Технології журналювання кіберінцидентів : сучасний стан і перспективи розвитку

Cyber Incident Logging Technologies: Current State and Future Directions

Святослав Василишин [A]

Corresponding author: доктор філософії з кібербезпеки, доцент кафедри захисту інформації, e-mail: Sviatoslav.I.Vasylyshyn@lpnu.ua, ORCID: 0000-0003-1944-2979

Ігор Власюк [A]

аспірант кафедри захисту інформації, e-mail: Ihor.D.Vlasiuk@lpnu.ua, ORCID: 0009-0008-3600-750X

Віталій Сусукайло [A]

доктор філософії з кібербезпеки, асистент кафедри захисту інформації, e-mail: vitalii.a.susukailo@lpnu.ua, ORCID: 0000-0003-4431-9964

Sviatoslav Vasylyshyn [A]

Corresponding author: PhD, Associate Professor, Department of Information Protection, e-mail: Sviatoslav.I.Vasylyshyn@lpnu.ua, ORCID: 0000-0003-1944-2979

Ihor Vlasiuk [A]

Postgraduate Student, Department of Information Protection, e-mail: Ihor.D.Vlasiuk@lpnu.ua, ORCID: 0009-0008-3600-750X

Vitalii Susukailo [A]

PhD, Assistant, Department of Information Protection, e-mail: vitalii.a.susukailo@lpnu.ua, ORCID: 0000-0003-4431-9964

[A] Національний університет "Львівська політехніка", м. Львів, Україна

[A] Lviv Polytechnic National University, Lviv, Ukraine

Received: June 14, 2025 | Revised: June 23, 2025 | Accepted: June 30, 2025

DOI: 10.33445/sds.2025.15.3.18

Мета роботи: аналіз проблем нормативного регулювання процедури журналювання кіберінцидентів в умовах України, огляді методів реалізації технологій аналізу журналів та способів оцінки їх ефективності, а також в аналізі основних напрямів застосування можливостей штучного інтелекту у вказаних цілях.

Метод дослідження: системно-порівняльний аналіз.

Результати дослідження: Окреслено та систематизовано основні проблеми з якими зустрічаються академічні дослідники, розробники програмного забезпечення та системні адміністратори при дослідженні проблеми ефективності автоматизації процесів журналювання кіберінцидентів, порівняльній оцінці функціональності і ефективності їх існуючих реалізацій, впровадження моделей штучного інтелекту у вказані процеси та виділено основні шляхи подолання вказаних проблем.

Теоретична цінність дослідження: Дослідження дозволяє на основі аналізу актуальних нормативних документів, останніх наукових і корпоративних публікацій та змісту репозитаріїв програмного забезпечення з відкритим кодом провести теоретичне узагальнення актуального рівня існуючих та перспективних технологій журналювання кіберінцидентів.

Практична цінність дослідження: Результати дослідження дозволяють фахівцям в галузі кібербезпеки систематизувати критерії оцінки в процесі вибору, розробки, апробації, системної інтеграції та аудиту засобів журналювання кіберінцидентів.

Цінність дослідження: В національному науковому просторі практично відсутні узагальнюючі публікації по проблемі комплексного аналізу стану і перспектив технологій журналювання кіберінцидентів, які б розглядали проблему в контексті українських реалій.

Обмеження дослідження: Для реалізації більшості окреслених підходів до вдосконалення та інтелектуалізації процедур журналювання кіберінцидентів треба враховувати потребу в значних фінансових і обчислювальних ресурсах.

Тип статті: оглядово-аналітичний.

Purpose: to analyze and evaluate the challenges of regulatory frameworks governing cyber incident logging procedures in the context of Ukraine, methods and approaches to implementing log analysis technologies, techniques for assessing their effectiveness, and the primary directions for applying artificial intelligence capabilities in these processes.

Method: comparative analysis.

Findings: The study outlines and systematizes the main challenges faced by academic researchers, software developers, and system administrators in investigating the efficiency of automating cyber incident logging processes, comparing the functionality and effectiveness of existing implementations, and integrating artificial intelligence models into these processes. It also identifies key approaches to overcoming these challenges.

Theoretical implications: The research provides a theoretical generalization of the current state and prospects of cyber incident logging technologies based on an analysis of relevant regulatory documents, recent scientific and corporate publications, and the content of open-source software repositories.

Practical implications: The findings enable cybersecurity professionals to systematize evaluation criteria for selecting, developing, testing, integrating, and auditing cyber incident logging tools.

Value: In the national academic landscape, there is a notable lack of comprehensive publications addressing the holistic analysis of the state and prospects of cyber incident logging technologies, particularly in the context of Ukrainian realities.

Future research: Implementing most of the proposed approaches to improving cyber incident logging procedures requires significant financial and computational resource.

Papertype: Review and analytical.

Ключові слова: кіберінциденти, технології журналювання, порівняння ефективності, машинне навчання, великі мовні моделі.

Key words: cyber incidents, logging technologies, efficiency comparison, machine learning, large language models.

Вступ

Аналіз вітчизняних наукових публікацій вказує за рідкісним винятками на практичну відсутність досліджень присвячених технологіям журналювання кіберінцидентів (ЖК).

Разом з тим цей важливий аспект кібербезпеки переживає різкий сплеск наукового інтересу в зв'язку з тими можливостями, які відкрив вибуховий розвиток технологій штучного інтелекту (ШІ) за останні 5-10 років. В світі спостерігається швидкий ріст досліджень спрямованих на підвищення функціональності і ефективності технологій журналювання кіберінцидентів на основі можливостей ШІ.

Дане дослідження має на меті дати відповіді на наступні питання, які виникають перед дослідниками та розробниками систем ЖК чи адміністраторами систем кіберзахисту, що мають вибрати оптимальне рішення з урахуванням специфіки діяльності і національних вимог до даного роду програмно-апаратного забезпечення:

які характеристики має мати застосунок, щоб відповідати вимогам міжнародних і національних стандартів та рекомендацій щодо ЖК;

які методи і технології використовуються в даний час в системах ЖК в програмній індустрії кібербезпеки та дослідницьких проєктах;

які методи використовуються для порівняльної оцінки ефективності існуючих та перспективних систем аналізу ЖК;

які програми чи програмні комплекси з функціями автоматичного аналізу журнальних даних представлені в репозитаріях з відкритим програмним кодом;

які функції ЖК можуть бути оптимізовані шляхом впровадження технологій ШІ, які методи для цього використовуються і за якими критеріями можна оцінити їх ефективність.

Теоретичні основи дослідження

Законодавство України визначає кіберінцидент наступним чином: **інцидент кібербезпеки** (далі – кіберінцидент) – подія або ряд несприятливих подій ненавмисного характеру (природного, технічного, технологічного, помилкового, у тому числі внаслідок дії людського фактора) та/або таких, що мають ознаки можливої (потенційної) кібератаки, які становлять загрозу безпеці систем електронних комунікацій, систем управління технологічними процесами, створюють імовірність порушення штатного режиму функціонування таких систем (у тому числі зриву та/або блокування роботи системи, та/або несанкціонованого управління її ресурсами), ставлять під загрозу безпеку (захищеність) електронних інформаційних ресурсів (*Про основні засади забезпечення кібербезпеки України, 2017).*

Журналювання кіберінцидентів – це постійний процес запису подій, пов'язаних із безпекою комп'ютерної системи, які відбуваються в IT-інфраструктурі організації. Процеси ЖК мають виключне значення у виявленні факту інциденту, аналізі його масштабу і небезпечності, в необхідному реагуванні та відновленні функціональності системи після інциденту. Правильно налаштовані процедури ЖК є не лише інструментом виявлення інциденту, але і засобом розуміння повного обсягу інциденту, проведення післяінцидентного аналізу та виявлення слабких місць в організації системи комп'ютерної безпеки (ACSC, 2024). Процедури журналювання є важливим компонентом реагування на кіберінциденти і регламентуються методичними рекомендаціями Державної служби спеціального зв'язку та захисту інформації України від 03.07.2023 (АДССЗІУ, 2023).

В більшості випадків ЖК є обов'язковим компонентом систем управління інформацією та подіями безпеки (SIEM – Security Information & Event Management). Першопочатково оригінальні платформи SIEM були саме інструментами керування журналами. Ці платформи забезпечували моніторинг та аналіз подій, пов'язаних з безпекою, у режимі реального часу.

Сам термін SIEM був запропонований в 2005 році консалтинговою компанією Gartner, Inc. і з того часу фактично перетворився в галузевий стандарт (Mokalled et al., 2019)

Надалі розвиток SIEM йшов шляхом поглибленої автоматичної обробки потоку журнальних даних, включаючи аналітику поведінки користувачів та об'єктів (UEBA – user and entity behavior analytics). Останні роки провідною тенденцією став пошук ефективних технік аналізу журнальних даних в цілях кібербезпеки на основі технологій машинного навчання (МН) (Опірський та ін., 2022; Beck et al., 2025; Landauer et al., 2023).

Постановка проблеми

Журналювання кіберінцидентів важливий і обов'язковий компонент будь якої системи забезпечення кібербезпеки, який має відповідати багатьом формальним і прикладним критеріям. Проте технології реалізації ЖК характеризуються високою різноманітністю і швидким прогресом, що вимагає постійного фокусування на даній проблемі цілої низки фахівців – дослідників, розробників і експлуатантів систем кібербезпеки, з врахуванням контексту їх застосування. Аналіз поточної ситуації в даній галузі є предметом даного дослідження.

Результати

Регламентація процесів ЖК

В сучасних умовах процедури журналювання кіберінцидентів стали об'єктом правового і рекомендаційного регулювання регіонального та галузевого характеру. Так Міжнародний стандарт управління інформаційною безпекою (СУІБ) – ISO/IEC 27001:2022 вимагає впровадження процедур для моніторингу, журналювання та аудиту подій інформаційної безпеки. Зокрема, розділ A.12.4 стосується журналювання подій, аналізу журналів та захисту журнальних даних від несанкціонованого доступу. (International Organization for Standardization, 2022, Clause A.12.4). Керівні принципи щодо процесу управління інцидентами інформаційної безпеки наводяться в стандарті ISO/IEC 27035 – 1:2023 (International Organization for Standardization, 2023).

Вимоги до журналювання кіберінцидентів регламентуються також національним і регіональним законодавством. В США основоположним документом з управління журналами безпеки комп'ютерів є посібник NIST SP 800-92 виданий Національним інститутом стандартів і технологій США (Kent & Souppaya, 2006). В Європейському співтоваристві ключовим документом є NIS2 Directive (Directive (EU) 2022/2555) – директива ЄС щодо безпеки мереж та інформації (European Parliament and Council of the European Union, 2022). Дана директива вимагає від операторів критичної інфраструктури та постачальників цифрових послуг впроваджувати заходи для виявлення та повідомлення про кіберінциденти. Такі організації повинні мати системи для фіксації інцидентів і передачі інформації регуляторним органам.

Існує також низка галузевих стандартів регіонального і національного масштабу, які регламентують процедури ЖК на об'єктах критичної інфраструктури, в фінансовій, медичній та інших чутливих сферах. Прикладом такого регламенту може бути PCI DSS (Payment Card Industry Data Security Standard (PCI DSS) — стандарт безпеки даних індустрії платіжних карток, розроблений Радою зі стандартів безпеки індустрії платіжних карток (Payment Card Industry Security Standards Council, 2022). В медичній галузі зразком такої регламентації є HIPAA (Health Insurance Portability and Accountability Act) закон США, що регламентує вимоги до захисту особистої інформації медичного характеру (HIPAA, 1996). Хоча фактично ця регламентація може мати характер національної чи корпоративної, внаслідок процесів глобалізації інформаційної мережі вона фактично стає міжнародною.

В світовому масштабі спостерігається тісна координація зусиль між національними урядовими центрами і установами кібербезпеки. щодо базових принципів процедур ЖК. Як

приклад можна навести рекомендаційну публікацію Австралійського центру кібербезпеки (ACSC) випущену у співпраці з урядовими установами з питань кібербезпеки восьми провідних країн. (Australian Cyber Security Centre, 2024).

В Україні ЖК регламентується серед іншого наступними документами різного статусного рівня:

Закон України "Про основні засади забезпечення кібербезпеки України" в редакції від 27.03.2025 (*Про основні засади забезпечення кібербезпеки України, 2017)* ;

Наказ Держспецзв'язку № 570 від 03.07.2023 "Про затвердження Методичних рекомендацій щодо реагування суб'єктами забезпечення кібербезпеки на різні види подій у кіберпросторі" (АДССЗІУ , 2023);

Наказ Держспецзв'язку та СБУ № 627/772 від 19.12.2024 "Деякі питання розробки, затвердження та погодження планів захисту об'єктів критичної інфраструктури за проєктною загрозою національного рівня "кібератака кіберінцидент" (АДССЗІУ та СБУ, 2024);

Нормативний документ Держспецзв'язку НД ТЗІ 3.6-006-24 (АДССЗІУ, 2023).

Ці документи визначають таксономію кіберінцидентів та порядок ведення репозитарію інформації про них, створення та зберігання журналів, рекомендації щодо реагування на різні види кіберінцидентів, встановлюють вимоги до кіберзахисту об'єктів критичної інфраструктури, журналювання подій для виявлення та аналізу інцидентів, визначають порядок вибору заходів захисту інформації, включаючи вимоги до журналювання подій безпеки. Є очевидним, що розробники і експлуатанти систем кібербезпеки повинні враховувати вимоги регулятивних документів в процесі і розробляти процедури аудиту на відповідність нормативним вимогам.

Класифікація видів ЖК

Процедури ЖК охоплюють самі різноманітні процеси на різних рівнях інформаційної інфраструктури організації і можуть мати на меті різні аспекти кібербезпеки. Для більш чіткого розуміння функціональної різноманітності ЖК варто розглянути підходи до класифікації процедур ЖК.

Класифікація видів і методів ЖК відбувається переважно за двома основними критеріями:

в залежності від архітектури генерації і збереження даних;

в залежності від видів джерел даних, завдань та технологій.

Згідно першого критерію розрізняють централізоване, розподілене та гібридне журналювання, які суттєво відрізняються за такими характеристиками як безпека, стійкість, масштабованість, складність, керованість, аналіз даних, вартість підтримки.

За другим критерієм виокремлюють наступні види журналювання:

системних подій;

подій комп'ютерної мережі;

апаратної активності кінцевих точок;

активності користувачів;

подій в хмарних середовищах;

безпеки системних застосунків;

кореляцій подій;

метаданних;

для відповідності стандартам безпеки.

На нашу думку дана класифікація може слугувати орієнтиром при розробці систем ЖК організацій з метою оцінки меж необхідної функціональності та порівняння з функціоналом існуючих пропрієтарних систем ЖК та їх аналогів з відкритим кодом. З наведеного вище слідує, що з огляду на масштаби масивів даних та різноплановість завдань процесу журналювання, система ЖК вимагає спрямування значних зусиль на виконання досить рутинних за змістом,

проте масштабних процедур. В цих умовах існує гостра потреба в автоматизації моніторингу та аналізу різних аспектів системи ЖК.

Методи аналізу журналів

Зростання масштабів і складності сучасних програмних систем, в тому числі систем інформаційного захисту, призводить до різкого збільшення обсягу журналів системних подій. Це веде до експоненційного зростання трудомісткості аналітичної роботи, а традиційний спосіб аналізу журналу експертами в ручному режимі стає практично неприйнятним.

Останнім часом академічні дослідження, а також процеси розробки програмного інструментарію почали орієнтуватися на методи текстового пошуку та аналітичних рішень на основі процедур машинного навчання (Landauer et al., 2023). Проте через неструктуровану природу журналів першочерговою проблемою постає оцінка записів журналу на можливість структурування даних для наступної автоматизації процедур аналізу (Ma et al., 2023). Цей процес прийнято визначати як процес нормалізації журналів. За визначенням NIST Glossary нормалізація журналів полягає в конвертації кожного поля даних журналу в певне представлення та їхній послідовній категоризації (Kent & Souppaya, 2006, p. A1). Це означає, що сирі дані журналів, які можуть мати різну структуру (наприклад, різні формати часу, ідентифікатори чи рівні деталізації), трансформуються в уніфікований вигляд, придатний для подальшого аналізу. Автоматизація цього процесу, забезпечує вищу швидкість і точність аналізу, особливо в умовах великих обсягів даних, що генеруються сучасними IT-системами.

Зрозуміло, що існує значний інтерес до підвищення ефективності цього завдання. Тому процеси автоматизованого парсингу і нормалізації журналів подій в комп'ютерних системах інтенсивно досліджуються як академічними науковцями, так і в програмній індустрії. В результаті була створена низка синтаксичних аналізаторів журналів, які побудовані на основі використання різних методів аналізу. Вказані системи мають різне цільове спрямування і охоплюють розподілені системи, суперкомп'ютери, операційні системи, мобільні системи, серверні програми та автономне програмне забезпечення.

Серед алгоритимічних методів аналізу журналів можна відмітити наступні:

SLCT (Simple Logfile Clustering Tool) та LogCluster – розширення на його основі, які використовують алгоритм кластеризації для наборів даних файлу журналу, що допомагає виявляти часті шаблони файлів журналу, створювати їх профілі та виділяти аномальний рядок файлу журналу. Репозиторії коду даного методу розміщені за адресою (https://ristov.github.io/slct/) та (https://ristov.github.io/logcluster/).

IPLoM (Interative Partitioning Log Mining) – алгоритм для видобутку кластерів із журналів подій, який шляхом ієрархічного поділу розподіляє дані журналів по кластерам та створює описи кластерів або формати рядків для кожного зі створених кластерів. Репозиторії коду методу розміщені за адресою (https://github.com/fluency03/iplom-java).

Spell (Streaming Parser for Event Logs), який використовує підхід на основі найдовшої загальної підпослідовності (LCS – longest common subsequence) для онлайн-потокового аналізу системних журналів подій. Репозиторії коду методу розміщені за адресою (https://github.com/nbigaouette/spell-rs).

Дані методи аналізу реалізовані у численних програмних застосунках, спрямованих на аналіз журналів, як у вигляді окремих інструментів, так і в складі багатофункціональних комплексних систем кіберзахисту. Абсолютна більшість таких інструментів має комерційний характер і відповідно недоступні для академічного дослідження. Поряд з цим існує низка програмних продуктів з відкритим кодом, які можуть слугувати для ознайомлення з методами реалізації технологій ЖК. Перевагою цього виду програм є їх абсолютна або умовна безкоштовність, оскільки багато академічних досліджень розвивається без серйозної фінансової підтримки.

Зокрема надзвичайно просунутою підсистемою журналювання володіє OSSEC (Open Source HIDS Security) – система з відкритим кодом для виявлення вторгнень на основі хосту (HIDS – host-based intrusion detection system) (https://www.ossec.net/). OSSEC здатна збирати та аналізувати журнали з таких сімейств операційних систем, як Linux, Windows, macOS та Unix-подібні системи. OSSEC підтримує аналіз журналів від багатьох популярних програмних продуктів та служб, які генерують журнали у сумісному з OSSEC форматі. Дана система здатна обробляти дані журналів вебсерверів і їх служб, мережевих і хмарних сервісів, баз даних і програм безпеки. Система OSSEC підтримує такі формати журналів як Syslog, Windows Event Logs, JSON та інші. Вона також має вбудовані правила для парсингу журналів таких популярних застосунків як Apache, MySQL, Cisco та інші. OSSEC також дозволяє створювати власні правила для парсингу журналів, згенерованих специфічними програмами або пристроями (Teixeira, et al., 2019). Підтримка широкого спектру систем та програм робить OSSEC універсальним інструментом для дослідження проблем кібербезпеки, зокрема аналізу журналів і розробки власних концептуальних моделей журналювання кіберінцидентів. Репозиторії коду системи розміщені за адресою (https://github.com/ossec/ossec-hids).

Ще одним прикладом системи виявлення вторгнень (HIDS) з відкритим кодом, є Wazuh (https://wazuh.com/). Проєкт розпочинався як відгалуження системи OSSEC, проте швидко набув популярності і отримав багато удосконалень, які надають цьому проєкт у суттєві переваги над оригінальною системою. Ці переваги стосуються таких сфер як:

масштабованість та надійність;
управління встановленням та конфігурацією;
виявлення вторгнень;
інтеграція з хмарними провайдерами;
відповідність нормативним вимогам;
функція реагування на інциденти;
виявлення вразливостей та оцінка конфігурації.

Стосовно процесу журналювання даних в системі Wazuh серед вдосконалень можна відмітити покращений механізм аналізу журналів з вбудованим декодуванням JSON та можливістю динамічного найменування полів, десятикратне збільшення максимального розміру повідомлення журналу, а також оновлений набір правил аналізу журналів та декодери повідомлень. (Stanković, Gajin, & Petrović, 2022). На даний час розглянутий проєкт розвивається більш інтенсивно в порівняння з вихідною системою OSSEC. Репозиторії коду системи розміщені за адресою (https://github.com/orgs/wazuh/repositories?type=all).

В розглянутих вище системах функція журналювання є лише складовою універсальної потужної HIDS і тісно пов’язана з іншими функціональними можливостями. Разом з тим серед програм з відкритим кодом існують спеціалізовані рішення спрямованні саме на процеси журналювання, маючи на меті удосконалити зручність і ефективність процесу керування всім комплексом процесів журналювання. Серед таких систем в першу чергу зазначимо проєкт Fluentd (https://www.fluentd.org/). З самого початку проєкт був налаштований на збір та обробку напівструктурованих або неструктурованих великих наборів даних. Поява такого роду програм є відображенням такого факту, що якщо на початках комп’ютерної ери процес журналювання слугував меті аналізу збоїв і пошуку помилок, що дозволяв розробникам здійснювати процес налагодження (debugging) програмних засобів, то в 21 столітті основним споживачем даних журналів стали вже машини, а не люди. Саме аналітичні програмні комплекси, в тому числі кібербезпекового спрямування стали основним споживачем журнальних даних самого різного походження з інфраструктури журналювання, яка не була першопочатково “машинно-орієнтована”. Тому Fluentd активно просуває і реалізує концепцію уніфікованого рівня ведення журналу (unified logging layer), яка дозволяє розширити і полегшити аналіз журналів (Tamura, 2014). Серед переваг Fluentd виокремлюють наступні:

низькі вимоги до системних ресурсів;

використання уніфікованого формату журналювання;

зручна архітектура;

можливість інтеграції із різними мовами програмування.

Слід відмітити, що наявність відкритого коду робить ці проєкти особливо цінними для завдань наукового дослідження, оскільки дозволяє оцінити і порівняти різні концепції автоматизації процесу збору і аналізу журнальних даних. Репозиторії коду системи розміщені за адресою (https://github.com/fluent/fluentd).

Оцінка ефективності процесів журналювання кіберінцидентів

Наявність різних реалізацій систем збору та аналізу журнальних даних неминуче ставить проблему порівняння їх ефективності в якості інструментів ЖК. Порівняння різних аспектів ефективності систем ЖК вимагає наявності спеціальних методів тестування та тестових шкал для такої оцінки. Для цієї мети науковим співтовариством була запропонована низка наборів журнальних даних, які представляють дані з журналів подій операційних систем, застосунків та дані про поведінку користувачів. Значна частина серед цих наборів спеціально налаштовані на перевірку здатності HIDS виявляти аномальну активність зловмисного характеру шляхом аналізу журналів.

Серед таких наборів, які широко використовуються дослідниками кібербезпеки, відмітимо “Набір даних для оцінки виявлення вторгнень” (CIC-IDS2017) Канадського інституту кібербезпеки (CIC) (Sharafaldin, Lashkari & Ghorbani, 2018), «Набір даних для виявлення вторгнень на базі хоста» (ADFA IDS) Університету Південного Уельсу в Канберрі (Creech, 2014), набір даних журналів Loghub 2.0 (Zhu et al., 2023), набір даних BETH Dataset (Highnam, Arulkumaran, Hanif & Jennings, 2021), набір журнальних даних LO2 (Bakhtin, 2025). У випадках, коли реальні дані важко отримати або вони чутливі, створюються синтетичні набори даних, які імітують поведінку системи та атаки. Також для перевірки ефективності систем аналізу ЖК використовуються **емулятори вторгнень** з відкритим кодом такі як Atomic Red Team, MITRE Caldera, Metasploit Framework (Landauer, Mayer, Skopik, Wurzenberger & Kern, 2024).

Останнім часом набори журнальних даних використовуються не лише з метою порівняльного тестування різних технологій ЖК, але і виступають в якості бази для машинного навчання компонентів HIDS. Проте останні публікації вказують на ризики релевантності такого підходу (Dube, 2024; Cantone, Marocco, & Bria, 2024).

Перспективи автоматизації ЖК методами штучного інтелекту

Очевидні успіхи систем штучного інтелекту, привели до сплеску досліджень з метою застосування цих досягнень до автоматизації процесів ЖК. Ці підходи ґрунтуються як на використанні можливостей великих лінгвістичних моделей (LLM) універсального призначення (Li et al., 2023; Beck, Landauer, Wurzenberger, Skopik & Rauber 2025), так і на вузькоспеціалізованих системах машинного навчання (ML) (Skopik, Landauer & Wurzenberger, 2021; Landauer, Onder, Skopik, & Wurzenberger, 2023).

Застосування технологій штучного інтелекту в цілях ЖК, за незначним виключенням є полем діяльності комерційних організацій, які із зрозумілих причин не схильні ділитись подробицями своїх пропрієтарних інтелектуальних технологій ЖК. Відповідно практично єдиним джерелом такої інформації є матеріали сайтів компаній та корпоративних блогів. Результати мета-аналізу змісту ресурсів корпоративних сайтів компаній, що пропонують інструменти ЖК з використанням технологій штучного інтелекту, представлені в Таблиці 1.

Таблиця 1 – Основні напрямки впровадження технологій штучного інтелекту в системи журналювання кіберінцидентів

Інтелектуальна функція	Назва продукту
Виявлення аномалій у трафіку мережі та прогнозування потенційних атак	Cisco Secure Network Analytics, Darktrace AI, Palo Alto Networks Cortex XDR.
Аналіз поведінки користувачів (UEBA) для виявлення відхилень від типових патернів	Securonix, Exabeam Fusion SIEM, CrowdStrike Falcon.
Виявлення кореляцій подій для створення патернів атак і автоматичного реагування	CrowdStrike Falcon, Rapid7 InsightIDR, Carbon Black.
Аналіз хмарних журналів для виявлення аномалій і автоматизації реагування	Sumo Logic Cloud SIEM, AWS CloudTrail, Azure Sentinel.
Виявлення патернів атак і автоматична класифікація загроз	Splunk, LogRhythm, Palo Alto Networks Prisma Cloud.
Автоматизація реагування на основі ШІ-аналітики та прогнозування	Rapid7 InsightIDR, Darktrace Antigena, Palo Alto Networks Cortex XSOAR.
Формування запитів і створення звітів природною мовою	Microsoft Sentinel , CrowdStrike Falcon,
Аналіз інцидентів та генерація рекомендацій	CrowdStrike Falcon, Security Operations – SecOps, Splunk

Зрозуміло, що перелічені функції істотно полегшують аналіз журналів подій з метою виявлення кіберінцидентів, але можуть істотно розрізнятися технологіями реалізації. В більшості випадків такі технології є пропрієтарними і не доступними для дослідження з науковими і академічними цілями. Тому дослідження підвищення ефективності журналювання кіберінцидентів на основі машинного навчання може спиратись лише на бази даних і технології навчання з відкритим кодом і ліцензією, що дозволяє без обмежень використовувати технології та моделі з науковими цілями (Габрильчук та ін., 2025).

Оскільки дані журналів подій можна розглядати як текст зі специфічною семантикою то виглядає перспективним використання технології трансформерів, взявши за основу принципи і архітектуру мовної моделі BERT (Bidirectional encoder representations from transformers) (Devlin, Chang, Lee & Toutanova, 2019). Дана технологія машинного навчання дозволяє представляти текст як послідовність векторів за допомогою самостійного навчання і виявилась досить гнучкою, універсальною та ефективною для реалізації задач в самих різних предметних галузях, в тому числі для потреб ЖК (Chen & Liao, 2022; Guo, Yuan & Wu, 2021; Tang & Guan, 2024; Chai et al., 2024). Достатньо актуальний огляд проблем використання машинного навчання (ML) з метою ЖК представлений Landauer et al., (2023).

Інший підхід щодо впровадження інтелектуальних функцій в системи ЖК полягає в використанні можливостей великих мовних моделей (Large Language Model – LLM). В нещодавньому великому огляді таких досліджень було ретельно проаналізовано 29 публікацій по даній проблемі (Beck et al.,2025). Підбірка включала як вже опубліковані дослідження, так і препринти. Автори роблять висновок, що даний підхід є життєздатним і має ряд переваг, які полягають в адаптивності до різних форматів журналів та їхній здатності узагальнювати дані на невидимих шаблонах. Разом з тим відмічені такі проблеми, як високі затрати обчислювальних ресурсів, схильність до “галюцинацій” та проблеми з інтерпретацією. Було показано, що парсери журналів на основі LLM часто зменшують потребу в ручному налаштуванні та маркуванні, виявлено підходи, які здатні суттєво підвищити ефективність та точність цих методів. Однак, було виявлено, що лише два з семи парсерів на основі LLM, явно перевершують парсери, що не базуються на LLM.

Оскільки більшість наукових проєктів обмежені за ресурсами і фінансуванням, то при розробці концептуальної моделі ЖК на основі підходів штучного інтелекту і етапів її реалізації, слід враховувати необхідність оптимізації процесу машинного навчання. В недавніх дослідженнях спрямованих на оцінку ефективності машинного навчання було запропоновано емпіричні закони масштабування, які оцінюють, як зміни в моделі та розмірі навчальних даних впливають на якість моделі. Було встановлено, що для найефективнішого масштабування мовних моделей, параметри та токени повинні зростати приблизно лінійно. Автори застосували цей закон масштабування для навчання моделі з 70 мільярдами параметрів під назвою Chinchilla, яка перевершила набагато більші та дорожчі моделі, такі як GPT-3 (Hoffmann et al., 2022). В результаті багато наступних LLM були навчені з врахуванням законів масштабування встановлених дослідниками на LLM Chinchilla (Sardana, Portes, Doubov & Frankle, 2023).

Висновки

Процеси журналювання кіберінцидентів (ЖК), які є обов'язковим компонентом систем кіберзахисту різного масштабу та відіграють особливе значення у виявленні факту кіберінциденту, аналізі та подоланні його наслідків, підлягають регламентації обов'язкового і рекомендаційного характеру на відповідність стандартам міжнародного і національного рівня. Така регламентація динамічно змінюється з врахуванням ландшафту технологій кіберзагроз і кіберзахисту, що вимагає постійного моніторингу і аудиту програмних засобів систем ЖК на відповідність регулятивним нормам.

Експоненційний ріст даних, що вимагають журналювання та аналізу в контексті забезпечення кібербезпеки, потребує пошуку технологій підвищення ефективності цих процесів та впровадження методів їх релевантної оцінки при порівнянні різних реалізацій технологій ЖК.

Недоступність для академічних досліджень комерційних пропрієтарних технологій змушує наукову спільноту приділяти особливу увагу технологіям з відкритим програмним кодом, надаючи особливу вагу можливостям їх використання для наукових завдань.

Магістральним напрямом розвитку технологій ЖК на даному етапі є впровадження методів штучного інтелекту, що вимагає ретельного аналізу балансу переваг і ризиків в порівнянні з традиційними алгоритмічними підходами.

Ресурсна затратність методів машинного навчання гостро ставить проблему оптимізації стратегій впровадження ефективних технологій штучного інтелекту в системні процеси журналювання кіберінцидентів.

Фінансування

Це дослідження не отримало конкретної фінансової підтримки.

Конкуруючі інтереси

Автори заявляють, що у них немає конкуруючих інтересів.

Список використаних джерел

1. Адміністрація Державної служби спеціального зв'язку та захисту інформації України. (2024). *Нормативний документ НД ТЗІ 3.6-006-24*. URL : https://cip.gov.ua/services/cm/api/attachment/download?id=66109
2. Адміністрація Державної служби спеціального зв'язку та захисту інформації України та Служба безпеки України. (2024). *Деякі питання розробки, затвердження та погодження планів захисту об'єктів критичної інфраструктури за проєкт ною загрозою національного рівня «кібератака кіберінцидент» (Наказ № 627/772 від 19*

грудня 2024 року). URL : https://cip.gov.ua/ua/news/spilnii-nakaz-sluzhbi-bezpeki-ukrayini-ta-administraciyi-derzhspeczv-yazku-vid-19-grudnya-2024-roku-627-772-deyaki-pitannya-rozrobki-zatverdzhennya-ta-pogodzhennya-planiv-zakhistu-ob-yektiv-kritichnoyi-infrastrukturi-za-proektnoyu-zagrozoyu-nacionalnogo-rivnya-kiberataka-kiberincident

3. Адміністрація Державної служби спеціального зв'язку та захисту інформації України (2023) Методичні рекомендації щодо реагування суб'єктами забезпечення кібербезпеки на різні види подій у кіберпросторі. Наказ Адміністрації Державної служби спеціального зв'язку та захисту інформації України № 570 від 03.07.2023 URL : https://zakon.rada.gov.ua/rada/show/v0570519-23#top
4. Габрильчук А. В., Сусукайло В. А., Курій Є. О., Василишин С. І. (2025). Дослідження кібератак з використанням машинного навчання на системи управління інформаційною безпекою. *Комп'ютерні системи та мережі*, *7(1),* 68-78. https://doi.org/10.23939/csn2025.01.068
5. Опірський І.Р., Сусукайло В., Василишин, С. (2022). Дослідження можливостей використання чатботів зі штучним інтелектом для дослідження журналів подій. *Захист інформації, 24(4),* 177-183. https://doi.org/10.18372/2410-7840.24.17380
6. *Про основні засади забезпечення кібербезпеки України (2017). Закон України № 2163-VIII від 05.10.2017 (зі змінами).* URL: https://zakon.rada.gov.ua/laws/show/2163-19#Text (Дата звернення: 23.05.2025)
7. Australian Cyber Security Centre (2024). *Best practices for event logging and threat detection.* Australian Signals Directorate. URL : https://www.cyber.gov.au/sites/default/files/2024-08/best-practices-for-event-logging-and-threat-detection.pdf
8. Bakhtin, A., Nyyssölä, J., Wang, Y., Ahmad, N., Ping, K., Esposito, M., ... & Taibi, D. (2025). LO2: Microservice API Anomaly Dataset of Logs and Metrics. *arXiv preprint arXiv:2504.12067*. https://doi.org/10.48550/arXiv.2504.12067
9. Beck, V., Landauer, M., Wurzenberger, M., Skopik, F., & Rauber, A. (2025). System Log Parsing with Large Language Models: A Review. *arXiv preprint arXiv:2504.04877*. https://doi.org/10.48550/arXiv.2504.04877
10. Cantone, M., Marocco, C., & Bria, A. (2024) Generalization Challenges in Network Intrusion Detection: A Study on CIC-IDS2017 and CSE-CIC-IDS2018 Datasets. In *1st INTERNATIONAL PhD SYMPOSIUM ON ENGINEERING AND SPORT SCIENCE* (p. 185).
11. Chai, X., Zhang, H., Zhang, J., Sun, Y., & Das, S. K. (2024). Log Sequence Anomaly Detection based on Template and Parameter Parsing via BERT. *IEEE Transactions on Dependable and Secure Computing*. https://doi.org/10.1109/TDSC.2024.3428538
12. Chen, S., & Liao, H. (2022). Bert-log: Anomaly detection for system logs based on pre-trained language model. *Applied Artificial Intelligence*, *36*(1), 2145642. https://doi.org/10.1080/08839514.2022.2145642
13. Creech, G. (2014). *Developing a high-accuracy cross platform host-based intrusion detection system capable of reliably detecting zero-day attacks* (Doctoral dissertation, UNSW Sydney). https://doi.org/10.26190/unsworks/16615
14. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186). https://doi.org/10.18653/v1/N19-1423
15. Dube, R. (2024). Faulty use of the cic-ids 2017 dataset in information security research. *Journal of Computer Virology and Hacking Techniques*, *20*(1), 203-211. http://dx.doi.org/10.1007/s11416-023-00509-7
16. European Parliament and Council of the European Union. (2022). *Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and*

Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive). URL : https://eur-lex.europa.eu/eli/dir/2022/2555/oj

17. Guo, H., Yuan, S., & Wu, X. (2021, July). Logbert: Log anomaly detection via bert. In *2021 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE. https://doi.org/10.48550/arXiv.2103.04475
18. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996). URL : https://www.hhs.gov/hipaa/index.html
19. Highnam, K., Arulkumaran, K., Hanif, Z., & Jennings, N. R. (2021). Beth dataset: Real cybersecurity data for unsupervised anomaly detection research. In *CEUR Workshop Proc* (Vol. 3095, pp. 1-12).
20. Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., ... & Sifre, L. (2022). Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*. https://doi.org/10.48550/arXiv.2203.15556
21. International Organization for Standardization. (2022). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO/IEC 27001:2022). URL : https://www.iso.org/standard/27001
22. International Organization for Standardization. (2023). Information technology — Information security incident management — Part 1: Principles and process (ISO/IEC 27035-1:2023). URL : https://www.iso.org/standard/78973.html
23. Kent, K., & Souppaya, M. (2006). *Guide to computer security log management* (NIST Special Publication 800-92). National Institute of Standards and Technology. URL : https://csrc.nist.gov/publications/detail/sp/800-92/final
24. Landauer, M., Mayer, K., Skopik, F., Wurzenberger, M., & Kern, M. (2024, December). Red team redemption: A structured comparison of open-source tools for adversary emulation. In *2024 IEEE 23rd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (pp. 117-128). IEEE. https://doi.org/10.48550/arXiv.2408.15645
25. Landauer, M., Onder, S., Skopik, F., & Wurzenberger, M. (2023). Deep learning for anomaly detection in log data: A survey. *Machine Learning with Applications*, *12*, 100470. https://doi.org/10.1016/j.mlwa.2023.100470
26. Li, Y., Huo, Y., Jiang, Z., Zhong, R., He, P., Su, Y., ... & Lyu, M. R. (2023). Exploring the effectiveness of llms in automated logging generation: An empirical study. *arXiv preprint arXiv:2307.05950*. https://doi.org/10.48550/arXiv.2307.05950
27. Ma, J., Liu, Y., Wan, H., & Sun, G. (2023). Automatic parsing and utilization of system log features in log analysis: A survey. *Applied Sciences*, *13*(8), 4930. http://dx.doi.org/10.3390/app13084930
28. Mokalled, H., Catelli, R., Casola, V., Debertol, D., Meda, E., & Zunino, R. (2019, June). The applicability of a siem solution: Requirements and evaluation. In *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 132-137). IEEE. DOI: 10.1109/WETICE.2019.00036 http://dx.doi.org/10.1109/WETICE.2019.00036
29. Payment Card Industry Security Standards Council (2022). *Payment Card Industry Data Security Standard (PCI DSS) version 4.0*. URL : https://www.pcisecuritystandards.org/document_library?category=pcidss&document=pci_dss
30. Sardana, N., Portes, J., Doubov, S., & Frankle, J. (2023). Beyond chinchilla-optimal: Accounting for inference in language model scaling laws. *arXiv preprint arXiv:2401.00448*. https://doi.org/10.48550/arXiv.2401.00448
31. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSp*, *1*(2018), 108-116. http://dx.doi.org/10.5220/0006639801080116

32. Skopik, F., Landauer, M., & Wurzenberger, M. (2021). Online log data analysis with efficient machine learning: A review. *IEEE Security & Privacy*, *20*(3), 80-90. https://doi.org/10.1109/MSEC.2021.3113275
33. Stanković, S., Gajin, S., & Petrović, R. (2022). A Review of Wazuh tool capabilities for detecting attacks based on log analysis. *No Nama Agent Integrity File Added Delete Modified*, *1*. URL : https://www.etran.rs/2022/zbornik/ICETRAN-22_radovi/068-RTI2.6.pdf
34. Tamura, K. (2014, August 6). *Unified logging layer: Turning data into action*. Fluentd. URL : https://www.fluentd.org/blog/unified-logging-layer
35. Tang, P., & Guan, Y. (2024). Log anomaly detection based on BERT. *Signal, Image and Video Processing*, *18*(8), 6431-6441. http://dx.doi.org/10.1007/s11760-024-03327-6
36. Teixeira, D., Assunção, L., Pereira, T., Malta, S., & Pinto, P. (2019). OSSEC IDS extension to improve log analysis and override false positive or negative detections. *Journal of Sensor and Actuator Networks*, *8*(3), 46. https://doi.org/10.3390/jsan8030046
37. Zhu, J., He, S., He, P., Liu, J., & Lyu, M. R. (2023, October). Loghub: A large collection of system log datasets for ai-driven log analytics. In *2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE)* (pp. 355-366). IEEE. https://doi.org/10.48550/arXiv.2008.06448

References

1. Administration of the State Service for Special Communications and Information Protection of Ukraine. (2024). Normative document of the State Service for Special Communications and Information Protection of Ukraine. *3.6-006-24*. Available from : https://cip.gov.ua/services/cm/api/attachment/download?id=66109
2. Administration of the State Service for Special Communications and Information Protection of Ukraine and the Security Service of Ukraine. (2024). Some issues of development, approval and coordination of plans for the protection of critical infrastructure facilities under the projected national-level threat “cyber attack, cyber incident” (Order No. 627/772 of December 19, 2024). Available from : https://cip.gov.ua/ua/news/spilnii-nakaz-sluzhbi-bezpeki-ukrayini-ta-administraciyi-derzhspeczv-yazku-vid-19-grudnya-2024-roku-627-772-deyaki-pitannya-rozrobki-zatverdzhennya-ta-pogodzhennya-planiv-zakhistu-ob-yektiv-kritichnoyi-infrastrukturi-za-proektnoyu-zagrozoyu-nacionalnogo-rivnya-kiberataka-kiberincident
3. Administration of the State Service for Special Communications and Information Protection of Ukraine (2023) Methodological recommendations on the response of cybersecurity entities to various types of events in cyberspace. Order of the Administration of the State Service for Special Communications and Information Protection of Ukraine No. 570 dated 03.07.2023 Available from : https://zakon.rada.gov.ua/rada/show/v0570519-23#top
4. Gabrylchuk A. V., Susukaylo V. A., Kuriy E. O., Vasylyshyn S. I. (2025). Research on cyberattacks using machine learning on information security management systems. Computer Systems and Networks,*7(1),* 68-78. https://doi.org/10.23939/csn2025.01.068
5. Opirsky I.R., Susukaylo V., Vasylyshyn, S. (2022). Research into the possibilities of using chatbots with artificial intelligence for researching event logs. Information Protection*, 24(4),* 177-183. https://doi.org/10.18372/2410-7840.24.17380
6. *On the Basic Principles of Ensuring Cybersecurity in Ukraine (2017). Law of Ukraine No. 2163-VIII of 05.10.2017 (as amended).* Available from : https://zakon.rada.gov.ua/laws/show/2163-19#Text (Дата звернення: 23.05.2025)
7. Australian Cyber Security Centre (2024). *Best practices for event logging and threat detection.* Australian Signals Directorate. Available from : https://www.cyber.gov.au/sites/default/files/2024-08/best-practices-for-event-logging-and-threat-detection.pdf

8. Bakhtin, A., Nyyssölä, J., Wang, Y., Ahmad, N., Ping, K., Esposito, M., ... & Taibi, D. (2025). LO2: Microservice API Anomaly Dataset of Logs and Metrics. *arXiv preprint arXiv:2504.12067*. https://doi.org/10.48550/arXiv.2504.12067
9. Beck, V., Landauer, M., Wurzenberger, M., Skopik, F., & Rauber, A. (2025). System Log Parsing with Large Language Models: A Review. *arXiv preprint arXiv:2504.04877*. https://doi.org/10.48550/arXiv.2504.04877
10. Cantone, M., Marocco, C., & Bria, A. (2024) Generalization Challenges in Network Intrusion Detection: A Study on CIC-IDS2017 and CSE-CIC-IDS2018 Datasets. In *1st INTERNATIONAL PhD SYMPOSIUM ON ENGINEERING AND SPORT SCIENCE* (p. 185).
11. Chai, X., Zhang, H., Zhang, J., Sun, Y., & Das, S. K. (2024). Log Sequence Anomaly Detection based on Template and Parameter Parsing via BERT. *IEEE Transactions on Dependable and Secure Computing*. https://doi.org/10.1109/TDSC.2024.3428538
12. Chen, S., & Liao, H. (2022). Bert-log: Anomaly detection for system logs based on pre-trained language model. *Applied Artificial Intelligence*, *36*(1), 2145642. https://doi.org/10.1080/08839514.2022.2145642
13. Creech, G. (2014). *Developing a high-accuracy cross platform host-based intrusion detection system capable of reliably detecting zero-day attacks* (Doctoral dissertation, UNSW Sydney). https://doi.org/10.26190/unsworks/16615
14. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186). https://doi.org/10.18653/v1/N19-1423
15. Dube, R. (2024). Faulty use of the cic-ids 2017 dataset in information security research. *Journal of Computer Virology and Hacking Techniques*, *20*(1), 203-211. http://dx.doi.org/10.1007/s11416-023-00509-7
16. European Parliament and Council of the European Union. (2022). *Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)*. Available from : https://eur-lex.europa.eu/eli/dir/2022/2555/oj
17. Guo, H., Yuan, S., & Wu, X. (2021, July). Logbert: Log anomaly detection via bert. In *2021 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE. https://doi.org/10.48550/arXiv.2103.04475
18. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996). URL : https://www.hhs.gov/hipaa/index.html
19. Highnam, K., Arulkumaran, K., Hanif, Z., & Jennings, N. R. (2021). Beth dataset: Real cybersecurity data for unsupervised anomaly detection research. In *CEUR Workshop Proc* (Vol. 3095, pp. 1-12).
20. Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., ... & Sifre, L. (2022). Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*. https://doi.org/10.48550/arXiv.2203.15556
21. International Organization for Standardization. (2022). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO/IEC 27001:2022). Available from : https://www.iso.org/standard/27001
22. International Organization for Standardization. (2023). Information technology — Information security incident management — Part 1: Principles and process (ISO/IEC 27035-1:2023). Available from : https://www.iso.org/standard/78973.html
23. Kent, K., & Souppaya, M. (2006). *Guide to computer security log management* (NIST Special Publication 800-92). National Institute of Standards and Technology. Available from : https://csrc.nist.gov/publications/detail/sp/800-92/final

24. Landauer, M., Mayer, K., Skopik, F., Wurzenberger, M., & Kern, M. (2024, December). Red team redemption: A structured comparison of open-source tools for adversary emulation. In *2024 IEEE 23rd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (pp. 117-128). IEEE. https://doi.org/10.48550/arXiv.2408.15645
25. Landauer, M., Onder, S., Skopik, F., & Wurzenberger, M. (2023). Deep learning for anomaly detection in log data: A survey. *Machine Learning with Applications*, *12*, 100470. https://doi.org/10.1016/j.mlwa.2023.100470
26. Li, Y., Huo, Y., Jiang, Z., Zhong, R., He, P., Su, Y., ... & Lyu, M. R. (2023). Exploring the effectiveness of llms in automated logging generation: An empirical study. *arXiv preprint arXiv:2307.05950*. https://doi.org/10.48550/arXiv.2307.05950
27. Ma, J., Liu, Y., Wan, H., & Sun, G. (2023). Automatic parsing and utilization of system log features in log analysis: A survey. *Applied Sciences*, *13*(8), 4930. http://dx.doi.org/10.3390/app13084930
28. Mokalled, H., Catelli, R., Casola, V., Debertol, D., Meda, E., & Zunino, R. (2019, June). The applicability of a siem solution: Requirements and evaluation. In *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 132-137). IEEE. DOI: 10.1109/WETICE.2019.00036 http://dx.doi.org/10.1109/WETICE.2019.00036
29. Payment Card Industry Security Standards Council (2022). *Payment Card Industry Data Security Standard (PCI DSS) version 4.0*. Available from : https://www.pcisecuritystandards.org/document_library?category=pcidss&document=pci_dss
30. Sardana, N., Portes, J., Doubov, S., & Frankle, J. (2023). Beyond chinchilla-optimal: Accounting for inference in language model scaling laws. *arXiv preprint arXiv:2401.00448*. https://doi.org/10.48550/arXiv.2401.00448
31. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSp*, *1*(2018), 108-116. http://dx.doi.org/10.5220/0006639801080116
32. Skopik, F., Landauer, M., & Wurzenberger, M. (2021). Online log data analysis with efficient machine learning: A review. *IEEE Security & Privacy*, *20*(3), 80-90. https://doi.org/10.1109/MSEC.2021.3113275
33. Stanković, S., Gajin, S., & Petrović, R. (2022). A Review of Wazuh tool capabilities for detecting attacks based on log analysis. *No Nama Agent Integrity File Added Delete Modified*, *1*. Available from : https://www.etran.rs/2022/zbornik/ICETRAN-22_radovi/068-RTI2.6.pdf
34. Tamura, K. (2014, August 6). *Unified logging layer: Turning data into action*. Fluentd. Available from : https://www.fluentd.org/blog/unified-logging-layer
35. Tang, P., & Guan, Y. (2024). Log anomaly detection based on BERT. *Signal, Image and Video Processing*, *18*(8), 6431-6441. http://dx.doi.org/10.1007/s11760-024-03327-6
36. Teixeira, D., Assunção, L., Pereira, T., Malta, S., & Pinto, P. (2019). OSSEC IDS extension to improve log analysis and override false positive or negative detections. *Journal of Sensor and Actuator Networks*, *8*(3), 46. https://doi.org/10.3390/jsan8030046
37. Zhu, J., He, S., He, P., Liu, J., & Lyu, M. R. (2023, October). Loghub: A large collection of system log datasets for ai-driven log analytics. In *2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE)* (pp. 355-366). IEEE. https://doi.org/10.48550/arXiv.2008.06448