УДК 004.75; 004.8
DOI: 10.31673/2409-7292.2025.026884

Абібулаєв А.Р., Піскозуб А.З.

АНАЛІЗ МОЖЛИВОСТЕЙ ПОКРАЩЕННЯ СТАНУ БЕЗПЕКИ ХМАРНОЇ ІНФРАСТРУКТУРИ ЗА ДОПОМОГОЮ NLP ТА ML

Разом із зростанням обсягів даних та складності мультихмарних середовищ забезпечення кібербезпеки хмарної інфраструктури стає дедалі важчим завданням. Традиційні підходи на основі статичних правил, сигнатурного аналізу та централізованих SIEM-систем виявляють обмежену ефективність при роботі з динамічними ресурсами й адаптивними атаками, такими як APT-кампанії, insider threat чи zero-day експлойти. Це зумовлює необхідність впровадження інтелектуальних механізмів аналізу та реагування, здатних оперативно корелювати різнорідні події та знижувати кількість хибнопозитивних спрацювань. Інтеграція технологій обробки природної мови (NLP) і машинного навчання (ML) відкриває нові можливості для автоматизації аналітики інцидентів, семантичного розбору журналів подій (далі – логів) і класифікації загроз за рівнем ризику. NLP-модулі дозволяють обробляти великі масиви неструктурованих текстових даних — журнали подій, повідомлення користувачів та конфігураційні файли — й ідентифікувати соціотехнічні шаблони атак. ML-алгоритми, у свою чергу, забезпечують виявлення аномалій із використанням класифікації, кластеризації та поведінкової аналітики (UEBA), що дозволяє прогнозувати потенційні атаки ще до їхньої реалізації. Сучасні концепції кіберзахисту, зокрема модель Zero Trust і принцип найменших привілеїв (PoLP), у поєднанні з підходом Security as Code створюють основу для динамічного контролю доступу та автоматизованого управління правами. Архітектурні рішення, що поєднують Cloud IAM, PAM і CIEM, доповнюються механізмами, керованими штучним інтелектом, для оцінки контексту запитів у реальному часі та автоматизованої перевірки надлишкових привілеїв. Це сприяє зменшенню часу реагування та підвищенню адаптивності політик безпеки. В рамках цього дослідження проведено систематичний огляд більше десяти сучасних наукових публікацій, що охоплюють практичні реалізації інтелектуальних DLP-систем, механізми автоматизованого виявлення загроз у AWS, Azure та GCP, а також підходи до інтеграції NLP/ML у CI/CD процеси та SOAR-платформи. Сформульовано вимоги до побудови адаптивних, контекстно-чутливих рішень із урахуванням масштабованості, інтерпретованого штучного інтелекту (Explainable AI) та дотримання етичних і правових норм (GDPR, ISO/IEC 27001). Результати дослідження доводять, що комбінований підхід на основі NLP і ML дозволяє значно знизити кількість хибнопозитивних спрацювань, скоротити середній час реагування на інциденти та підвищити точність виявлення складних загроз. Отримані висновки будуть корисними для IT-відділів, інженерів із безпеки та DevOps-команд, які прагнуть оптимізувати процеси кіберзахисту в динамічних мультихмарних середовищах.

Ключові слова: кібербезпека, хмарні технології, NLP, ML, Zero Trust, Security as Code, UEBA, DLP.

Вступ

У сучасних умовах стрімкої цифрової трансформації хмарні обчислення стали критично важливим компонентом інформаційної інфраструктури як приватного, так і державного секторів. Однак із зростанням обсягів даних, які обробляються в хмарних середовищах, та розширенням векторів атак, забезпечення кібербезпеки в хмарі стає дедалі складнішим завданням. Традиційні методи захисту, засновані на статичних правилах, часто виявляються недостатніми для виявлення складних та адаптивних загроз, таких як APT-атаки, інсайдерські загрози або цілеспрямовані кампанії з викрадення даних [1].

У зв’язку з цим, зростає інтерес до застосування технологій машинного навчання (ML) та обробки природної мови (NLP) для підвищення ефективності систем захисту. Завдяки здатності до аналізу великих обсягів структурованих і неструктурованих даних, ML та NLP дають змогу автоматично виявляти аномалії, класифікувати інциденти, розпізнавати шкідливі шаблони в логах, мережевому трафіку та навіть у текстах політик безпеки або звітів про події.

Особливої актуальності ця тема набуває в умовах мультихмарного розгортання IT-інфраструктури [2], де системи безпеки повинні забезпечувати контроль доступу, виявлення витоків даних і реагування на інциденти в реальному часі на основі різнорідної інформації. Наявні дослідження показують, що поєднання можливостей NLP і ML сприяє створенню адаптивних систем безпеки, здатних працювати в умовах невизначеності та високої динаміки загроз. Загалом, інтеграція NLP та ML у сферу кібербезпеки хмарних середовищ відкриває нові перспективи для проактивного захисту, аналітики інцидентів і оптимізації прийняття рішень у системах захисту інформації.

Аналіз літературних джерел та формулювання проблеми

Останні роки характеризуються активним розвитком досліджень у галузі інтеграції технологій обробки природної мови (NLP) та машинного навчання (ML) для забезпечення безпеки хмарних інфраструктур. Як зазначено у джерелах [1, 3], основним чинником такого інтересу є недостатня ефективність традиційних підходів, таких як сигнатурний аналіз та статичні політики безпеки, в умовах динамічних мультихмарних середовищ. Автори відзначають, що сучасні загрози потребують не лише реактивних заходів, а й проактивного виявлення та реагування на інциденти у реальному часі, що може бути досягнуто лише за допомогою адаптивних, інтелектуальних рішень [4, 8]. Автори у джерелах [2, 5] акцентують на тому, що NLP-технології здатні ефективно обробляти великі обсяги неструктурованих даних, таких як журнали подій, повідомлення користувачів та конфігурації, що дозволяє оперативно виявляти ознаки загроз і порушення політик безпеки. Зокрема, такі технології застосовуються для автоматичної класифікації текстових повідомлень за рівнем ризику, що значно знижує кількість хибно-позитивних спрацювань та навантаження на аналітиків безпеки [7, 18].

Окремо наголошується роль ML у прогнозуванні загроз та поведінковому аналізі. У джерелах [6, 7] підкреслюється, що системи поведінкового аналізу користувачів (UEBA) дозволяють формувати профілі типової активності, що є критично важливим для виявлення як зовнішніх атак, так і інсайдерських загроз. Це дає змогу оперативно реагувати на аномальні дії, що не характерні для звичайної поведінки користувача [12, 14].

Крім того, у дослідженнях [3, 9, 10] розглядається інтеграція ML та NLP у контексті захисту даних та контролю доступу. Автоматизовані DLP-системи, які використовують NLP та ML, демонструють високу ефективність у виявленні спроб несанкціонованої передачі конфіденційної інформації, завдяки здатності моделі аналізувати зміст документів і комунікацій, навіть якщо вони не збігаються із стандартними шаблонами [13, 16].

Важливим напрямом є також інтеграція інтелектуальних систем з концепцією Zero Trust. Автори [4, 11, 22] зазначають, що ML-алгоритми ефективно доповнюють політики Zero Trust, забезпечуючи динамічну перевірку контексту запитів доступу та постійний моніторинг активності користувачів, що суттєво знижує ймовірність несанкціонованих дій [23].

Однак, попри значні досягнення, існуючі дослідження вказують на проблеми масштабування та адаптації моделей ML і NLP до нових і непередбачуваних ситуацій, а також на питання пояснюваності рішень (explainability), що є критично важливим для практичного впровадження інтелектуальних систем безпеки [15, 17, 21]. Таким чином, аналіз сучасних літературних джерел підтверджує необхідність подальшого розвитку інтелектуальних систем на основі NLP і ML з урахуванням принципів Zero Trust, для забезпечення ефективного і адаптивного захисту хмарних середовищ від складних і динамічних загроз.

Мета та завдання дослідження

У рамках цього дослідження автори ставлять за мету системно узагальнити та проаналізувати існуючі підходи до інтеграції методів NLP та ML в системи кібербезпеки хмарної інфраструктури, з акцентом на їхній потенціал у контексті сучасних викликів та обмеження інтеграції у систему кібербезпеки хмарної інфраструктури з урахуванням сучасних викликів, таких як масштабованість, динамічність загроз, децентралізованість середовищ і поява Shadow IT. На думку авторів, саме формування концептуальної основи для інтеграції цих технологій у системи кіберзахисту хмарної інфраструктури дозволить розробити підходи до побудови адаптивних, контекстно чутливих засобів виявлення загроз і захисту даних, орієнтованих на сучасні моделі мультихмарного розгортання.

Для досягнення поставленої мети в рамках дослідження передбачено вирішення таких завдань:

- здійснити огляд існуючих досліджень щодо застосування NLP та ML у сфері хмарної безпеки;

© Абібулаєв А.Р., Піскозуб А.З. (2025). Аналіз можливостей покращення стану безпеки хмарної інфраструктури за допомогою NLP та ML. Сучасний захист інформації, 2(62), 124–140. https://doi.org/10.31673/2409-7292.2025.026884

- проаналізувати потенціал і обмеження використання NLP-моделей у задачах виявлення аномалій, класифікації подій безпеки та обробки логів;
- визначити роль ML-алгоритмів у виявленні загроз і попередженні витоків даних у мультихмарних середовищах;
- систематизувати архітектурні та методологічні підходи до інтеграції NLP/ML-рішень у хмарні системи безпеки;
- сформулювати вимоги до побудови комплексного підходу до кіберзахисту на основі інтеграції із NLP/ML.

У рамках дослідження розглядається завдання системного аналізу можливостей застосування NLP і ML для підвищення ефективності кіберзахисту хмарної інфраструктури. Особливу увагу приділено виявленню практичних сценаріїв використання таких технологій у задачах класифікації подій безпеки, виявлення аномалій, обробки журналів подій, автоматичного формування політик доступу та запобігання витокам даних.

Для досягнення цілей дослідження було проведено критичний огляд більше двадцяти сучасних наукових публікацій [1-25], у яких розглядаються методи й моделі, що ґрунтуються на NLP і ML у контексті хмарної безпеки. Також враховано практичні приклади реалізації інтелектуальних DLP-систем, засобів контролю доступу, систем аналізу журналів і автоматизованого управління ризиками.

Завданням цієї статті є не лише огляд існуючих рішень та прогалин в захисті хмарної інфраструктури, а і формулювання вимог до побудови сучасної безпекової архітектури із інтеграцією NLP та ML.

Результати досліджень

Традиційна архітектура кібербезпеки, що базується на сигнатурному аналізі, фільтрації трафіку та контрольованому доступі, історично демонструвала ефективність у середовищах з передбачуваною топологією та централізованими обчислювальними ресурсами. До таких інструментів відносять системи управління інформацією та подіями безпеки (SIEM), міжмережеві екрани (WAF, NGFW), системи управління ідентичністю та доступом (IAM), а також класичні механізми записи контролю доступу (Access Control List, ACL). Проте в умовах хмарної парадигми, що передбачає динамічну масштабованість, автоматизоване розгортання ресурсів, тимчасові і гнучкі політики доступу, ці засоби часто виявляються неефективними або надмірно обмежуючими.

По-перше, більшість класичних засобів базуються на заздалегідь визначених правилах або сигнатурах. Виявлення нових, нетипових загроз, які не мають чітко описаних ознак (так звані zero-day атаки), є малоймовірним. Стандартні WAF-рішення можуть ефективно блокувати SQL-ін’єкції або XSS-атаки, однак виявлення обфускованих або соціотехнічних вторгнень поза їх межами можливе лише частково або зовсім відсутнє [3]. Це також стосується і SIEM-рішень: хоч вони й збирають великий обсяг даних, аналітика часто є поверхневою і не включає контекст, необхідний для коректного реагування.

По-друге, фрагментованість контролю між різними рівнями хмарної інфраструктури призводить до утворення “сліпих зон”. В ситуаціях, коли події відбуваються на стику кількох сервісів - наприклад, між API-шлюзом, безсерверним компонентом і хмарною базою даних - класичні SIEM/WAF не мають повної видимості й не здатні коректно корелювати події. Внаслідок цього атаки, які відбуваються пошарово і в умовах високої автоматизації, залишаються непоміченими.

Третім обмеженням є висока частота хибно-позитивних спрацювань, що спостерігається у статичних політиках безпеки. Наприклад, при виявленні “незвичної” активності користувача система може блокувати роботу навіть при легітимному запиті (наприклад, під час планового оновлення або резервного копіювання), що ускладнює продуктивну роботу команд DevOps/SecOps, викликає обурення користувачів і змушує адміністраторів “послаблювати” правила, що своєю чергою підвищує ризики компрометації.

Четвертим ключовим недоліком є затримка в реагуванні. У багатьох системах сповіщення про інциденти надходить уже після того, як атака відбулася. У дослідженні [4] зазначено, що середній час реагування на складні інциденти у традиційних системах безпеки перевищує 6–12 годин, що в умовах безперервної хмарної доступності є критичною втратою часу, що є особливо актуальним для атак, які розвиваються у кілька етапів – з початковим проникненням, рухом всередині мережі (lateral movement) і ексфільтрацією даних.

Окремо варто згадати і про відсутність семантичного аналізу, що робить традиційні системи неспроможними інтерпретувати події на рівні змісту. Зауважимо, що SIEM може зафіксувати факт звернення до певного API, але не здатна визначити, чи було запитано критичні дані, чи йшлося про банальну перевірку доступності. Нажаль, це ще більше ускладнює реагування, оскільки не дозволяє оцінити вагу події без втручання людини.

Робимо однозначний висновок, що класичні інструменти безпеки, незважаючи на їхню важливість як базової лінії захисту, не можуть самостійно забезпечити необхідну гнучкість, швидкодію та контекстність в умовах сучасної хмарної інфраструктури. Це створює необхідність у доповненні їх інтелектуальними системами, побудованими на основі NLP та ML, що дозволяють підвищити рівень автоматизації, виявляти складні загрози й адаптуватись до умов реального часу.

Теоретичні засади застосування NLP та ML у захисті хмарної інфраструктури

Інтеграція технологій NLP і ML у сферу кібербезпеки хмарної інфраструктури відкриває нові горизонти у забезпеченні адаптивного та інтелектуального захисту інформаційних ресурсів. На відміну від традиційних інструментів, заснованих на фіксованих правилах, NLP та ML забезпечують можливість аналізу великого обсягу різнорідної інформації, виявлення закономірностей, аномалій і прогнозування загроз у динамічних середовищах.

Авторами зроблено виклад основних теоретичних положень, що лежать в основі використання NLP і ML у контексті захисту хмарних систем. Було розглянуто функціональні принципи роботи NLP у сфері безпеки, підходи до виявлення інцидентів на основі машинного навчання, а також типи загроз, які покриваються такими рішеннями.

Принципи роботи NLP у сфері кібербезпеки

Обробка природної мови (Natural Language Processing, NLP) у кібербезпеці охоплює сукупність алгоритмічних методів, що дозволяють аналізувати неструктуровані текстові дані для виявлення загроз, аномалій та ознак порушення політик безпеки. Основу таких підходів становить здатність NLP-технологій до синтаксичного та семантичного аналізу логів, повідомлень, звітів про події безпеки, службових повідомлень, політик доступу, а також контенту, що генерується користувачами [5, 6].

Ключовими задачами NLP у сфері безпеки є:

- автоматична класифікація текстових повідомлень за ступенем ризику;
- виявлення шаблонів поведінки та індикаторів компрометації у текстових логах;
- семантичний аналіз запитів доступу до хмарних ресурсів;
- виявлення потенційно небезпечних інструкцій у документах або конфігураціях.

Однією з переваг NLP є можливість обробки великої кількості подій у режимі реального часу з урахуванням контексту, що дозволяє підвищити точність систем виявлення загроз і зменшити кількість хибно-позитивних спрацювань. Зокрема, використання моделей на кшталт BERT або GPT дозволяє не лише зіставляти ключові терміни, а й інтерпретувати їхній контекст у межах інформаційного потоку.

У багатьох сучасних реалізаціях NLP інтегрується в процес автоматизації аналітики інцидентів (SOAR-платформи) або обробки логів SIEM-систем, що значно розширює їх функціональні можливості [7]. NLP також використовується для побудови чатботів і віртуальних асистентів безпеки, здатних інтерпретувати природньомовні запити аналітиків і генерувати відповіді, знижуючи когнітивне навантаження на команди реагування [8].

Окрім традиційного використання NLP у сфері лог-аналізу, моніторингу та семантичної інтерпретації запитів, ця технологія відкриває нові перспективи у виявленні проявів Shadow IT, зокрема шляхом аналізу поведінки користувачів, вхідних запитів до хмарних сервісів і внутрішніх комунікацій, що містять згадки про сторонні або несанкціоновані застосунки [9] (рис. 2).

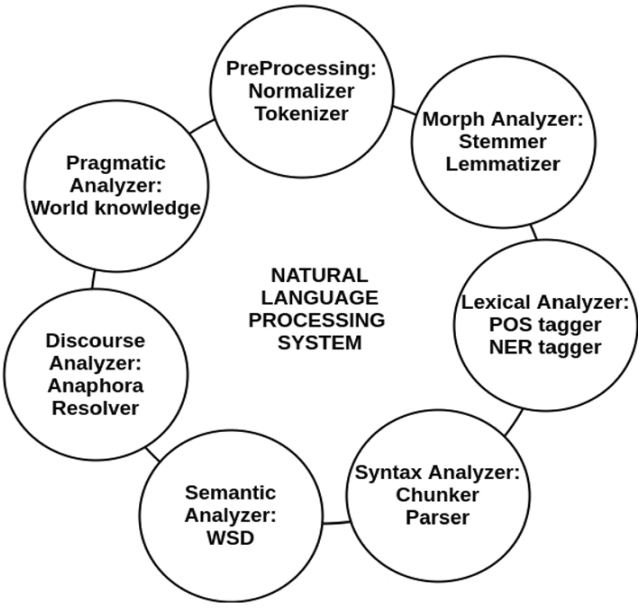

Рис. 1. Візуалізація компонентів NLP

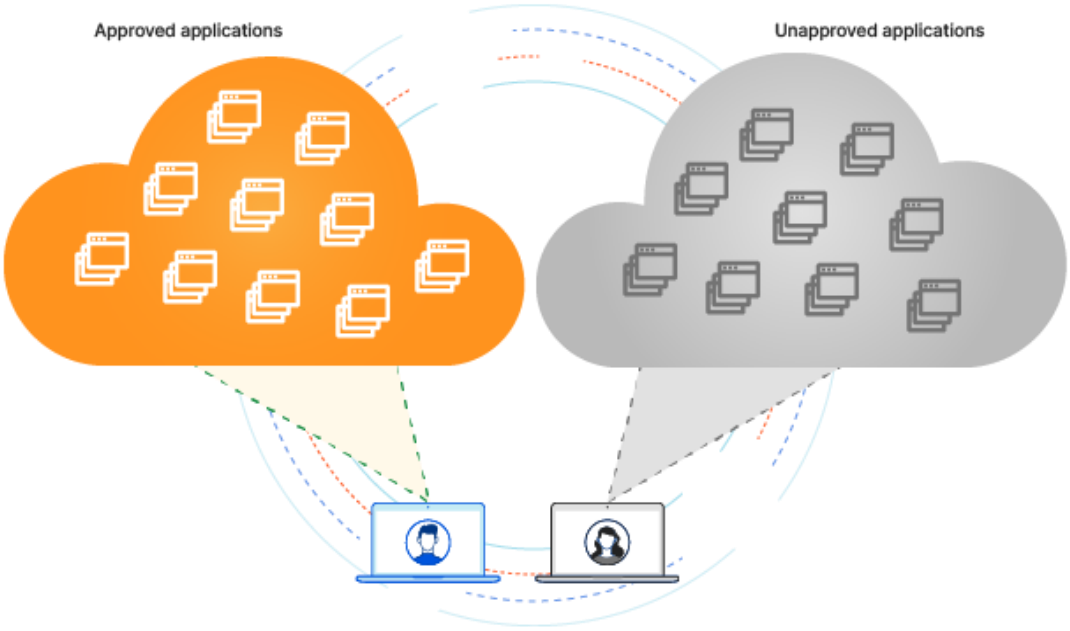

Рис. 2. Додатки із перспективи Shadow-IT

Застосування NLP у таких сценаріях дозволяє:

- виявляти натяки на використання незареєстрованих сервісів або зовнішніх хмарних середовищ;
- визначати шаблони порушення політик доступу, що маскуються під звичайну активність;

- автоматично створювати інформативні звіти для команд безпеки, використовуючи природномовні пояснення;
- покращувати ефективність автоматизованих опитувальників і форм зворотного зв'язку при аудиті політик і доступів.

Ще однією важливою сферою, де NLP активно інтегрується в сучасні хмарні рішення, є управління секретами. У поєднанні з системами на кшталт HashiCorp Vault, NLP може використовуватись для:

- аналізу текстових конфігурацій, сценаріїв CI/CD та IaC (наприклад, Terraform, Ansible) з метою виявлення в них відкритих секретів, API-ключів або hardcoded credentials [10];
- контекстної класифікації вмісту змінних середовища, які часто є джерелами витоків конфіденційних даних;
- автоматизації генерації політик доступу на основі описових запитів адміністраторів безпеки;
- інтеграції з чатботами, які можуть пояснювати політики управління секретами, допомагати в ротації ключів та моніторити зміни в хмарному середовищі.

У сучасних реалізаціях NLP тісно пов'язується з підходом "Security as Code" [11], де усі аспекти безпеки (включаючи управління доступами, перевірку конфігурацій, моніторинг активності тощо) виражаються у вигляді коду. Використання NLP у цьому контексті дозволяє:

- трансформувати неструктуровані запити (від користувачів, devops-фахівців) у формальні security-політики;
- забезпечити семантичну перевірку IaC-файлів на відповідність стандартам безпеки;
- інтегрувати з CI/CD процесами для автоматичної генерації рекомендацій на основі безпекових метрик.

Таким чином, NLP стає не лише інструментом реактивного аналізу, а й активним компонентом екосистеми захисту, що взаємодіє з платформами безперервного розгортання, політиками управління секретами та механізмами обмеження Shadow IT. Це відкриває перспективу створення самонавчальних систем безпеки, які не тільки реагують на події, але й прогнозують їх, формуючи дійсно інтелектуальний підхід до безпеки хмарних систем.

Можливості ML у виявленні та попередженні загроз

Машинне навчання є ключовим компонентом сучасних систем виявлення загроз у хмарній інфраструктурі, оскільки дозволяє ефективно аналізувати великі обсяги даних, виявляти приховані закономірності та прогнозувати потенційні атаки на основі поведінкових аномалій. Алгоритми ML забезпечують проактивний підхід до безпеки: замість реагування на вже виявлену загрозу, система вчиться розпізнавати загрозу ще до її реалізації.

Класифікація, кластеризація та навчання з підкріпленням є найпоширенішими підходами в контексті хмарної безпеки. Зокрема:

- моделі класифікації (напр. random forest, XGBoost, SVM) ефективно використовуються для визначення типу атаки за зібраними характеристиками мережевого трафіку або журналів;
- алгоритми кластеризації, як-от DBSCAN чи k-means, дозволяють виявляти аномальні шаблони поведінки у великому масиві активності користувачів або в телеметричних даних;
- гібридні моделі можуть поєднувати ознаки відомих атак із можливістю навчання на нових вхідних даних, що важливо для захисту від zero-day загроз.

Особливу роль відіграють системи поведінкового аналізу користувачів (User and Entity Behavior Analytics, UEBA), які на основі історичних даних та моделей ML дозволяють формувати профілі типових дій, а потім виявляти їх відхилення. Це дає змогу виявляти як зовнішні атаки, так і інсайдерську активність без потреби в ручному налаштуванні правил.

Хмарні провайдери також активно інтегрують ML у свої сервіси безпеки: Amazon GuardDuty, Azure Sentinel і Google Chronicle застосовують ML для обробки сигналів з різних джерел і виявлення відхилень у режимі реального часу. У поєднанні з автоматизованими

механізмами реагування, це дозволяє значно зменшити час на ліквідацію інцидентів. Завдяки гнучкості, адаптивності та здатності до самонавчання, ML-моделі стають основою для побудови ефективних систем запобігання витокам даних (DLP), захисту від фішингу, виявлення ботнет-активності та оцінки ризиків у багатокомпонентному середовищі хмари.

Типи атак, які покриваються NLP/ML-рішеннями

Використання технологій NLP та ML у сфері кібербезпеки хмарної інфраструктури дозволяє ефективно виявляти та запобігати широкому спектру атак, у тому числі тих, що залишаються малопомітними або не виявляються традиційними системами, побудованими на фіксованих сигнатурах.

Серед основних типів атак, які можуть бути виявлені завдяки застосуванню NLP/ML, виділяють [12]:

- фішингові атаки, зокрема на основі аналізу тексту електронних листів або повідомлень у месенджерах за допомогою NLP-алгоритмів, що розпізнають соціотехнічні шаблони поведінки;
- атаки типу insider threat – моделі ML здатні виявити відхилення в поведінці користувача, які не притаманні його типовому профілю, що дозволяє вчасно ідентифікувати внутрішню загрозу;
- атаки zero-day та невідомі експлойти, які ML-системи можуть виявити за поведінковими характеристиками, навіть якщо сигнатури атак ще не відомі;
- атаки на основі ненормативного доступу до ресурсів (privilege escalation, horizontal/vertical movement) – поєднання UEBA та NLP-аналізу логів дозволяє ідентифікувати невластиві шаблони використання прав;
- DDoS та ботнет-атаки – за рахунок кластеризації мережевих шаблонів, ML-системи виявляють аномалії в трафіку та підозрілу координацію запитів;
- витоки даних (data exfiltration) – NLP-моделі виявляють спроби передати конфіденційну інформацію за межі дозволених каналів, зокрема у текстовому вигляді (наприклад, через e-mail або чат).

Також активно розвивається напрям автоматичної генерації правил безпеки за допомогою LLM-моделей, які здатні не лише аналізувати події, а й генерувати відповідні політики доступу, сценарії для SOAR-систем та рекомендації щодо протидії конкретним загрозам [13].

Завдяки гнучкості та контекстній чутливості, NLP і ML дозволяють покрити як класичні, так і еволюціонуючі загрози, особливо в умовах динамічної мультихмарної інфраструктури, де традиційні підходи до виявлення вже не дають належної ефективності.

Сучасні підходи до інтеграції NLP/ML у захист хмарної інфраструктури

З розвитком мультихмарних та гібридних інфраструктур кібербезпека стикається з новими викликами: складністю міжплатформеної взаємодії, масштабованістю загроз та потребою в реальному часі адаптувати політики захисту до динамічного середовища. У відповідь на ці виклики дослідницька спільнота та індустрія розробляють рішення, що базуються на технологіях ML та NLP.

Застосування NLP дозволяє автоматизувати аналіз логів, політик безпеки та повідомлень про події. ML, у свою чергу, забезпечує глибоку аналітику поведінки користувачів, адаптивне реагування на атаки, класифікацію інцидентів та оцінку ризиків. Об’єднання цих технологій дає змогу будувати інтелектуальні системи, здатні до самонавчання, прогнозування загроз і автономного реагування.

Автоматизоване виявлення загроз

Автоматизоване виявлення загроз є однією з найважливіших задач у сфері хмарної безпеки. Умови мультихмарних середовищ вимагають високошвидкісної обробки великого обсягу подій, що унеможливлює ручний аналіз усіх інцидентів. Застосування NLP/ML

© Абібулаєв А.Р., Піскозуб А.З. (2025). Аналіз можливостей покращення стану безпеки хмарної інфраструктури за допомогою NLP та ML. Сучасний захист інформації, 2(62), 124–140. https://doi.org/10.31673/2409-7292.2025.026884

дозволяє реалізувати механізми раннього виявлення загроз, що забезпечують проактивний підхід до кіберзахисту.

У дослідженні [14] продемонстровано, як NLP-моделі можуть обробляти текстові журнали подій, класифікуючи їх за рівнем загрози. Застосування моделей типу TF-IDF та BERT забезпечує якісне виявлення аномалій навіть за наявності слабкоструктурованих вхідних даних, що дозволяє зменшити навантаження на аналітиків SOC та покращити пріоритезацію інцидентів.

У роботі [15] запропоновано онтологічну модель представлення знань у сфері кібербезпеки, створену за допомогою методів обробки природної мови та машиного навчання. Дослідники розробили предметно-орієнтовану онтологію, що включає 18 основних класів, зокрема Attacker, Exploit, Vulnerability, Software, Risk тощо, і визначає 33 різні типи зв'язків між ними (наприклад, exploits, performs, generates, involves) (рис. 3).

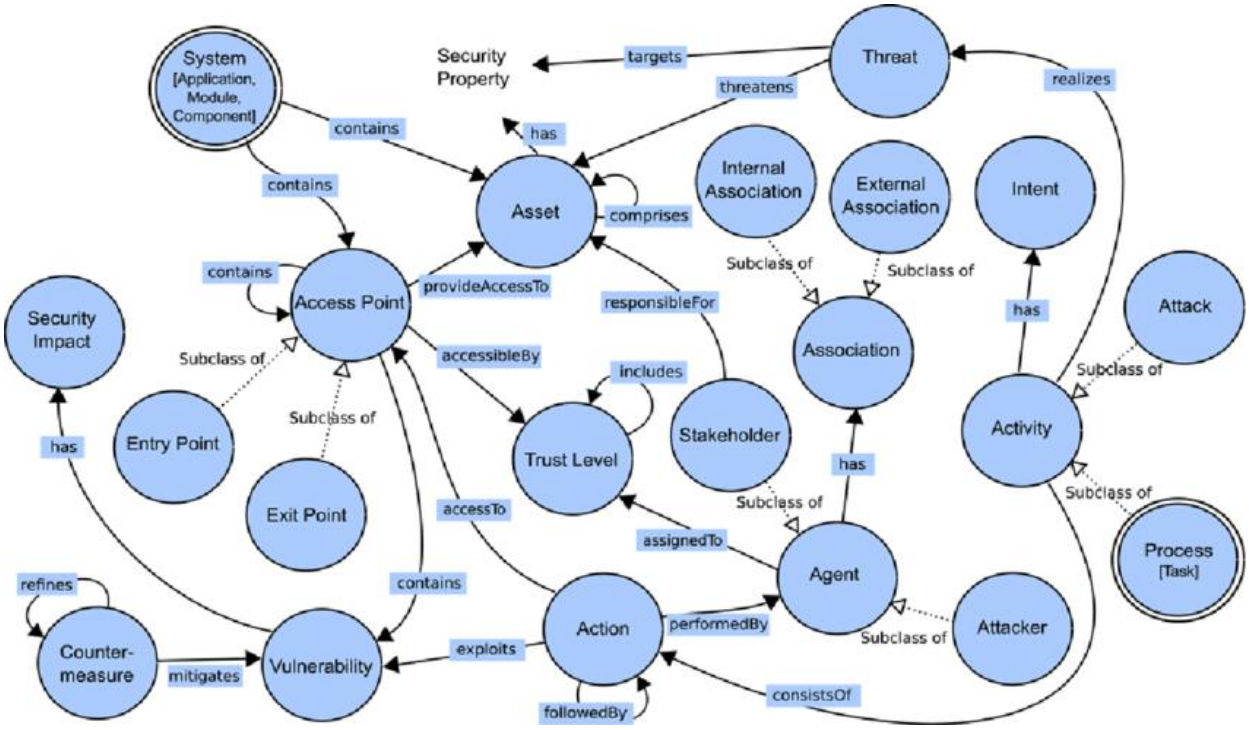

Рис. 3. Онтологічна модель сутностей кібербезпеки та їхніх взаємозв'язків [15]

Архітектура цієї моделі відображає причинно-наслідкові взаємодії між акторами атак, вразливостями, подіями та їхніми наслідками. На основі побудованої онтології була навчена система іменованого розпізнавання сутностей та екстракції відношень у текстах. Це дало змогу автоматизувати семантичний аналіз інцидентів безпеки документів і підвищити точність виявлення ключових елементів інформації.

У публікації [16] розглянуто впровадження ML-алгоритмів для аналізу поведінкових шаблонів у системах ERP, що функціонують у хмарному середовищі. Алгоритми класифікації на основі історичних даних дозволили виявляти нестандартну активність користувачів і запускати автоматизовані процедури реагування. Водночас модель з урахуванням контексту бізнес-операцій дозволила знизити кількість хибно-позитивних спрацювань.

Практичний кейс [17] (рис. 4) демонструє, що автоматизоване виявлення загроз на основі ML у поєднанні з NLP дозволяє реалізовувати адаптивну відповідь на загрози, що постійно еволюціонують. В особливості – фішингові атаки, витоки даних, botnet-активність - системи на базі ML класифікують за кількома рівнями ризику, що полегшує рішення для блокування або карантину в автоматичному режимі.

Деякі підходи, як зазначено у [18], фокусуються на автономному навчанні – модель самостійно адаптується до нових типів загроз без необхідності повного перенавчання. Це значно зменшує час між появою загрози та її детекцією.

Узагальнюючи, автоматизоване виявлення загроз за допомогою NLP та ML дозволяє суттєво підвищити швидкість, точність і ефективність захисту в умовах хмарної динаміки. Ключовими перевагами є масштабованість, здатність до адаптації, зниження кількості хибнопозитивних спрацювань і скорочення часу на реакцію.

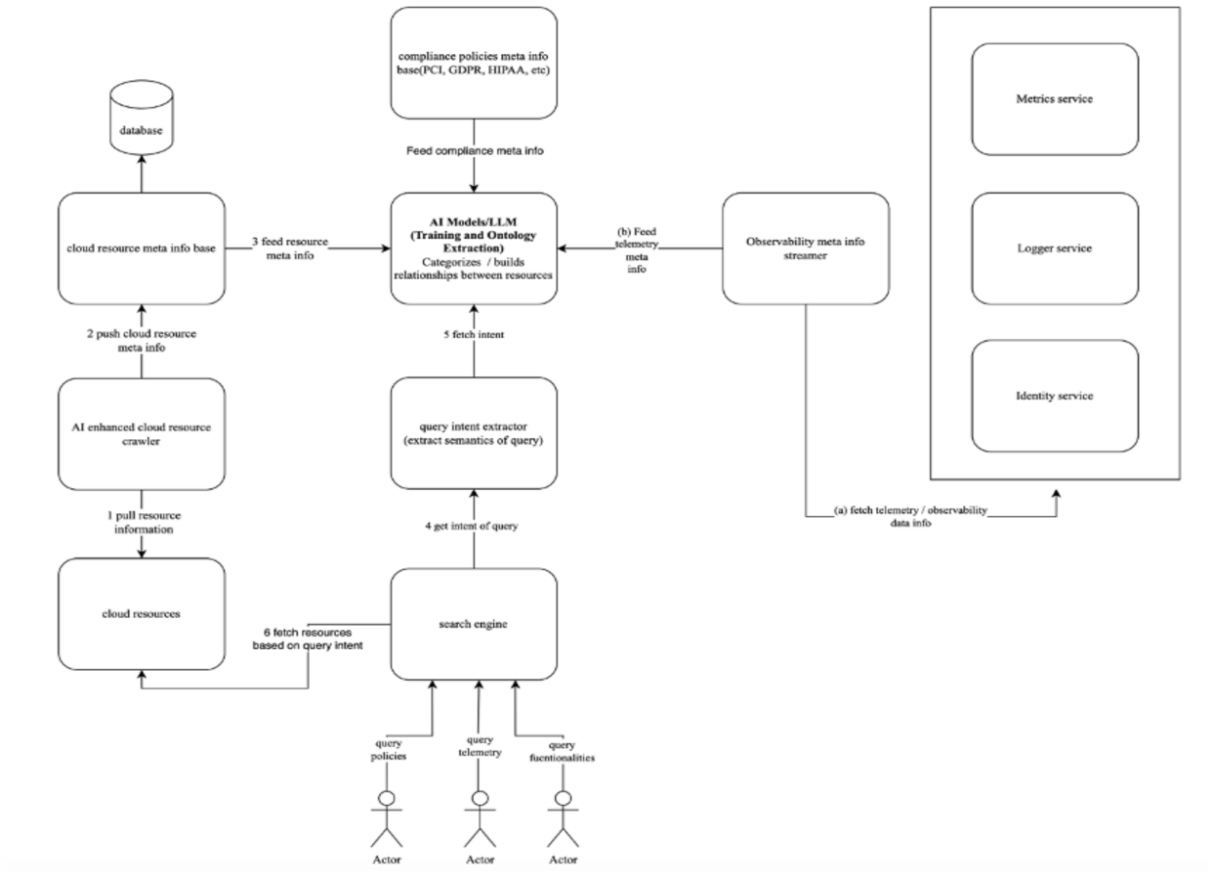

Рис. 4. Інтеграція NLP/ML у безпеку хмарної інфраструктури [17]

Захист даних і DLP

Захист даних у хмарних середовищах, особливо в мультихмарній архітектурі, потребує не лише контролю доступу, а й механізмів виявлення і запобігання витокам інформації (Data Loss Prevention, DLP) (рис. 5).

На думку авторів, традиційні DLP-системи, які ґрунтуються на сигнатурах та ручному налаштуванні правил, виявляються малоефективними в умовах динамічного переміщення, реплікації та шифрування даних у хмарі. І саме інтеграція ML та NLP дозволяє реалізувати інтелектуальні DLP-рішення, здатні до адаптації та самонавчання.

У архітектурному рішенні [19] детально розглянуто стратегії AI-орієнтованого DLP для мультихмарних середовищ. Автоматизоване сканування, класифікація даних, контекстна оцінка ризику та моніторинг дій користувачів здійснюються в режимі реального часу. NLP-алгоритми дозволяють ідентифікувати конфіденційні дані навіть у неструктурованих джерелах, зокрема в текстах повідомлень, документах або логах. Одним із ключових елементів таких систем є фазова обробка: спершу дані ідентифікуються, класифікуються за рівнем чутливості, після чого запускаються ML-моделі для прогнозування ймовірності витоку на основі поведінкових ознак.

Наведемо приклад: якщо користувач починає копіювати велику кількість даних із захищених ресурсів на незареєстровані пристрої або виводить їх через нестандартні канали, система блокує дію і повідомляє команду безпеки.

Дослідження [20] описує практичну реалізацію DLP з використанням гібридної моделі: NLP модуль виявляє PII (особисті ідентифікаційні дані) у документах, а ML-модуль навчається на попередніх інцидентах і сигналізує про ризикову поведінку. У результаті вдалось зменшити кількість витоків на понад 40% у порівнянні з традиційною DLP-системою в аналогічному середовищі.

Також важливим напрямом є відповідність нормативним вимогам (GDPR, HIPAA, ISO/IEC 27001). Рішення, описане в дослідженні [21], інтегрує механізми постійного моніторингу за допомогою ML для динамічного оновлення політик відповідно до змін у нормативній базі.

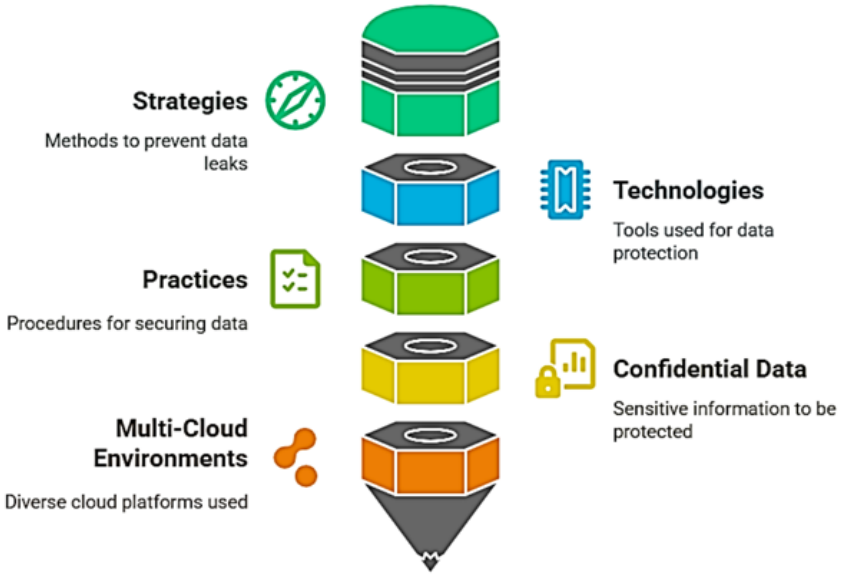

Рис. 5. Компоненти DLP у мультихмарному середовищі [19]

Можемо зробити висновок, що інтеграція NLP та ML у DLP-модулі дозволяє забезпечити адаптивний, контекстно-залежний захист даних із мінімальним втручанням людини, що особливо важливо в умовах високої динаміки хмарних інфраструктур.

Інтелектуальні системи контролю доступу

Контроль доступу є базовим елементом архітектури інформаційної безпеки, особливо в умовах хмарних середовищ, де ресурси масштабуються динамічно, а користувачі можуть отримувати доступ з будь-якої точки світу. Традиційні моделі контролю доступу, зокрема RBAC (рольова модель) або ACL (списки контролю доступу), не здатні оперативно враховувати контекст дій, шаблони поведінки або рівень ризику доступу. Тому дедалі більшої популярності набувають інтелектуальні системи доступу, які інтегрують технології ML та NLP [21].

У статті [22] підкреслено важливість динамічного, контекстно-залежного контролю, який адаптується до поточних умов користувача - його ролі, пристрою, геолокації, типу даних і характеру запиту. Такі системи використовують ML-моделі для побудови поведінкових профілів, які постійно оновлюються на основі активності користувача.

NLP, у свою чергу, застосовується для аналізу запитів у природній мові, наприклад, при зверненнях до чатботів технічної підтримки, де можлива спроба непрямого отримання конфіденційної інформації. Алгоритми NLP допомагають виявити приховані наміри у тексті запиту та заблокувати доступ до критичних даних (рис. 6).

Авторами публікації [23] розглянуто приклад розширеного контролю доступу у Smart Grid-інфраструктурі, де ML використовується для прогнозування потенційних зловживань, а правила доступу генеруються або уточнюються на основі даних про інциденти. Таке рішення дозволяє системі не лише реагувати, а й передбачати зміну прав доступу ще до виникнення загрози.

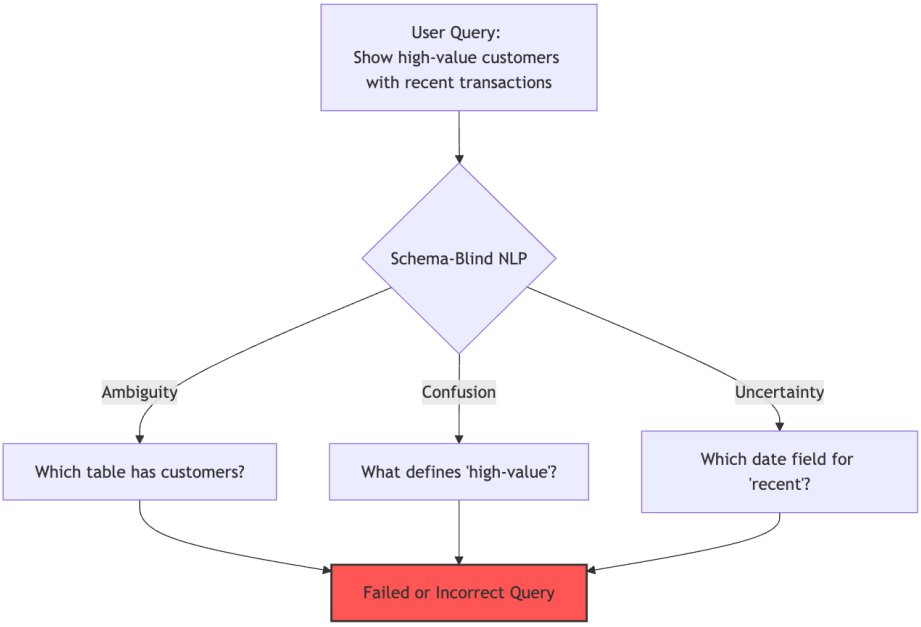

Рис. 6. Контроль доступу до критичних даних за допомогою NLP

У дослідженні [24] описано інтеграцію механізмів ідентифікації та авторизації з елементами штучного інтелекту. Наприклад, система IAM оцінює ризик доступу в реальному часі: якщо запит містить невідповідності (новий пристрій, незвичний час доби, підозрілий маршрут мережі), ML-модель може заборонити доступ або запросити додаткову автентифікацію. Такий підхід відомий як Adaptive Access Control або Risk-Based Access Control, і він лежить в основі концепції Zero Trust – коли жоден запит не вважається автоматично довіреним, навіть з внутрішньої мережі.

Завдяки впровадженню ML та NLP інтелектуальні системи доступу забезпечують не лише гнучкість і контекстну обізнаність, а й проактивність у запобіганні несанкціонованому доступу, що є критично важливим для захисту розподілених хмарних середовищ.

Приклади застосування у провідних компаніях та системах

Впровадження технологій ML та NLP у кіберзахист хмарної інфраструктури вже не є винятком, а стало пріоритетним напрямом розвитку IT-безпеки у провідних компаніях світу. Практичні реалізації таких підходів демонструють значну ефективність у виявленні загроз, захисті даних та автоматизації рутинних процесів (табл. 1).

Amazon Web Services (AWS) реалізувала низку сервісів, таких як Amazon Macie, що використовує ML для виявлення PII у сховищах Amazon S3, та Amazon GuardDuty, який виявляє загрози шляхом аналізу журналів подій, мережевого трафіку та поведінки користувачів.

У Google Cloud платформа Chronicle забезпечує виявлення атак на основі поведінкових аномалій. Вона інтегрує великомасштабну обробку даних із моделями ML, які аналізують мільярди подій щодня, виявляючи шаблони поведінки, що притаманні цілеспрямованим атакам [13].

У роботі [5] подано приклад автоматизованої платформи для тестування безпеки Cloud RAN на основі ML/NLP. Система автоматично формує сценарії тестування загроз, інтерпретує лог-файли та визначає слабкі місця в архітектурі на основі попереднього навчання.

Таблиця 1

Порівняння сценаріїв застосування ML для забезпечення безпеки та відповідності серед провайдерів хмарних сервісів [25]

Cloud Service	**ML Application**	**Key Features**	**Compliance Benefits**
Microsoft Azure	Azure Sentinel	Real-time anomaly detection	Proactive threat mitigation
Google Cloud	DLP API	Data classification and analysis	GDPR and CCPA compliance
Amazon Web Services	AWS GuardDuty	Automated threat detection	Industry-standard compliance
Financial Services	Fraud Detection	Transaction monitoring	AML compliance and fraud prevention
Healthcare	Anomaly Detection	Patient data protection	HIPAA compliance and data privacy

Компанії, що працюють у галузі фінансових технологій, активно впроваджують ML у модулі поведінкового контролю доступу, адаптивної аутентифікації та оцінки ризиків. Такі рішення реалізовано, зокрема, у платформах Oracle Cloud та Microsoft Azure, де ML використовується у Azure Sentinel – SIEM-платформі з підтримкою аналітики подій безпеки на основі штучного інтелекту [7].

У статті [14] показано кейс впровадження гібридної системи захисту даних на основі NLP, Blockchain і Smart Contracts. Таке поєднання технологій дозволило забезпечити контроль доступу до критичних даних і виявляти спроби маніпуляцій або витоків у децентралізованому середовищі.

Інтеграція моделей виявлення з хмарними платформами також активно розвивається. У [17] проаналізовано сервіс Amazon GuardDuty, який використовує машинне навчання для виявлення спроб сканування портів, змін у шаблонах мережевих викликів, підозрілих звернень до API та неодноманітності в логах авторизації. Завдяки інтеграції з CloudTrail та VPC Flow Logs, GuardDuty формує аналітичну базу для адаптивного контролю доступу і автоматичного реагування (наприклад, ізоляції інстансу).

Ключовою перевагою ML-підходів у виявленні загроз є можливість працювати із неструктурованими, динамічними та непередбачуваними сценаріями, які не піддаються ручному опису. Це дозволяє не лише виявляти загрози, а й розставляти пріоритети в їх обробці (threat scoring), знижуючи навантаження на аналітиків SOC та покращуючи швидкодію реагування.

Водночас, ряд досліджень [21], [25] звертають увагу на обмеження: моделі можуть страждати на ефект перенавчання, потребують постійного оновлення наборів даних, а іноді – й пояснення висновків, що ускладнює впровадження в організаціях, де прийняття рішень має бути обґрунтованим. У таких випадках актуальною є інтеграція з методами XAI (explainable AI), які дозволяють деталізувати причину того чи іншого попередження.

Інтеграція ML/NLP-технологій у процеси виявлення вразливостей демонструє значну перевагу порівняно з традиційними та ручними методами. Як видно з рис. 7, рівень виявлення загроз за допомогою NLP, що є складовою AI, та ML значно перевищує результати ручного аналізу та оцінювання на основі CVSS. Наприклад, при виявленні SQL-ін’єкцій, ML та NLP досягають понад 78% точності, у той час як ручне виявлення — лише близько 60%, а фреймворк загальної системи оцінки вразливостей (Common Vulterability Scoring System,

CVSS) — не перевищує 10%. Схожі тенденції спостерігаються і для XSS, CSRF та небезпек із вкладеними файлами (File Inclusion) — у всіх випадках NLP/ML-детекція стабільно демонструє найвищі показники.

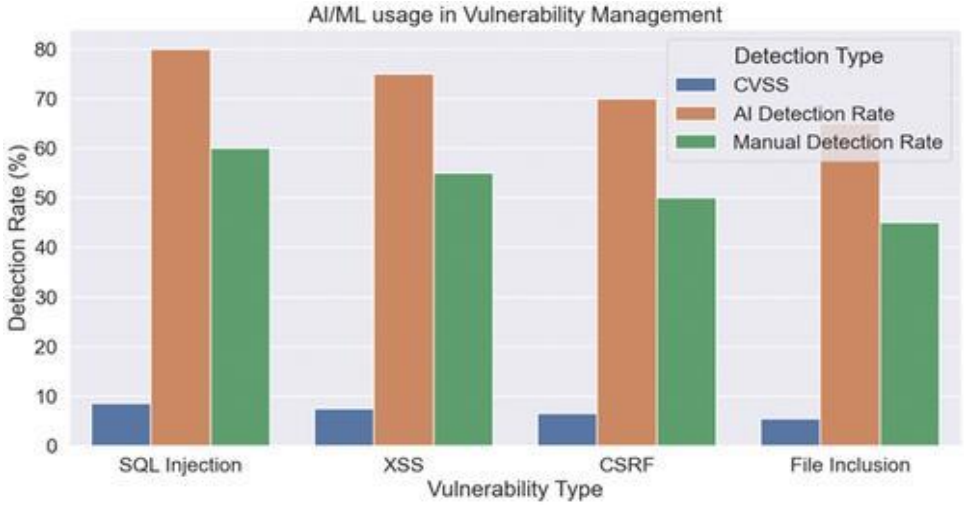

Рис. 7. Співвідношення можливостей розпізнавання людиною та AI/ML [26]

Ці результати підкреслюють ефективність ML/NLP у виявленні складних типів вразливостей, які часто залишаються поза увагою при ручному перегляді або стандартному скорингу. Крім того, використання цих технологій дозволяє оперативно аналізувати великий обсяг даних у реальному часі, що є особливо важливим у мультихмарному середовищі з високим рівнем автоматизації. У цьому контексті впровадження ML/NLP-моделей у системи управління вразливостями не тільки підвищує якість захисту, а ще і знижує навантаження на аналітиків, які раніше змушені були вручну переглядати численні журнали та сигнали безпеки.

Можемо зробити висновок, що провідні хмарні платформи та великі корпорації демонструють широкий спектр застосувань інтелектуальних технологій, які поєднують ефективність, масштабованість і адаптивність до сучасних викликів у сфері кіберзахисту.

Аналіз проблем, обмежень і напрямів вдосконалення

Попри стрімкий розвиток інтелектуальних рішень у сфері хмарної безпеки, застосування технологій NLP та ML супроводжується низкою викликів. Більшість існуючих підходів демонструють високі результати в лабораторних умовах або на синтетичних даних, проте на практиці стикаються з проблемами масштабованості, адаптації до нових загроз, а також дотриманням етичних та правових норм. Важливим аспектом є також потреба у забезпеченні точності моделей при обмеженості навчальних даних або у разі появи zero-day атак.

Проблеми точності моделей

Одна з головних проблем у застосуванні ML та NLP для захисту хмарної інфраструктури полягає в забезпеченні стабільно високої точності моделей при реальному використанні. Попри вражаючі результати на контрольованих наборах даних, моделі часто демонструють зниження якості на нових або непередбачуваних даних, зокрема в умовах змінної топології хмарної архітектури, появи нових типів атак або зміни поведінки користувачів [3, 6].

У дослідженні [5] наголошується, що ефективність навіть добре навчених моделей значно знижується в разі появи нестандартних форматів журналів подій або фрагментів тексту, які не містять ключових слів, що були визначені під час навчання. Це є критичною проблемою для NLP-підходів, які опираються на семантичну інтерпретацію повідомлень без урахування прихованих контекстів.

Крім того, автори [12] вказують на ефект перенавчання, коли моделі добре працюють на відомих шаблонах атак, але виявляють нові загрози з низькою ймовірністю. Нажаль, це

створює оманливе враження ефективності в тестовому середовищі, і не гарантує реальної користі у виробничих системах.

Ще однією проблемою є дисбаланс класів: дані про інциденти безпеки мають суттєво меншу репрезентацію порівняно зі звичайною (нормальною) активністю. Як наслідок, моделі можуть недооцінювати рідкісні, але критичні інциденти. У статті [14] запропоновано використання методів перенавчання, зваженого навчання та підходів SMOTE для вирішення цієї проблеми, але автори вказують на обмежену ефективність без якісного ручного тюнінгу.

Проблеми також виникають при генерації інтерпретованих результатів: багато ML-моделей, зокрема deep learning, є «чорними скриньками», і без належного пояснення їх висновків аналітики не можуть ефективно використати результат. У відповідь на це, як зазначено авторами у [12], почали впроваджуватись Explainable AI (XAI)-рішення, проте вони часто спрощують складну логіку до рівня, недостатнього для прийняття відповідальних рішень.

Незважаючи на значний прогрес, точність моделей залишається обмежувальним чинником при впровадженні NLP/ML у хмарну безпеку. Надалі необхідно поєднувати методи машинного навчання з механізмами валідації, адаптації та інтерпретації, аби забезпечити надійну, перевірену і стабільну продуктивність у динамічному середовищі.

Етичні та правові аспекти

Інтеграція ML та NLP у системи кібербезпеки хмарної інфраструктури супроводжується не лише технічними, а й значними етичними та правовими викликами. З одного боку, ці технології надають нові можливості для виявлення загроз і підвищення ефективності захисту, з іншого – породжують ризики зловживань, упередженості моделей та порушення прав людини.

Однією з ключових етичних проблем є відсутність прозорості в роботі інтелектуальних систем. Як зазначено у [15], складні моделі (особливо на основі глибокого навчання) часто не мають інтерпретованих механізмів прийняття рішень. У контексті кібербезпеки це може призводити до санкціонування дій без чіткої аргументації або пояснення, що суперечить принципам підзвітності.

Ще одним етичним викликом є упередженість алгоритмів (algorithmic bias). Наприклад, моделі можуть враховувати фактори, які непрямо корелюють із дискримінаційними ознаками, що призводить до нерівного трактування користувачів або фальшивих звинувачень у ризикованій поведінці. У дослідженні [15] підкреслюється необхідність використання fairness-aware ML-технологій, здатних виявляти та усувати упередженість у даних і моделях.

З правової точки зору, особливу увагу привертає питання відповідності нормативним актам, зокрема GDPR, CCPA, ISO/IEC 27001. Використання NLP для аналізу текстових повідомлень, листування або логів потребує суворого контролю над приватністю та збереженням персональних даних [6]. Згідно з нормами GDPR, навіть частковий аналіз особистої інформації без згоди користувача є порушенням, а отже системи мають реалізовувати вбудовані механізми обмеження доступу до персональних даних (Personally Identifiable Information, PII).

У статті [14] розглянуто підхід до захисту даних на основі смарт-контрактів і блокчейн-технологій, що дозволяє формалізувати правила доступу до даних, зафіксувати згоду користувача та забезпечити відстежуваність змін. Це є прикладом архітектури, яка одночасно відповідає принципам «приватність за замовченням» і вимогам до автоматизованого обліку взаємодії з критичними даними.

Також важливим аспектом є використання ML/NLP у автоматизованих рішеннях, що впливають на права або свободи користувачів – наприклад, блокування акаунтів, ізоляція доступу до ресурсів, ініціювання дій з кіберзахисту без участі людини. У таких випадках відповідно до регуляторних норм повинна бути можливість апеляції, доступ до пояснення рішень та людське втручання у процес прийняття остаточного рішення [15]. Однозначно, що

© Абібулаєв А.Р., Піскозуб А.З. (2025). Аналіз можливостей покращення стану безпеки хмарної інфраструктури за допомогою NLP та ML. Сучасний захист інформації, 2(62), 124–140.
https://doi.org/10.31673/2409-7292.2025.026884

розробка та впровадження інтелектуальних систем кіберзахисту має супроводжуватись не лише технічною валідністю, але й дотриманням етичних принципів, прозорості, захисту персональних даних та правової відповідності.

Потреба в узгодженні з Zero Trust / IAM / CIEM

Інтеграція інтелектуальних систем на основі NLP та ML у хмарну безпеку не може бути повноцінною без узгодження з сучасними концепціями управління доступом і довірою. Зокрема, мова йде про принцип Zero Trust Architecture (ZTA), моделі IAM (Identity and Access Management) та більш динамічний підхід CIEM (Cloud Infrastructure Entitlement Management).

Zero Trust передбачає відсутність апріорної довіри до будь-якого користувача або процесу, незалежно від розташування в мережі. У цьому контексті ML-алгоритми відіграють критичну роль у динамічній перевірці контексту доступу, оцінці ризику в реальному часі та прийнятті рішення про авторизацію [6 ,13].

У роботі [7] підкреслюється, що традиційні IAM-системи часто не справляються з постійними змінами в хмарному середовищі – нові сервіси, тимчасові користувачі, зовнішні інтеграції. Тут на допомогу приходить CIEM, який, використовуючи ML, дозволяє постійно переглядати права доступу, виявляти надлишкові привілеї та створювати політики least privilege на основі поведінки.

Особливої актуальності узгодження з Zero Trust набуває при масштабному впровадженні автоматизованих засобів виявлення та реагування на загрози, побудованих на ML/NLP. Без прив'язки до механізмів ідентифікації та контролю доступу такі рішення можуть виявлятися неефективними або навіть небезпечними – наприклад, за відсутності перевірки запиту користувача або аутентифікації [4].

У дослідженні [12] описано механізм об'єднання моделей ризику з IAM-рішеннями. Якщо рівень ризику запиту є високим (на основі ML), система ініціює перевірку другорядними засобами (MFA, або запит до адміністратора). Таким чином, ML працює не ізольовано, а в координації з політиками безпеки. Узгодження з CIEM також дозволяє впроваджувати автоматичну ревізію прав доступу, що актуально для мультихмарних середовищ, де ручне управління привілеями стає недоцільним. ML-моделі аналізують історію доступу, визначають критичні відхилення і пропонують зміни в конфігурації.

Розуміємо, що для досягнення високої ефективності та узгодженості систем кіберзахисту, заснованих на ML/NLP, необхідно інтегрувати їх у рамки Zero Trust, IAM та CIEM - як з точки зору політик, так і інтерфейсів передачі даних і прийняття рішень.

Висновки

У статті здійснено системний аналіз сучасних підходів до застосування технологій обробки природної мови (NLP) та машинного навчання (ML) у контексті забезпечення кібербезпеки хмарної інфраструктури. Проведене дослідження показало, що впровадження NLP/ML у сферу хмарної безпеки відкриває нові можливості для автоматизованого виявлення загроз, адаптивного контролю доступу, захисту чутливої інформації та загальної оптимізації процесів кіберзахисту.

На основі критичного аналізу літератури, практичних кейсів та існуючих архітектурних рішень зроблено такі ключові висновки:

1. NLP ефективно використовується для обробки логів, повідомлень, політик доступу та інших текстових даних з метою виявлення загроз, індикаторів компрометації та соціотехнічних шабонів.
2. ML забезпечує виявлення аномалій, класифікацію інцидентів, поведінковий аналіз та оцінку ризиків, зокрема в контексті zero-day атак, фішингу, DLP та контролю доступу.
3. Успішна інтеграція NLP/ML-рішень спостерігається в багатьох провідних хмарних платформах — AWS, Google Cloud, Microsoft Azure — де такі технології стали складовою частиною автоматизованих засобів виявлення та реагування на загрози.

4. Інтелектуальні DLP-системи, засновані на NLP/ML, демонструють підвищену здатність виявляти витоки чутливої інформації в мультихмарному середовищі.

5. Системи адаптивного контролю доступу, зокрема IAM і CIEM, значно виграють від інтеграції з ML/NLP — забезпечується динамічна оцінка ризику запитів, автоматичне виявлення надлишкових привілеїв, генерація рекомендацій на основі поведінкових шаблонів.

6. Застосування NLP/ML супроводжується низкою обмежень: зниженням точності, перенавчанням моделей, неможливістю пояснити рішення, упередженістю алгоритмів, необхідністю дотримання етичних норм та правової відповідності.

7. Концепція Zero Trust вимагає глибокої інтеграції з NLP/ML-рішеннями, зокрема в частині побудови політик доступу, адаптивної аутентифікації та автоматизованої ревізії прав.

8. Узгодження з підходами типу «Security as Code» та підтримка Explainable AI дозволяють забезпечити прозорість і контрольованість процесів кіберзахисту.

9. Інтеграція ML/NLP-технологій у процеси виявлення вразливостей демонструє значну перевагу порівняно з традиційними та ручними методами – SQL-ін’єкції, XSS, CSRF та небезпеки із вкладеними файлами (File Inclusion) визначаються на 20-30% точніше за допомогою NLP/ML.

Отже, майбутнє кібербезпеки хмарних систем значною мірою залежить від ефективності впровадження інтелектуальних рішень, заснованих на NLP і ML. Подальші дослідження мають бути зосереджені на покращенні точності, зниженні енерго-затратності моделей, побудові етичних та прозорих архітектур, які можна масштабувати у великі середовища. Особливу увагу слід приділяти координації між ML/NLP-модулями та системами управління доступом, а також дотриманню вимог до приватності та аудиту рішень.

Перелік посилань

1. K.C. Sunkara, K. Narukulla, AI Enhanced Ontology Driven NLP for Intelligent Cloud Resource Query Processing Using Knowledge Graphs, Independent Research Report, IEEE Senior Members, Raleigh/San Jose, USA (2023). doi: 10.48550/arXiv.2502.18484.

2. Rajendra Muppalaneni, Anil Chowdary Inaganti and Nischal Ravichandran, AI-Enhanced Data Loss Prevention (DLP) Strategies for Multi-Cloud Environments, Journal of Computing Innovations and Applications, 2(2), pp. 1–13. (2024). Available at: https://ciajournal.com/index.php/jcia/article/view/9 (Accessed: 10 May 2025).

3. Jaya J. Application of Deep Learning in Cloud Security. Deep Learning Approaches to Cloud Security. (2022). doi: 10.1002/9781119760542.ch12

4. J.S. Nimbhorkar, AI Enabled Cloud RAN Test Automation: Automatic Test Case Prediction Using Natural Language Processing and Machine Learning Techniques, M.Sc. Thesis, KTH Royal Institute of Technology, Ericsson AB, Stockholm (2023). URN: urn:nbn:se:kth:diva-340090

5. T.K. Vashishth, V. Sharma, B. Kumar, S. Chaudhary, R. Panwar, Enhancing Cloud Security: The Role of Artificial Intelligence and Machine Learning, In: IGI Global, Chapter 4 (2024). doi: 10.4018/979-8-3693-1431-9.ch004.

6. R.K. Jha, Strengthening Smart Grid Cybersecurity: An In-Depth Investigation into the Fusion of Machine Learning and Natural Language Processing, J. Trends Comput. Sci. Smart Technol. 5(3) (2023) 284–301. doi: 10.36548/jtcsst.2023.3.005.

7. Y.I. Alzoubi, A. Mishra, A.E. Topcu, Research trends in deep learning and machine learning for cloud computing security, Artif. Intell. Rev. 57 (2024) 132. doi: 10.1007/s10462-024-10776-5.

8. Martseniuk, Y., Partyka, A., Harasymchuk, O., Nyemkova, E., Karpinski, M. Shadow IT risk analysis in public cloud infrastructure (2024) CEUR Workshop Proceedings, 3800, pp. 22-31. URN: urn:nbn:de:0074-3800-2.

9. Martseniuk, Y., Partyka, A., Harasymchuk, O., Shevchenko, S. Universal centralized secret data management for automated public cloud provisioning (2024) CEUR Workshop Proceedings, 3826, pp. 72-81. URN: urn:nbn:de:0074-3826-1.

10. Volodymyr Khoma, Aziz Abibulaiev, Andrian Piskozub, and Taras Kret. Comprehensive Approach for Developing an Enterprise Cloud Infrastructure (2024) CEUR Workshop Proceedings, 3654, pp. 201-215. URN: urn:nbn:de:0074-3654-7.

11. S.R. Mamidi, The Role of AI and Machine Learning in Enhancing Cloud Security, J. Artif. Intell. Gen. Sci. 3(1) (2024). doi: 10.5281/zenodo.10987665.

12. J. Wang, AI/ML-Powered Cybersecurity and Cloud Computing Strategies for Optimized Business Intelligence in ERP Cloud, ResearchGate (2023). doi: 10.13140/RG.2.2.27926.66882.

13. K. Rangappa, A.K.B. Ramaswamy, M. Prasad, S.A. Kumar, A Secure Cloud Service for Managing User's Crucial Data Using NLP, Blockchain, and Smart Contracts, Preprints.org (2024). doi: 10.20944/preprints202409.1738.v1.

14. Buttar AM, Shahzad F, Jamil U. Conversational AI: Security Features, Applications, and Future Scope at Cloud Platform. Conversational Artificial Intelligence, (2024). doi: 10.1002/9781394200801.ch3.

15. T.-M. Georgescu, Natural Language Processing Model for Automatic Analysis of Cybersecurity-Related Documents, Symmetry 12(3) (2020) 354. doi: 10.3390/sym12030354.

16. Belal MM, Sundaram DM. Comprehensive review on intelligent security defences in cloud: Taxonomy, security issues, ML/DL techniques, challenges and future trends. Journal of King Saud University-Computer and Information Sciences. (2022). doi: 10.1016/j.jksuci.2022.08.035.

17. J. Wang, AI/ML-Powered Cybersecurity and Cloud Computing Strategies for Optimized Business Intelligence in ERP Cloud, ResearchGate (2023). doi: 10.13140/RG.2.2.27926.66882.

18. Nina P, Ethan K. AI-driven threat detection: Enhancing cloud security with cutting-edge technologies. International Journal of Trend in Scientific Research and Development, Volume-4, pp.1362-1374. (2019). Available at: https://www.ijtsrd.com/papers/ijtsrd29520.pdf (Accessed 12 May 2025).

19. Z. Kilhoffer and M. Bashir, Cloud Privacy Beyond Legal Compliance: An NLP Analysis of Certifiable Privacy and Security Standards, IEEE Cloud Summit, Washington, DC, USA, pp. 79-86, (2024). doi: 10.1109/Cloud-Summit61220.2024.00020.

20. Sunkara KC, Narukulla K. AI Enhanced Ontology Driven NLP for Intelligent Cloud Resource Query Processing Using Knowledge Graphs, (2025). doi: 10.48550/arXiv.2502.18484.

21. Mamidi SR. The Role of AI and Machine Learning in Enhancing Cloud Security. Journal of Artificial Intelligence General science (JAIGS), (2024). doi: 10.60087/jaigs.v3i1.161.

22. D. M. Rakgoale, H. I. Kobo, Z. Z. Mapundu and T. N. Khosa, A Review of AI/ML Algorithms for Security Enhancement in Cloud Computing with Emphasis on Artificial Neural Networks, 4th International Multidisciplinary Information Technology and Engineering Conference (IMITEC), Vanderbijlpark, South Africa, pp. 329-336, (2024). doi: 10.1109/IMITEC60221.2024.10851076.

23. Talati, N. D. V., Scalable AI and data processing strategies for hybrid cloud environments, World Journal of Advanced Research and Reviews, 10(3), pp. 482–492, (2021), doi: 10.30574/wjarr.2021.10.3.0289.

24. Al Saidat MR, Yerima SY, Shaalan K. Advancements of SMS Spam Detection: A Comprehensive Survey of NLP and ML Techniques. Procedia Computer Science, (2024). doi: 10.1016/j.procs.2024.10.198.

25. H. Aldawsari, S.A. Kouchay, Integrating AI and Machine Learning Algorithms in Cloud Security Frameworks for Enhanced Proactive Threat Detection and Mitigation, J. Eng. Technol. Manag. 74 (2024). Available at: https://ciajournal.com/index.php/jcia/article/view/9 (Accessed: 11 May 2025).

26. Mohamed, N., Current trends in AI and ML for cybersecurity: A state-of-the-art survey. Cogent Engineering, 10(2), (2023). doi: 10.1080/23311916.2023.2272358.

Надійшла 18.05.2025