation Forest. Навчання та валідація моделі здійснювалися на базі новітнього багаторівневого крос-архітектурного датасету CIC-YNU-IoTMal 2026, який містить повні профілі мережевої (PCAP) та системної (SAR) активності пристроїв під управлінням ОС OpenWrt.

Ключові слова: кібербезпека, IoT-периферія, приховані кібератаки, загрози нульового дня, статистичне профілювання, однокласове машинне навчання, Isolation Forest, CIC-YNU-IoTMal 2026, мінімізація ознак.

Стрімке масштабування інфраструктури Інтернету речей (IoT), кількість пристроїв у якій за прогнозами аналітиків Statista перевищить 40,6 мільярда одиниць до 2034 року, супроводжується радикальною зміною стратегій зловмисників. Сучасний ландшафт кіберзагроз характеризується переходом від деструктивних високоінтенсивних впливів до імплементації прихованих (stealthy) та низькоінтенсивних (low-rate) кібератак. Головною метою таких зловмисних дій є тривале латентне закріплення скомпрометованого периферійного пристрою у внутрішньому контурі мережі організації або домашньої інфраструктури [1, 2].

Відповідно до масштабного дослідження компаній Bitdefender та NETGEAR [2], опублікованого наприкінці 2026 року, концепція сприйняття IoT-заліза виключно як кінцевої цілі атаки повністю застаріла. Сучасні загрози трансформували вразливе обладнання периферії (маршрутизатори, смарт-телевізори, мережеві сховища) на постійні інфраструктурні плацдарми. Зокрема, фіксується масове приховане використання обчислювальних ресурсів та легітимної IP-репутації житлових домогосподарств операторами резидентних проксі-серверів і stealth-ботнетів (таких як Aisuru, Kimwolf, JackSkid або платформи BADBOX 2.0). Такі шкідливі модулі інтегруються безпосередньо у прошивки або легітимні SDK безкоштовних додатків, штучно підтримуючи низьку інтенсивність використання каналу зв'язку, що дозволяє їм залишатися

непоміченими системами безпеки протягом багатьох місяців [1, 2].

Актуальність виявлення низькоінтенсивних аномалій загострюється майже відсутністю апаратних ресурсів для захисту IoT-периферії. Жорсткі обмеження центральних процесорів (CPU) та критично малий обсяг оперативної пам'яті (RAM) кінцевих вузлів унеможливлюють встановлення повнофункціональних антивірусних агентів чи великовагових систем моніторингу трафіку. Оскільки низькоінтенсивні атаки навмисно розмиваються у часі та імітують легітимні транзакції, сигнатурні методи детекції виявляються повністю безсилими. Це зумовлює гостру науково-практичну потребу в розробці альтернативних підходів, які на основі статистичного профілювання та оптимізованих методів машинного навчання здатні ідентифікувати мікрозміни у поведінці пристроїв безпосередньо на межі мережі (edge), мінімізуючи апаратні накладні витрати.

Для забезпечення високої достовірності результатів дослідження та імітації реального середовища функціонування Інтернету речей, у роботі застосовано новітній багаторівневий датасет CIC-YNU-IoTMal 2026 [3]. Головною перевагою цього набору даних є його крос-архітектурна та мультимодальна структура, сформована шляхом моніторингу понад 10 000 реальних зразків IoT-мальварі (Mirai, Gafgyt, DarkNexus тощо) на чотирьох поширених процесорних архітектурах периферії: ARM, MIPS, MIPSEL та x86 [3].

Розглянемо етапи запропонованої технології виявлення низькоінтенсивних та прихованих кібератак на IoT-пристрої на основі статистичного профілювання та методів машинного навчання.

Етап 1. Ресурсно-оптимізоване статистичне профілювання (Resource-Aware Profiling). Зміст цього етапу полягає в здійсненні попередньої обробки багаторівневого набору даних CIC-YNU-IoTMal 2026 з метою виокремлення профілів легітимної поведінки (Benign). Для мінімізації обчислювальних та апаратних витрат виключаються ресурсоємні системні журнали STRACE. Початковий вектор ознак формується шляхом агрегації метрик мережевого трафіку (PCAP) та низькорівневих показників операційної системи OpenWrt (SAR) у фіксованих часових вікнах тривалістю 2–5 секунд. Такий підхід забезпечує узагальнення поведінкових характеристик пристрою та підвищує можливість виявлення слабко виражених часових відхилень, характерних для низькоінтенсивних атак.

Етап 2. Статистична стабілізація та нормалізація ознак (Feature Stabilization). На цьому етапі застосовується статистичний фільтр Variance Threshold для автоматичного видалення константних ознак із нульовою дисперсією. Відібрані параметри масштабуються за допомогою алгоритму RobustScaler, який використовує медіану та інтерквартильний розмах (IQR). Це забезпечує зменшення впливу окремих екстремальних значень на процес масштабування та сприяє формуванню статистично узгодженого простору ознак для подальшого моделювання легітимної поведінки IoT-пристроїв.

Етап 3. Математичне моделювання еталонного профілю (Isolation Forest Core Training). На цьому етапі реалізується концепція однокласового навчання (One-Class Learning), за якої модель навчається виключно на даних легітимної поведінки (Benign). На основі сформованих статистичних профілів будується ансамбль ізоляційних дерев алгоритму Isolation Forest. Модель визначає ступінь аномальності кожного спостереження на основі довжини шляху його ізоляції в деревах ансамблю, формуючи відповідний показник Anomaly Score. Спостереження, поведінкові характеристики яких істотно відрізняються від еталонного профілю, отримують вищу оцінку аномальності, що створює передумови для виявлення низькоінтенсивних і прихованих кібератак.

Етап 4. Крос-архітектурне калібрування порогу детекції (Threshold Optimization). На цьому етапі виконується експериментальне оцінювання моделі на тестових вибірках, що містять як легітимну активність, так і сценарії низькоінтенсивних атак, зокрема приховану активність ботнетів та несанкціоновані мережеві сесії. На основі аналізу розподілу Anomaly Score визначається оптимальний поріг детекції τ за критерієм мінімізації частоти хибнопозитивних спрацьовувань (False Positive Rate, FPR) із цільовим обмеженням $FPR<1\%$. Додатково проводиться експериментальна перевірка працездатності розробленого підходу на апаратних архітектурах ARM, MIPS, MIPSEL та x86 з метою оцінювання його кросплатформної сумісності та стабільності показників детекції.

Етап 5. Скорочення простору ознак та інтеграція в периферійне програмне забезпечення (Edge C-Code Generation). На цьому етапі за допомогою методу рекурсивного усунення ознак (Recursive Feature Elimination, RFE) здійснюється скорочення початкового простору параметрів до компактного набору з 5–8 найбільш інформативних метрик.

Розроблений модуль адаптується для інтеграції в програмні середовища OpenWrt/FreeRTOS, забезпечуючи можливість автономного виявлення аномальної активності безпосередньо на периферійних IoT-пристроях у режимі, наближеному до реального часу.

Таким чином, запропонована технологія уможливлює інтеграцію інтелектуальних модулів безпеки безпосередньо у прошивки обмежених у ресурсах периферійних пристроїв, гарантуючи автономну детекцію прихованих кібератак у реальному часі.

Перелік використаної літератури

1. Number of Internet of Things (IoT) connections worldwide from 2022 to 2023, with forecasts from 2024 to 2034 (in billions) [Electronic resource] / Statista Research Department. – May 26, 2025. – URL: statista.com (дата звернення: 30.10.2026).

2. The 2026 IoT Security Landscape Report [Electronic resource] / Bitdefender, NETGEAR. – October 22, 2026. – 33 p. – URL: bitdefender.com (дата звернення: 30.10.2026).

3. CIC-YNU-IoTMal: A comprehensive multilayer dataset for static and dynamic analysis of IoT malware behavior / S. Dadkhah, O. D. Okey, S. A. Maret et al. // Information Systems. – 2026. – Vol. 140. – 102722. – ISSN 0306-4379. – DOI: https://doi.org/10.1016/j.is.2026.102722.