КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (33), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.33.1276
УДК 004.056 (045)

Вавіленкова Анастасія Ігорівна
доктор технічних наук, професор, завідувач кафедри кібербезпеки
Національна академія Служби безпеки України, м. Київ, Україна
ORCID: 0000-0002-9630-4951
vavilenkovaa@gmail.com

ВЕЛИКІ МОВНІ МОДЕЛІ ДЛЯ ОПТИМІЗАЦІЇ ЗАДАЧ У СФЕРІ ПОШУКУ ТА ЗАХИСТУ ІНФОРМАЦІЇ

Анотація. Стрімкий розвиток цифрових технологій, глобальна інформатизація суспільства та безперервне зростання обсягів даних створюють нові виклики у сфері пошуку, оброблення та захисту інформації. Щоденно інформаційні системи генерують терабайти даних, серед яких містяться як корисні відомості, так і потенційні ознаки кіберзагроз, аномалій та атак. Традиційні підходи до аналізу інформації дедалі частіше виявляються недостатньо ефективними через обмеженість можливостей обробки неструктурованих даних та необхідність залучення значних людських ресурсів.

Поява великих мовних моделей стала одним із найважливіших досягнень сучасного штучного інтелекту. Завдяки здатності аналізувати природну мову, розуміти контекст, узагальнювати великі масиви інформації та генерувати змістовні відповіді, LLM відкривають нові можливості для автоматизації процесів інформаційного пошуку та забезпечення інформаційної безпеки. Сучасні дослідження демонструють, що великі мовні моделі можуть ефективно застосовуватися для виявлення вразливостей, аналізу шкідливого програмного забезпечення, розслідування кіберінцидентів, обробки журналів подій безпеки та автоматизації діяльності центрів операційної безпеки.

У статті досліджено можливості використання великих мовних моделей для оптимізації процесів пошуку, аналізу, оброблення та захисту інформації в сучасному цифровому середовищі. Розглянуто особливості функціонування великих мовних моделей, принципи їх побудови та механізми застосування в інформаційно-комунікаційних системах. Проаналізовано основні напрями використання LLM у сфері інформаційного пошуку, автоматизованого аналізу текстових даних, виявлення інформаційних загроз, моніторингу кіберінцидентів, пошуку вразливостей програмного забезпечення та підтримки процесів прийняття рішень у сфері кібербезпеки. Особливу увагу приділено інтеграції великих мовних моделей із системами управління подіями інформаційної безпеки, системами виявлення вторгнень, технологіями розвідки кіберзагроз та механізмами Retrieval-Augmented Generation. Визначено переваги та обмеження використання LLM у задачах пошуку та захисту інформації. Обґрунтовано перспективи подальшого розвитку інтелектуальних систем кіберзахисту на основі великих мовних моделей.

Ключові слова: великі мовні моделі (LLM); технології штучного інтелекту; кібербезпека; пошук інформації; кіберінцидент; кіберзагроза; захист інформації.

ВСТУП

Сучасний етап розвитку інформаційних технологій характеризується експоненційним зростанням обсягів даних, що генеруються та обробляються у цифровому середовищі. Традиційні методи обробки текстової інформації, що базуються на статистичних підходах та ключових словах, демонструють обмежену ефективність в умовах семантичної неоднозначності природної мови та постійно еволюціонуючих кіберзагроз.

Особливої актуальності набуває використання LLM у зв'язку зі зростанням складності сучасних кіберзагроз, появою атак на основі штучного інтелекту та необхідністю оперативного реагування на інциденти безпеки. Використання великих мовних моделей дозволяє значно підвищити швидкість аналізу інформації, скоротити час реагування на кіберінциденти та підвищити якість прийняття управлінських рішень у сфері кібербезпеки. Революційним проривом у вирішенні цих проблем стала поява великих мовних моделей (Large Language Models, LLM) – нейромережевих архітектур, здатних розуміти контекст, генерувати текст та виконувати складні завдання обробки природної мови.

Постановка проблеми. Сучасні інформаційні системи функціонують в умовах постійного зростання обсягів даних, складності інформаційних потоків та підвищення інтенсивності кіберзагроз. Традиційні системи інформаційного пошуку та засоби захисту інформації характеризуються обмеженими можливостями щодо обробки великих обсягів неструктурованих даних, виявлення складних взаємозв'язків між подіями та оперативного реагування на нові кіберзагрози. У таких умовах традиційні методи пошуку інформації та захисту інформаційних ресурсів стикаються з низкою проблем: низька ефективність аналізу великих обсягів неструктурованої інформації; складність виявлення прихованих взаємозв'язків між подіями безпеки; значні часові витрати на аналіз журналів подій та повідомлень про інциденти; недостатня автоматизація процесів кіберзахисту; потреба у висококваліфікованих спеціалістах для аналізу кіберзагроз; швидка еволюція методів проведення кібератак.

Зростання кількості кібератак, збільшення обсягів журналів подій, а також необхідність швидкого прийняття рішень потребують використання нових інтелектуальних підходів. Одним із перспективних напрямів розв'язання зазначених проблем є застосування великих мовних моделей, які здатні виконувати семантичний аналіз текстової інформації, автоматизувати процеси пошуку та класифікації даних, виявляти аномалії та забезпечувати підтримку фахівців у сфері кібербезпеки. Проте існує необхідність дослідження особливостей інтеграції LLM у системи пошуку та захисту інформації, а також оцінювання ефективності їх використання.

У зв'язку з цим виникає необхідність розроблення нових інтелектуальних підходів, здатних автоматизувати процеси пошуку інформації, виявлення загроз та підтримки прийняття рішень у сфері інформаційної безпеки. Одним із найбільш перспективних підходів є використання великих мовних моделей.

Аналіз останніх досліджень і публікацій. Проблемам застосування великих мовних моделей у кібербезпеці присвячено значну кількість сучасних досліджень. Так, у роботі [1] Н. Джаффал, М. Альханафсеха та Д. Мохайсена проведено комплексний аналіз використання LLM у задачах виявлення загроз, оцінювання вразливостей та реагування на кіберінциденти. Автори відзначають, що великі мовні моделі забезпечують більш глибоке контекстне розуміння інформації порівняно з традиційними методами машинного навчання.

Систематичний огляд, виконаний J. Zhang та співавторами [2], демонструє зростання кількості наукових досліджень, присвячених інтеграції LLM у системи інформаційної безпеки. Автори зазначають, що використання генеративного штучного інтелекту сприяє автоматизації процесів аналізу журналів подій, класифікації загроз і підтримки прийняття рішень.

Питання безпеки самих великих мовних моделей розглядаються у роботі [3], де наведено класифікацію основних загроз, пов'язаних із функціонуванням LLM, включаючи атаки на етапі навчання, атаки через маніпулювання підказками та ризики автономних агентних систем.

У дослідженні [4] представлено результати порівняння різних великих мовних моделей під час виконання завдань цифрової криміналістики та аналізу кіберінцидентів. Отримані результати свідчать про ефективність поєднання LLM із технологією Retrieval-Augmented Generation (RAG), що дозволяє значно підвищити точність виявлення індикаторів компрометації.

Робота [5] присвячена застосуванню технологій RAG для адаптації великих мовних моделей до динамічного середовища кіберзагроз. Автори доводять, що використання зовнішніх баз знань забезпечує покращення процесів аналізу нових вразливостей та актуальних атак.

У дослідженні [6] проаналізовано перспективи впровадження великих мовних моделей у діяльність центрів операційної безпеки. Встановлено, що LLM можуть використовуватися для автоматизації процесів тріажу інцидентів, аналізу повідомлень систем SIEM та підтримки фахівців з інформаційної безпеки. Незважаючи на значну кількість досліджень, питання комплексного використання великих мовних моделей для одночасної оптимізації процесів пошуку інформації та забезпечення її захисту залишаються недостатньо дослідженими, що визначає актуальність подальших наукових розробок.

Мета статті. Метою статті є дослідження можливостей використання великих мовних моделей для оптимізації задач пошуку та захисту інформації, аналіз сучасних напрямів їх застосування в інформаційних системах, визначення переваг, недоліків та перспектив розвитку технологій на основі LLM у сфері інформаційної безпеки.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

LLM – це нейромережі, що складаються з мільярдів параметрів, які тренуються на величезних обсягах даних з мережі Інтернет. Модель не просто копіює текст, а вгадує наступне логічне слово чи речення на основі отриманого контексту. Моделі здатні виявляти приховані зв'язки та логіку завдяки

архітектурі глибокого навчання. Таким чином, LLM здатні прогнозувати наступні слова, будувати логічні зв'язки, підтримувати діалог, виконувати аналітичні завдання, генерувати код, створювати документи [7]. Серед найвідоміших LLM ChatGPT, Claude, Gemini, LLaMA.

Функціонування сучасних сервісів, побудованих на технологіях генеративного штучного інтелекту, являє собою складний багатоступеневий процес, у межах якого здійснюється послідовне перетворення вхідної інформації, її інтелектуальний аналіз, формування нових даних та надання результатів користувачеві (рис. 1).

1. Збір вхідних даних. Початковим етапом роботи будь-якого сервісу генеративного штучного інтелекту є отримання вхідних даних від користувача або зовнішніх інформаційних систем. На цьому етапі здійснюється первинне приймання інформації через відповідний інтерфейс взаємодії, який може бути реалізований у вигляді чат-інтерфейсу, голосового помічника, мобільного застосунку, веб-сервісу або програмного інтерфейсу API. У випадку мультимодальних систем одночасно можуть використовуватися декілька джерел інформації, що забезпечує більш повне розуміння контексту задачі. Крім безпосереднього запиту користувача, система може враховувати додаткову інформацію, зокрема попередню історію взаємодії, налаштування користувача, часові характеристики, географічне розташування, зовнішні бази знань та контекстні параметри середовища функціонування. Таким чином, на першому етапі відбувається формування сукупності вхідних даних, необхідних для подальшого інтелектуального аналізу.

2. Відображення ознак. Оскільки нейронні мережі не можуть безпосередньо працювати з інформацією у природному вигляді, наступним етапом є перетворення даних у внутрішнє математичне представлення. Для текстових моделей початково виконується токенізація, тобто розбиття тексту на окремі елементи (токени), які можуть відповідати словам, частинам слів, символам або їх комбінаціям. Кожному токену ставиться у відповідність унікальний числовий ідентифікатор.

Рис.1. Механізм функціонування систем, що використовують генеративний штучний інтелект

Після цього здійснюється перехід від дискретного представлення до багатовимірного векторного простору за допомогою механізму Embedding. Вектори ознак формують числовий опис семантичних властивостей інформації та дозволяють моделі визначати ступінь подібності між окремими елементами. У випадку обробки зображень використовуються згорткові або трансформерні механізми виділення ознак, які перетворюють піксельне представлення у набір числових параметрів. Аналогічно аудіосигнали можуть перетворюватися у спектрограми або латентні вектори. У результаті цього етапу вся інформація переходить до математичного простору, придатного для подальшої роботи нейронних мереж.

3. Навчання патернів та формування знань. Одним із найважливіших етапів є навчання моделі, під час якого відбувається формування внутрішніх закономірностей та накопичення статистичних знань. У процесі навчання модель аналізує великі масиви даних, які можуть містити текстові документи, програмний код, зображення, аудіоматеріали, наукові публікації, енциклопедичні ресурси та інші джерела інформації. Використовуючи алгоритми оптимізації, нейронна мережа поступово встановлює взаємозв'язки між окремими об'єктами, виявляє приховані закономірності, причинно-наслідкові залежності та статистичні структури даних. У великих мовних моделях процес навчання ґрунтується на прогнозуванні наступного токена. У ході багаторазових ітерацій параметри мережі коригуються таким чином, щоб мінімізувати функцію втрат і підвищити точність прогнозування. У випадку використання

вже готових сервісів етап навчання виконується заздалегідь, а користувач взаємодіє з попередньо навченою моделлю.

4. Використання навчених моделей. Після завершення процесу навчання формується готова модель, яка містить знання у вигляді мільйонів або мільярдів параметрів нейронної мережі. Параметри моделі представляють собою вагові коефіцієнти, які зберігають статистичні закономірності, отримані під час навчання. Саме вони визначають поведінку системи та забезпечують здатність до виконання різноманітних інтелектуальних завдань. У процесі виконання запитів модель переходить у режим інференсу, під час якого відбувається використання накопичених знань без зміни основних параметрів нейронної мережі.

5. Визначення простору намірів. Після отримання запиту система повинна інтерпретувати наміри користувача та сформувати внутрішнє розуміння поставленої задачі. На цьому етапі виконується семантичний аналіз вхідного повідомлення, виявлення ключових понять, встановлення логічних зв'язків та визначення контексту взаємодії. Система оцінює предметну область запиту; тип необхідного результату; ціль користувача; контекст попереднього діалогу; обмеження та додаткові умови. У великих мовних моделях простір намірів формується завдяки механізму Self-Attention, який дозволяє враховувати взаємозв'язки між окремими словами та фрагментами тексту.

На основі сформованого контексту система визначає, чи необхідно надати пояснення, створити текст, згенерувати програмний код, виконати переклад, проаналізувати документ, створити зображення, виконати пошук інформації, розв'язати математичну задачу. Фактично на цьому етапі формується внутрішнє семантичне представлення задачі.

6. Генерація контенту. Генерація результату є центральним етапом функціонування систем генеративного штучного інтелекту. У великих мовних моделях процес генерації здійснюється послідовно. Для кожної позиції тексту модель прогнозує ймовірності появи наступного токена на основі попереднього контексту. Після визначення найбільш імовірного токена він додається до вже сформованої послідовності, після чого процедура повторюється. Завдяки механізму багатоголової уваги (Multi-Head Attention) модель здатна враховувати складні семантичні залежності між словами та підтримувати логічну цілісність тексту.

7. Декодування латентних ознак. Сформоване внутрішнє представлення результату має математичний характер і потребує перетворення у формат, зрозумілий людині. У текстових моделях виконується декодування послідовності токенів у звичайний текст. Для цього числові ідентифікатори перетворюються у слова та символи відповідно до словника токенізатора. У системах генерації зображень латентні вектори проходять через декодер, який відновлює кольорові значення пікселів та просторову структуру зображення. Для аудіогенераторів виконується реконструкція звукової хвилі, а в системах генерації програмного коду формується синтаксично коректний текст мовою програмування. Таким чином, здійснюється перехід від прихованого простору ознак до кінцевого представлення результату.

8. Уточнення та фільтрація. Перед передачею результату користувачеві більшість сучасних сервісів виконують додаткову перевірку та оптимізацію сформованого контенту. На цьому етапі можуть застосовуватися алгоритми модерації контенту, фільтрація небажаних відповідей, механізми виявлення токсичних висловлювань, перевірка відповідності політикам безпеки, контроль якості, ранжування альтернативних варіантів відповіді, постобробка результатів.

9. Візуалізація результату. Після завершення всіх етапів обробки сформований результат подається користувачеві у відповідному форматі. Форма представлення залежить від типу сервісу та може бути реалізована у вигляді текстової відповіді у чаті, згенерованого зображення, аудіофайлу, відеоматеріалу, програмного коду, структурованого документа, JSON-відповіді через API, інтерактивних елементів інтерфейсу. Візуалізація результатів є завершальним етапом циклу інференсу та забезпечує взаємодію між штучним інтелектом і користувачем.

10. Зворотний зв'язок користувача. Завершальним етапом функціонування сервісу є отримання зворотного зв'язку від користувача.

Одним із найперспективніших напрямів використання великих мовних моделей у сфері кібербезпеки є їх інтеграція з платформами управління подіями та інцидентами інформаційної безпеки SIEM) центрами операційної безпеки SOC та системами виявлення вторгнень IDS. Стрімке збільшення обсягів телеметричних даних, журналів подій, мережевого трафіку та повідомлень безпеки обумовлює необхідність автоматизації процесів аналізу та підтримки прийняття рішень [8].

1) Сучасні SIEM-платформи, такі як Splunk, IBM QRadar, Microsoft Sentinel, ArcSight або Elastic Security, здійснюють централізований збір, зберігання та кореляцію інформації про події безпеки, що надходять від серверів, робочих станцій, мережевого обладнання, міжмережевих екранів, антивірусних систем та інших джерел.

У великих корпоративних мережах кількість подій може досягати десятків або сотень мільйонів записів щоденно. Значна частина повідомлень не має критичного характеру, проте їх аналіз потребує значних ресурсів та високої кваліфікації персоналу.

Інтеграція великих мовних моделей із SIEM-системами дозволяє автоматизувати низку функцій [9]:

- аналіз журналів подій;
- класифікацію інцидентів;
- визначення пріоритетності загроз;
- автоматичне формування аналітичних звітів;
- пояснення причин спрацювання правил кореляції;
- генерацію рекомендацій щодо реагування;
- створення запитів до баз даних журналів;
- автоматизацію розслідування кіберінцидентів.

Однією з найбільш важливих задач є інтерпретація повідомлень безпеки. Традиційні системи генерують значну кількість технічних повідомлень, які потребують додаткового аналізу. Великі мовні моделі здатні перетворювати складні записи журналів подій у зрозумілі текстові пояснення, що значно полегшує роботу аналітиків.

Наприклад, модель може проаналізувати журнал Windows Event Log, зіставити його з базою знань MITRE ATT&CK та сформувати висновок щодо можливого застосування техніки Privilege Escalation або Credential Dumping.

Крім того, LLM можуть автоматично генерувати запити мовами SPL, KQL або SQL для пошуку подій у SIEM-платформах. Це значно прискорює проведення розслідувань та зменшує навантаження на фахівців.

2) Security Operations Center являє собою спеціалізований підрозділ, відповідальний за цілодобовий моніторинг подій безпеки, виявлення загроз та реагування на кіберінциденти. Основними проблемами функціонування SOC є надмірна кількість сповіщень, високий рівень хибнопозитивних спрацьовувань, дефіцит кваліфікованих спеціалістів, значний час реагування на інциденти, складність аналізу великих обсягів телеметричної інформації [10].

Великі мовні моделі можуть виконувати роль інтелектуального помічника аналітика SOC та брати участь у всіх стадіях життєвого циклу реагування на інциденти.

Під час надходження повідомлення про потенційну загрозу модель аналізує тип інциденту, джерело події, потенційний вплив, критичність активу, можливі індикатори компрометації, відповідність технікам MITRE ATT&CK. На основі отриманої інформації формується пріоритет інциденту та рекомендації щодо подальших дій.

Застосування LLM дозволяє автоматизувати процес тріажу інцидентів, який традиційно виконується аналітиками першої лінії SOC. Це забезпечує скорочення часу реагування та підвищення ефективності роботи центрів моніторингу безпеки. Перспективним напрямом є поєднання великих мовних моделей із платформами SOAR, що забезпечує автоматизоване виконання окремих дій у відповідь на виявлені загрози.

Наприклад, використання Ollama – одного із найпопулярніших безкоштовних інструментів із відкритим вихідним кодом, який дозволяє запускати LLM локально на власному комп'ютері (macOS, Windows, Linux), без використання хмарних сервісів та інтернету (рис. 2). Оскільки всі обчислення відбуваються локально, то дані, промпти та комерційні таємниці нікуди не відправляються після завантаження LLM.

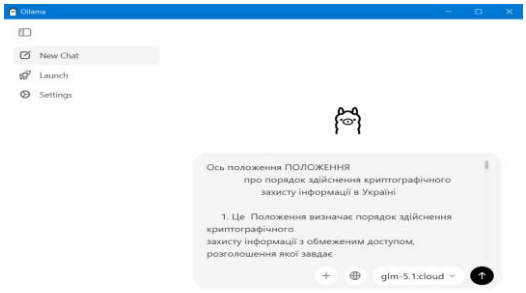

Рис.2. Інструмент ШІ з відкритим кодом Ollama

Задаємо, наприклад, запит щодо аналізу логів Windows для виявлення аномальної активності. LLM обробляє системні події, зокрема *Event ID 4624* (успішний вхід у систему), перевіряє IP-адреси та

імена користувачів. Результат аналізу логів Windows (рис. 3) показує, що LLM здатна виділити як нормальні події, так і потенційні загрози.

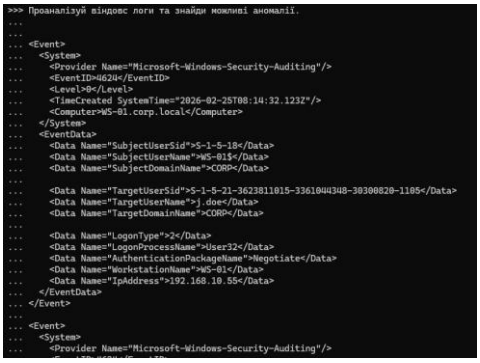

Рис. 3. Результат аналізу логів Windows

3) Системи виявлення вторгнень призначені для аналізу мережевого трафіку та виявлення ознак несанкціонованої активності. Класичні IDS-системи використовують сигнатурний аналіз, статистичні методи, машинне навчання та поведінковий аналіз. Проте традиційні підходи часто характеризуються високим рівнем хибнопозитивних результатів та обмеженою здатністю до виявлення нових типів атак.

Використання великих мовних моделей дозволяє реалізувати більш глибокий контекстний аналіз мережевих подій. LLM можуть виконувати аналіз мережевих журналів; інтерпретацію пакетних трасувань; класифікацію аномальної поведінки; визначення можливих сценаріїв атак; аналіз причин інциденту; автоматичне формування висновків. Особливу ефективність демонструють гібридні системи, у яких алгоритми машинного навчання виконують первинне виявлення аномалій, а великі мовні моделі забезпечують інтелектуальну інтерпретацію результатів [11].

Використання великих мовних моделей супроводжується низкою загроз для конфіденційності та інформаційної безпеки, оскільки такі системи працюють із великими обсягами текстових даних, які можуть містити персональну, фінансову або корпоративну інформацію. Однією з основних проблем є ризик витоку конфіденційних даних під час взаємодії користувача з генеративними моделями. Якщо працівники компанії вводять у систему внутрішню документацію, вихідний код, персональні дані клієнтів або комерційну таємницю, існує небезпека, що ця інформація може бути збережена, використана для навчання моделей або випадково відтворена у відповідях іншим користувачам. Особливо високий рівень ризику характерний для використання публічних хмарних сервісів, де організація не має повного контролю над інфраструктурою обробки даних.

Ще однією серйозною загрозою є можливість маніпуляції моделями шляхом спеціально сформованих запитів, відомих як Prompt Injection або Jailbreaking. Зловмисники можуть намагатися змусити систему обійти внутрішні обмеження, розкрити приховані інструкції або надати доступ до конфіденційної інформації. Крім того, великі мовні моделі схильні до так званих «галюцинацій» – генерації неправдивої, неточної або вигаданої інформації, яка може сприйматися як достовірна. Це створює ризики поширення помилкових даних, юридичних порушень та репутаційних втрат. Саме тому використання LLM у корпоративному середовищі потребує комплексного підходу до захисту даних, контролю якості контенту та впровадження спеціалізованих механізмів безпеки [11].

Одним із найважливіших заходів захисту конфіденційності під час роботи з LLM є вимкнення використання введених даних для подальшого навчання моделі. У багатьох публічних сервісах генеративного штучного інтелекту історія чатів та введені користувачем дані можуть використовуватися для вдосконалення алгоритмів. Це створює потенційний ризик потрапляння конфіденційної інформації до навчальних наборів даних. Для бізнесу така ситуація є особливо небезпечною, оскільки працівники можуть ненавмисно вводити службову інформацію, фінансові звіти або внутрішню документацію. Тому під час використання платформ, таких як OpenAI та Google, рекомендується вимикати історію чатів або функції навчання на користувацьких даних у налаштуваннях акаунта. Це знижує ризик використання інформації для подальшого тренування моделей. Такий підхід є базовою вимогою політики кібербезпеки під час роботи з генеративним ШІ у корпоративному середовищі.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Проведене дослідження свідчить, що великі мовні моделі стають одним із ключових інструментів цифрової трансформації процесів пошуку та захисту інформації. Їх використання дозволяє значно

підвищити ефективність аналізу інформаційних потоків, автоматизувати процеси виявлення кіберзагроз, аналізу вразливостей та реагування на інциденти інформаційної безпеки.

Встановлено, що інтеграція великих мовних моделей із системами кіберзахисту забезпечує підвищення швидкості обробки даних, покращення якості прийняття рішень та зменшення навантаження на фахівців з інформаційної безпеки. Перспективним напрямом розвитку є поєднання LLM із технологіями Retrieval-Augmented Generation, системами розвідки кіберзагроз, платформами SIEM та інтелектуальними агентами кіберзахисту.

Разом із тим широке впровадження великих мовних моделей потребує вирішення низки проблем, пов'язаних із забезпеченням достовірності результатів, захистом конфіденційної інформації, стійкістю до атак на моделі та підвищенням рівня пояснюваності їх функціонування [6]. Подальші наукові дослідження мають бути спрямовані на створення безпечних, адаптивних та надійних інтелектуальних систем підтримки кібербезпеки нового покоління, здатних ефективно функціонувати в умовах постійного ускладнення сучасного кіберпростору.

Перспективи подальшого розвитку інтелектуальних систем кіберзахисту на основі великих мовних моделей (LLM) пов’язані з переходом від традиційних засобів виявлення загроз до автономних і проактивних систем безпеки, здатних аналізувати, прогнозувати та нейтралізувати кіберзагрози в режимі реального часу.

Великі мовні моделі вже сьогодні демонструють значний потенціал у сфері кібербезпеки завдяки здатності обробляти великі обсяги неструктурованих даних, аналізувати журнали подій, мережевий трафік, звіти про інциденти та технічну документацію. Подальший розвиток цих технологій дозволить створювати інтелектуальні системи, які автоматично виявлятимуть аномалії, визначатимуть ознаки кібератак та пропонуватимуть оптимальні сценарії реагування.

Одним із ключових напрямів розвитку є створення автономних агентів кіберзахисту. Такі системи зможуть не лише аналізувати інформацію, а й самостійно виконувати захисні дії: блокувати підозрілий трафік, ізолювати заражені вузли мережі, змінювати правила міжмережевих екранів та координувати роботу різних засобів захисту. Це дозволить значно скоротити час реагування на інциденти та зменшити навантаження на фахівців з інформаційної безпеки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Jaffal, N. O., Alkhanafseh, M., & Mohaisen, D. (2025). Large language models in cybersecurity: A survey of applications, vulnerabilities, and defense techniques. *AI*, *6*(9), Article 216. https://doi.org/10.3390/ai6090216
2. Zhang, J., Bu, H., Wen, H., Liu, Y., Fei, H., et al. (2025). When LLMs meet cybersecurity: A systematic literature review. *Cybersecurity, 8*, Article 55. https://doi.org/10.1186/s42400-025-00361-w
3. Security concerns for large language models: A survey. (2025). *Journal of Information Security and Applications, 95*, Article 104284. https://doi.org/10.1016/j.jisa.2025.104284
4. Evaluating three large language models for security incident detection and analysis. (2025). *Procedia Computer Science, 269*, 465-473. https://doi.org/10.1016/j.procs.2025.08.299
5. Borah, A., Alam, M. T., & Rastogi, N. (2025). *Adapting large language models to emerging cybersecurity using retrieval augmented generation* (arXiv:2510.27080). arXiv. https://arxiv.org/abs/2510.27080
6. Habibzadeh, A., Feyzi, F., & Atani, R. E. (2025). *Large language models for security operations centers: A comprehensive survey* (arXiv:2509.10858). arXiv. https://arxiv.org/abs/2509.10858
7. Kopytin, O. V. (2023). *Machine learning in cybersecurity systems*. FOP Panov.
8. Dovhan, O. D. (2022). *Cybersecurity and artificial intelligence: Modern information protection technologies*. Lira-K.
9. Zhezhnych, P. I. (2021). *Data analysis and artificial intelligence in cyberspace*. Lviv Polytechnic Publishing House.
10. Das, S., Kim, J., Jang, J., & Alazab, M. (2024). *Large language models for cyber security: A systematic literature review* (arXiv:2405.04760). arXiv. https://arxiv.org/abs/2405.04760
11. Al-Hawawreh, M., Sitnikova, E., Aboutorab, N., & Alazab, M. (2025). When large language models meet cybersecurity: A systematic literature review. *Cybersecurity, 8*(1), Article 18.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (33), 2026
ISSN 2663 – 4023

Anastasiia Vavilenkova
Doctor of Technical Sciences, Professor, Head of the Department of Cyber Security
National Academy of the Security Service of Ukraine, Kyiv, Ukraine
ORCID: 0000-0002-9630-4951
vavilenkovaa@gmail.com

LARGE LANGUAGE MODELS FOR OPTIMIZING TASKS IN INFORMATION RETRIEVAL AND PROTECTION

Abstract. The rapid development of digital technologies, the global informatization of society, and the continuous growth in data volumes create new challenges in the field of information retrieval, processing, and protection. Information systems generate terabytes of data every day, including both valuable information and potential indicators of cyber threats, anomalies, and attacks. Traditional approaches to information analysis are increasingly proving insufficient due to their limited capability to process unstructured data and the need for significant human involvement.

The emergence of large language models (LLMs) has become one of the most significant achievements of modern artificial intelligence. Owing to their ability to analyze natural language, understand context, summarize vast amounts of information, and generate meaningful responses, LLMs provide new opportunities for automating information retrieval processes and enhancing information security. Recent studies demonstrate that large language models can be effectively applied to vulnerability detection, malware analysis, cyber incident investigation, security log processing, and the automation of Security Operations Center (SOC) activities.

This article investigates the potential of large language models for optimizing information retrieval, analysis, processing, and protection processes in the modern digital environment. The study examines the characteristics of large language models, the principles underlying their construction, and the mechanisms of their application within information and communication systems. The main areas of LLM utilization in information retrieval, automated text analysis, information threat detection, cyber incident monitoring, software vulnerability discovery, and decision-support processes in cybersecurity are analyzed. Particular attention is paid to the integration of large language models with Security Information and Event Management (SIEM) systems, intrusion detection systems, cyber threat intelligence technologies, and Retrieval-Augmented Generation (RAG) mechanisms. The advantages and limitations of employing LLMs in information retrieval and protection tasks are identified. Furthermore, the prospects for the further development of intelligent cyber defense systems based on large language models are substantiated.

Keywords: Large Language Models (LLMs); Artificial Intelligence Technologies; Cybersecurity; Information Retrieval; Cyber Incident; Cyber Threat; Information Protection.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Jaffal, N. O., Alkhanafseh, M., & Mohaisen, D. (2025). Large language models in cybersecurity: A survey of applications, vulnerabilities, and defense techniques. *AI*, *6*(9), Article 216. https://doi.org/10.3390/ai6090216
2. Zhang, J., Bu, H., Wen, H., Liu, Y., Fei, H., et al. (2025). When LLMs meet cybersecurity: A systematic literature review. *Cybersecurity*, *8*, Article 55. https://doi.org/10.1186/s42400-025-00361-w
3. Security concerns for large language models: A survey. (2025). *Journal of Information Security and Applications*, *95*, Article 104284. https://doi.org/10.1016/j.jisa.2025.104284
4. Evaluating three large language models for security incident detection and analysis. (2025). *Procedia Computer Science*, *269*, 465-473. https://doi.org/10.1016/j.procs.2025.08.299
5. Borah, A., Alam, M. T., & Rastogi, N. (2025). *Adapting large language models to emerging cybersecurity using retrieval augmented generation* (arXiv:2510.27080). arXiv. https://arxiv.org/abs/2510.27080
6. Habibzadeh, A., Feyzi, F., & Atani, R. E. (2025). *Large language models for security operations centers: A comprehensive survey* (arXiv:2509.10858). arXiv. https://arxiv.org/abs/2509.10858
7. Kopytin, O. V. (2023). *Machine learning in cybersecurity systems*. FOP Panov.

8. Dovhan, O. D. (2022). *Cybersecurity and artificial intelligence: Modern information protection technologies*. Lira-K.
9. Zhezhnych, P. I. (2021). *Data analysis and artificial intelligence in cyberspace*. Lviv Polytechnic Publishing House.
10. Das, S., Kim, J., Jang, J., & Alazab, M. (2024). *Large language models for cyber security: A systematic literature review* (arXiv:2405.04760). arXiv. https://arxiv.org/abs/2405.04760
11. Al-Hawawreh, M., Sitnikova, E., Aboutorab, N., & Alazab, M. (2025). When large language models meet cybersecurity: A systematic literature review. *Cybersecurity*, *8*(1), Article 18.

Отримано редакцією журналу / Received: 03.03.26
Прорецензовано / Revised: 15.03.26
Схвалено до друку / Accepted: 25.06.26