Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра кібербезпеки
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр
(назва освітнього ступеня)

на тему: Використання штучного інтелекту для ефективного реагування на інциденти у SIEM системи

Виконав: студент VI курсу, групи СБм-61
спеціальності 125 Кібербезпека
(шифр і назва спеціальності)

(підпис) *Кубарич З.П.* (прізвище та ініціали)

Керівник (підпис) *Скарга-Бандурова І.С.* (прізвище та ініціали)

Нормоконтроль (підпис) *Лечаченко Т. А.* (прізвище та ініціали)

Завідувач кафедри (підпис) *Загородна Н.В.* (прізвище та ініціали)

Рецензент (підпис) (прізвище та ініціали)

Тернопіль
2023

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра Кібербезпеки
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Загородна Н.В.
(підпис) (прізвище та ініціали)
«____» _____ 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

Студенту Кубаричу Захару Петровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Використання штучного інтелекту для ефективного реагування на інциденти у SIEM системі

Керівник роботи Скарга-Бандурова Інна Сергіївна, д.т.н., проф., проф. кафедри КБ
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «_16_» _листопада__ 2023 року № _4/7-1061_____

2. Термін подання студентом завершеної роботи 22 грудня 2023р.

3. Вихідні дані до роботи

4. Зміст роботи (перелік питань, які потрібно розробити): Вступ, 1 Штучний інтелект в SIEM системах: стан і перспективи, 1.1 Аналіз існуючого стану SIEM систем з інтеграцією інтелекту, 1.2 Оцінка потреб і викликів для сучасних систем безпеки, 1.3 Розгляд потенціалу Штучного інтелекту в контексті реагування на інциденти, 1.4 Висновки до першого розділу, 2 Теоретичні основи ШІ в системи кібербезпеки, 2.1 Огляд ключових теорій штучного інтелекту, що застосовуються у SIEM системах, 2.2 Дослідження переваг і можливих ризиків Застосування ШІ в кібербезпеці, 2.3 Аналіз методів машинного навчання для вдосконалення роботи SIEM систем, 2.4 Огляд перспектив використання великих мовних моделей в SIEM, 2.5 Висновки до другого розділу, 3 Практичне застосування штучного інтелекту в SIEM, 3.1 Проектування моделі ШІ для оптимізації реагування на інциденти, 3.1.1 Постановка задачі, 3.1.2 Налаштування системи Wazuh, 3.1.2.1 Особливості системи Wazuh, 3.1.2.2 Конфігурація Wazuh, 3.1.3 Реалізація доступу до інтерфейсу керування Wazuh, 3.1.4 Налаштування скрипту для автоматичного сканування мережі за допомогою утиліти Nmap, 3.1.5 Інтеграція ChatGPT, 3.2 Виявлення сканування портів за допомогою ChatGPT, 3.2.1 Запит на фільтацію даних в контексті пошуку подій в SIEM, 3.2.2 Запит SPLUNK, 3.2.3 Запит в Wazuh, 3.4 Обмеження використання ChatGPT в SIEM, 3.4 Обмеження використання ChatGPT в SIEM, 3.5 Висновки до третього розділу, 4 Безпека життєдіяльності, Основи охорони праці, 4.1 Охорона праці, 4.2 Шкідливий вплив іонізуючого випромінювання, 4.3 Висновки до четвертого розділу, Висновок, Перелік використаних джерел, Додатки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Осухівська Г.М., к.т.н., доцент		
Безпека в надзвичайних ситуаціях	Клепчик В.М., старший викладач з адміністративно-господарської роботи та будівництва		

7. Дата видачі завдання 2 листопада 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	16.11.2023	*Виконано*
2.	Підбір наукових джерел ШІ в SIEM системах	17.11.2023-20.11.2023	*Виконано*
3.	Переклад та опрацювання наукових джерел про		
	ШІ в SIEM системах	21.11.2023-23.11.2023	*Виконано*
4.	Виконання дослідження щодо аналізу існуючого стану		
	SIEM систем з ШІ	24.11.2023-27.11.2023	*Виконано*
5.	Оформлення розділу «Штучний інтелект в SIEM		
	системах: стан і перспективи»	28.11.2023-30.11.2023	*Виконано*
6.	Оформлення розділу «Теоретичні основи інтеграції		
	ШІ в системи кібербезпеки»	01.12.2023-02.12.2023	*Виконано*
7.	Оформлення розділу «Практичне застосування	03.12.2023-08.12.2023	*Виконано*
	Штучного інтелекту в SIEM системах»		
8.	Виконання завдання до підрозділу «Охорона праці»	09.12.2023-10.12.2023	*Виконано*
9.	Виконання завдання до підрозділу «Безпека в надзви-		
	чай них ситуаціях»	10.12.2023-11.12.2023	*Виконано*
10.	Оформлення кваліфікаційної роботи	14.12.2023-15.12.2023	*Виконано*
11.	Нормоконтроль	20.12.2023-21.11.2023	*Виконано*
12.	Перевірка на плагіат	12.12.2023	*Виконано*
13.	Попередній захист кваліфікаційної роботи	.12.2023	
14.	Захист кваліфікаційної роботи	26.12.2023	

Студент _____ Кубарич З. П.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Скарга-Бандурова І.С
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Використання штучного інтелекту для ефективного реагування на інциденти у SIEM системі // Кваліфікаційна робота освітнього рівня «Магістр» // Кубарич Захар Петрович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра кібербезпеки, група СБм-61 // Тернопіль, 2023 // С. 98, рис. – 10, табл. – 3, додат. – 3, бібліогр. – 36.

КЛЮЧОВІ СЛОВА: КІБЕРБЕЗПЕКА, SIEM, ШТУЧНИЙ ІНТЕЛЕКТ,

Штучний інтелект розширює можливості SIEM-систем в управлінні інформацією та подіями безпеки, пропонуючи безпрецедентну автоматизацію, точність і швидкість.

Традиційні заходи безпеки еволюціонують, а штучний інтелект надає можливості для проактивного виявлення загроз і реалізації прогностичних моделей безпеки.

У дослідженні були розглянуті основні теорії штучного інтелекту, які можуть бути застосовні до SIEM систем і підкреслено баланс між використанням можливостей штучного інтелекту та усуненням пов'язаних з ним ризиків, таких як: якість даних і конфіденційність.

Було продемонстровано процес впровадження мовленнєвої моделі штучного інтелекту до SIEM системи для надання інформації про наявні загрози у системі.

1

0

3

3LLM

3

6

ЗМІСТ

ВСТУП

Актуальність роботи. У сучасному світі, де кібербезпека стає все більш важливою в умовах зростаючих кіберзагроз, значення ефективного управління інформаційною безпекою набуває особливої ваги. Системи управління інформацією та подіями безпеки (SIEM) відіграють ключову роль у забезпеченні цієї безпеки, але зі зростанням складності та об'ємів даних, їх ефективність може бути обмежена. В цьому контексті, інтеграція штучного інтелекту (ШІ) і методів машинного навчання (МН) може стати рішенням, яке здатне значно покращити здатність SIEM систем виявляти та реагувати на кіберзагрози в реальному часі. Штучний інтелект може допомогти ідентифікувати складні шаблони поведінки та аномалії, які можуть вказувати на потенційні загрози, тим самим збільшуючи швидкість і точність відповіді на інциденти.

Важливо також зазначити, що інтеграція ШІ і МН у системи SIEM має не тільки технічний, але й стратегічний аспект, оскільки вона впливає на загальну структуру управління безпекою і робочі процеси в організаціях. Ефективне впровадження таких технологій вимагає глибокого розуміння їх потенціалу та викликів, пов'язаних з їх інтеграцією в існуючі системи.

Останнім часом великі мовні моделі (LLM), які стали доступними для широкої аудиторії, привернули значну увагу через їх потенціал широкого застосування. Ці моделі пропонують зручне спілкування за допомогою природної мови і здатні надавати відповіді, які демонструють глибоке розуміння на рівні людського мислення. У сфері кібербезпеки, наприклад, ChatGPT має потенціал стати цінним інструментом, який може поліпшити навчання та комунікаційні можливості щодо кібератак.

Метою роботи є огляд існуючих теорій ШІ, які використовуються в SIEM системах та реалізація інтегрування великої мовленнєвої моделі

C
h
a
t
G
P
T
у SIEM системах шляхом інтеграції LLM у SIEM.

Згідно мети визначено наступні завдання:

– Провести аналіз існуючого стану SIEM систем з інтеграцією ШІ
– Оцінити потреби та виклики для SIEM систем
– Розглянути основні теорії ШІ, що застосовуються в SIEM системах
– Визначити перспективи використання LLM у SIEM системах
– Розробити ефективні інтеграції LLM у SIEM системи для виявлення та реагування на інциденти кібербезпеки.

Об’єкт дослідження: ШІ у поєднанні з SIEM системою.

Предметом дослідження: Використання LLM, а саме ChatGPT, у SIEM системі.

Наукова новизна: Розробка методів використання LLM у SIEM системах, з можливістю подальшого розвитку такого підходу для інших систем.

Практичне значення: Спроектована модель використання ШІ у SIEM системах може бути вдосконалена для відповідності конкретним задачам. Розробка та впровадження нових підходів і технологій на основі LLM, що можуть значно підвищити ефективність SIEM систем для виявлення та протидії кіберзагрозам. Це може включати покращення систем автоматичного моніторингу, аналізу даних безпеки та швидкого реагування на інциденти.

Апробація результатів магістерської роботи: Окремі результати роботи доповідались на XI науково-технічній конференції «Інформаційні моделі, системи та технології», Тернопіль, ТНТУ, 7 – 8 грудня 2023 р.

Публікації: За темою роботи з викладенням її основних результатів опубліковано 1 наукова праця, що являє собою тези в збірнику матеріалів науково-практичних конференцій (див. Додаток А).

ШТУЧНИЙ ІНТЕЛЕКТ В SIEM СИСТЕМАХ: СТАН І ПЕРСПЕКТИВИ

У цьому розділі розглянуто роль штучного інтелекту у системах Security Information and Event Management (SIEM), які є важливими інструментами в сучасному світі кібербезпеки. SIEM системи використовуються для своєчасного виявлення та ефективного реагування на потенційні кіберзагрози і уразливості. Вони здійснюють збір та аналіз великої кількості даних з різних джерел у мережі організації, включаючи сервери, мережеве обладнання та програмне забезпечення. Це дозволяє командам кібербезпеки виявляти аномальну поведінку або незвичайні події, що можуть свідчити про безпекові інциденти. Оглянуто інтеграції штучного інтелекту та машинного навчання у ці системи, які сприяють автоматизації та підвищенню точності виявлення загроз. Розглянуто, як завдяки глибокому аналізу даних та автоматизації процесів SIEM системи допомагають організаціям швидко реагувати на інциденти, тим самим мінімізуючи потенційні ризики та наслідки кібератак.

1.1 Аналіз існуючого стану SIEM систем з інтеграцією штучного інтелекту

— це набір технологій безпеки, призначених для допомоги компаніям у своєчасному розпізнаванні та ефективному усуненні потенційних кіберзагроз і уразливостей. Важливою особливістю таких систем є їхня здатність збирати та аналізувати велику кількість даних з різних джерел у мережі організації, включаючи сервери, мережеве обладнання та програмне забезпечення. Це дозволяє командам кібербезпеки виявляти аномальну поведінку або незвичайні події, які можуть свідчити про безпекові інциденти.

Однією з ключових переваг SIEM є інтеграція штучного інтелекту та машинного навчання, які сприяють автоматизації та підвищенню точності виявлення загроз. Ці технології дозволяють аналізувати поведінкові шаблони користувачів та мережевого трафіку, підвищуючи ефективність виявлення внутрішніх та зовнішніх загроз. Завдяки здатності SIEM до глибокого аналізу даних та автоматизації процесів, організації мають можливість швидко реагувати на інциденти, мінімізуючи потенційні ризики та наслідки кібератак. Таким чином, SIEM є важливим інструментом у сучасному світі кібербезпеки, забезпечуючи компанії необхідними засобами для захисту своїх мереж та даних від різноманітних кіберзагроз.

Традиційні SIEM системи вже давно стали наріжним каменем зусиль з кібербезпеки, допомагаючи консолідувати, співвідносити та аналізувати дані про безпеку з різних джерел. Однак зі зростанням складності кіберзагроз та обсягів даних про безпеку традиційні SIEM системи не встигають за ними. SIEM на основі штучного інтелекту — це вдосконалена форма системи, яка використовує можливості ШІ і МН для вирішення багатьох проблем минулого.

SIEM на основі штучного інтелекту — це технологія, яка не тільки автоматизує складні процеси агрегації та нормалізації даних, але й дозволяє проактивно виявляти загрози та реагувати на них завдяки машинному навчанню та предиктивній аналітиці. Навчаючись на минулих даних і закономірностях безпеки, шучний інтелект у SIEM може прогнозувати і виявляти потенційні загрози ще до того, як вони відбудуться. Крім того, він може автоматизувати процес реагування на інциденти, тим самим мінімізуючи наслідки порушень безпеки. По суті, SIEM системи з ШІ забезпечує інтелектуальний, автоматизований і проактивний підхід до виявлення загроз і реагування на них.

У контексті кібербезпеки агрегація даних означає процес збору даних про безпеку з різних джерел, включаючи мережеві пристрої, сервери, бази

даних, додатки тощо. Ці дані можуть включати журнали, дані про події, розвіддані про загрози та інші типи інформації, пов'язаної з безпекою.

Нормалізація, з іншого боку, полягає в перетворенні цих необроблених даних про безпеку в послідовний, стандартизований формат. Цей процес має вирішальне значення для того, щоб система SIEM могла точно аналізувати і співвідносити дані, незалежно від їхнього джерела. Однак SIEM системи з штучним інтелектом вирізняється своєю здатністю автоматизувати ці процеси. Використовуючи ШІ та МН, система може швидше сортувати дані, а також інтелектуально агрегувати і нормалізувати дані про безпеку, тим самим значно скорочуючи час і зусилля, необхідні для виконання цих завдань.

Машинне навчання та розпізнавання шаблонів дозволяють системам вчитися на основі минулих даних і шаблонів безпеки, що дає змогу виявляти аномалії та потенційні загрози, які традиційні SIEM системи, що покладаються виключно на підписи та критичні моменти з самих джерел журналів, можуть пропустити.

Наприклад, SIEM система з штучним інтелектом може використовувати алгоритми машинного навчання для аналізу історичних даних про безпеку, виявлення закономірностей і тенденцій та створення шаблонів "нормальної" поведінки. Потім вона може безперервно відстежувати поточні дані про безпеку, порівнюючи їх з цім шаблоном, що дозволяє їй виявляти будь-які відхилення або аномалії, які можуть вказувати на потенційну загрозу.

Крім того, завдяки розпізнаванню шаблонів така система може виявляти кореляцію в журналах, пов'язану з відомими загрозами або векторами атак. Ця можливість дозволяє SIEM виявляти потенційні загрози та сповіщати про них майже в режимі реального часу, тим самим значно скорочуючи час на виявлення та реагування.

У разі виявлення загрози або порушення безпеки швидке та ефективне реагування має вирішальне значення для мінімізації наслідків. SIEM на основі штучного інтелекту використовує можливості автоматизації для оптимізації та прискорення процесу реагування на інциденти. Вона може автоматично запускати оповіщення, виконувати заздалегідь визначені дії реагування або навіть організовувати складні робочі процеси реагування. ЕМ на основі штучного інтелекту може надати командам безпеки детальну інформацію про загрозу, яка допоможе їм приймати обґрунтовані рішення та вживати ефективних заходів.

Однією з таких систем, що використовують штучний інтелект є хмарний SIEM в галузі, який представляє SIEM нового масштабу. Він об'єднує комбіновані можливості всіх продуктів Exabeam: хмарне зберігання даних, швидке отримання даних, надшвидке виконання запитів, потужну поведінкову аналітику та автоматизацію, яка змінює спосіб виконання аналітиками своєї роботи. Exabeam Fusion дозволяє аналітикам запускати свої наскрізні робочі процеси TDIR з єдиної площини управління, яка виконує автоматизацію завдань, що вимагають ручної роботи.

Exabeam була заснована у 2013 році і базується у Фостер-Сіті, штат Каліфорнія. Компанія є дітищем піонера кібербезпеки Шломо Крамера, який був одним із засновників Check Point та Imperva. Він також володіє великими частками в Cato Networks і Sumo Logic, серед інших компаній, що займаються кібербезпекою та IT-послугами. Для створення компанії Крамер зібрав команду керівників з Imperva та Sumo Logic. Exabeam є приватною компанією, і Крамер не є її єдиним власником - інвесторами також є низка ключових IT-венчурних фондів. Першим продуктом Exabeam була система UEBA, яка була задумана як надбудова, яку компанії могли придбати для покращення роботи вже встановленої SIEM. У 2017 році компанія

розширила свої послуги і запустила власний SIEM. SIEM позиціонується як платформа для розвідки безпеки наступного покоління.

Користувачі відзначили його легке налаштування і розширену функціональність за розумну ціну. Його похвалили за простоту і зручність у використанні, як комплексний інструмент SIEM і SOAR, що використовується для розслідувань, ведення журналів, звітності про відповідність вимогам, оповіщення та продажу квитків InfoSec.

Серед ключових переваг - швидкий час пошуку, доступ до інтерфейсу командного рядка для усунення несправностей, відповідна ліцензійна модель для поглинання великих обсягів журналів і хороша апаратна підтримка для локального розгортання. Також було відзначено його стабільність і додаткові функції пошуку.

Було підкреслено його простий графічний інтерфейс і швидкий час відгуку, а також можливість зберігати журнали за 7 років з можливістю миттєвого пошуку.

– це хмарне рішення, призначене для автоматизації процесів виявлення загроз і реагування на них, мінімізуючи при цьому втому від оповіщень і помилкові спрацьовування для команд SOC. Воно також пропонує готові звіти для підтримки різних вимог щодо відповідності та аудиту, таких як PCI-DSS, HIPAA, SOX та GDPR.

Основна сила Exabeam полягає в обробці даних, особливо в управлінні інформацією про безпеку (Security Information Management, SIM) в аспекті SIEM. Три основні фази стратегії Exabeam - це Exabeam Data

Серед переваг - підтримка робочих процесів реагування на інциденти, сценаріїв та автоматизації, корисні функції запитів для великих наборів даних та можливості звітування про відповідність вимогам. Однак йому не вистачає можливостей моніторингу мережі в реальному часі, і він не був спочатку розроблений як інструмент SIEM.

Ці огляди в сукупності малюють картину Exabeam Fusion як надійного, настроюваного і зручного рішення SIEM і SOAR. Воно відмінно справляється з обробкою даних і розширеною аналітикою, пропонуючи цінні функції для реагування на інциденти, звітності про дотримання нормативних вимог і мінімізації помилкових спрацьовувань. Хоча він отримує високі оцінки за свої функції запитів і можливості автоматизації, деякі огляди вказують на відсутність моніторингу мережі в реальному часі і на те, що його початковий дизайн не був спеціально розроблений для

Exabeam Fusion використовує технології ШІ та МН, щоб розширити свої можливості як рішення для управління інформацією та подіями безпеки, а також для оркестрування, автоматизації та реагування на події безпеки. Exabeam Fusion використовує ШІ та МН для аналітики поведінки користувачів, автоматичного реагування на інциденти. Завдяки тому що, система постійно навчається зменшуються кількість хибно позитивних спрацювань.

Exabeam Fusion використовує ШІ і МН для аналізу дій користувачів і пристроїв, встановлюючи базові шаблони нормальної поведінки і виявляючи аномалії. Це включає в себе присвоєння балів ризику діям і використання понад 750 гістограм поведінкових моделей для виявлення незвичайних шаблонів, які можуть вказувати на загрози безпеці.

Система використовує штучний інтелект для виявлення складних загроз, таких як атаки на основі облікових даних, інсайдерські загрози та дії з вимогами викупу, які можуть бути пропущені іншими інструментами. Алгоритми штучного інтелекту постійно навчаються на минулих інцидентах, підвищуючи свою точність і адаптивність у виявленні загроз.

Можливості штучного інтелекту Exabeam Fusion поширюються на автоматизацію реагування на виявлені загрози. Він використовує Smart Timelines на основі машинного навчання для автоматичного збору доказів,

оцінки ризиків і зведення інформації в цілісну історію для проведення розслідувань. Така автоматизація скорочує час і зусилля, необхідні для реагування на інциденти, та підвищує ефективність операцій з безпеки.

Використання штучного інтелекту для зменшення кількості хибних спрацьовувань у таких системах, як Exabeam Fusion, є значним досягненням у сфері кібербезпеки. Хибні спрацьовування, коли законні дії помилково позначаються як загрози, можуть забирати значний час і ресурси команд безпеки. Це досягається завдяки автоматичному аналізу поведінкових шаблонів користувачів та постійному навчанню. Система також може аналізувати загрози залежно від контексту повязаних з ними подій. ШІ може співвідносити розрізнені події в мережі, щоб виявити шаблони загроз.

Впровадження такої системи має як свої плюси так і мінуси. До переваг можна додати зменшення кількості помилкових спрацьовувань, вивчаючи звичайні моделі поведінки, система може зменшити кількість помилкових спрацьовувань, дозволяючи командам безпеки зосередитися на реальних загрозах; змога краще виявляти наявні загрози; автоматично реагувати на інциденти.

З недоліків можна визначити наступні: високі вимоги до досвіду фахівців, така система може вимагає конкретних знань для налаштування для ефективного управління та роботи; висока вартість, впровадження такої системи у декілька разів дорожче ніж використання традіційних SIEM систем; залежність від великої кількості даних та їхньої якості, ефективність ШІ в системі значною мірою залежить від якості та кількості даних, які він обробляє, низька якість даних може призвести до неточного аналізу; потенційно велика надія на ШІ, існує ризик надмірної залежності від автоматизованих процесів, що може призвести до недооцінки важливості людського нагляду в кібербезпеці; вимоги до постійного оновлення, алгоритми ШІ постійно потребують оновлень, щоб залишатися ефективними у боротьбі з кіберзагрозами, що розвиваються.

Оцінка потреб і викликів для сучасних систем безпеки

Для застосування алгоритмів штучного інтелекту необхідна велика кількість даних, а також значні та ефективні апаратні ресурси. Моделі машинного навчання зазвичай розробляються і навчаються проти конкретних кібератак. Модель не може ефективно виявляти різноманітні атаки або протистояти кібератакам, що еволюціонують. Виявлення дій, які раніше не спостерігалися, може виявитися складним завданням, а такі дії технічно суттєво відрізняються від своїх попередників [3]. Моделі зазвичай навчаються на попередніх ознаках у наборі даних, тому найновіші атаки можуть можуть бути непомічені класифікаторами, що призводить до зниження рівня виявлення та хибних спрацьовувань.

Тож основним викликом є потреба у наявності якісних даних. Алгоритми штучного інтелекту базуються на даних, отже від того наскільки вони якісні залежить і те, наскільки якісно буде працювати штучний інтелект. І тут постає питання про те які конфеденційні дані можна додавати для навчання, адже штучний інтелект теж є можливою атаки. Забезпечення якості даних і вирішення проблем конфіденційності, особливо при роботі з конфіденційною інформацією, є серйозним викликом.

При розробці заходів кібербезпеки слід враховувати атаки на сам алгоритм штучного інтелекту. У літературі описані різні типи атак, що мають на меті обдурити моделі, подаючи неправдиві вхідні дані. Нижче наведені деякі з атак, які наразі доступні в літературі [3]:

- Метод швидкого градієнтного знаку (FGSM)
- Багатокроковий бітовий координатний підйом (BCAk)
- Багатокрокове бітове градієнтне сходження (BGAk)
- Генеративні змагальні мережі (GAN)
- Атаки Карліні та Вагнера (C&W)

Дослідники шукають способи захисту від таких атак, деякі з них можна підсумувати наступним чином [4, 5]:

– Захисна дистиляція додає гнучкості процесу класифікації алгоритму, завдяки чому модель стає менш вразливою до експлуатації.

– Витискання ознак виконує згладжуючі перетворення вхідних ознак для усунення збурень, спричинених противником.

– У змагальному навчанні мінімізуються помилки класифікації, спричинені змагальними прикладами, шляхом введення до набору даних вхідних даних, які містять змагальні збурення з правильними вихідними мітками.

– При градієнтному маскуванні зменшується чутливість моделі до малих збурень у вхідних даних шляхом обчислення похідних першого порядку.

– Ансамблеві методи підвищують міцність за рахунок спільного навчання декількох класифікаторів.

– Модифікація процесу навчання та вхідних даних може підвищити надійність глибокої мережі шляхом безперервного введення нових типів змагальних зразків під час виконання змагального навчання. Цей метод, однак, вимагає достатньої виразності та високої якості навчальних прикладів.

– Модифікація мережі. Дослідники запропонували введення глибоких контрактивних мереж, використання градієнтної регуляризації та біологічно натхненних рішень.

– Використання додаткової мережі. Універсальні збурення були запропоновані для боротьби з атаками супротивника, додаючи окрему навчену мережу до оригінальної моделі.

Національний інститут стандартів і технологій підготував звіт [5] про систематику і термінологію змагального машинного навчання (AML). У ньому визначено проблеми безпеки для ШІ і, зокрема, для машинного

навчання - потенціал використання чутливості моделей зловмисниками для негативного впливу на продуктивність класифікації та регресії машинного навчання. AML займається "розробкою алгоритмів машинного навчання, які можуть протистояти викликам безпеки, вивченням можливостей зловмисників і розумінням наслідків атак" [5]. Вони дотримуються ризик-орієнтованого підходу, в результаті чого систематика AML узгоджується з трьома вимірами оцінки ризиків AML (атаки, захист і наслідки). Узагальнену систематику зображено на рис. 1.1.

Не менш важливим та складним у реалізації є питання постійного розвитку штучного інтелекту. Кіберзагрози не стоять на місці, вони постійно розвиваються, використовуючи нові технології та методології для обходу традиційних заходів безпеки. Щоб ефективно протистояти цим еволюціонуючим загрозам, SIEM системи з використанням штучного інтелекту повинні втілювати в собі адаптивність і маневреність. Загрози кібербезпеці швидко змінюються, і те, що вважається аномалією або індикатором загрози, може змінюватися. Як наслідок, моделі штучного інтелекту повинні регулярно перенавчатися і поповнюватися новими даними, щоб підтримувати свою ефективність.

Коригування моделей штучного інтелекту також передбачає налаштування параметрів і алгоритмів для кращого узгодження з мінливим характером мережевих середовищ і поведінкою користувачів. Оскільки організації впроваджують нові технології та змінюють свої операційні моделі, SIEM системи з ШІ повинні адаптуватися до цих змін, щоб підтримувати високий рівень точності виявлення загроз.

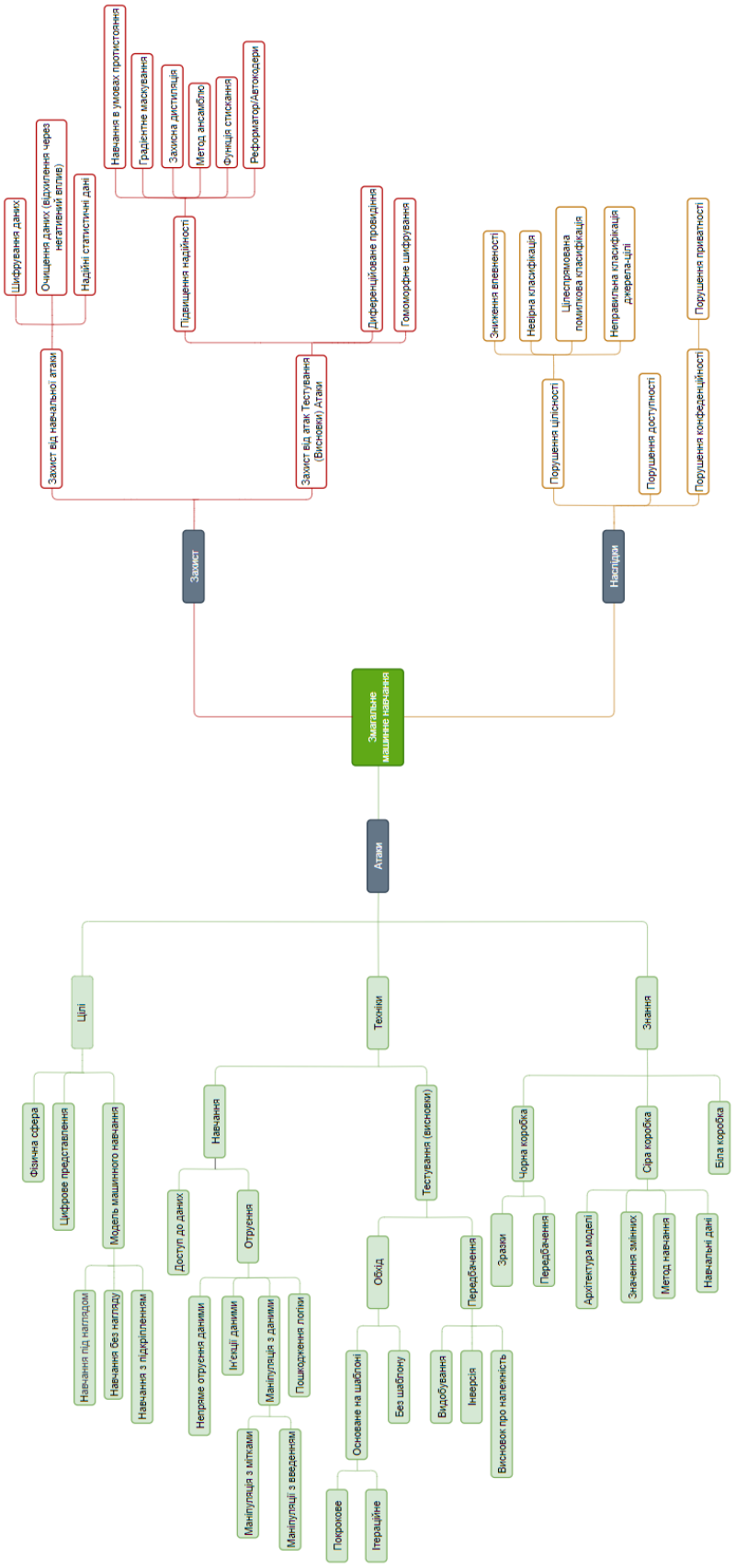


Рисунок 1.1 — Систематика атак, захистів та наслідків у змагальному машинному навчанні (Перекладено з [5]).

По суті, безперервне навчання, оновлення та коригування моделей ШІ мають вирішальне значення для забезпечення того, щоб SIEM системи залишалися ефективними проти постійно мінливої тактики кіберсупротивників.

1.3 Розгляд потенціалу штучного інтелекту в контексті реагування на інциденти

В основі застосування ШІ в реагуванні на інциденти лежить його здатність аналізувати великі обсяги даних з безпрецедентною швидкістю. Традиційне реагування на інциденти часто передбачає ручний перегляд великої кількості журналів і сповіщень для виявлення та усунення загроз. Цей процес не тільки забирає багато часу, але й схильний до людських помилок і недогляду. Штучний інтелект, з його передовими можливостями обробки даних, може швидко проаналізувати ці набори даних, виявляючи закономірності та аномалії, які можуть вислизнути навіть від найдосвідченіших фахівців з безпеки. Такий швидкий аналіз може значно скоротити час між виявленням загрози та її усуненням, що є критично важливим фактором для мінімізації потенційних збитків. Перелік деяких систем і засобів ШІ які вони використовують надано у табл. 1.1.

Крім того, алгоритми машинного навчання ШІ можуть адаптуватися і розвиватися на основі нової інформації з минулих інцидентів. Таке безперервне навчання дозволяє поступово підвищувати точність і ефективність реагування на загрози. На відміну від статичних систем, заснованих на правилах, системи реагування на інциденти, керовані штучним інтелектом, можуть швидко та постійно розвиватися у можливості реагування на новітні загрози , включаючи експлойти "нульового дня" і складні шкідливі програми, які традиційні методи можуть пропустити.

Ще однією значною перевагою ШІ в реагуванні на інциденти є зменшення кількості хибних спрацьовувань. Хибні спрацьовування або нешкідливі дії, помилково позначені як загрози, можуть перевантажити команди безпеки, що призводить до втоми і ризику пропустити справжні загрози. Здатність ШІ навчатися і розрізняти нормальну поведінку мережі і справжні аномалії може призвести до більш цілеспрямованого і ефективного реагування на інциденти, гарантуючи, що команди безпеки сконцентрують свої зусилля там, де вони найбільш необхідні.

Таблиця 1.1 – Засоби ШІ використовувані в SIEM

SIEM	Засоби ШІ	Використання AI	Основні Моделі ШІ
Splunk	Splunk User Behavior Analytics, Splunk UBA	Аналіз поведінки користувачів, виявлення аномалій, ризиків та уразливостей, розпізнавання атак та вторгнень.	Нейромережі, статистичні методи, кореляція
IBM QRadar	Watson for Cyber Security, QRadar Network Insights	Виявлення загроз, аналіз потоків мережевого трафіку, розпізнавання подій та аномалій.	Машинне навчання, алгоритми класифікації
ArcSight (Micro Focus)	ArcSight Interset	Виявлення загроз, вивчення моделей користувацької поведінки, аналіз потоків подій.	Методи машинного навчання, кластеризація
LogRhythm	LogRhythm SmartResponse, LogRhythm UEBA	Автоматизована реакція на загрози, виявлення аномалій у поведінці користувачів.	Машинне навчання, алгоритми кластеризації
Elastic SIEM	Elastic Machine Learning	Детектування аномалій, прогнозування інцидентів, виявлення подій, що потребують уваги.	Методи машинного навчання, алгоритми кластеризації
SolarWinds Security Event Manager	SolarWinds Threat Monitor	Автоматизоване виявлення та відновлення інцидентів, аналіз залежностей подій та виявлення вразливостей.	Машинне навчання, алгоритми кластеризації

Продовження таблиці 1.1.

SIEM	Засоби ШІ	Використання AI	Основні Моделі ШІ
AlienVault USM (AT&T Cybersecurity)	Open Threat Exchange (OTX), AlienVault OTX DirectConnect	Виявлення загроз, обмін інформацією про загрози, виявлення вразливостей.	Нейромережі, ансамбль моделей
McAfee Enterprise Security Manager	McAfee Behavioral Analytics	Виявлення аномалій, ризиків та загроз у реальному часі.	Алгоритми машинного навчання
Graylog	N/A (Can integrate with third-party ML tools)	Інтеграція з іншими інструментами машинного навчання, виявлення аномалій у журналах.	Варіативно в залежності від інтеграції
Trustwave SIEM	Trustwave NAC, Trustwave Threat Detection and Response (TDR)	Автоматичне виявлення загроз, розпізнавання зловмисного коду та аномалій у мережі.	Методи машинного навчання, кореляція
Netsurion SIEM	Netsurion EventTracker	Автоматизоване виявлення та відновлення інцидентів, виявлення аномалій у мережі.	Машинне навчання, алгоритми кластеризації
ManageEngine Log360	Zoho Analytics	Аналіз журналів, виявлення інцидентів, аудит та відновлення.	Нейромережі, алгоритми кластеризації
Awake Security Adversarial Modeling	Awake Security Adversarial Modeling	Виявлення зловмисницької активності, моделювання загроз та інтеграція зі сторонніми системами.	Машинне навчання, нейромережі
Cisco Stealthwatch	Cisco Stealthwatch	Виявлення аномалій у мережі, аналіз мережевого трафіку та потенційно небезпечних з'єднань.	Машинне навчання, аналітика потоків даних
Flowmon NBAD	Flowmon NBAD	Виявлення аномалій у мережі, аналіз мережевого трафіку, ідентифікація небезпечних або невірних активностей.	Машинне навчання, аналітика потоків даних
Wazuh	Wazuh Machine Learning	Детектування аномалій, виявлення зловмисної активності та інших	

Використання мовленнєвих моделей ШІ у рамках SIEM систем значно оптимізує процес аналізу інцидентів. Ці технології здатні швидко обробляти та інтерпретувати великі обсяги даних, виявляючи аномалії та надаючи фахівцям з кібербезпеки структуровані, зрозумілі висновки. Це, в свою чергу, сприяє пришвидшенню процесу прийняття рішень та визначенню необхідності негайного втручання. Ключовим аспектом є можливість мовленнєвих моделей не просто фіксувати аномалії, а й надавати їх опис в доступній формі, зменшуючи потребу в ручному перегляді та аналізі великої кількості безпекових даних.

Висновки до першого розділу

У цьому розділі було проведено аналіз важливості та ролі штучного інтелекту в системах SIEM. Було розглянуто потенціал функціональності SIEM систем, визначено, що інтеграція ШІ значно підвищує ефективність виявлення та реагування на кіберзагрози. Автоматизація процесів та здатність до глибокого аналізу даних з різних джерел допомагають організаціям оперативно виявляти аномальну поведінку і незвичайні події, що можуть свідчити про безпекові інциденти. Були розглянуті переваги використання ШІ в SIEM системах, такі як покращення точності виявлення внутрішніх та зовнішніх загроз, а також зниження часу реагування на інциденти. Надано огляд деяких SIEM систем і засобів ШІ які вони використовують.

ТЕОРЕТИЧНІ ОСНОВИ ІНТЕГРАЦІЇ ШІ В СИСТЕМИ КІБЕРБЕЗПЕКИ

У цьому розділі буде розглянуто теоретичні основи та використання штучного інтелекту в системах кібербезпеки, зокрема у SIEM системах. Розглядаються SIEM системи на основі штучного інтелекту, які використовують можливості ШІ та машинного навчання для вирішення цих проблем, забезпечуючи автоматизацію, прогнозування загроз, та ефективне реагування на інциденти. Виділено перспективи використання великих мовленнєвих моделей у SIEM системах.

2.1 Огляд ключових теорій штучного інтелекту, що застосовуються у SIEM системах

SIEM включають в себе різні теорії штучного інтелекту, щоб розширити свої можливості у виявленні та управлінні загрозами безпеці. Розглянемо основні теорії.

Теорія Демпстера-Шейфера (DST) відіграє ключову роль у вдосконаленні SIEM, зокрема у виявленні шахрайства в системах мобільних грошових переказів. Як пояснюють [6] DST допомагає агрегувати докази з розрізнених джерел даних для підвищення точності та надійності механізмів виявлення шахрайства. Ця теорія, що ґрунтується на математичних основах теорії ймовірності та доказів, пропонує більш гнучкий підхід порівняно з традиційними ймовірнісними методами. Вона дозволяє поєднувати докази з різних джерел, тим самим сприяючи всебічному аналізу потенційної шахрайської діяльності.

У контексті послуг мобільних грошових переказів система DST добре справляється з невизначеністю та частковими знаннями, які є поширеними в динамічній і часто непрозорій сфері фінансових транзакцій. Присвоюючи ступені вірогідності різним доказам та інтегруючи їх у функцію

вірогідності, DST ефективно оцінює ймовірність шахрайських дій. Цей підхід значно покращує процес прийняття рішень в SIEM системах, дозволяючи їм виявляти і реагувати на складні схеми шахрайства з більшою точністю і ефективністю.

Алгоритми виявлення аномалій, зокрема, глобальні алгоритми k-Nearest Neighbors (k-NN), відіграють важливу роль у сфері систем управління інформацією та подіями безпеки, як підкреслюється в дослідженні [7]. Ці алгоритми особливо добре виявляють неправильні конфігурації та аномалії в даних, що є важливим для забезпечення цілісності та безпеки інформаційних систем.

Глобальний алгоритм k-NN працює шляхом оцінки близькості точок даних у заданому наборі даних. У контексті систем SIEM він аналізує події безпеки та журнали, щоб виявити відхилення від нормальних патернів, які вказують на потенційні порушення безпеки або неправильні конфігурації. Алгоритм функціонує шляхом обчислення відстані між кожною точкою даних та її k найближчими сусідами в просторі ознак. Аномалії ідентифікуються на основі припущення, що вони матимуть значно відмінні характеристики порівняно з більшістю точок даних, а отже, лежатимуть далі від своїх найближчих сусідів.

Цей підхід особливо вигідний, оскільки не вимагає попереднього маркування даних, яке часто є непрактичним або недоступним у реальних умовах. Це дозволяє алгоритму адаптивно навчатися на основі даних, що робить його високоефективним у динамічних середовищах, де можуть з'являтися нові типи атак або неправильних конфігурацій. Здатність глобального алгоритму k-NN працювати без попередньо визначених правил або шаблонів дозволяє створити більш гнучкий і надійний механізм виявлення, що має вирішальне значення для кіберзагроз, що постійно змінюються.

Більше того, інформація, яку надають ці алгоритми виявлення аномалій, виходить за рамки простої ідентифікації порушень. Вони можуть запропонувати цінну інформацію про основні процеси та поведінку в системі, допомагаючи адміністраторам зрозуміти природу аномалій та розробити ефективні стратегії реагування. Ця технічна можливість значно підвищує операційну ефективність систем SIEM у захисті від складних і нових загроз безпеці [7].

Алгоритм RETE є основоположним методом у у формулюванні правил для виявлення атак TCP SYN flood, як описано в роботі [8]. Цей алгоритм є критично важливим компонентом експертних систем, заснованих на правилах, де його основною функцією є ефективна обробка і зіставлення великих наборів правил з швидкозмінним набором вхідних даних, що є типовим сценарієм для моніторингу мережевої безпеки.

Атаки TCP SYN flood, що характеризуються переповненням цільової системи SYN-пакетами з метою виснаження її ресурсів, вимагають складних механізмів виявлення через їхню часто приховану та розподілену природу. Алгоритм RETE вирішує цю проблему, дозволяючи швидко і динамічно формулювати правила, які можуть адаптуватися до мінливих моделей таких атак. Він працює шляхом побудови мережі вузлів, які представляють умови та дії правил. Коли з'являється нова інформація або події (наприклад, дані про мережевий трафік), вони поширюються цією мережею, запускаючи відповідні правила для виявлення потенційних атак.

Однією з ключових переваг алгоритму RETE є його ефективність в обробці модифікацій та переоцінок правил, що є критично важливим у швидкозмінному середовищі мережевої безпеки. Ця ефективність досягається завдяки використанню підходу зіставлення шаблонів, який мінімізує надлишкові оцінки, що значно прискорює процес виявлення потенційних атак SYN-флуду. Здатність алгоритму швидко обробляти і зіставляти складні набори правил з об'ємними і різноманітними вхідними

даними робить його незамінним інструментом в арсеналі SIEM-систем для проактивного виявлення і пом'якшення наслідків TCP SYN flood-атак [8].

Генетичне програмування (ГП) і штучні нейронні мережі (ШНМ) представляють собою передові обчислювальні парадигми, що розширюють можливості самоадаптації, зокрема, у вивченні кореляційних правил для виявлення багатокрокових атак. [9] наголошують на застосуванні цих методів для розпізнавання складних моделей атак, які розгортаються в кілька етапів.

ГП, натхненний біологічною еволюцією, працює шляхом еволюції популяції рішень-кандидатів (тобто правил кореляції) за допомогою процесів, що імітують природний відбір і генетику. У контексті SIEM-систем ГП ітеративно вдосконалює ці правила, адаптуючи їх до нових загроз, оцінюючи їх ефективність у виявленні багатокрокових атак. Цей еволюційний підхід дозволяє динамічно генерувати надійні та тонкі правила виявлення, що відповідають мінливому ландшафту кіберзагроз.

ШНМ, з іншого боку, імітують роботу людського мозку, обробляючи інформацію за допомогою взаємопов'язаних вузлів (нейронів). Вони досягають успіху в розпізнаванні образів, навчаючись на великих масивах даних для виявлення тонких кореляцій і аномалій, що вказують на витончені атаки. Інтеграція ШНМ в SIEM-системи дає їм можливість аналізувати складні патерни даних, підвищуючи їхні прогностичні можливості у визначенні етапів багатоетапних атак.

Синергія ГП та ШНМ в SIEM-системах сприяє створенню самоадаптивної структури. Ця структура здатна автономно розвивати свої механізми виявлення, гарантуючи, що система залишається ефективною проти просунутих, багатоступеневих кіберзагроз, таким чином підтримуючи високий рівень мережевої безпеки та стійкості.

Високорівневі мережі Петрі (HLPN) та мова Z є ключовими в архітектурному проектуванні та аналізі складних систем управління

інформацією та подіями безпеки, таких як OSTORM, як детально описано в системи підтримували критичні властивості безпеки, включаючи конфіденційність і цілісність даних про події.

HLPN розширює класичний фреймворк Petri Net, інтегруючи концепції програмування високого рівня, що дозволяє моделювати складні системи зі складною поведінкою і потоками даних. У системах SIEM HLPN використовуються для представлення та аналізу динамічних взаємодій між різними компонентами, особливо приділяючи особливу увагу потоку та обробці даних про події. Їх графічна і математична природа полегшує точне моделювання паралельних процесів і виявлення потенційних вразливостей безпеки, забезпечуючи надійну обробку даних і цілісність процесів.

Мова Z, формальна мова специфікацій, доповнює HLPN у моделюванні систем SIEM. Вона забезпечує строгі рамки для визначення властивостей і поведінки системи за допомогою теорії множин і логіки предикатів першого порядку. У контексті SIEM така специфікація використовується для визначення та перевірки відповідності системи вимогам безпеки, таким як конфіденційність та цілісність даних. Цей формальний підхід гарантує, що кожен аспект дизайну системи буде ретельно вивчений і перевірений на відповідність заздалегідь визначеним критеріям безпеки, таким чином посилюючи загальний рівень безпеки системи SIEM.

Штучні імунні системи (ШІС), як дослідили [11], є новим обчислювальним підходом, що застосовується в SIEM, надихаючись механізмами біологічної імунної системи. Ці системи вміють напівавтоматично генерувати правила кореляції подій, що є критично важливою функцією у виявленні та реагуванні на інциденти безпеки.

ШІС працюють, імітуючи адаптивні та саморегулюючі властивості імунної системи людини. У контексті систем SIEM це означає здатність

вчитися і розвиватися у відповідь на нові і нові загрози безпеці. Розпізнаючи закономірності та аномалії в поведінці системи, подібно до того, як біологічна імунна система виявляє патогени і реагує на них, ШІС можуть динамічно генерувати і коригувати правила кореляції для кращого виявлення і управління подіями безпеки.

Ця біологічна методологія дозволяє системам SIEM постійно адаптувати свої стратегії виявлення, підвищуючи ефективність і результативність у виявленні складних і витончених загроз безпеці. Напівавтоматичний характер генерації правил в ШІС зменшує ручну роботу і підвищує швидкість реагування та стійкість системи до кіберзагроз, що еволюціонують.

Техніка баг-бару [12] є інноваційним методом, що розширює їхні можливості класифікувати загрози безпеці та обчислювати їхню критичність. Цей метод значно розширює можливості предиктивної аналітики SIEM систем, забезпечуючи більш тонкий підхід до оцінки загроз.

Центральним елементом цього методу є його здатність розшарування загрози на основі їхньої серйозності та потенційного впливу. Вона використовує системний підхід для оцінки різних атрибутів виявлених загроз, таких як їх походження, характер і потенційна шкода. Присвоюючи кожній загрозі оцінку критичності, метод полегшує визначення пріоритетів реагування, спрямовуючи увагу в першу чергу на найсерйозніші загрози.

Інтеграція цього методу в комерційні SIEM системи пропонує додатковий рівень до існуючих механізмів виявлення загроз. Він збагачує набір даних системи, забезпечуючи додатковий вимір аналізу загроз, що має вирішальне значення для точного і своєчасного прогнозування загроз. Структурований підхід техніки баг-бару до класифікації загроз не лише допомагає швидко виявляти критичні загрози, але й покращує загальний

процес прийняття рішень в операціях з кібербезпеки, забезпечуючи більш надійну та проактивну позицію оборони [12].

Фреймворк Generic Event Translator (GET) [13] є складним інструментом, який долає розрив між фізичною та логічною сферами безпеки. Цей фреймворк особливо добре справляється з обробкою та аналізом різнорідних даних, використовуючи розуміння бізнес-процесів для виявлення проблем безпеки.

В основі фреймворку GET лежить його здатність контекстуалізувати та інтерпретувати різноманітні типи даних, що походять з різних фізичних та логічних джерел. Зіставляючи ці дані з конкретними бізнес-процесами, фреймворк досягає всебічного розуміння того, як розрізнені події безпеки пов'язані з більш широким операційним контекстом. Цей цілісний підхід має вирішальне значення для виявлення тонких, але важливих аномалій безпеки, які в іншому випадку можуть залишитися невиявленими при ізольованому аналізі.

Інтеграція GET в SIEM системи дає їм можливість ефективно управляти і співвідносити дані на різних рівнях безпеки. Це дає змогу проводити більш детальний та обґрунтований аналіз подій безпеки, сприяючи ранньому виявленню потенційних порушень або вразливостей. Поєднуючи аналіз подій безпеки зі знаннями бізнес-процесів, фреймворк GET розширює можливості стратегічного реагування систем SIEM, гарантуючи, що заходи безпеки є ефективними і відповідають цілям організації [13].

2.2 Дослідження переваг і можливих ризиків застосування ШІ в кібербезпеці

У сфері кібербезпеки інтеграція ШІ знаменує собою трансформаційний прогрес у боротьбі з цифровими загрозами. Системи,

керовані штучним інтелектом, демонструють неабияку майстерність у виявленні та пом'якшенні широкого спектру кіберзагроз - від вторгнень шкідливого програмного забезпечення до витончених фішингових кампаній. Ці алгоритми ШІ використовують методи розпізнавання образів і виявлення аномалій, що дозволяє їм розпізнавати і реагувати на порушення в мережевому трафіку і поведінці користувачів, які часто свідчать про зловмисну діяльність. Ця здатність особливо ефективна проти атак шкідливого програмного забезпечення та мережевих вторгнень, де швидке виявлення має вирішальне значення. Крім того, здатність штучного інтелекту обробляти природну мову підвищує його ефективність у виявленні фішингових і спам-повідомлень, які часто характеризуються тонкими лінгвістичними і структурними особливостями, які традиційні фільтри можуть не помітити. Механізми швидкого автоматизованого оповіщення, вбудовані в системи штучного інтелекту, забезпечують своєчасне сповіщення про інциденти безпеки, тим самим сприяючи швидкому та ефективному реагуванню на них. Цей технологічний стрибок у практиці кібербезпеки не лише підвищує ефективність виявлення загроз, але й значно зменшує вікно вразливості, пропонуючи надійний механізм захисту в цифровому світі [14].

ШІ суттєво впливає на кібербезпеку, автоматизуючи критичні завдання, покращуючи таким чином виявлення загроз та реагування на інциденти. У сфері кібербезпеки роль ШІ виходить за рамки простого виявлення; він активно бере участь в автоматизації повторюваних і трудомістких завдань. Ця автоматизація має вирішальне значення для вдосконалення розвідки загроз і стратегій реагування на інциденти. Алгоритми штучного інтелекту вміють просіювати величезні масиви даних і виявляти потенційні загрози, такі як нові підписи шкідливих програм або аномальну поведінку мережі, які можуть вислизнути від традиційних систем виявлення. Автоматизуючи ці процеси, ШІ не лише пришвидшує

виявлення загроз кібербезпеці, але й мінімізує ймовірність людського нагляду.

Крім того, інтеграція ШІ в кібербезпеку звільняє ІТ-фахівців від рутинних завдань моніторингу, дозволяючи їм сконцентруватися на більш стратегічних ініціативах. Такий перерозподіл людських ресурсів призводить до загального покращення стану кібербезпеки. Здатність штучного інтелекту навчатися і адаптуватися до кіберзагроз, що змінюються, гарантує, що автоматизовані системи залишатимуться ефективними з часом, що робить їх безцінним активом в умовах динамічного ландшафту кіберзагроз. Ефективність і результативність, які забезпечує ШІ в автоматизації завдань кібербезпеки, перетворюються на більш надійний захист від все більш складних [15].

У фінансовому секторі штучний інтелект відіграє важливу роль у розширенні цифрової інклюзії шляхом посилення заходів кібербезпеки, зокрема у виявленні ризиків, управлінні ними та запобіганні шахрайству забезпеченням безпеки цифрових фінансових транзакцій і послуг. Системи на основі штучного інтелекту використовують складні алгоритми, які аналізують шаблони транзакцій і поведінку клієнтів, що дозволяє їм виявляти аномалії, які можуть свідчити про шахрайські дії. Ця здатність має вирішальне значення для виявлення та запобігання фінансовому шахрайству, наприклад, несанкціонованим транзакціям або крадіжкам персональних даних, які поширені в цифровому банкінгу та фінансових послугах.

Крім того, предиктивна аналітика штучного інтелекту допомагає в управлінні ризиками. Обробляючи величезні обсяги даних, системи штучного інтелекту можуть прогнозувати потенційні ризики кібербезпеки та пропонувати проактивні заходи для їхнього зменшення. Такий прогностичний підхід є безцінним у фінансовому секторі, де вартість і

наслідки кіберінцидентів можуть бути значними. Здатність штучного інтелекту постійно навчатися та адаптуватися до нових загроз гарантує, що фінансові установи залишатимуться стійкими до нових кіберризиків.

Крім того, штучний інтелект сприяє підвищенню якості обслуговування клієнтів у сфері цифрових фінансових послуг, забезпечуючи безпечну та безперешкодну взаємодію. Цей прогрес не лише зміцнює довіру до цифрових фінансових платформ, а й сприяє ширшій фінансовій інклюзії, роблячи послуги доступнішими та безпечнішими для населення.

Інтеграція ШІ в кібербезпеку, посилюючи захисні механізми, парадоксальним чином створює нові вразливості, зокрема, вразливість до складних атак. Системи штучного інтелекту з їхніми складними алгоритмами і процесами прийняття рішень на основі даних можуть використовуватися або краще сказати піддваатися маніпуляціям з боку зловмисників, що призводить до потенційних порушень безпеки. Ця вразливість пов'язана насамперед з непрозорою природою алгоритмів ШІ, особливо в моделях глибокого навчання, де процес прийняття рішень не завжди є прозорим або таким, що піддається інтерпретації.

Відповіддю на цей виклик є розробка методів пояснюваного штучного інтелекту (XAI). Мета XAI - зробити процеси прийняття рішень в системах ШІ прозорими і зрозумілими для людини-оператора. Така прозорість має вирішальне значення для виявлення та усунення потенційних слабких місць у системах кібербезпеки, керованих ШІ, якими можуть скористатися зловмисники. Зрозумілість ШІ не лише підвищує довіру до цих систем, але й дозволяє фахівцям з кібербезпеки розуміти і прогнозувати поведінку ШІ в різних сценаріях, що дає їм змогу посилити захист від потенційних маніпуляцій з боку супротивників.

Таким чином, хоча ШІ значно розширює можливості кібербезпеки, для усунення нових вразливостей, які він створює, вкрай необхідні надійні

методи пояснення. Забезпечення того, щоб системи ШІ в кібербезпеці були не лише ефективними, але й прозорими та підзвітними, має вирішальне значення для захисту від сучасних атак супротивника [14].

Існують атаки які виключно націлені на зміну даних які використовуються для навчання. Ці атаки на ШІ часто включають такі тактики, як вороже машинне навчання, коли зловмисники подають неправдиві вхідні дані в моделі ШІ, щоб змусити їх працювати несправно або видавати помилкові результати.

Зловмисники також можуть використовувати залежність систем ШІ від даних, маніпулюючи даними, які використовуються для навчання, щоб спотворити процес навчання ШІ, що призводить до упередженого або помилкового прийняття рішень. Ще один вектор загроз - використання інтерпретованості та прозорості ШІ-моделі. Зловмисники можуть провести зворотню інженерію моделей які використовуються, щоб виявити їхні операційні моделі, тим самим виявляючи вразливості, які можна використати.

Ці загрози вимагають від фахівців з кібербезпеки не лише зосереджуватися на традиційних методах захисту, але й розробляти нові стратегії та технології для захисту систем які використовують ШІ від цих унікальних загроз. Розробка і розгортання ШІ в контексті кібербезпеки повинні супроводжуватися надійними заходами для виявлення і пом'якшення наслідків атак, спеціально спрямованих на вразливі місця алгоритмів, щоб ці системи залишалися стійкими до кіберзагроз, що постійно змінюються [17].

Потрібно враховувати виникнення безлічі правових і економічних проблем, особливо в сфері визнання авторства і захисту прав інтелектуальної власності. Ці виклики пов'язані з унікальними характеристиками систем штучного інтелекту, такими як їхня здатність навчатися, адаптуватися і самостійно приймати рішення. У контексті

кібербезпеки, де алгоритми ШІ часто створюють або модифікують контент, визначення чітких меж авторства стає складним завданням. Ця невизначеність викликає серйозні юридичні питання щодо права власності та розподілу відповідальності у випадках зловживання або порушення.

Крім того, використання ШІ в розробці рішень з кібербезпеки або в створенні програмного забезпечення, яке може містити запатентовані методології, ускладнює права інтелектуальної власності. Традиційні системи інтелектувальної власності не повністю пристосовані для врахування нюансів творів, створених за допомогою ШІ, що призводить до потенційних прогалин у захисті та правозастосуванні. Ситуація ще більше ускладнюється глобальним характером кіберзагроз і цифровим середовищем, де межі юрисдикції часто розмиті.

Вирішення цих правових та економічних проблем вимагає переоцінки та можливого реформування існуючої правової бази, щоб пристосувати її до тонкощів застосування ШІ в кібербезпеці. Забезпечення того, щоб закони йшли в ногу з технологічним прогресом, має вирішальне значення для підтримки надійного захисту прав інтелектуальної власності та чіткого визначення авторства в мінливому ландшафті рішень з кібербезпеки на основі ШІ [18].

Інтеграція штучного інтелекту в такі сектори, як охорона здоров'я, хоча і пропонує значні переваги з точки зору ефективності та предиктивної аналітики, водночас викликає значні побоювання щодо конфіденційності та безпеки даних. Ці побоювання пов'язані з великими вимогами до даних, які пред'являються до систем штучного інтелекту, особливо в охороні здоров'я, де йдеться про конфіденційну особисту інформацію про стан здоров'я. Залежність ШІ від великих масивів даних для навчання і роботи часто вимагає широкого обміну та агрегації даних, що потенційно збільшує ризик витоку даних і несанкціонованого доступу до них.

Складність алгоритмів ШІ та непрозорість процесів прийняття рішень загострюють проблеми конфіденційності. Пацієнти та медичні працівники можуть не знати, як дані використовуються, обробляються або поширюються в системах штучного інтелекту, що призводить до невизначеності щодо контролю над даними та права власності на них. Ситуація ще більше ускладнюється різними правилами і стандартами конфіденційності даних у різних юрисдикціях, що створює клаптикову тканину вимог до дотримання законодавства, в якій медичним організаціям доводиться орієнтуватися.

Для вирішення цих проблем необхідна надійна система управління даними, яка забезпечить етичне використання ШІ в охороні здоров'я. Це передбачає впровадження суворих заходів безпеки даних, таких як шифрування та контроль доступу, а також забезпечення прозорості та зрозумілості алгоритмів ШІ для побудови довіри між користувачами. Крім того, регуляторні органи повинні розробити чіткі рекомендації та стандарти щодо конфіденційності даних у додатках зі штучним інтелектом, збалансувавши потребу в інноваціях, керованих даними, із захистом індивідуальних прав на приватність [19].

2.3 Аналіз методів машинного навчання для вдосконалення роботи SIEM систем

У сфері кібербезпеки методології машинного навчання революціонізують операційну динаміку SIEM систем. Ключовою сферою, де МН досягає успіху, є мінімізація ручного введення даних і посилення можливостей проактивного полювання на загрози. Ця інновація в першу чергу ґрунтується на автоматизованому вилученні та аналізі унікальної поведінки мережевих комунікацій на безлічі кінцевих точок - процесі, який традиційно був трудомістким і схильним до помилок через залучення

людини. Суть цього підходу полягає в здатності аналізувати та розуміти складні патерни мережевого трафіку. Використовуючи передові алгоритми МН, SIEM системи тепер можуть самостійно аналізувати мережеві потоки даних, виявляючи аномальні або потенційно зловмисні дії, які в іншому випадку можуть бути непомічені людиною. Таке автоматизоване спостереження охоплює широкий спектр кінцевих точок мережі, забезпечуючи комплексний захист.

Однією з найважливіших переваг цього методу є його проактивний характер. Замість того, щоб реагувати на загрози після інциденту, SIEM системи з підтримкою МН тепер здатні виявляти і нейтралізувати загрози на випередження. Така зміна парадигми від реактивного до проактивного полювання на загрози має величезне значення для кібербезпеки, оскільки суттєво скорочує час на виявлення та реагування на потенційні порушення безпеки.

Інтеграція МН в SIEM, як підкреслюють [20], не тільки підвищує ефективність виявлення загроз, але й трансформує формат управління мережевою безпекою. Автоматизуючи складні процеси аналізу та переходячи до управління загрозами на випередження, SIEM системи на основі МН встановлюють новий стандарт пильності та стійкості кібербезпеки.

Інтеграція кореляційних правил для виявлення загроз у режимі реального часу є значним досягненням у системах управління інформацією та подіями безпеки. Ця методика, заснована на складних алгоритмах, полегшує миттєвий аналіз об'ємних даних журналів, що є фундаментальним аспектом сучасних систем кібербезпеки. Ретельно вивчаючи записи журналів у режимі реального часу, ці правила кореляції вміло виявляють закономірності та аномалії, які вказують на потенційні загрози безпеці.

Суть цього підходу полягає в його здатності ефективно аналізувати великі і часто складні дані журналів, що генеруються різними мережевими

пристроями і додатками. Правила кореляції ретельно розроблені, щоб визначити конкретні шаблони і послідовності подій, які вказують на кіберзагрози, починаючи від спроб вторгнення і закінчуючи інсайдерськими загрозами і навіть сучасними постійними загрозами (APT). Цей аналіз у режимі реального часу гарантує, що потенційні інциденти безпеки виявляються майже миттєво, тим самим значно зменшуючи вікно можливостей для зловмисників.

Проактивний характер цієї методології дозволяє швидко реагувати на інциденти, що є критично важливою вимогою в цифровому ландшафті, який швидко розвивається. Впровадження цих правил кореляції в SIEM системах підвищує ефективність центрів управління безпекою, надаючи їм гнучкість і точність, необхідні для боротьби з кіберзагрозами, складність яких постійно зростає. Ця можливість виявлення загроз у режимі реального часу допомагає забезпечити цілісність і безпеку інформаційних систем.

У сфері систем SIEM використання генетичного програмування (ГП) та штучних нейронних мереж (ШНМ) знаменує собою значний стрибок вперед у механізмах кіберзахисту. Ці передові обчислювальні методи використовуються для розробки та вдосконалення правил кореляції, що дозволяє системам SIEM адаптивно навчатися та ідентифікувати складні схеми атак, особливо в контексті багатоетапних кібератак.

Генетичне програмування, позичаючи ідеї в біологічної еволюції, ітеративно генерує і тестує популяцію правил, кожне з яких представляє собою потенційне рішення для виявлення загроз. Завдяки процесам, подібним до природного відбору та мутації, ГП розвиває ці правила протягом наступних поколінь, відточуючи найефективніші стратегії для виявлення складних послідовностей атак. Цей адаптивний процес особливо вправний у виявленні прихованих зв'язків у великих масивах даних, що має вирішальне значення для виявлення складних, багатоступеневих кіберзагроз.

Доповнюючи ГП, штучні нейронні мережі пропонують надійну основу для розпізнавання і класифікації образів. Імітуючи структуру і функції біологічних нейронних мереж, ШНМ здатні навчатися і узагальнювати вхідні дані, що робить їх винятково придатними для розшифровки нюансів поведінки, які вказують на сучасні кібератаки.

Штучні нейронні мережі, доповнюючи ГП, пропонують надійну основу для розпізнавання та класифікації образів. Імітуючи структуру та функції біологічних нейронних мереж, ШНМ здатні навчатися та узагальнювати вхідні дані, що робить їх винятково придатними для розшифровки нюансів поведінки, які вказують на сучасні кібератаки.

Як підкреслюють [9], ГП і ШНМ надають SIEM системам динамічний, контекстно-орієнтований інтелект. Цей симбіоз еволюційних алгоритмів і нейронного навчання формує потужний захист від дедалі складнішого ландшафту кіберзагроз, посилюючи прогностичні та адаптивні можливості систем SIEM у виявленні та пом'якшенні наслідків багатоетапних атак.

Мережі глибоких переконань (Deep Belief Networks, DBN), складний клас алгоритмів машинного навчання, стали потужним інструментом для вдосконалення систем управління інформацією та подіями безпеки. Характеризуючись глибокою архітектурою та імовірнісним графічним моделюванням, DBN пропонують тонкий підхід до виявлення потенційно скомпрометованих хостів у мережі. Їх ефективність, як продемонстрували значно перевершує ефективність традиційних алгоритмів і систем, заснованих на правилах, в додатках кібербезпеки.

В основі DBN лежить багаторівнева структура нейронних мереж, кожна з яких будує все більш абстрактне представлення вхідних даних. Такий ієрархічний підхід до навчання дозволяє DBN розпізнавати складні і тонкі закономірності, що вказують на загрози кібербезпеці, які можуть вислизнути від простіших, неглибоких моделей. Сильною стороною DBN в

цьому контексті є їхня здатність моделювати складні, багатовимірні дані, що робить їх особливо вправними в аналізі безлічі сигналів і поведінки, що генеруються мережевими хостами.

Ефективність DBN у виявленні скомпрометованих хостів зумовлена їхніми потужними можливостями вилучення та класифікації ознак. Обробляючи вхідні дані за допомогою декількох рівнів нелінійних перетворень, DBN можуть виявляти приховані структури і залежності в даних, що має вирішальне значення для виявлення аномальних або зловмисних дій. Такий підхід до глибокого навчання забезпечує значну перевагу над традиційними методами, які часто покладаються на заздалегідь визначені правила або моделі поверхневого навчання, яким бракує глибини та адаптивності DBN.

Таким чином, мережі глибоких переконань представляють собою зміну парадигми в області систем SIEM, пропонуючи безпрецедентну майстерність у виявленні скомпрометованих хостів.

Важливу роль відіграють алгоритми виявлення аномалій без нагляду, зокрема, глобальний підхід k-Nearest Neighbours (k-NN). Ці алгоритми вміло виявляють неправильні конфігурації мережі та аномалії, тим самим посилюючи інфраструктуру безпеки SIEM-систем. Робота [7] є прикладом застосування таких методів для посилення заходів кібербезпеки.

Алгоритм k-NN в контексті неконтрольованого навчання працює шляхом вимірювання схожості або відстані між точками даних у заданому наборі даних. У сфері SIEM це означає аналіз мережевого трафіку і журналів для виявлення шаблонів або поведінки, які відхиляються від норми. "k" в k-NN означає кількість найближчих сусідів, з якими порівнюється точка, і алгоритм класифікує дані на основі того, наскільки тісно вони узгоджуються з цими сусідами. Глобальний аспект цього підходу означає розгляд всього набору даних, що забезпечує всебічний аналіз.

Що відрізняє виявлення аномалій без нагляду - це його здатність працювати без попередньо визначених міток або категорій. Це особливо корисно в кібербезпеці, де загрози постійно розвиваються і нові типи атак можуть не вписуватися у відомі категорії. Аналізуючи дані в неконтрольований спосіб, алгоритм k-NN може виявити незвичайні шаблони, які можуть вказувати на порушення безпеки, наприклад, неправильну конфігурацію або нову форму вторгнення.

Цей метод покращує системи SIEM, надаючи рівень інтелекту, здатний автономно аналізувати величезні обсяги даних для виявлення потенційних інцидентів безпеки. Його застосування не лише допомагає у ранньому виявленні загроз, але й сприяє загальній стійкості інфраструктури кібербезпеки до постійних змін у кіберзагрозах.

Впровадження керованого позитивного і немаркованого навчання з використанням метод опорних векторів (SVM) знаменує собою значний крок у точному виявленні хостів з високим ступенем ризику. Автори [22] підкреслюють ефективність цього підходу, зокрема його виняткову точність класифікації, що є критично важливим фактором в операціях з кібербезпеки.

Контрольоване навчання з використанням SVM - це спеціалізована форма машинного навчання, де навчальні дані складаються лише з позитивних прикладів і великого набору немаркованих прикладів, які можуть бути як позитивними, так і негативними. Цей підхід особливо підходить для таких сценаріїв, як кібербезпека, де отримання маркованих негативних прикладів (тобто, які не є загрозами) може бути складним або непрактичним. У цьому контексті SVM застосовуються завдяки їхній надійній класифікації, яка ефективно розрізняє хости з високим рівнем ризику (позитивні) і нормальні (немарковані) хости.

Сила SVM полягає в їхній здатності знаходити оптимальну гіперплощину, яка максимізує різницю між позитивними і немаркованими

зразками. Використовуючи функції ядра, SVM можуть працювати у високорозмірному просторі ознак, що дозволяє їм обробляти складні та нелінійні взаємозв'язки, які часто присутні в даних з кібербезпеки. Ця характеристика робить SVM особливо ефективними в середовищах з асиметричним розподілом даних, типовим для систем SIEM.

Інтеграція навчання керованого PU з SVM в SIEM системах дозволяє більш тонко і точно ідентифікувати хости з високим рівнем ризику. Ця методологія підвищує рівень безпеки, дозволяючи виявляти потенційні загрози на ранній стадії, що дає змогу вчасно та ефективно реагувати на них. Точність і адаптивність цього підходу гарантують, що SIEM системи краще орієнтуються в мінливому ландшафті кіберзагроз, підтримуючи надійний захист від потенційних вторгнень.

Штучні імунні системи, натхненні біологічною імунною системою, представляють собою новий підхід у сфері кібербезпеки. Як продемонстрували [11], ШІС вміють напівавтоматично генерувати правила кореляції подій, що є ключовим процесом для підвищення ефективності та результативності систем SIEM.

Основна концепція ШІС в кібербезпеці полягає в імітації здатності біологічної імунної системи вивчати і запам'ятовувати сигнатури патогенів. Аналогічно, ШІС в SIEM системах вчаться розпізнавати шаблони мережевих загроз і аномалій. Цей біологічний підхід дозволяє системам виявляти та адаптуватися до нових кіберзагроз, подібно до того, як імунна система людини еволюціонує для боротьби з новими патогенами.

Однією з найважливіших функцій ШІС в SIEM є генерація правил кореляції. Ці правила є життєво важливими для виявлення взаємозв'язків між різними подіями безпеки, що уможливлює раннє виявлення складних багатоетапних кібератак. Напівавтоматична генерація цих правил за допомогою ШІС забезпечує динамічну та адаптивну систему безпеки,

зменшуючи залежність від попередньо визначених статичних правил, які можуть не враховувати постійне зростання витонченості та складності атак.

ШІС вдосконалюють SIEM системи, надаючи їм форму штучного інтелекту, здатного до безперервного навчання та адаптації. Це не тільки підвищує точність і швидкість виявлення загроз, але й значно зменшує навантаження на аналітиків з безпеки, які тепер можуть зосередитися на більш стратегічних завданнях. Таким чином, впровадження ШІС в SIEM системи знаменує собою значний прогрес у сфері кібербезпеки, пропонуючи більш стійкий та інтелектуальний підхід до захисту цифрових інфраструктур.

Застосування нейронних мереж і метода опорних векторів (SVM) для аналізу мережевих вторгнень є значним технологічним досягненням у сфері кібербезпеки. Як підкреслюють [23], ці методи мають вирішальне значення для точного аналізу та ідентифікації мережевих вторгнень, а вибір відповідних функцій ефективності та алгоритмів навчання є ключовим для досягнення високої точності класифікації.

Нейронні мережі, з їх здатністю навчатися і моделювати складні взаємозв'язки в даних, особливо підходять для виявлення тонких і складних шаблонів вторгнень в мережевому трафіку. Їх багаторівнева архітектура, яка імітує зв'язки нейронів людського мозку, дозволяє обробляти великі набори даних з декількома входами, що робить їх ідеальними для динамічного і часто неоднозначного характеру даних про мережеву безпеку.

У тандемі, SVM та нейронні мережі пропонують надійний та ефективний механізм класифікації. Відомі своєю високою точністю в задачах бінарної класифікації, SVM функціонують шляхом створення гіперплощини, яка оптимально розділяє різні класи даних. Таке розділення має вирішальне значення для розрізнення нормальної поведінки мережі та потенційних вторгнень. Ефективність SVM у виявленні вторгнень ще більше підвищується завдяки використанню функцій ядра, які дозволяють

їм обробляти нелінійні взаємозв'язки даних, що є загальною характеристикою в наборах даних мережевої безпеки.

Синергія нейронних мереж і SVM в аналізі вторгнень полягає в їхніх взаємодоповнюючих сильних сторонах: досвіді нейронних мереж у розпізнаванні образів і здатності SVM класифікувати складні дані з високою точністю. Ретельне налаштування їхніх робочих функцій і алгоритмів навчання має важливе значення для максимізації їхньої ефективності. Такий комплексний підхід надає фахівцям з кібербезпеки складний і тонкий інструментарій, що значно розширює їхні можливості для виявлення та реагування на різноманітні мережеві вторгнення.

Перелік деяких моделей ШІ та машинного навчання та перспективних завдань в SIEM надано у табл. 2.1.

Таблиця 2.1 – Моделі ШІ з перспективними завданнями SIEM

Модель	Моніторинг мережі	Виявлення подій безпеки	Аналіз журналу	Політика та відповідність
Дерева рішень	Аналіз мережевої діяльності на основі правил	Категоризація подій на основі правил	Шляхи прийняття рішень для інтерпретації журналу	Виконання правил відповідності
	Виявлення аномалій на основі traffic	Виявлення незвичайних моделей поведінки	Журнали кластеризації для виявлення аномалій	Виявлення та попередження на основі правил
	Групування подібних мережевих дій	Кластеризація подій безпеки	Групування журналів для полегшення аналізу	Перевірки відповідності та групування зразків
Випадковий ліс	Прогнозне моделювання поведінки мережі	Ансамбль навчання для виявлення подій	Аналіз атрибутів журналу для аналізу	Прогноз дотримання політики
	Візуалізація мережевої поведінки кластерів	Виявлення закономірностей у послідовності подій	Аналіз кластерів журналів для отримання інформації	Виявлення аномалій для перевірки відповідності

Продовження таблиці 2.1

Модель	Моніторинг мережі	Виявлення подій безпеки	Аналіз журналу	Політика та відповідність
	Візуалізація багатовимірних даних журналу	Виявлення аномалій у зменшених розмірах	Кластеризація журналів для інтуїтивного аналізу	Візуалізація відповідності та звітність
	Зменшення розмірності для аналізу журналу	Ідентифікація корельованих атрибутів журналу	Зменшення розмірності журналу для аналізу	Розрахунок оцінки відповідності
	Послідовний аналіз мережевої діяльності	Аналіз залежних від часу моделей подій	Розуміння зв'язків у часових журналах	Виявлення порушень комплаєнсу з часом
	Виявлення закономірностей у мережевому трафіку	Розпізнавання складних подійних структур	Витяг ознак із даних журналу	Глибока перевірка пакетів на відповідність
	Семантичний аналіз змісту журналу	Контекстно-залежна категоризація подій	Отримання значущої інформації з журналів	Аналіз журналів на наявність порушень політики
	Взаємодія природної мови та запити	Контекстуальне розуміння та відповідь	Текстова інтерпретація журналу	Роз'яснення та роз'яснення політики

Огляд перспектив використання великих мовних моделей в SIEM

Нещодавний успіх великих мовних моделей (LLM), які було надано спільнотам користувачів викликав велкий інтерес до їх широкого впровадження. Дослідники з усіх галузей наразі вивчають потенціал і обмеження LLM. З ними можна дуже зручно спілкуватися, використовуючи підказки природної мови, і вони здатні до вражаючих міркувань на людському рівні у відповідь.

Як зазначається в [24], найбільші та найефективніші LLM навчаються на наборах даних із трильйонів слів за допомогою сканування Інтернет-тексту, книг та інших текстових джерел. Враховуючи величезний обсяг цих наборів даних, природньо очікувати, що на додаток до здатності моделей

міркувати на загальні людські теми, LLM пройшли підготовку зі звітами та ресурсами з кібербезпеки, а також із кодом, пов’язаним із загрозами, і кодом захисту. Ймовірно, вони поглинули багато джерел загальнодоступної кіберінформації. Як наслідок, LLM такі як ChatGPT мають можливість полегшити життя як фахівцям із кібербезпеки, так і кіберзлочинцям. Незважаючи на те, що ChatGPT все ще перебуває на стадії бета-тестування, він ілюструє нову технологію, яка, ймовірно, матиме позитивний вплив на індустрію кібербезпеки. Проте немає жодних сумнівів, що цей інструмент також використовуватимуть зловмисники.

ChatGPT, розроблений OpenAI, є варіантом моделі Generative Pre-trained Transformer (GPT), спеціально розробленої для генерації людського тексту на основі підказок. Вона є частиною сімейства моделей GPT-3, яке являє собою значний прогрес в області обробки природної мови.

ChatGPT використовує архітектуру на основі трансформатора, нейромережевого дизайну. Ця архітектура характеризується механізмами самоуваги, які дозволяють моделі зважувати важливість різних слів у реченні або послідовності, що сприяє її розумінню та можливостям генерації. Модель чудово справляється з різними завданнями NLP, такими як генерація тексту, ведення бесіди, узагальнення, переклад і відповіді на запитання. Вона може генерувати зв'язний і контекстуально релевантний текст з декількох абзаців.

ChatGPT має широке застосування, включаючи чат-боти для обслуговування клієнтів, створення контенту, допомоги в програмуванні, освітні інструменти тощо.

GPT навчається на різноманітних текстах з Інтернету. Процес навчання включає неконтрольоване навчання на великому масиві текстів з подальшим доопрацюванням під час контрольованого навчання з використанням людського зворотного зв'язку для покращення релевантності та безпеки моделі.

Така система має свої сильні сторони, такі як: високоякісна генерація тексту; контекстне розуміння для продовження діалогу; універсальність у різних сферах та мовах; здатність генерувати творчий і технічний контент.

Проте у ChatGPT є вагомі недоліки: іноді видає правдоподібні, але неправильні або безглузді відповіді; обмежене розуміння контексту за межами певного текстового вікна; можливі упередження в мовній моделі через навчальні дані.

Спеціальне навчання або точне налаштування на спеціалізованих наборах даних може зробити ChatGPT більш ефективним для конкретних галузей або випадків використання. ChatGPT є значним кроком вперед у здатності ШІ розуміти і генерувати людську мову, пропонуючи широкий спектр застосувань і продовжуючи розвиватися разом з досягненнями в дослідженнях ШІ.

Для експертів з кібербезпеки ChatGPT міг би представити інструменти кібербезпеки, які покращують можливості навчання та комунікації навколо кібератак. Однак це також означає необхідність розробки процедур для виявлення нових і посилених загроз, створених

Ці припущення спонукають вивчити здатність LLM як створювати так і зменшувати загрози кібербезпеці. Це в свою чергу викликає цілу низку питань, частина яких знайшла відповіді у [25], а на решту ще необхідно відповісти, а саме:

и може ШІ надати інформацію про загрози?

и можна використовувати таку систему для підвищення кваліфікації початківців у загрозах?

и може ШІ інтерпретувати інформацію, зібрану з командного рядка?

аскільки такий процес може бути повністю автоматизований?

ким буде потенційний вплив?

кі можливості можуть з’явитися в міру розвитку як LLM, так і методів їх використання?

Це дослідження має на меті відповісти на перші 2 запитання.

2.5 Висновки до другого розділу

У цьому розділі було розглянуто ключові теорії ШІ, які використовуються у SIEM системах. Розглянуто переваги та ризики, що несе інтеграція та використання ШІ у такі системи. Виділено, які моделі машинного навчання можуть використовуватися для різних подій безпеки. Також віділено перспективи використання великих мовленнєвих моделій, таких як ChatGPT.

ПРАКТИЧНЕ ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В SIEM СИСТЕМАХ

У цьому розділі розглягуто два сценарії практичного використання штучного інтелекту в системах SIEM. Розглянуто таке практичне питання, як використання та інтеграція великої мовленнєвої моделі штучного інтелекту можуть оптимізувати процеси реагування на інциденти в кібербезпеці, підвищуючи ефективність та точність виявлення загроз. Ми зосереджуємо LLM на процесі прийняття рішень, який контролює виконання та інтерпретацію команд або інструментів на терміналі командного рядка.

3.1 Проектування моделі ШІ для оптимізації реагування на інциденти

3.1.1 Постановка задачі

Процес прийняття рішення щодо суб’єкта загрози зазвичай вимагає від людини розуміння інформації, яка повертається під час виконання команди. Наприклад, базовий інструмент сканування розвідки, такий як Nmap, у деяких випадках може відповідати сотнями рядків тексту, що містить IP-адреси, відкриті порти та програми, запущені на хості. Їх потрібно прочитати та інтерпретувати, щоб прийняти рішення про наступну команду. Традиційно для аналізу відповіді, вилучення імен хостів, IP-адрес, відкритих портів, запущених служб тощо часто використовується низький рівень шуму сигналу. співвідношення в таких скануваннях з точки зору інформації, яка може призвести до успішного використання. Це трудомісткий процес ручного сканування багатьох рядків виводу або розробки аналізатора для кожного інструменту; де потрібен додатковий спеціальний код для інтерпретації вмісту з огляду на мету суб’єкта загрози.

Використання SIEM-систем може бути складним для новачків у цій галузі. Реагування на інциденти кібербезпеки часто вимагає втручання не тільки спеціалістів, але й звичайних користувачів, тому зрозумілість інформації про інциденти стає ключовою. І в цьому аспекті, інтеграція мовної моделі такої як ChatGPT в SIEM-систему відкриває нові можливості. А саме, розглянемо два наступних питання:

– Чи може ChatGPT надати інформацію про загрози?

– Чи можна використовувати таку систему для підвищення кваліфікації початківців у загрозах?

Для відповіді на ці запитання необхідно виконати наступне:

– Налаштування SIEM системи Wazuh для моніторингу безпеки та відповіді на інциденти.

– Реалізація доступу до інтерфейсу керування Wazuh за допомогою Ubuntu, що дозволить використовувати цю систему для доступу до інтерфейсу керування Wazuh

– Розробка та налаштування скрипту на Python, який буде ініціювати автоматичне сканування мережі за допомогою утиліти Nmap. Це сканування дозволить ідентифікувати відкриті порти та інші потенційні вектори кібератак.

– Інтеграція мовленнєвої моделі ChatGPT. Ця інтеграція має на меті надання зрозумілих та детальних пояснень щодо виявлених потенційних ризиків та небезпек, пов'язаних з відкритими портами. ChatGPT може також пропонувати рекомендації щодо кроків для усунення виявлених уразливостей або зменшення ризиків.

Налаштування системи Wazuh

Wazuh - це рішення для моніторингу безпеки з відкритим вихідним кодом, яке пропонує систему управління інформацією та подіями безпеки

(SIEM). Воно призначене для забезпечення комплексної видимості IT-безпеки організації шляхом агрегування, зберігання, управління та аналізу даних журналів, що допомагає визначити стан безпеки систем та мереж

3.1.2.1 Особливості системи Wazuh

До ключових особливостей цієї системи потрібно віднести її основну функцію – агрегація та аналіз логів. Основна функціональність Wazuh включає агрегування та аналіз даних журналів з різних джерел для виявлення потенційних інцидентів безпеки або вразливостей.Wazuh забезпечує моніторинг логів системи та додатків, мережевого трафіку та дій користувачів в режимі реального часу. Він генерує сповіщення, коли виявляє дії, які можуть вказувати на інцидент безпеки. Система оснащена правилами для виявлення типових атак, загроз і системних збоїв, а також може бути розширена за допомогою користувацьких правил відповідно до специфіки середовища.

має можливості моніторингу цілісності файлів, які виявляють зміни у файлах та конфігураціях. Ця функція має вирішальне значення для дотримання різних правил і стандартів. Регулярно сканує відстежувані системи на наявність вразливостей, надаючи цінну інформацію для управління виправленнями та оцінки ризиків та допомагає підтримувати відповідність різним стандартам, таким як PCI DSS, GDPR та HIPAA, надаючи детальні звіти та журнали. Гнучкість і масштабованість Wazuh роблять його підходящим вибором для малих і великих підприємств. Його відкритий вихідний код дозволяє налаштовувати та інтегрувати з іншими інструментами, що розширює його можливості. Крім того, підтримка спільноти та постійний розвиток системи сприяють його надійності та актуальності у швидкозмінному ландшафті кібербезпеки.

Для розширеного використання Wazuh можна інтегрувати з іншими інструментами, такими як Elastic Stack, для покращеної візуалізації та аналітики даних. Поєднання функцій безпеки Wazuh з потужними можливостями пошуку та візуалізації даних Elastic Stack створює надійне рішення SIEM, яке може адаптуватися до різноманітних потреб сучасних IT-інфраструктур.

Інтеграція штучного інтелекту в систему Wazuh може значно розширити її можливості для моніторингу безпеки. ШІ можна використовувати для підвищення точності виявлення загроз і зменшення кількості помилкових спрацьовувань, як показав приклад системи SIEM з штучним інтелектом, яка перевершила традиційні методи машинного навчання у виявленні кіберзагроз, ефективно розрізняючи правдиві та хибні сповіщення [27].

Крім того, інструменти зі штучним інтелектом в SIEM можуть досягти високих показників відгуку на критичні оповіщення і вкрай низьких показників помилкових спрацьовувань, що значно знижує втому від оповіщень [28]. Таке застосування ШІ може підвищити здатність Wazuh відстежувати величезні обсяги даних і виявляти складні патерни зловмисної діяльності, сприяючи більш ефективним і результативним операціям з кібербезпеки. Крім того, методи на основі ШІ, включаючи байєсівські та онтологічні механізми, можуть бути впроваджені в SIEM системи для посилення моніторингу безпеки та прийняття рішень щодо вибору контрзаходів. Така інтеграція потенційно може перетворити Wazuh SIEM на більш динамічний, оперативний та прогностичний інструмент безпеки, що значно покращить стан кібербезпеки організації.

Для прикладу розглянемо можливості моніторингу команд Wazuh у поєднанні з Nmap для періодичного сканування сервісів на відкритих портах кінцевих точок. Модуль моніторингу команд Wazuh дозволяє виконувати задані команди на відстежуваних кінцевих точках, надаючи

можливість збирати важливу інформацію або виконувати заплановані завдання. Вихідні дані, що генеруються цими командами, записуються у вигляді журналу. Такі дані можна проаналізувати, щоб виявити потенційні загрози безпеці або отримати цінну інформацію про поведінку вашої мережі.

Конфігурація Wazuh

Для того, щоб забезпечити періодичне запитування служби відкритих портів кінцевих точок використовуємо можливість моніторингу команд Wazuh у поєднанні з Nmap. Модуль моніторингу команд Wazuh дозволяє виконувати певні команди на контрольованих кінцевих точках, забезпечуючи спосіб збору важливої інформації або виконання запланованих завдань. Результати, створені цими командами, фіксуються як дані журналу. Ці дані можна проаналізувати щоб виявити потенційні загрози безпеці або отримати цінну інформацію про поведінку мережі.

Для виконання конфігурації було виконано наступні кроки:

- Встановлення серверу Wazuh у віртуальному середовищі
- Встановлення операційної системи Ubuntu
- Перевірка функціонування системи

Для початку, нам потрібно встановити сервер Wazuh у віртуальному середовищі. Завантажуємо готовий віртуальний образ, який доступний на сайті документації проекту Wazuh. Цей образ являє собою віртуальну машину у форматі Open Virtual Appliance (OVA), яку можна легко імпортувати до VirtualBox або інших систем віртуалізації, що підтримують формат OVA. Важливо врахувати, що ця віртуальна машина призначена лише для 64-бітних систем і не забезпечує високий рівень доступності чи масштабованості за замовчуванням, хоча це можна налаштувати через розподілене розгортання.

Перед тим як імпортувати віртуальну машину Wazuh на хост-систему, потрібно забезпечити виконання таких умов:

– операційна система хоста повинна бути 64-розрядною;

– апаратна віртуалізація повинна бути;

– на хост-системі має бути встановлена платформа віртуалізації, наприклад, VirtualBox.

Стандартна віртуальна машина Wazuh має встановлені певні технічні специфікації "з коробки" (табл.3.1).

Таблиця 3.1 — Технічні характеристики потрібні для Wazuh VM

Компонент	CPU (ядер)	RAM (ГБ)	Жорсткий диск (ГБ)
Wazuh v4.7.0 OVA			

Реалізація доступу до інтерфейсу керування Wazuh

Після налаштування віртуальної машини Wazuh, наступним кроком є встановлення операційної системи Ubuntu. Ця операційна система використовуватиметься як кінцева точка, на якій буде розміщено агента Wazuh. Агент Wazuh - це компонент, який встановлюється на кінцевих точках для збору даних моніторингу і відправки їх до центрального сервера Wazuh для аналізу та обробки.

Встановлення Ubuntu як кінцевої точки дозволить використовувати цю систему для доступу до інтерфейсу керування Wazuh, звідки ви можете керувати налаштуваннями безпеки, моніторингом і відповіддю на інциденти.

Ubuntu є популярною та широко використовуваною операційною системою з великою кількістю підтримки та документації, що робить її хорошим вибором для встановлення агента Wazuh.

Щоб перевірити функціонування системи, спочатку потрібно її запустити. Для цього запускаємо віртуальну машину з встановленим Wazuh і входимо за допомогою стандартних (за замовчуванням) облікових даних. Після успішного входу, для з'єднання з панеллю керування Wazuh через Ubuntu, знадобиться IP-адреса віртуальної машини. Цю адресу можна отримати за допомогою команди **ip a**, яку потрібно виконати в терміналі машини з Wazuh (рис.3.1).

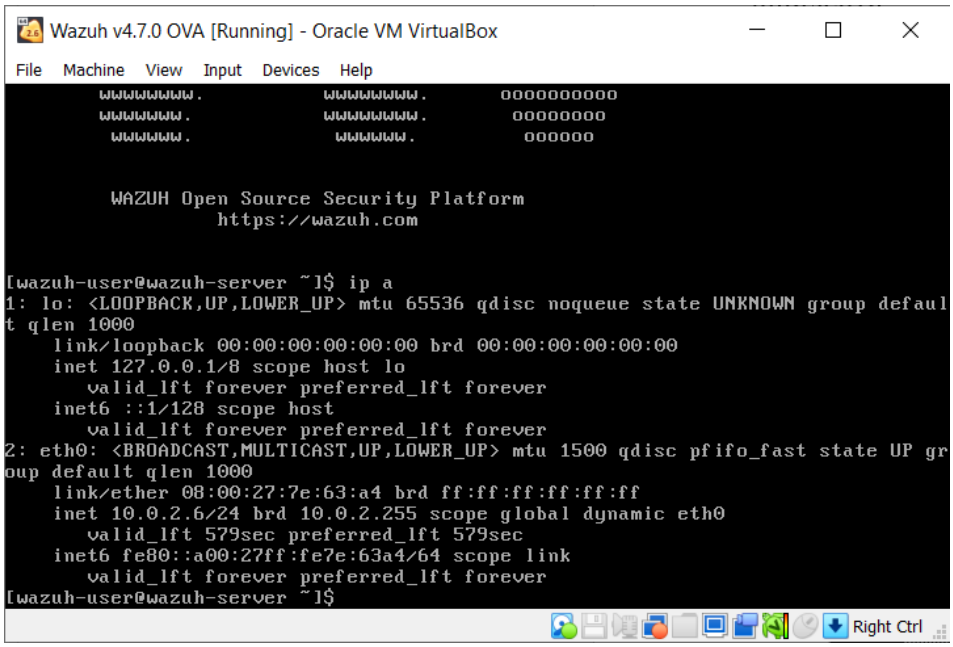


Рисунок 3.1 — Зображення IP-адреса віртуальної машини Wazuh

Потрібно перевірити мережеві налаштування на віртуальній машині Ubuntu, щоб впевнитися, що вона може досягати сервера Wazuh. Якщо мережевий доступ обмежений через налаштування брандмауера або інші мережеві політики, можливо потрібно буде внести зміни, щоб забезпечити безперебійний зв'язок.

Для доступу до панелі управління Wazuh потрібно ввести в адресну строку браузера IP-адресу сервера, де вона розгорнута, або IP-адресу віртуальної машини, як у цьому випадку та ввести стандартні облікові дані. З доступом до веб-інтерфейсу Wazuh, можна керувати налаштуваннями

системи безпеки, переглядати алерти та логи, а також налаштовувати правила і політики безпеки (рис. 3.2).

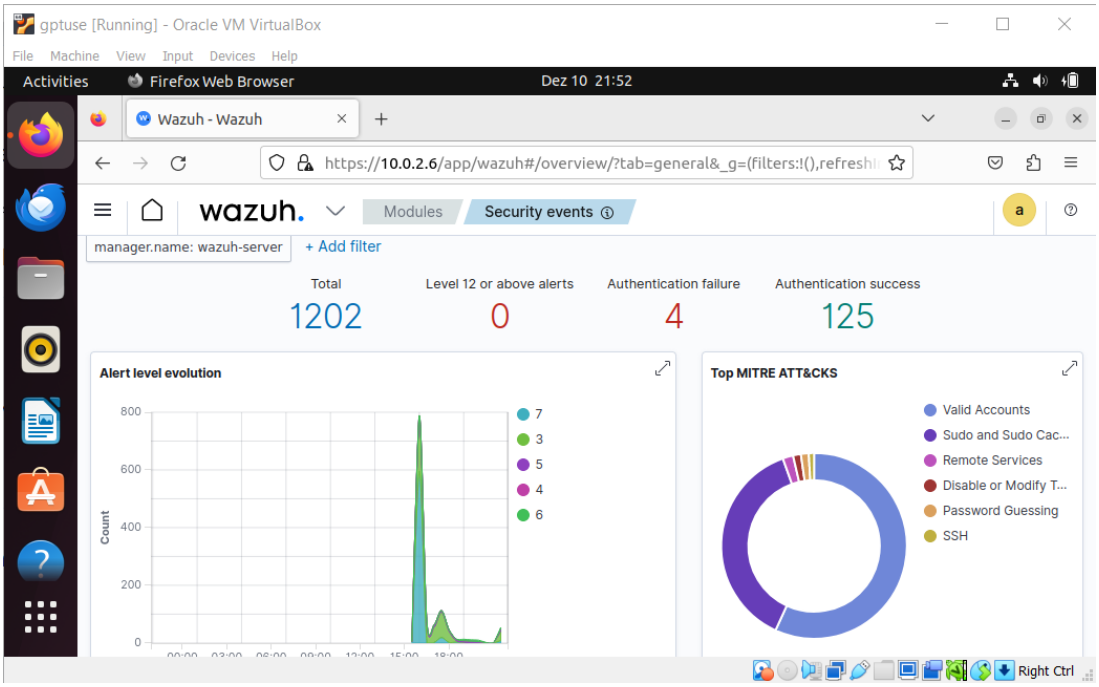


Рисунок 3.2 — Вигляд панелі управління Wazuh

Для того щоб встановити агента Wazuh на цій системі достатньо перейти в вкладинку «Agents», пролистати вниз та обрати кнопку «Deploy агента, потрібно його для зв'язку з сервером Wazuh, виконуючи конфігураційні кроки, надані в інструкції. Також потрбіно переконатися, що агент працює коректно і відправляє дані на сервер, перевіривши його статус на панелі управління.

3.1.4 Налаштування скрипту для автоматичного сканування мережі за допомогою утиліти Nmap

Розглянемо процес виконання сканування мережі за допомогою Nmap, використовуючи Python, щоб виявити відкриті порти на пристроях з операційною системою Ubuntu. Nmap (network mapper) — це сканер

безпеки з відкритим кодом, який використовується для дослідження мережі та аудиту безпеки. Він визначає кінцеві точки та служби в мережі та надає комплексну карту мережі.

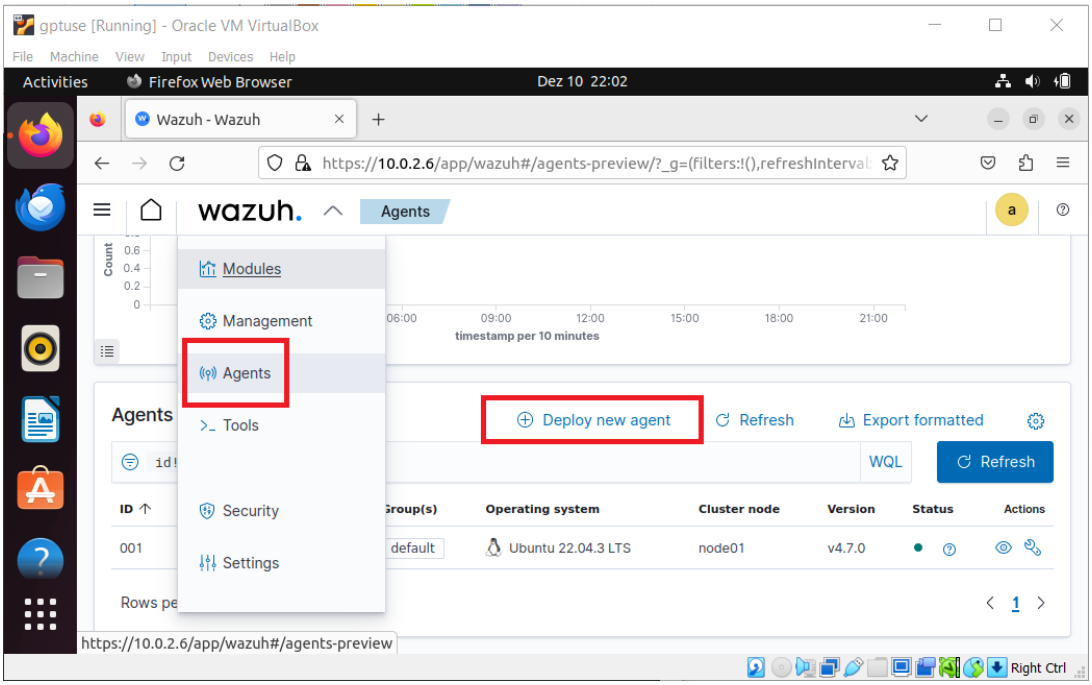


Рисунок 3.3 — Інформаційна панель активних агентів Wazuh

В якості базового, в роботі використано скрипт Python [29] (Додаток Б, лістинг 3.1), який дозволяє провести детальне мережеве сканування на цільовому пристрої.

Цей скрипт здатний витягувати різноманітну інформацію, таку як імена хостів, використовувані протоколи, а також перелік відкритих портів. Окрім цього, скрипт також може визначати типи служб, що працюють на цих портах, та надавати загальний огляд безпеки мережевого середовища.

Конфігурація автоматичного виконання скрипту через Wazuh виконується за допомогою (Лістинг 3.2). Скрипт запускається кожні 3600 секунд (1 година). Додамо виконання цього скрипта в автоматичні завдання агента Wazuh у файл ossec.conf

Лістинг 3.2 — Автоматичне виконання скрипту

Лістинг 3.3 представляє правило для обробки журналів, які генерує скрипт nmap, щоб Wazuh міг отримувати результати сканування nmap. Це правило визначає, як отримувати та обробляти дані з цих журналів, враховуючи відкриті порти та сервіси. У файл local_rules.xml додаємо наступне:

Лістинг 3.3 — Правило для отримання журналів nmap

Перевіряємо результат у модулі Security events. На рис. 3.4 показано сповіщення, які генеруються на панелі інструментів Wazuh, коли виконується сканування за допомогою Nmap на кінцевих точках Ubuntu і

Побачити згенеровані сповіщення можна на вкладці «Security Event».

	Time	agent.name	rule.description	rule.level	rule.id
›	Dec 10, 2023 @ 16:18:21.105	GPTuse	NMAP: Host scan. Port 631 is open and hosting the C UPS 2.4 service.	3	100100
›	Dec 10, 2023 @ 16:12:20.683	GPTuse	NMAP: Host scan. Port 631 is open and hosting the C UPS 2.4 service.	3	100100

Рисунок 3.4. — Сповіщення про відкриті порти

Інтеграція ChatGPT

Інтеграція ChatGPT у Wazuh SIEM може запропонувати кілька покращень, зокрема, в обробці природної мови (NLP) та автоматизованому реагуванні. Вдосконалені алгоритми NLP ChatGPT можуть аналізувати та інтерпретувати неструктуровані дані з журналів, сповіщень та звітів, надаючи більш інтуїтивно зрозумілі висновки та узагальнення. Така інтеграція дозволить ефективніше обробляти величезні обсяги текстових даних, що дасть змогу швидше виявляти потенційні загрози та аномалії.

Крім того, ChatGPT можна використовувати для автоматизованого реагування на інциденти, де він може генерувати сценарії або команди на основі аналізу оповіщень про загрози безпеці, оптимізуючи процес реагування. Ця інтеграція також може допомогти у створенні більш зручних інтерфейсів для Wazuh SIEM, дозволяючи користувачам взаємодіяти з системою за допомогою природної мови, тим самим знижуючи бар'єр для ефективного використання SIEM.

Інтеграція Wazuh ChatGPT — це конфігурація, яка дозволяє Wazuh спілкуватися з системою штучного інтелекту ChatGPT. Користувачі можуть покращити моніторинг безпеки та можливості реагування на інциденти, використовуючи потужність обробки природної мови. Інтегрувавши Wazuh із ChatGPT, ми можемо створити інтерфейс чат-бота, який взаємодіє з платформою Wazuh для виконання завдань, пов’язаних із безпекою, або надання додаткової інформації.

Розглянемо можливість використання штучного інтелекту для надання інформації про відкритий порт, а саме чи є це загрозою та що за сервіс на ньому розташований.

Але спочатку потрібно додати ще два правила, які будуть допомагати створювати практичний запит для ChatGPT:

Лістинг 3.4 — правила для побудови запиту для ШІ

Правило з ідентифікатором 100101 спрацьовує після успішного сканування Nmap на контрольованій кінцевій точці за умови, що на ній є один або декілька відкритих портів зі знайденим сервісом.

Правило з ідентифікатором 100103 спрацьовує після успішного сканування Nmap на контрольованій кінцевій точці з умовою, що є один або декілька відкритих портів без знайденого сервісу.

Створемо скрипт інтеграції під назвою /var/ossec/integrations/custom-chatgpt.py і використаємо наведений нижче скрипт Python (Додаток В, лістинг 3.5) до файлу custom-chatgpt.py. Наведений нижче скрипт Python фіксує відкриті порти на кінцевій точці та надсилає їх до ChatGPT, щоб отримати інформацію про відкриті сервіси та минулі вразливості.

Також потрібно дадати конфігурацію до сервера Wazuh у файл

Лістинг 3.6 — Використання ключа API ChatGPT

```
…T3BlbkFJcyY1SQSqXunxq30wj5Rl></api_key>
```

Щоб переглянути згенеровані сповіщення прейдімо на вкладку «Security Event». У такому сповіщенні ми бачимо які самі дані перейшли до ChatGPT та його відповідь у полі data.chatgpt.choices (рис. 3.5).

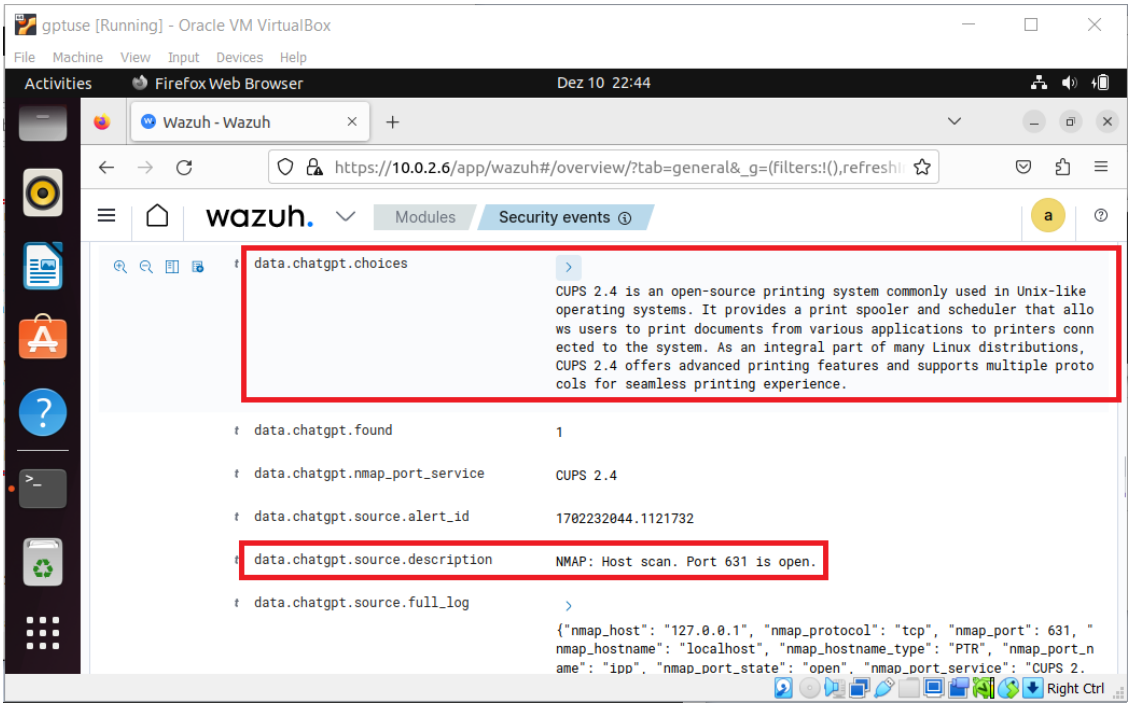


Рисунок 3.5 — Результат роботи алгоритму

Після проведення цього експерименту було розроблено функціональний алгоритм, який автоматизує запити до ChatGPT та інтегрує відповідь мовної моделі ШІ в журнали Wazuh. Ця інтеграція посилює заходи кібербезпеки завдяки використанню передових аналітичних можливостей ШІ. Скрипт, який використовується для автоматизованого запуску

сканування відкритих портів утилітою Nmap, ефективно передає всі виявлені відкриті порти до SIEM систем. Згодом активується правило для використання ШІ для обробки цих даних.

Таке застосування ШІ значно прояснює розуміння існуючих проблем кібербезпеки без необхідності великого пошуку документації або онлайн-рішень. Наприклад, випадок з відкритим портом 631, який використовується для сервісу CUPS, був прояснений за допомогою ШІ. Він надає важливу інформацію про сервіс, зокрема про те, чи становить відкритий порт загрозу. ШІ підтверджує, що поточна версія сервісу є найновішою, але попереджає про вразливості у версіях до 2.0.3, наголошуючи на необхідності регулярного оновлення та пильного обслуговування.

Використання моделі штучного інтелекту ChatGPT у такий спосіб дає переваги швидкого реагування, особливо в інтерпретації проблем кібербезпеки. Однак він не позбавлений недоліків. Хоча створення такої системи, як показано у прикладі, є економічно ефективним, розгортання штучного інтелекту, може спричинити значні фінансові та репутаційні ризики. Тому перед прийняттям остаточного рішення слід використовувати різні методи оцінки ефективності рішення, щоб переконатися, що обраний підхід відповідає конкретним потребам організації.

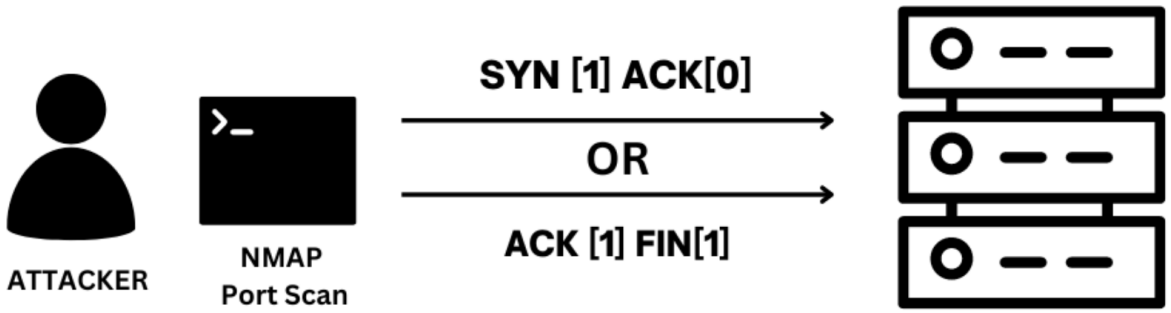
Виявлення сканування портів за допомогою ChatGPT

Запит SIEM для виявлення сканування портів за допомогою ChatGPT, ця спеціальна технологія допомагає групі безпеки створити певне базове правило виявлення для кількох атак.

В цій частині роботи було створено і перевірено три різні правила виявлення для різних інструментів SIEM [30].

Зазвичай зловмисники використовують сканування портів для виявлення вразливих відкритих портів на цільових серверах.

Базовий метод сканування портів, який надсилає послідовність пакетів до кожного з 65 536 портів одночасно і передбачає тристороннє рукостискання з використанням прапора SYN, відповіді SYN-ACK і прапора ACK для визначення вразливих служб.



Р

и

3.2.1 Запит для фільтрації даних в контексті пошуку подій в SIEM

у

Наведений у лістингу 3.7 запит SIEM в основному перевіряє тристороннє рукостискання TCP-з’єднання, сканування порту і містить комбінацію пакета SYN (використовується для ініціювання TCP-з’єднання) із невстановленим прапором ACK[0] або FIN-пакету (використовується для закриття TCP-з’єднання) або підключення із встановленим прапором Elasticsearch Query DSL, який є мовою запитів для пошукового движка

6

Перший блок "bool" - це складена умова, яка включається в "should". У цьому випадку, блок вимагає, щоб були виконані дві умови: (1) TCP-прапор з флагом SYN повинен дорівнювати 1 (це зазвичай означає початок TCP-з'єднання) і (2) TCP-прапор з флагом ACK повинен дорівнювати 0 (це

о

т

а

означає, що прапор ACK не встановлений). Другий блок "bool" також включається в "should". Умови в цьому блоку вимагають: (1) TCP-прапор з флагом FIN повинен дорівнювати 1 (це зазвичай означає завершення TCP-з'єднання) і (2) TCP-прапор з флагом ACK повинен дорівнювати 1 (це означає, що прапор ACK встановлений). Отже, цей запит вибирає дані, які відповідають умовам, пов'язаним з TCP-прапорами SYN та FIN у мережевому трафіку.

Лістинг 3.7 —Запит для фільтрації даних в SIEM

Таким чином, цей запит шукає документи, які мають вказані умови, пов'язані з TCP-флагами. Якщо будь-яка з цих умов виконується для документу, то цей документ буде включений у результати запиту. Приклад записів, які задовольняють умовам надано на рис. 3.7.

Відсутність підтвердження (ACK=0) після синхронізації (SYN=1) може вказувати на потенційний SYN-скан. SYN-скан часто

використовується зловмисниками для визначення доступних портів на системі без встановлення повного TCP-з'єднання.

Наявність підтвердження (ACK=1) після фіналізації (FIN=1) може вказувати на фінську атаку (FIN/ACK-атаку). Це може бути спроба визначити відкриті порти або вразити систему через закриття активних TCP-з'єднань.

Важливо відзначити, що це всього лише можливі індикатори атак або аномальної активності. Перед тим, як робити подальші висновки, важливо враховувати контекст і можливі інші фактори, такі як інтенсивність та тривалість цих подій.


```
{
    "timestamp": "2023-01-01T12:00:00",
    "tcp": {
        "flags": {
            "syn": 1,
            "ack": 0
        }
    }
},
{
    "timestamp": "2023-01-01T12:00:05",
    "tcp": {
        "flags": {
            "syn": 1,
            "ack": 0
        }
    }
},
{
    "timestamp": "2023-01-01T12:00:10",
    "tcp": {
        "flags": {
            "fin": 1,
            "ack": 1
        }
    }
},
```


Рисунок 3.6 – Приклад записів що включені у результати запиту

3.2.2 Запит SPLUNK

SPLUNK-запит має на меті виявлення потенційних порушень безпеки, таких як сканування портів.

Лістинг 3.8 – Запит SPLUNK

Результатом буде таблиця, яка включає два стовпці: src_ip (IP-адреса джерела) та dest_ip (цільовий IP-адреса), а також стовпець count, який вказує кількість відповідних записів. Умова where count > 100 обмежує результати лише тими випадками, де кількість входжень більше 100. Такий фільтр може використовуватися для визначення активності, яка може свідчити про масштабні атаки або потенційні проблеми з безпекою.

Важливо відзначити, що результати SPLUNK-запиту будуть залежати від наявності відповідних подій у системі журналювання та правильного визначення поля sourcetype та tcp.flags в логах. Також, перед запуском подібних запитів, слід ретельно аналізувати контекст системи та впевнитися, що цей метод виявлення відповідає потребам в безпеці.

Припустимо, що є дані журналу, які виглядають як на рисунку 3.7:

timestamp	sourcetype=tcp	src_ip	dest_ip	tcp.flags.syn	tcp.flags.ack	tcp.flags.fin
2023-01-01T12:00:00	tcp	192.168.1.1	10.0.0.1	1	0	0
2023-01-01T12:00:05	tcp	192.168.1.2	10.0.0.1	1	0	0
2023-01-01T12:00:10	tcp	192.168.1.3	10.0.0.1	0	1	1
2023-01-01T12:00:15	tcp	192.168.1.4	10.0.0.1	1	0	0
...						

Рисунок 3.7 — Журнали мережевого трафіку в SPLUNK

В журналі подій вказані IP-адреси джерела (src_ip) та цільова IP-адреса (dest_ip), а також прапорці TCP-пакетів: syn, ack, та fin.

Якщо використовувати SPLUNK-запит наданий вище, результат буде виглядати, як на рисунку 3.8.

src_ip	dest_ip	count
192.168.1.1	10.0.0.1	150
192.168.1.2	10.0.0.1	110
...		

Рисунок 3.8 — Вигляд журналів після виконання запиту

Цей результат показує, що є певні пари IP-адрес, для яких кількість виявлених подій (де syn=1 та ack=0, або fin=1 та ack=1) перевищує 100, що може вказувати на потенційні проблеми або атаки.

Запит в Wazuh

Щоб створити правило для виявлення подібних ситуацій в системі Wazuh, використаємо мову правил Wazuh. Нижче наведений приклад правила для виявлення ознак сканування портів у Wazuh (лістинг 3.9)

Лістинг 3.9 – Запит в системі Wazuh

Де if_sid вказує на те, що це правило базується на існуючому правилі з id 550, яке, можливо, визначає подібні ситуації;

match містить умови, що вказують на можливий SYN скан або FIN/ACK атаку;

description надає короткий опис того, що виявлено;

group вказує на групи, до яких відноситься правило (firewall, attack,

Це правило перевіряє, чи є в TCP-пакеті прапорець SYN, і якщо так, чи прапорець ACK встановлено в 0.

Якщо ці умови виконуються, генерується тривога "Suspicious SYN scan detected". Те саме стосується атаки з прапорцем FIN та ACK.

Аномалії, які можуть виникнути, включають у себе:
егітимні аномалії, наприклад, деякі специфічні мережеві операції чи конфігурації які можуть впливати на TCP-пакети.
alse Positives (Помилкові позитиви). Правило може спрацьовувати на тимчасові аномалії, що виникають з-за регулярних змін у мережевій активності, і спричиняти помилкові позитиви.
alse Negatives (Помилкові негативи). Якщо атака ведеться способом, який не спричиняє прапорців SYN або FIN/ACK, то правило може не виявити цю атаку.

Важливо зазначити, що правило перевіряє тільки певні комбінації прапорців, і можливо, не охоплює всі можливі атаки або ненормальні ситуації. Усі ці аспекти підкреслюють важливість тонкого налаштування та валідації правил, а також їхню регулярну перевірку та адаптацію до змін у середовищі.

ChatGPT може бути використаний для покращення та розширення правил щодо детекції аномалій у системах безпеки. Він може використовуватися для:

Легітимних аномалій - при виникненні нових мережевих або системних подій, які можуть призвести до виникнення помилкових сигналів, ChatGPT може аналізувати ці події, враховуючи контекст та інші фактори, і визначити, чи є вони легітимними. Він може допомогти у розумінні та оцінці нових обставин.

alse Positives (Помилкові позитиви) – ChatGPT може допомогти перевіряти велику кількість подій та аналізувати їх контекст, враховуючи попередні поведінки системи. Він може визначити, чи є деякі події випадковими або специфічними для конкретного середовища, що допомагає зменшити кількість помилкових сигналів.

Шляхом аналізу попередніх атак та їх характеристик, ChatGPT потенційно може допомогти визначити нові способи ведення атак, які можуть бути пропущені звичайними правилами. Він може виявити невизначені або менш відомі атаки, допомагаючи покращити детектори.

Крім того, ChatGPT може надати додатковий контекст для адаптації правил у реальному часі. Це може включати в себе динамічне змінення правил на основі зміни ситуацій, що допомагає більш гнучко реагувати на нові загрози.

Таким чином, можна зробити висновок, що ChatGPT може бути використаний як інструмент для аналізу та розуміння складних сценаріїв у кібербезпеці, враховуючи велику кількість інформації та забезпечуючи контекстне розуміння подій. Інтеграція дозволяє ChatGPT взаємодіяти з Nmap і Wazuh, використовуючи обробку природної мови для надання інтелектуальної допомоги.

3.4 Обмеження використання ChatGPT в SIEM

Обмеження інтеграції ChatGPT в SIEM системи стають очевидними, якщо врахувати поточну структуру обробки ШІ та природу кіберзагроз.

Кожна взаємодія з ChatGPT обробляється як дискретна подія. Це означає, що для таких додатків, як написання програмного коду або встановлення правил для автоматизованого реагування на інциденти, результат не завжди може бути негайно функціональним або контекстуально точним. Часто потрібне втручання людини для уточнення та перевірки цих результатів, що може бути суттєвим недоліком у сценаріях, чутливих до часу.

Відповіді ChatGPT генеруються безвідносно до попередніх взаємодій. Такий ізольований підхід може призвести до неадекватності, особливо при обробці складних або постійних кіберзагроз. Такі інциденти, як правило, вимагають тонкого розуміння контексту, що змінюється, з чим сучасні моделі штучного інтелекту, такі як ChatGPT, можуть мати труднощі. У випадках тривалих або складних кібератак, коли безперервність і розуміння еволюції атаки мають вирішальне значення, автономні відповіді ChatGPT можуть бути не тільки недостатніми, але й потенційно призвести до помилкових стратегій реагування. Крім того, така система повинна включати в себе надійні заходи безпеки, щоб не стати вектором для атак. Оскільки системи штучного інтелекту стають все більш інтегрованими в кібербезпеку, вони також стають потенційними цілями. Забезпечення цілісності ШІ та його захисту від маніпуляцій і експлуатації має першорядне значення. Це передбачає впровадження надійного шифрування даних, безпечних методів кодування, регулярний аудит безпеки і, можливо, навіть розробку спеціальних захисних механізмів для ШІ.

Ефективність різних моделей штучного інтелекту, інтегрованих в системи управління інформацією та подіями безпеки (SIEM), залишається неоднозначною до моменту розгортання повноцінної системи. Ця невизначеність створює значні ризики для організацій, які інвестують значні ресурси в інфраструктуру кібербезпеки, очікуючи відчутних результатів.

З технічної точки зору, інтеграція ШІ в SIEM системи передбачає складні алгоритмічні взаємодії та можливості аналізу даних. Ефективність цих моделей ШІ може бути непередбачуваною через такі змінні, як різноманітність кіберзагроз, складність мережевих середовищ і постійно мінливий характер векторів атак. Ці системи штучного інтелекту часто навчаються на історичних даних, які можуть не повністю відображати поточні або майбутні кіберзагрози, що призводить до потенційних прогалин у можливостях виявлення та реагування.

Висновки до третього розділу

У розділі було продемонстровано інтеграцію LLM ChatGPT у Wazuh для отримання аналізу результатів сканування утилітою Nmap. В отриманій системі відбувається аналіз цілі на наявність відкритих портів, що потім передається у LLM та обробляється нею. Як результат отримуємо інформацію про те для чого використовується відкритий порт та чи це становить загрозу. Також, було створено і перевірено три різні правила виявлення для різних інструментів SIEM. Виявлено основні обмеження використання ChatGPT.

4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

хорона праці

Інтеграція штучного інтелекту для ефективного реагування на інциденти у SIEM системі потребує проведення багато робочого часу спеціаліста за комп'ютером, тому необхідно дотримуватись норм роботи для працівників. Норми по охороні праці та техніки безпеки при роботі за електронно обчислювальною технікою регламентуються НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями» [31], вимогами та нормами державно-санітарної служби при роботі з дисплеями комп'ютерної техніки ДСанПІН 3.3.2.007-98 [32], Санітарними вимогами шуму, інфразвуку та ультразвуку на виробництві ДСН 3.3.6.037-99 [33] та Санітарними мікрокліматичними вимогами до приміщень на виробництві ДСН 3.3.6.042-99 [34].

НПАОП 0.00-7.15-18 [31] — описує вимоги безпеки та охорони праці для працівників, які використовують екранні пристрої у своїй роботі. Робочі місця працівників з екранними пристроями мають бути спроектовані так і мати такі розміри, щоб працівники мали простір для зміни робочого положення та рухів. Для забезпечення безпеки та захисту здоров'я працівників усе випромінювання від екранних пристроїв має бути зведене до гранично допустимого рівня (вплив на людину факторів довкілля - шуму, вібрації, забруднювачів, температури тощо, який не спричиняє соматичних або психічних розладів, а також змін стану здоров'я, працездатності, поведінки, що виходять за межі пристосувальних реакцій) з погляду безпеки та охорони праці працівників. Освітлення робочого місця працівника з екранними пристроями має створювати відповідний контраст між екраном і навколишнім середовищем (з урахуванням виду роботи) та

відповідати вимогам ДСанПІН 3.3.2.007-98.

Роботодавець повинен за рахунок тривалості робочої зміни організувати внутрішні регламентовані перерви для відпочинку відповідно до Державних санітарних правил і норм роботи з візуальними дисплейними терміналами електронно-обчислювальних машин ДСанПІН 3.3.2.007-98. Обладнання і організація робочого місця працівників мають забезпечувати відповідність конструкції всіх елементів робочого місця та їх взаємного, розташування ергономічним вимогам з урахуванням характеру і особливостей трудової діяльності.

ДСанПІН 3.3.2.007-98 [32] є державним санітарним нормативом, який описує стандарти охорони праці для працівників, що користуються комп'ютерними дисплеями. Він включає вимоги до ергономіки робочих місць, забезпечуючи правильне розташування моніторів, клавіатур, та іншого обладнання для мінімізації фізичного дискомфорту. Також норматив описує стандарти освітлення, які допомагають знижувати втому очей та запобігають відблискам, що можуть спричинити зоровий дискомфорт. Документ містить вказівки щодо контролю рівнів шуму та електромагнітного випромінювання, щоб забезпечити безпечне та комфортне робоче середовище.

Також роботодавець або замовник повинен поінформувати працівників про умови праці та наявність на робочих місцях небезпечних та шкідливих виробничих факторів (фізичних, хімічних, біологічних, психофізіологічних), які виникають під час роботи з екранними пристроями та ще не являються виправленими, а також про можливі наслідки впливу цих шкідливих факторів на здоров’я працівників відповідно до вимог статті 5 Закону України „Про охорону праці”.

ДСН 3.3.6.037-99 [33] визначає санітарні норми виробничого шуму, ультразвуку та інфразвуку, встановлюючи класифікації цих факторів, методи їх гігієнічної оцінки, параметри, які потребують нормування, та

вимоги до вимірювань на робочих місцях. Ці норми є обов'язковими для всіх підприємств та організацій, незалежно від форм власності, які проектують, виготовляють або експлуатують обладнання, механізми та інструменти, що є джерелами шуму, ультразвуку, та інфразвуку. ДСН 3.3.6.037-99 [33] забезпечує здорове робоче середовище шляхом встановлення максимально допустимих рівнів шуму, ультразвуку та інфразвуку, вимог до організації робочих місць, обмеження тривалості впливу цих факторів на працівників, а також рекомендує заходи щодо зниження впливу негативних акустичних факторів, які включають використання захисних засобів, технічні заходи щодо зниження шуму на джерелі, та організаційні заходи.

ДСН 3.3.6.042-99 [34] — це нормативний документ, який встановлює санітарні вимоги до мікроклімату в приміщеннях на виробництві. Він визначає оптимальні та допустимі параметри температури, вологості, швидкості руху повітря та інших факторів мікроклімату, що мають велике значення для забезпечення комфортних та безпечних умов праці. Враховуючи специфіку різних видів діяльності та особливості робочих процесів, цей документ спрямований на створення умов, які сприяють збереженню здоров'я та продуктивності працівників. Мікроклімат в приміщенні інтегратора з електронно- обчислювальною машиною потрібно підтримувати на сталому рівні та відповідно нормам описаним в ДСН 3.3.6.042-99 [34].

Згідно з наказом Міністерства охорони здоров'я України від 14 лютого 2012 року № 107 [35], роботодавець зобов'язаний забезпечувати проведення медичних оглядів працівників певних категорій за свій рахунок. Цей наказ встановлює порядок медичних оглядів, визначаючи, які категорії працівників підлягають обов'язковим медичним оглядам, їх періодичність, та які медичні показники повинні бути оцінені. Це забезпечує контроль за станом здоров'я працівників, особливо тих, хто працює в умовах

потенційного ризику для здоров'я.

Виходячи з цього, можна використовувати такі методи впливу на мотиви, які стимулюють безпечну поведінку працівників: установити працівникам чітку мету щодо дотримання правил безпеки; створити умови для можливості досягнення цієї мети. Під час виконання кваліфікаційної роботи було дотримано усіх вищевказаних норм щодо охорони праці.

кідливий вплив іонізуючого випромінювання

Іонізуючі випромінювання знаходять широке використання в різних галузяхпромисловості. Їх використовують для автоматичного контролю технологічних процесів, контролю якості виробів, зварних швів, структури металів тощо. Длявиробництва електроенергії на атомних електростанціях необхідне ядерне пальне, виробництво якого, починаючи від добування уранової руди і закінчуючи виготовленням та транспортуванням паливних елементів, призводить до опромінення персоналу. Незначні додаткові дози опромінення працівники отримують від таких техногенних джерел, як теплові електростанції (підвищена активність їх відходів та аерозолів), підприємств, які пов’язані з видобуванням та переробкою корисних копалин, а також різноманітних приладів та обладнання з джерелами випромінювання, що знаходять широке використання у промисловості і сільськогосподарському виробництві. Основним документом, що встановлює радіаційно-гігієнічні регламенти для забезпечення прийнятих рівнів опромінення, є Норми радіаційної безпеки України (НРБУ-97) [36].

НРБУ-97 регламентують опромінення людини джерелами іонізуючого випромінювання в умовах:

– Н

о – медичної практики;

р – радіаційних аварій;

м

а

л

– опромінення техногенно-підсиленими джерелами природного походження.

Відповідно до цього НРБУ-97 встановлено чотири групи радіаційно-гігієнічних регламентів:

– перша – обмежує опромінення від ядерно-радіаційних об’єктів;

– друга – обмежує опромінення людей від медичних джерел;

– третя – обмежує опромінення в умовах радіаційних аварій;

– четверта – обмежує опромінення від техногенно підсилених джерелприродного походження.

Враховуючи різнобічні наслідки опромінення людей іонізуючим випромінюванням, їх нормування здійснюється залежно від категорії людей, що опромінюються, а також від чутливості органів тіла людини, на які діє іонізуюче випромінювання.

Виділяють наступні категорії:

– особи з числа персоналу, які постійно чи тимчасово працюють безпосередньо з джерелами іонізуючого випромінювання;

– особи з числа персоналу, які безпосередньо не зайняті роботою з джерелами іонізуючого випромінювання, але у зв’язку з розташування робочих місць в приміщеннях та на промислових майданчиках об’єктів з радіаційно ядерними технологіями можуть отримувати додаткове опромінення;

– все населення.

Для осіб категорій А та Б НРБУ-97 [36] встановлюються ліміти річних ефективних доз зовнішнього опромінення, а також ліміти річних еквівалентних доз зовнішнього опромінення окремих органів і тканин людини. Аналогічні ліміти вводяться і для критичних груп осіб категорії В.

Є також обмеження стосовно швидкості накопичення дози для жінок дітородного віку та вагітних жінок, підвищеного опромінення в

непередбачуваних ситуаціях та інші. Крім лімітів дози опромінення, встановлюють допустимі рівні (ДР): потужності дози зовнішнього опромінення, забруднення поверхонь, надходження радіонукліду через органи дихання тощо, які визначають виходячи ізнаведених лімітів дози опромінення. З метою зниження рівнів опромінення населення Міністерство охорони здоров'я України запроваджує рекомендованірівні медичного опромінення.

Медичне опромінення – це опромінення працівників при медичних обстеженнях чи лікуванні. Опромінення повинно бути обґрунтованим і призначеним тільки лікарем для досягнення корисних діагностичних та терапевтичних ефектів, які неможливо отримати іншими методами діагностики талікування. Рекомендовані рівні медичного опромінення та детальні вимоги до обмеження і контролю за опроміненням пацієнтів регламентуються окремими спеціальними документами Міністерства охорони здоров'я України. Для радіометричного і дозиметричного контролю використовуються:

– дозиметри – для вимірювання зовнішніх потоків радіоактивного випромінювання;

– радіометри – для вимірювання рівнів забруднення навколишнього середовища; індивідуальні

– дозиметри – для індивідуального контролю.

Серед індивідуальних дозиметрів найбільше розповсюджені прилади, в яких використовують іонізаційні (за величиною іонізації середовища, через яке пройшло випромінювання) та фотографічні (за величиною опроміненняфотографічної плівки іонізуючим випромінюванням) методи виміру. У приладах для контролю потужності дози випромінювання широко застосовують іонізаційний та сцинтиляційний методи (за інтенсивністю світлових спалахів, що виникають внаслідок люмінесценції в деяких речовинах під час проходження через них іонізуючих випромінювань).

При роботі з джерелами іонізуючих випромінювань здійснюють контроль іоцінку параметрів радіаційного фактора відповідно до НРБУ-97 [36]. При дотриманні контрольних рівнів умови праці на даному робочому місці оцінюютьсяяк допустимі. У разі їх перевищення оцінка шкідливості та небезпечності за радіаційним фактором здійснюється органами Держсанепіднагляду. Засоби та заходи захисту від іонізуючих випромінювань поділяють на організаційні, технічні, санітарно-гігієнічні та лікувально-профілактичні.

Як правило, ефективний захист від іонізуючого випромінювання досягається при одночасному комплексному використанні зазначених заходів та засобів. При їх виборі враховуються особливості джерел випромінювання. Так, основними заходами, направленими на захист від альфа- та бета-випромінювань, є заходи, що націлені на недопущення накопичення альфа- і бета-активних ізотопів в організмі людини та забруднення шкіри: використання спеціального одягу та взуття, протипилових респіраторів, обезпилення повітря, вологе прибирання помешкань, недопущення вживання радіоактивно забруднених харчових продуктів, води та інші.

При роботі з джерелами гама- та рентгенівського випромінювання захист персоналу досягається шляхом зниження активності джерел випромінювання, обмеження часу роботи з ними, збільшення відстані до джерел, екранування джерела іонізуючого випромінювання або зони знаходження людини.

Також у випадку такої радіаційної аварії забруднюється навколишнє середовище, люди можуть отримати травму у вигляді потужної дози опромінення.Призвести аварію на підприємстві може також якщо активна реакційна речовина знаходиться у роботі та це відбувається незаконно. Це може привезти до опромінення жителів та перевищити межу дози опромінення. Частинки з цього випромінювання можуть залишати сліди на

дихальній системі на травній системі людського організму. Також ці елементи можуть бути у водних каналах, які постачають питну воду людям.

На підприємстві де проводяться роботи з радіаційними речовинами обов'язково мають вживатись заходи проти радіації. Протирадіаційні захисти це така система правових, організаційних норм та санітарної гігієни. До переліку таких захистів можна включити медичні заходи для забезпечення радіаційної безпеки персоналу та проектно-конструкторські. Для організації заходів проти іонізації опромінювання підприємство має ввести обов'язкові методи щоб подбатипро безпеку працюючого персоналу. До таких методів можуть належати заходи які обмежують допуск працівників до джерел які випромінюють радіацію. До таких працівників можемо віднести таких, які не підходять за віком, за статтю та працівники які вже отримали дозу випромінення.

Підприємство мусить створити сприятливі умови що дотримуються встановлених норм та вимог для працівників та застосовувати індивідуальні засоби для захисту працівника цього підприємства. Організація повинна контролювати рівні опромінювання та вести інформаційну систему про стан радіації на підприємстві та призначених місць для праці.

На підприємстві повинні бути проведені заходи щодо організації безпеки для робіт які проводяться у радіаційних ділянках а саме:

– організація роботи нарядів та розпоряджень; -організація та перевірка пропусків до робочих місць;

– оформлення контролю за процесом виконання роботи;

– введення примусового часу на перерву та вчасне закінчення робочого процесу.

До фізичних норм захисту проти радіації існують перешкоди поширення іонізації опромінень. Для поширення дози випромінювання може бути ряд перешкод, залежать вони від кількості годин, перешкоджати може дистанція , також перешкодою може бути чисельність.

Реалізувати заходи проти радіації за певний відрізок часу можливо, тим щопрацівники , які працюють з іонізованими випромінюванням можуть виконувати вчасно свою роботу, відповідно керівництво може за якісну роботу зменшити кількість робочих днів у тижні. Цим самим вони застереженням вони зменшать знаходження працівників у зоні випромінювання та відповідно буде менше контактування з радіаційними приладами.

Захистити працівників за допомогою відстані підприємство може шляхом доцільного розміщення приміщення, правильно розставити та розрахувати робочі місця для працівників а також забезпечити приладами, які зможуть контактувати, керувати робочим процесом з технікою яка має радіаційний вплив на відстані. Слугувати захистом може покриття свинцем меблів які присутні у приміщенні (двері, вікна, робочі столи), створення перекриття між поверхами та перегородки. Працівникам обов’язково має бути виданий спеціальний одяг ,такі як фартухи, шапочки та рукавиці зшиті з просвинцевої тканини.

Розміщення робочих місць повинно мати правильний розрахунок на загальну кімнату, не робити перенабір та забезпечити відповідним та необхідним обладнанням робочі кабінети.

исновки до четвертого розділу

У результаті аналізу вимог щодо охорони праці було визначено, які саме нормативні документи потрібно використовувати для забезпечення ефективного та безпечного робочого місця. Під час виконання кваліфікаційної роботи з використання штучного інтелекту для ефективного реагування на інциденти у SIEM системі було дотримано усіх вищевказаних норм щодо охорони праці.

Було описано деталі основного документом, що встановлює

радіаційно- гігієнічні регламенти для забезпечення прийнятих рівнів опромінення – Норми радіаційної безпеки України (НРБУ-97).

ВИСНОВОК

Отже, у результаті аналізу існуючого стану SIEM систем з інтеграцією ШІ, можна виділити, що наразі існуючі системи тільки починають розвиватися, проте наявні рішення вже впливають на основні процеси у захисті кібербезпеки. Машинне навчання дедалі частіше використовується для автоматизації таких завдань, як аналіз трафіку, виявлення аномалій у шаблоні поведінки користувачів, створення звітів про наявні загрози.

Проте використання ШІ потребує наявності якісних даних, які будуть використовуватися для навчання, що є основною задачею у реалізації SIEM систем з ШІ. При розробці заходів кібербезпеки слід враховувати атаки на сам алгоритм штучного інтелекту, такі як FGSM, C&W, тощо.

Алгоритм RETE відіграє ключову роль у створенні правил для виявлення TCP SYN flood атак. Водночас, алгоритми для виявлення аномалій, як, наприклад, k-Nearest Neighbors (k-NN), є важливими у контексті систем управління інформацією та подіями безпеки, як це підкреслено у дослідженні. Такі алгоритми ефективно виявляють некоректні налаштування та відхилення в даних, що є критичним для забезпечення надійності та безпеки інформаційних систем.

Найбільші та найефективніші LLM навчаються на наборах даних із трильйонів слів за допомогою сканування Інтернет-тексту, книг та інших текстових джерел. Таким чином, LLM пройшли підготовку зі звітами та ресурсами з кібербезпеки, а також із кодом, пов’язаним із загрозами, і кодом захисту. Як наслідок, LLM такі як ChatGPT мають можливість аналізувати знайдені кіберзагрози та надавати звіт про них за лічені секунди. Така система більш чітко класифікує знання про конкретні кіберінциденти та може використовуватися для навчання або підвищення кваліфікації початківців у кібербезпеці.

Продемонстровано спосіб інтеграції LLM ChatGPT у SIEM систему Wazuh та інші. Така інтеграція дозволяє автоматично отримувати

інформацію, про наявну проблему та аналізувати чи наразі вона становить вразливість для системи.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

D. Molina-Markham, and Julian T. Sexton, "'Draft NISTIR 8269, A Taxonomy and Terminology of 20 Adversarial Machine Learning."' NIST, Oct. 2019.

еннадій Андрощук. ШТУЧНИЙ ІНТЕЛЕКТ: ЕКОНОМІКА, ІНТЕЛЕКТУАЛЬНА ВЛАСНІСТЬ, ЗАГРОЗИ. Теорія і практика інтелектуальної власності. 2021. URL:

https://doi.org/10.48550/arXiv.2310.06936.

ДСН 3.3.6.037–99. Санітарні норми виробничого шуму, ультразвуку та

Постанова 01.12.99 № 42 «Санітарні норми мікроклімату виробничих приміщень ДСН 3.3.6.042-99». URL:

Про затвердження Порядку проведення медичних оглядів працівників певних

Про введення в дію Державних гігієнічних нормативів "Норми радіаційної

Додаток А

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ПУЛЮЯ

М А Т Е Р І А Л И

XI НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

«ІНФОРМАЦІЙНІ МОДЕЛІ, СИСТЕМИ ТА ТЕХНОЛОГІЇ»

13-14 грудня 2023 року

ТЕРНОПІЛЬ
2023

УДК 004.056
Кубарич З.П., Скарга-Бандурова І.С., д.т.н., проф.
Тернопільський національний технічний університет імені Івана Пулюя, Україна

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ЕФЕКТИВНОГО РЕАГУВАННЯ НА ІНЦИДЕНТИ У SIEM СИСТЕМІ

Z.P. Kubarych, I.S. Skarga-Bandurova, DSc, Prof

USING ARTIFICIAL INTELLIGENCE FOR EFFECTIVE INCIDENT RESPONSE IN SIEM SYSTEM

Зростаюча кількість і складність кібератак підкреслюють нагальну потребу в інноваційних рішеннях для посилення безпеки цифрової інфраструктури. Security Incident and Event Management (SIEM) традиційно покладаються на звичайні механізми портів і переадресації для накопичення журналів подій та їхнього аналізу на основі встановлених сценаріїв. Однак, з розвитком штучного інтелекту (ШІ) і швидкою зміною методологій атак, ці традиційні методи стикаються з істотними перешкодами, особливо при обробці даних у режимі реального часу. Разом з тим, ШІ є багатообіцяючою технологією, яка може значно покращити кібербезпеку та реагування на інциденти.

Мета даного дослідження полягає у розумінні впливу ШІ та розвитку SIEM на основі ШІ для покращення виявлення та обробки інцидентів загрози безпеці. У роботі будуть розглянуті засоби керування на основі штучного інтелекту, які можуть посилити заходи кібербезпеки та дати можливість організаціям ефективно реагувати на загрози.

Системи SIEM на основі штучного інтелекту використовують алгоритми машинного навчання для аналізу великих обсягів даних у режимі реального часу, що дозволяє організаціям ефективніше виявляти загрози та реагувати на них [1]. Завдяки безперервному моніторингу та аналізу мережевого трафіку, поведінки користувачів і системних журналів ці системи можуть виявляти аномалії, шаблони та ознаки компрометації, які можуть залишитися непоміченими аналітиками [2]. Цей проактивний підхід покращує можливості виявлення загроз і реагування на них, скорочуючи час між вторгненням і ліквідацією. Отже, щоб посилити заходи кібербезпеки та дати можливість організаціям залишатися попереду можливих загроз, в роботі заплановано виконання наступних задач:

1. Огляд літератури з технологій штучного інтелекту використовуваних для поліпшення виявлення інцидентів у SIEM системах.
2. Оцінка та порівняння ефективності провідних SIEM-рішень на основі штучного інтелекту за набором певних параметрів.
3. Вивчення можливостей автоматизації реагування на інциденти на основі штучного інтелекту
4. Розробка і тестування алгоритму автоматичного реагування на потенційну кібер загрозу.

Література

1. Bandr Siraj Fakiha. 2020. Effectiveness of Security Incident Event Management (SIEM) System for Cyber Security Situation Awareness. International Journal of Forensic Medical and Toxicological Sciences. [online] Available at: https://medicopublication.com/index.php/ijfmt/article/view/11587/10679.

2. National Institute of Standards and Technology (NIST). 2020. Securing Manufacturing Industrial Control Systems: Behavioral Anomaly Detection. [online] Available at: https://csrc.nist.gov/pubs/ir/8219/final

Додаток Б

Лістинг 3.1 — Сканування портів за допомогою nmap

```
#!/var/ossec/framework/python/bin/python3
# Copyright (C) 2015-2023, Wazuh Inc.

import nmap
import time
import json
import platform

# The function to perform network scan on a host endpoint
def scan_subnet(subnet):
    nm = nmap.PortScanner()
    nm.scan(subnet)
    results = []

    for host in nm.all_hosts():
        for proto in nm[host].all_protocols():
            if proto not in ["tcp", "udp"]:
                continue

            lport = list(nm[host][proto].keys())
            lport.sort()
# Iterate over each port for the current host and protocol
            for port in lport:
                hostname = ""
                json_output = {
                    'nmap_host': host,
                    'nmap_protocol': proto,
                    'nmap_port': port,
                    'nmap_hostname': "",
                    'nmap_hostname_type': "",
                    'nmap_port_name': "",
                    'nmap_port_state': "",
                    'nmap_port_service': ""
                }
# Get the first hostname and it’s type
                if nm[host]["hostnames"]:
                    hostname =
nm[host]["hostnames"][0]["name"]
                    hostname_type =
nm[host]["hostnames"][0]["type"]
                    json_output['nmap_hostname'] = hostname
                    json_output['nmap_hostname_type'] =
hostname_type
# Get the port name if available
                if 'name' in nm[host][proto][port]:
```

```
                    json_output['nmap_port_name'] =
nm[host][proto][port]['name']
# Get the port state if available
                if 'state' in nm[host][proto][port]:
```

Продовження лістингу 3.1

```
                    json_output['nmap_port_state'] =
nm[host][proto][port]['state']
# Get the port service version if available
                if 'product' in nm[host][proto][port] and
'version' in nm[host][proto][port]:
                    service = nm[host][proto][port]['product']
+ " " + nm[host][proto][port]['version']
                    json_output['nmap_port_service'] = service

                results.append(json_output)
    return results

# The function to append the scan results to the active
response log file
def append_to_log(results, log_file):
    with open(log_file, "a") as active_response_log:
        for result in results:
            active_response_log.write(json.dumps(result))
            active_response_log.write("\n")
# Specify the address(es) to scan
subnets = ['127.0.0.1']
# path of the log file
if platform.system() == 'Windows':
    log_file = "C:\\Program Files (x86)\\ossec-agent\\active-
response\\active-responses.log"
elif platform.system() == 'Linux':
    log_file = "/var/ossec/logs/active-responses.log"
else:
    log_file = "/Library/Ossec/logs/active-responses.log"

for subnet in subnets:
    results = scan_subnet(subnet)
    append_to_log(results, log_file)
    time.sleep(2)
```

Додаток В

Лістинг 3.5 — Скрипт для створення запиту в ChatGPT

```
#!/var/ossec/framework/python/bin/python3
# Copyright (C) 2015-2023, Wazuh Inc.
# ChatGPT Integration template by @WhatDoesKmean

import json
import sys
import time
import os
from socket import socket, AF_UNIX, SOCK_DGRAM

try:
    import requests
    from requests.auth import HTTPBasicAuth
except Exception as e:
    print("No module 'requests' found. Install: pip install
requests")
    sys.exit(1)

# Global vars
debug_enabled = False
pwd =
os.path.dirname(os.path.dirname(os.path.realpath(__file__)))

print(pwd)
#exit()

json_alert = {}
now = time.strftime("%a %b %d %H:%M:%S %Z %Y")
# Set paths
log_file = '{0}/logs/integrations.log'.format(pwd)
socket_addr = '{0}/queue/sockets/queue'.format(pwd)

def main(args):
    debug("# Starting")
    # Read args
    alert_file_location = args[1]
    apikey = args[2]
    debug("# API Key")
    debug(apikey)
    debug("# File location")
    debug(alert_file_location)

    # Load alert. Parse JSON object.
    with open(alert_file_location) as alert_file:
        json_alert = json.load(alert_file)
    debug("# Processing alert")
```

```
    debug(json_alert)
    # Request chatgpt info
    msg = request_chatgpt_info(json_alert,apikey)
    # If positive match, send event to Wazuh Manager
    if msg:
        send_event(msg, json_alert["agent"])

def debug(msg):
    if debug_enabled:
        msg = "{0}: {1}\n".format(now, msg)
    print(msg)
    f = open(log_file,"a")
    f.write(str(msg))
    f.close()


def collect(data):
  nmap_port_service = data['nmap_port_service']
  choices = data['content']
  return nmap_port_service, choices


def in_database(data, nmap_port_service):
  result = data['nmap_port_service']
  if result == 0:
    return False
  return True


def query_api(nmap_port_service, apikey):
  # Calling ChatGPT API Endpoint
  headers = {
        'Authorization': 'Bearer ' + apikey,
        'Content-Type': 'application/json',
    }

  json_data = {
        'model': 'gpt-3.5-turbo',
        'messages': [
            {
                'role': 'user',
                'content': 'In 4 or 5 sentences, tell me about
this service and if there are past vulnerabilities: ' +
nmap_port_service,
            },
        ],
    }

  response =
requests.post('https://api.openai.com/v1/chat/completions',
headers=headers, json=json_data)
```

```
  if response.status_code == 200:
      # Create new JSON to add the port service
      ip = {"nmap_port_service": nmap_port_service}
      new_json = {}
      new_json = response.json()["choices"][0]["message"]
      new_json.update(ip)
      json_response = new_json

      data = json_response
      return data
  else:
      alert_output = {}
      alert_output["chatgpt"] = {}
      alert_output["integration"] = "custom-chatgpt"
      json_response = response.json()
      debug("# Error: The chatgpt encountered an error")
      alert_output["chatgpt"]["error"] = response.status_code
      alert_output["chatgpt"]["description"] =
json_response["errors"][0]["detail"]
      send_event(alert_output)
      exit(0)


def request_chatgpt_info(alert, apikey):
    alert_output = {}
    # If there is no port service present in the alert. Exit.
    if not "nmap_port_service" in alert["data"]:
        return(0)

    # Request info using chatgpt API
    data = query_api(alert["data"]["nmap_port_service"],
apikey)
    # Create alert
    alert_output["chatgpt"] = {}
    alert_output["integration"] = "custom-chatgpt"
    alert_output["chatgpt"]["found"] = 0
    alert_output["chatgpt"]["source"] = {}
    alert_output["chatgpt"]["source"]["alert_id"] =
alert["id"]
    alert_output["chatgpt"]["source"]["rule"] =
alert["rule"]["id"]
    alert_output["chatgpt"]["source"]["description"] =
alert["rule"]["description"]
    alert_output["chatgpt"]["source"]["full_log"] =
alert["full_log"]
    alert_output["chatgpt"]["source"]["nmap_port_service"] =
alert["data"]["nmap_port_service"]
    nmap_port_service = alert["data"]["nmap_port_service"]
```

```
    # Check if chatgpt has any info about the
nmap_port_service
    if in_database(data, nmap_port_service):
      alert_output["chatgpt"]["found"] = 1
    # Info about the port service found in chatgpt
    if alert_output["chatgpt"]["found"] == 1:
        nmap_port_service, choices = collect(data)

        # Populate JSON Output object with chatgpt request
        alert_output["chatgpt"]["nmap_port_service"] =
nmap_port_service
        alert_output["chatgpt"]["choices"] = choices

        debug(alert_output)

    return(alert_output)


def send_event(msg, agent = None):
    if not agent or agent["id"] == "000":
        string = '1:chatgpt:{0}'.format(json.dumps(msg))
    else:
        string = '1:[{0}] ({1}) {2}-
>chatgpt:{3}'.format(agent["id"], agent["name"], agent["ip"]
if "ip" in agent else "any", json.dumps(msg))

    debug(string)
    sock = socket(AF_UNIX, SOCK_DGRAM)
    sock.connect(socket_addr)
    sock.send(string.encode())
    sock.close()


if __name__ == "__main__":
    try:
        # Read arguments
        bad_arguments = False
        if len(sys.argv) >= 4:
            msg = '{0} {1} {2} {3} {4}'.format(now,
sys.argv[1], sys.argv[2], sys.argv[3], sys.argv[4] if
len(sys.argv) > 4 else '')
            debug_enabled = (len(sys.argv) > 4 and sys.argv[4]
== 'debug')
        else:
            msg = '{0} Wrong arguments'.format(now)
            bad_arguments = True

        # Logging the call
        f = open(log_file, 'a')
        f.write(str(msg) + '\n')
        f.close()
```

```
    if bad_arguments:
        debug("# Exiting: Bad arguments.")
        sys.exit(1)

    # Main function
    main(sys.argv)

except Exception as e:
    debug(str(e))
    raise
```