Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет __Комп'ютерної інженерії та управління_____
(повна назва)

Кафедра Автоматизації проектування обчислювальної техніки
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти____другий (магістерський)_____
(рівень вищої освіти)

Модель машинного навчання
для атрибуції кібератак
(тема)

Виконав: студент 2 курсу, групи СКСм-22-2
Сторожко А.В._
(прізвище, ініціали)
Спеціальність *123 _Комп'ютерна інженерія*
Тип програми ____освітньо-професійна_____
(освітньо-професійна або освітньо-наукова)
Освітня програма_ *Спеціалізовані*
_комп'ютерні _системи
(повна назва освітньої програми)

Керівник роботи _доц. Адамов О. С._
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ _Чумаченко С.В
(підпис) (прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет_____Комп'ютерної інженерії та управління _____

Кафедра_____Автоматизації проектування обчислювальної техніки_____

Рівень вищої освіти____другий (магістерський)_____

Спеціальність_____123_ Комп'ютерна інженерія _____
(шифр і назва)

Тип програми _____Освітньо-професійна_____
(освітньо-професійна або освітньо-наукова)

Освітня програма _____Спеціалізовані комп'ютерні системи_____
(повна назва

ЗАТВЕРДЖУЮ:
Зав.кафедри_____
(підпис)
"____" _____ 20 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові _____Сторожку Андрію В'ячеславовичу_____
(прізвище, ім'я, по батькові)

1. Тема роботи (проекту) Модель машинного навчання для атрибуції кібератак
затверджена наказом по університету від "06"____11_____2023 р. № _607 Ст_____
2. Термін подання студентом роботи до екзаменаційної комісії ____15.01.2024_____
3. Вихідні дані до роботи (проекту)_____
Звіти з кібератак_____
Тактики та техніки використані для кіберзлочину_____
Існуючі методи машинного навчання для атрибуції _____
Аналіз хешів, файлів, ip-адрес_____ _____
Мова програмування Pyton_____
Середовище розробки Anaconda IDE_____
4. Перелік питань, що потрібно опрацювати у роботі _____
Поняття атрибуції кібератак
Існуючі тактики та методи кібератак
Аналіз атак хакерських груп
Вибір платформи та інструменті для реалізації програмного продукту
Розробка моделі машинного навчання для атрибуції кібератак

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) __13 слайдів_____

_____

_____

_____

_____

6. Консультанти розділів роботи (проекту)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

7. Дата видачі завдання 07.11.2023_____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи (проекту)	Термін виконання етапів проекту (роботи)	Примітка
1	Видача теми проекту, узгодження і затвердження	07.11.2023 - 15.05.2023	
2	Аналіз обраної теми	15.11.2023 - 20.11.2023	
3	Огляд тактик та методів атрибуції	20.11.2023 - 22.11.2023	
4	Огляд моделей машинного навчання	22.11.2023 - 25.11.2023	
5	Вибір платформи для програмної реалізації моделі	25.11.2023 - 30.11.2023	
6	Програмна розробка та тестування моделі	31.11.2023 - 15.12.2023	
7	Оформлення пояснювальної записки	15.12.2023 - 02.01.2024	
8	Перевірка виконаного проекту керівником	02.01.2024 - 06.01.2024	
9	Захист проекту	09.01.2024 -26.01.2024	

Студент _____
(підпис)

Керівник роботи (проекту) _____ доц. Адамов О.С.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка до курсової роботи: 59 сторінки, 9 рисунків, 5 формул, 12 джерел.

PYTON, ANACONDA, JUPITER NOTEBOOK, CYBERATTACK, INDICATOR OF COMPROMISE, METHODS OF ATTRIBUTION, MACHINE LEARNING

Об’єктом дослідження даного кваліфікаційної роботи є модель машинного навчання для атрибуції кібератак за допомогою індикаторів компрометації та розробка штучного інтелекту для виконання поставленої мети.

Метою є отримання навичок роботи з індикаторами компрометації, тобто файли, хеші, логи, ip-адреси, артефакти файлової системи тощо. Створення програмного продукту у середовищі Anaconda, а точніше у додатку Jupyter Notebook.

Методи дослідження – розробка моделі штучного інтелекту за допомогою якого відбувається обробка інформаційних даних після кібератаки та порівняння з існуючими індикаторами компрометації.

Програмну частину розроблено за допомогою мови Pyton у додатку Jupyter Notebook.

ABSTRACT

The report contains 59 pages, 9 illustration, 5 formulas, 12 reference sources.

PYTON, ANACONDA, JUPITER NOTEBOOK, CYBERATTACK, INDICATOR OF COMPROMISE, METHODS OF ATTRIBUTION, MACHINE LEARNING

The subject of this qualifying is the machine learning model for cyber-attack attribution using compromise indicators and the development of artificial intelligence to accomplish the set goal.

The goal is to acquire skills in working with compromise indicators, such as files, hashes, logs, IP addresses, file system artifacts, etc., and to create a software product in the Anaconda environment, specifically in the Jupyter Notebook application.

Research methods include model developing artificial intelligence for processing information data after a cyber-attack and comparing it with existing compromise indicators.

The software part was developed using the Python language in the Jupyter Notebook application.

ЗМІСТ

ПРЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

IOC – Indicators of Compromise

ML – Машинне навчання

ANN – штучна нейронна мережа

APT – Advanced Persistent Threat

ВСТУП

У сучасному цифровому світі, де кіберзлочини стають все поширенішими, атрибуція кібератак стає надзвичайно важливою для розуміння та боротьби з цією загрозою. Атрибуція визначає хто здійснює кібератаку і які її мотиви.

Атрибуція кібератак базується на зборі, аналізі та інтерпретації цифрових слідів, залишених хакерами під час їхньої діяльності. Це може включати технічні дані, такі як IP-адреси, використані програми та атаки, методи вторгнень і навіть структуру коду. Крім того, атрибуція може використовувати розвідувальні дані, які виявляють мотивації, методику роботи та характеристики конкретних хакерських груп.

Процес атрибуції кібератаки складний і вимагає високої експертизи в галузі кібербезпеки. Він включає у себе співпрацю різних організацій, таких як кіберполіція, розвідувальні служби, приватні кібербезпекові компанії та академічні дослідницькі групи. Ці організації обмінюються інформацією, аналізують дані та використовують сучасні методи аналітики для встановлення зв'язків і знайдення індикаторів, які допомагають встановити авторство атаки. Атрибуція кібератаки має велике значення не лише для правоохоронних органів, але й для промислових компаній, урядових установ та громадських організацій. Вона дозволяє вжити відповідних заходів для уникнення подібних атак у майбутньому. Проте, атрибуція кібератаки також може бути складною задачею через використання різноманітних технік маскування та фальсифікації. Хакери можуть залишати навмисні сліди, що вказують на інші країни або групи, або навпаки, намагатися замаскувати свої дії. Такі спроби ускладнюють процес атрибуції та потребують більш глибокого аналізу та експертизи.

Ідентифікація хакерських угрупувань допомагає розкрити загрозу, встановити винуватців та запобігти майбутнім атакам. Зрозуміння процесу

атрибуції та постійне вдосконалення методів аналізу є критичними для ефективної боротьби з кіберзлочинцями та забезпечення цифрової безпеки в сучасному світі.

1.АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ. ПОСТАНОВКА ЗАДАЧІ

Атрибуція кібератак є важливим аспектом в галузі кібербезпеки. Вона означає визначення авторства та мотивацій захищених кіберінцидентів шляхом аналізу різних індикаторів компрометації. Індикатори компрометації включають технічні дані, логи, поведінкові шаблони та інші аспекти, які можуть надати важливу інформацію про хакерські групи та їхні методи.

Мета атрибуції кібератак полягає в ідентифікації зловмисників, розумінні їхніх мотивацій, цілей та зв'язків з іншими кібератаками. Це допомагає організаціям розуміти загрози, виявляти нові атаки та розробляти ефективні стратегії захисту. Однак атрибуція кібератак є складним завданням, оскільки зловмисники активно використовують методи компрометації та приховування свого сліду.

Завдання полягає в розробці ефективних та автоматизованих методів для встановлення походження кібератак, ідентифікації зловмисників та виявлення їхніх мотивацій.

Кібератаки становлять серйозну загрозу для сучасного цифрового світу. Зловмисники використовують різноманітні техніки, такі як фішинг, злам паролів, вразливості програмного забезпечення та інші, для незаконного доступу до комп'ютерних систем, крадіжки конфіденційних даних, шпигунство та інші злочинні діяння. У контексті атрибуції кібератак, важливо визначити, з якої країни, організації або групи походять ці атаки, ідентифікувати їхні мотивації та збирати докази для подальшого юридичного переслідування. Індикатори компрометації, такі як аномалії в мережевому трафіку, артефакти вірусів, IP-адреси та підписи атак, можуть бути використані для виявлення та аналізу цих кібератак.

Рисунок 1.1 – Активність кібератак за секторами. Державне підприємство «Українські спеціальні системи» 2022 р.

Основна мета полягає у розробці методології та алгоритмів машинного навчання для атрибуції кібератак на основі індикаторів компрометації.

Задачі включають:

— збір та аналіз індикаторів компрометації. Розробка механізмів для збору та обробки індикаторів компрометації, включаючи мережевий трафік, журнали подій, вразливості програмного забезпечення та інші джерела даних. Аналіз індикаторів компрометації допоможе виявити аномалії та підозрілі активності, пов'язані з кібератаками;

— розробка моделей штучного інтелекту. Використання технік машинного навчання, нейронних мереж та інших методів штучного інтелекту для створення моделей, які здатні класифікувати, кластеризувати та атрибувати кібератаки. Ці моделі можуть виявляти підписи атак, розпізнавати характеристики зловмисників та прогнозувати їхні мотивації;

— розробка системи атрибуції кібератак. Побудова інтегрованої системи, яка використовує зібрані індикатори компрометації та моделі штучного

інтелекту для проведення атрибуції кібератак. Система повинна забезпечувати ідентифікацію можливих зловмисників, визначення їхньої географічної локації, причини та способу атаки;

— експериментальне оцінювання та валідація. Проведення експериментальних досліджень для оцінки ефективності розроблених методів атрибуції кібератак. Валідація повинна здійснюватися на реальних даних про кібератаки та порівнюватися з існуючими методами атрибуції;

— розробка інтерфейсу користувача. Розробка зручного та інтуїтивно зрозумілого інтерфейсу, що дозволяє користувачам взаємодіяти з системою атрибуції кібератак, відображати результати аналізу та надавати звіти про виявлені атаки та їхніх характеристики;

— Забезпечення безпеки та конфіденційності. Розробка механізмів захисту від зловмисних атак, а також забезпечення конфіденційності зібраних даних та результатів аналізу. Застосування криптографічних методів, контролю доступу та інших технік безпеки є важливим аспектом розробки системи атрибуції.

Розвиток моделей машинного навчання для атрибуції кібератак на основі індикаторів компрометації є актуальним та важливим напрямком досліджень в галузі кібербезпеки. Використання штучного інтелекту дозволяє автоматизувати та поліпшити процес атрибуції, забезпечуючи швидкість та точність визначення авторства кібератак. Результати цього дослідження можуть бути використані для покращення захисту інформаційних систем та ефективної боротьби з кіберзагрозами.

Існує багато інструментів дослідження, які використовуються в області кібербезпеки для проведення досліджень, виявлення атак та аналізу компрометації. Ось кілька популярних інструментів:

— Wireshark. Це один з найпоширеніших інструментів аналізу мережевого трафіку. Він дозволяє перехоплювати, аналізувати та

відстежувати пакети мережі, що дозволяє виявити аномальну або шкідливу активність;

— Nmap. Це інструмент сканування портів та виявлення мережевих хостів. Він дозволяє досліджувати мережі, виявляти вразливості та знаходити можливі шляхи атак;

— Metasploit. Це фреймворк для розробки та використання експлойтів. Він дозволяє проводити тестування на проникнення, імітувати атаки та виявляти вразливості в системах;

— Snort. Це система виявлення вторгнень (IDS), яка аналізує мережевий трафік та виявляє підозрілу або шкідливу активність на основі заданих правил;

— Volatility. Це інструмент для аналізу пам'яті, який дозволяє виявляти та досліджувати шкідливе програмне забезпечення, витоки інформації та інші аномалії в оперативній пам'яті комп'ютерів;

— Maltego. Це інструмент для візуалізації та аналізу інформації про різні об'єкти (людей, організації, системи тощо). Він дозволяє збирати та вивчати дані з різних джерел для розуміння взаємозв'язків та проведення досліджень;

— YARA. Це інструмент для створення та використання правил виявлення шкідливого програмного забезпечення. Він дозволяє сканувати файли та системи на наявність підозрілих відповідностей правилам;

— OSINT (Open-Source Intelligence) інструменти. Це набір інструментів та технік для збору, аналізу та використання відкритої інформації з відкритих джерел. Вони дозволяють досліджувати веб-сайти, соціальні мережі, форуми тощо для знаходження інформації про загрози та зловмисників.

Це не повний перелік інструментів дослідження, що використовуються в кібербезпеці. Вибір конкретного інструменту залежить від конкретних потреб, типу дослідження та характеристик системи, що аналізується.

В сучасному контексті кібербезпеки використання моделей машинного навчання для атрибуції кібератак є стратегічно обґрунтованим. Атрибуція, що визначається як ідентифікація та приписування винуватців кіберзлочинності, стає невід'ємною частиною стратегії протидії високотехнологічним загрозам.

Моделі машинного навчання відзначаються ефективністю у роботі з великими обсягами даних, які є ключовим аспектом у галузі кібербезпеки. Це дозволяє аналізувати та виявляти унікальні патерни в поведінці кіберзлочинців.

Рисунок 2.1 – Звіт з опрацьованих даних у 2022 році Державної Служби Спеціального Зв’язку та Захисту Інформації України

Використання глибокого навчання та інших методів машинного навчання дозволяє виявляти складні закономірності, які можуть залишатися непоміченими традиційними методами аналізу. Це особливо актуально в умовах постійно зростаючого рівня складності та виразності кібератак.

Перевагою є також можливість аналізувати дані в режимі реального часу, що дозволяє оперативно реагувати на небезпеки та вживати заходів безпеки. Автоматизація процесу атрибуції за допомогою моделі машинного навчання зменшує трудомісткість завдання та сприяє більш швидкому реагуванню на кіберзагрози.

Також слід враховувати, що моделі машинного навчання дозволяють аналізувати великі обсяги даних, які можуть містити важливі інформації про характеристики атак та їхні атрибути.

Використання моделі машинного навчання для атрибуції кібератак розглядається як перспективний напрямок, оскільки вони полегшують не лише процес визначення винуватців, а й роблять його більш ефективним у змінних умовах кібербезпеки.

2. АТРИБУЦІЯ КІБЕРАТАК

2.1 Методи атрибуції

Атрибуція кібератак відноситься до процесу встановлення або ідентифікації походження або авторства кібератаки. Цей процес включає визначення особи, групи або країни, які стоять за конкретною кібератакою. Атрибуція може бути складною і вимагати значної кількості доказів і аналізу. Існує кілька методів та підходів до атрибуції кібератак, проте важливо зазначити, що вона не є завжди абсолютною і точною.

Зв'язати кібератаки з їхніми авторами - це досить складний і многогранний процес, який включає в себе широкий спектр технічних, аналітичних та дослідницьких методів.

Аналіз коду та сигнатур програм є одним із ключових підходів. Дослідники ретельно аналізують програмний код або сигнатури зразків шкідливих програм, спробуючи виявити схожість з попередніми атаками або групами. Це включає виявлення особливостей коду, використання подібних алгоритмів або методів шифрування.

Аналіз стилістики програмування також є важливим. Дослідники аналізують стиль програмістів, включаючи коментарі, структуру коду, форматування та навіть унікальність підходу до написання програм. Це може допомогти виявити індивідуальний стиль автора чи групи.

Використання метаданих, таких як IP-адреси, географічні дані серверів, також важливо. Ці дані можуть допомогти встановити зв'язок між різними кібератаками та визначити загальні патерни атак.

Більш технологічні підходи включають використання інтелектуальних систем та машинного навчання для аналізу великих обсягів даних. Ці системи можуть автоматично виявляти патерни та зв'язки між кібератаками, що робить їхнє виявлення більш ефективним.

Розвідка та розслідування - ще один важливий аспект. Дослідники можуть вивчати звіти про кібератаки, здійснювати співпрацю з урядовими органами та іншими структурами, а також використовувати різні джерела інформації для отримання деталей про можливих авторів або групи.

Крім цього, аналіз інфраструктури, методів, сигнатур та модус-операнді можуть розкрити використовувані зловмисниками технічні аспекти та спільні характеристики атак.

Цей комплексний аналіз дозволяє виявляти спільні риси та зв'язки між різними кібератаками, роблячи можливим встановлення зв'язків між ними та виявлення потенційних авторів для подальшого забезпечення кібербезпеки.

Зв'язати кібератаки з їхніми авторами вимагає багатошарового підходу, який поєднує технічні, аналітичні, та інтелектуальні методи.

Технічний аналіз коду та сигнатур програм включає уважний огляд програмного коду або вірусних сигнатур зразків шкідливих програм для виявлення схожості з попередніми атаками або групами. Це може включати перевірку алгоритмів, методів шифрування, та підписів програм.

Аналіз стилістики програмування полягає у вивченні стилів програмістів, їхніх підходів до створення програм, включаючи аналіз коментарів, форматування коду, та унікальність підходу до програмування.

Використання метаданих, таких як IP-адреси, географічні дані серверів, дозволяє встановити зв'язок між різними кібератаками та розкрити загальні патерни злочинної діяльності.

Розробка та використання інтелектуальних систем та машинного навчання дозволяють автоматизувати аналіз великих обсягів даних, що сприяє виявленню складних зв'язків та патернів між кібератаками.

Додатково, розвідка та розслідування включають в себе аналіз звітів про кібератаки, співпрацю з урядовими органами та іншими структурами, а також використання різноманітних джерел інформації для отримання вичерпної картини про можливих авторів чи групи.

Деякі методи, що використовуються для атрибуції кібератак, включають:

— аналіз коду та сигнатур. Дослідники аналізують код або сигнатури використовуваних програм або зразків шкідливих програм, щоб знайти спільні риси з попередніми атаками або групами;

— аналіз стилістики. Цей метод використовується для виявлення унікальних стилів програмістів або груп, що можуть розкрити їхніх авторів. Він включає аналіз використовуваних коментарів, змінних імен, форматування коду та інші стилістичні особливості;

— використання гібридних атак. Деякі атаки можуть включати комбінацію технік з різних джерел або використовувати зразки атак, які були виявлені раніше. Це може служити доказом, що той самий актор або група стоять за атаками;

— аналіз метаданих. Вивчення метаданих, таких як IP-адреси, географічні розташування серверів або інші деталі інфраструктури, може допомогти встановити зв'язки між різними кібератаками;

— використання інтелектуальних систем. Деякі організації використовують інтелектуальні системи аналізу даних та машинного навчання для виявлення патернів і асоціацій між кібератаками. Ці системи можуть аналізувати великі обсяги даних та виявляти зв'язки між різними атаками або акторами;

— використання розвідки та розслідувань. Розвідка та розслідування можуть включати збір відкритої інформації, вивчення звітів про кібератаки, співпрацю зі спеціальними службами та іншими організаціями для встановлення зв'язків та ідентифікації потенційних авторів;

— аналіз мовного стилю. Дослідники можуть аналізувати мовні особливості текстових повідомлень, електронних листів або коду, щоб виявити індивідуальність автора. Це може включати використання

унікальних слів, фраз, граматичних конструкцій або стилістичних особливостей;

— використання відомих атак. Деякі атаки можуть використовувати вже відомі методи або зразки, пов'язані з певними хакерськими групами або країнами. Порівняння таких атак з вже відомими може допомогти встановити асоціації та зробити припущення про авторство;

— використання інформації зі зовнішніх джерел. Дослідники можуть використовувати інформацію з відкритих джерел, таких як заяви від урядових органів, спеціальних служб або інших кібербезпекових компаній, щоб знайти зв'язки між атаками та потенційними авторами;

— аналіз інфраструктури. Дослідники можуть аналізувати інфраструктуру, використовану для кібератак, таку як IP-адреси, домени, хостинг-провайдери тощо. Це може розкрити зв'язки з попередніми атаками або ідентифікувати групи, які використовують спільні інфраструктурні компоненти;

— аналіз сигнатур та модус-операнді. Дослідники можуть аналізувати унікальні сигнатури атак або використані модус-операнді (спосіб дії) для встановлення зв'язку з попередніми атаками або відомими хакерськими групами. Це включає аналіз використовуваних алгоритмів шифрування, методів ідентифікації, способів виконання атак та інших технічних деталей;

— аналіз соціально-інженерних аспектів. Крім технічних аспектів, атрибуція кібератак може включати аналіз соціально-інженерних методів, використаних в атаках. Це може охоплювати вивчення використаних маніпулятивних технік, фішингових атак, впливу на людський фактор та інші аспекти, що допомагають ідентифікувати можливих авторів атак;

— співпраця зі спеціалістами та правоохоронними органами. Деякі серйозні кібератаки можуть вимагати співпраці зі спеціалістами в

області кібербезпеки та правоохоронними органами. Це може включати обмін інформацією, спільні розслідування та аналіз злочинів, що допомагають встановити походження атаки та визначити можливих винуватців;

— використання розвідувальних методів. Дослідники можуть використовувати методи розвідки, такі як збір відкритої інформації, вивчення геополітичного контексту, аналіз розробок та зв'язків між країнами та групами, щоб встановити можливі зв'язки між кібератаками та акторами;

— форензичний аналіз. Форензичний аналіз є одним із ключових методів атрибуції кібератак і включає детальне дослідження комп'ютерних систем, мереж та пристроїв, що були скомпрометовані в результаті атаки. Цей процес включає збір, аналіз та інтерпретацію цифрових доказів, що можуть допомогти встановити початок, спосіб виконання та наслідки кібератаки. Форензичний аналіз може включати розслідування дисків, відновлення видаленої інформації, аналіз системних журналів, розбирання коду, аналіз мережевого трафіку та багато іншого;

— створення профілів атак. Цей підхід використовується для створення профілів хакерських груп або індивідуальних хакерів на основі їхнього попереднього вчинку. Вивчення і аналіз раніше виконаних кібератак може розкрити характеристики, такі як використані методи, інструменти, цілі, мотивації та стиль, що допомагають визначити можливого автора атаки;

— контекстуальний аналіз. Контекстуальний аналіз передбачає врахування різноманітних факторів, таких як політична обстановка, геополітичні конфлікти, економічні інтереси, відносини між країнами тощо. Це дозволяє розуміти можливі мотивації та цілі хакерів, а також

встановити можливі зв'язки між кібератакою та конкретними акторами або країнами;

— використання відкритої інформації та розвідка. Використання відкритої інформації, такої як публічні заяви, звіти, джерела новин, форуми, соціальні медіа та інші відкриті джерела, може надати цінну інформацію про хакерські групи або індивідуальних хакерів. Розвідка може включати відстеження активності хакерів в інтернеті, збір відомостей про їхніх можливих співробітників або партнерів, а також вивчення методів, які вони використовують;

— гібридний підхід. Комбінування різних методів та підходів є ефективним способом досягнення точніших результатів атрибуції кібератак. Гібридний підхід включає поєднання технічного аналізу, форензичного дослідження, контекстуального аналізу та спеціалізованої експертної оцінки для знаходження спільних показників та встановлення зв'язків між атаками та авторами;

— узагальнення та навчання на основі досвіду. Інформація, накопичена в результаті атрибуції попередніх кібератак, може бути використана для узагальнення та навчання на основі досвіду. Розуміння попередніх шаблонів, методів та характеристик може сприяти швидшому виявленню та атрибуції майбутніх атак;

— зовнішні індикатори. Зовнішні індикатори, такі як IP-адреси, домени, використовувані хостинг-провайдери, шаблони атак, вразливості, використані інструменти та інші характеристики, можуть дати підказку щодо можливих авторів кібератак. Аналіз цих індикаторів може допомогти встановити зв'язок між різними атаками та ідентифікувати спільних акторів;

— мовний аналіз. Аналіз мови, використаної в хакерських повідомленнях, вірусах, фішингових листах та інших кібератаках, може бути корисним при атрибуції. Особливості граматики, орфографії, стилістики та

лексики можуть вказувати на конкретну мову чи географічний регіон, з якого походить автор атаки;

— зворотний інжиніринг. Застосування зворотного інжинірингу до шкідливого програмного забезпечення, використаного в кібератаках, може розкрити цінну інформацію про його походження. Аналіз внутрішньої структури програми, використання алгоритмів, захистних механізмів та інших технічних характеристик може вказати на певні групи або розробників програмного забезпечення;

— співставлення з попередніми інцидентами. Порівняння нових кібератак з попередніми інцидентами може допомогти встановити зв'язок між ними. Виявлення схожих методів, сигнатур, індикаторів або стилів може вказувати на причетність до певних груп або повторне використання засобів та інфраструктури;

— моделювання загроз. Моделювання та симуляція можливих сценаріїв кібератак можуть допомогти виявити характеристики та індикатори, які специфічні для конкретних акторів або груп. Це дозволяє провести експерименти з різними сценаріями та зібрати додаткові дані для аналізу та атрибуції;

— мережевий аналіз. Аналіз мережевого трафіку та журналів може надати цінну інформацію про характеристики кібератаки. Вивчення шаблонів спілкування, використання протоколів, перехоплення пакетів та аналіз мережевої активності можуть допомогти визначити походження та характеристики атаки;

— співпраця та обмін інформацією. Співпраця між різними організаціями, в тому числі компаніями, державними установами та громадськими організаціями, є важливим елементом атрибуції кібератак. Обмін інформацією про інциденти, підписи атак, характеристики та методи може допомогти виявити спільні зразки та ідентифікувати авторів;

— експертна оцінка. Експертна оцінка використовує знання та досвід кваліфікованих фахівців у галузі кібербезпеки. Це може включати аналіз поведінки атак, технічних деталей, контексту та інших факторів, що допомагають зробити висновки щодо авторства атаки;

— юридичний аналіз. Юридичний аналіз включає вивчення правових аспектів та законодавства, пов'язаного з кібербезпекою та кіберзлочинністю. Це може включати розгляд правових рамок, попередніх судових рішень та прецедентів, що стосуються подібних кібератак. Юридичний аналіз може надати додаткову підтримку при атрибуції та розслідуванні;

— моніторинг та інформаційна зброя. Після атрибуції кібератаки важливо продовжувати моніторинг активності та збирати нові дані. Інформаційна зброя, така як встановлення пасток, розголошення підроблених даних або розкриття інформації, може допомогти встановити додаткові зв'язки та зібрати додаткові докази.

Ці методи та підходи використовуються для атрибуції кібератак та допомагають встановити зв'язок між атаками та їх авторами. Важливо зазначити, що це далеко не усі методи бо атрибуція кібератаки є складним процесом і не завжди може бути досягнута з абсолютною впевненістю.

Рисунок 2.1 – Метод сканування мережевих служб. 2022 рік, Державна Служба Спеціального Зв’язку та Захисту Інформації України

Рисунок 2.2 – аналіз метаданих на початок 2022 року. Державна Служба Спеціального Зв’язку та Захисту Інформації України

Використання даних методів та підходів створює глибокий та комплексний аналіз, який дозволяє виявити зв'язки та ідентифікувати можливих авторів кібератак, важливий крок у підвищенні рівня кібербезпеки.

2.2 Сервіси звітності з кібератак

Глобальна база знань MITRE ATT&CK є джерелом тактик та методів хакерських атак. Дана база містить у собі реальні спостереженнях атак. У даній роботі використовується як джерело звітності для ідентифікації хакерських груп.

Рисунок 2.1 – Техніки хакерських атак з фреймворку MITRE ATT&CK

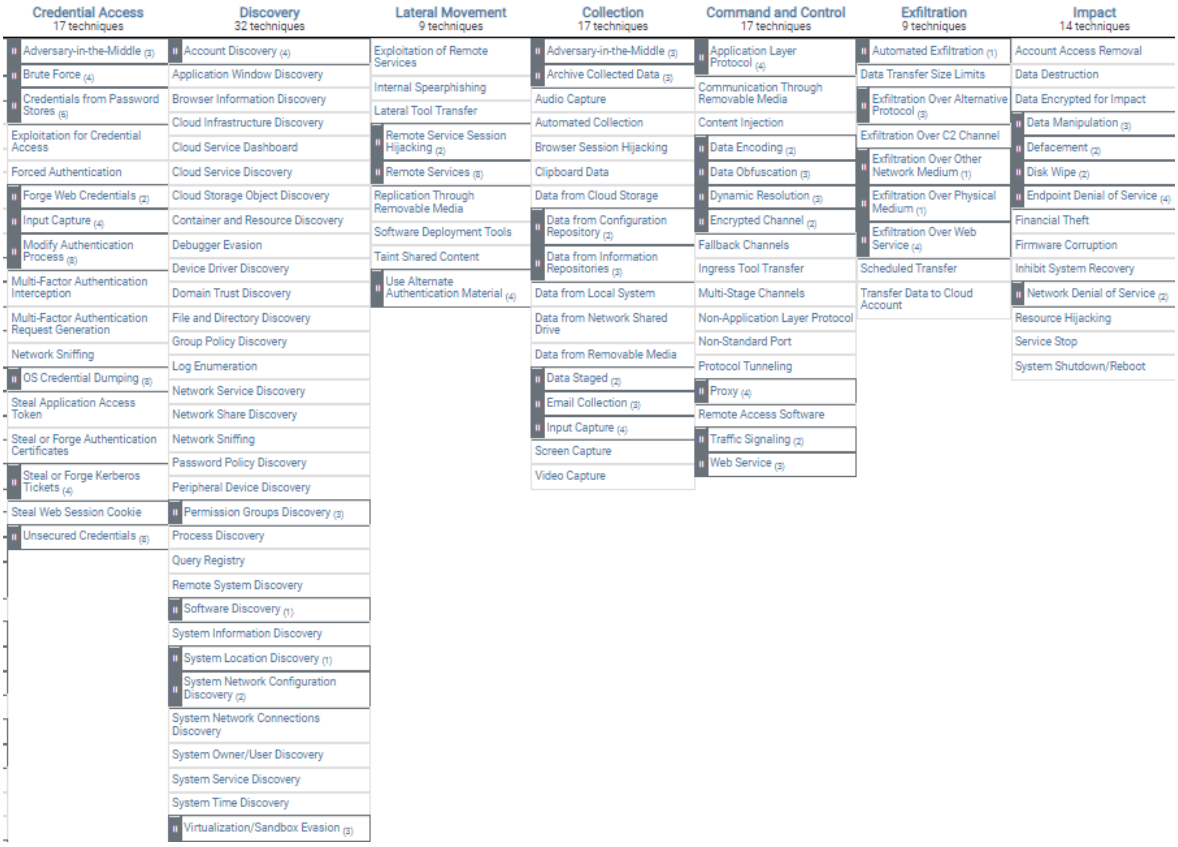


Рисунок 2.2 – Техніки хакерських атак з фреймворку MITRE ATT&CK (продовження)

Даний сервіс також надає звітність у форматі таблиці. Що надає зручний спосіб організації даних в окремих колонках (дата, час, тип атаки, метод що використовувався тощо).

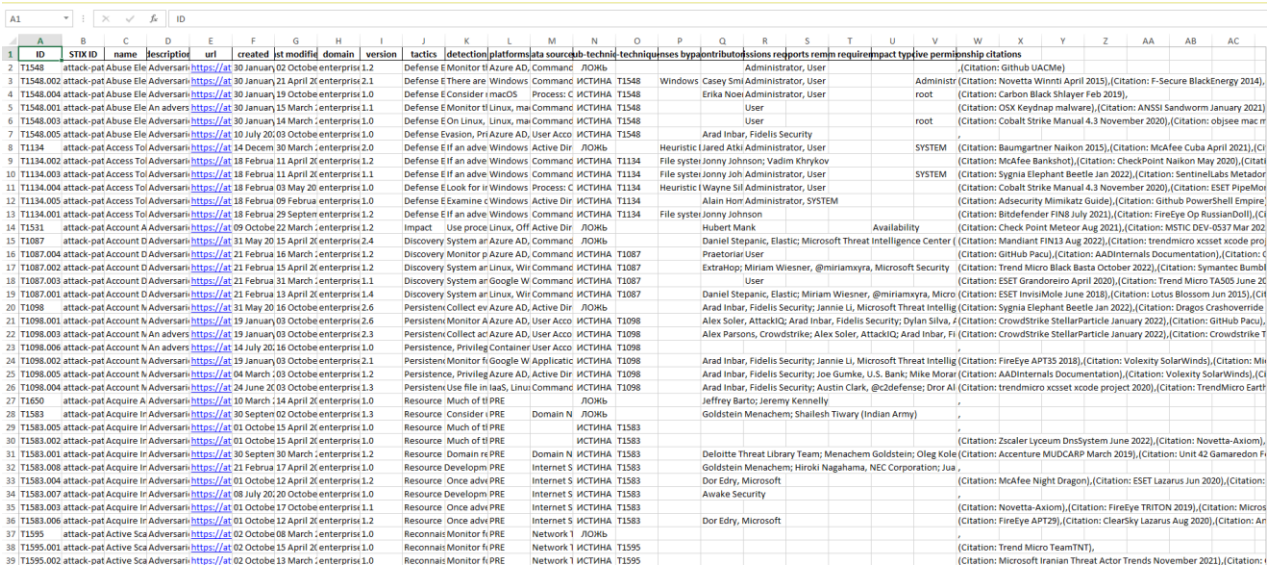
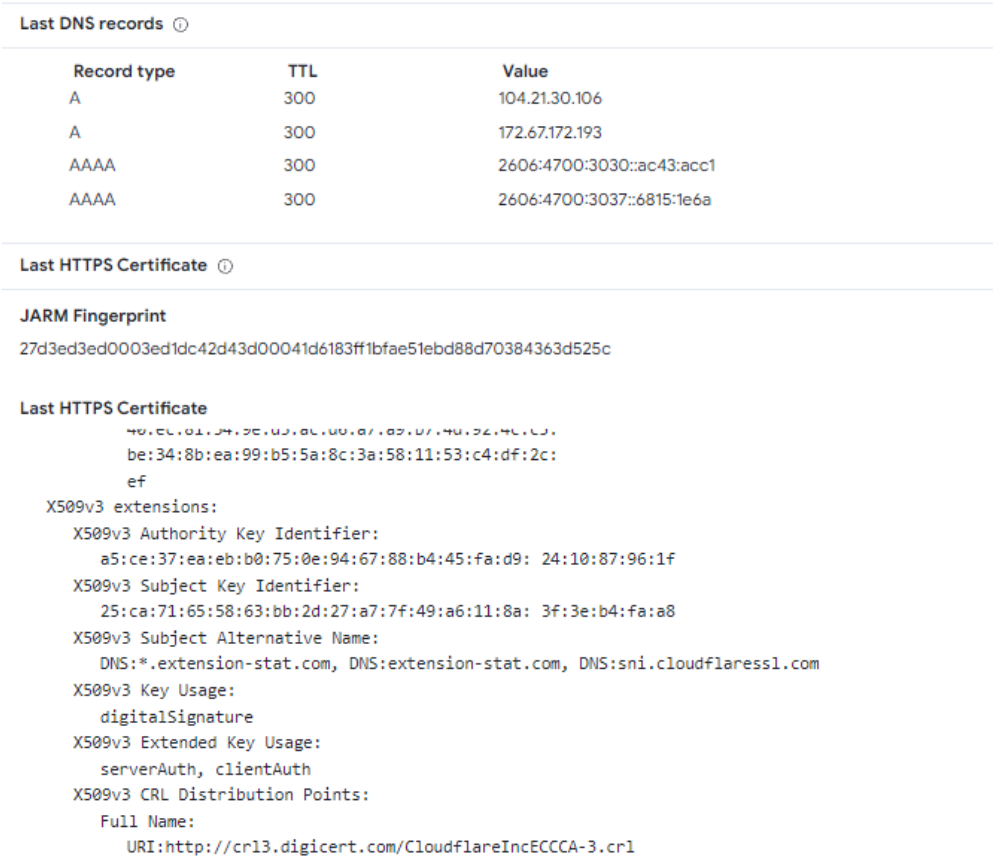


Рисунок 2.3 – Приклад таблиці звітності з кібератаки

Також у роботі був використана онлайн платформа VirusTotal яка надає інструменти для виявлення потенційних загроз. Фреймворк включає у себе мультидвигунне сканування яке дозволяє одночасно сканувати файли або URL-адреси за допомогою кількох антивірусних продуктів. Важливо виділити що платформа надає деталізований аналіз про поведінку файлів, URL-адрес, виявлену загрозу та інші аспекти.



Last DNS records

Record type	TTL	Value
A	300	104.21.30.106
A	300	172.67.172.193
AAAA	300	2606:4700:3030::ac43:acc1
AAAA	300	2606:4700:3037::6815:1e6a

Last HTTPS Certificate

JARM Fingerprint

27d3ed3ed0003ed1dc42d43d00041d6183ff1bfae51ebd88d70384363d525c

Last HTTPS Certificate

```
            be:34:8b:ea:99:b5:5a:8c:3a:58:11:53:c4:df:2c:
            ef
    X509v3 extensions:
        X509v3 Authority Key Identifier:
            a5:ce:37:ea:eb:b0:75:0e:94:67:88:b4:45:fa:d9: 24:10:87:96:1f
        X509v3 Subject Key Identifier:
            25:ca:71:65:58:63:bb:2d:27:a7:7f:49:a6:11:8a: 3f:3e:b4:fa:a8
        X509v3 Subject Alternative Name:
            DNS:*.extension-stat.com, DNS:extension-stat.com, DNS:sni.cloudflaressl.com
        X509v3 Key Usage:
            digitalSignature
        X509v3 Extended Key Usage:
            serverAuth, clientAuth
        X509v3 CRL Distribution Points:
            Full Name:
              URI:http://crl3.digicert.com/CloudflareIncECCCA-3.crl
```

Рисунок 2.4 – Звітність VirusTotal з аналізу URL-посилання «details.extension-stat.com»

2.3. Аналіз метаданих та мережевих слідів

Цей аспект атрибуції є критичним у встановленні зв'язків між різними кіберінцидентами та виявленні авторів цих атак через використання різноманітних мережевих даних та метаданих.

Перш за все, метадані можуть включати в себе різноманітну інформацію про роботу мережі, таку як IP-адреси, часові мітки, маршрутизація пакетів, деталі про апаратне забезпечення і багато іншого. Вони можуть надати унікальний погляд на те, як саме атака відбувалася, включаючи шляхи поширення, джерела та призначення трафіку, що є дорогоцінною інформацією для атрибуції.

Наприклад, аналіз IP-адрес може допомогти встановити походження атак, визначити географічні регіони, з яких вони були ініційовані, або навіть виявити інфраструктуру, яку використовували зловмисники. Ретельний аналіз маршрутизації трафіку може розкрити шляхи, якими атака розповсюджувалася, що допомагає виявити спільні пункти між різними інцидентами.

Часові мітки в даних можуть вказати на регулярність атак або певні взірці активності, що може вказувати на конкретний режим роботи зловмисників або групи. Ця інформація може стати важливим показником для встановлення мотивації чи цілей атаки.

Важливою частиною аналізу метаданих є ідентифікація зв'язків між різними кібератаками. Порівняльний аналіз цих даних може виявити подібність між атаками, встановити збіги чи спільні характеристики, що можуть свідчити про використання спільних інструментів, інфраструктури або навіть зв'язок між різними злочинними групами.

Додатково, аналіз мережевих слідів включає інструменти та методи, що вивчають поведінку трафіку, методи шифрування, характеристики пакетів, що можуть бути використані для аналізу способів атаки та ідентифікації особливостей комунікації між хакерами та їхніми цілями.

Загалом, аналіз метаданих та мережевих слідів є важливою частиною атрибуції кібератак, оскільки надає широкий спектр інформації, що може допомогти виявити зв'язки, розкрити технічні деталі та встановити можливість авторства цих атак.

2.4 Соціально-політичні аспекти атрибуції

Розгляд соціально-політичних аспектів у контексті атрибуції кібератак відображає важливість розуміння не лише технічних деталей, а й ширших соціальних і політичних впливів на ці атаки. Цей підпункт дозволяє розглянути кібербезпеку через призму міждержавних відносин, геополітичних конфліктів та політичних амбіцій.

Соціально-політичні мотивації можуть визначати цілі кібератак. Наприклад, цілеспрямовані атаки на критичну інфраструктуру можуть мати геополітичний мотив, спрямований на підкреслення влади чи реалізацію політичних стратегій. Аналіз таких атак в контексті геополітичних конфліктів може розкрити міжнародні амбіції держав чи груп.

Політичні та легітимні фактори також впливають на кіберактивність груп ініціаторів атак. Підтримка з боку держави, внутрішня чи міжнародна політика, або навіть прямий вплив різних режимів можуть стимулювати кібератаки та впливати на їхній обсяг та складність.

Додатково, кібератаки можуть використовуватися для впливу на політичні рішення чи громадську думку. Маніпулювання інформацією, кібершпигунство та атаки на системи виборів можуть мати за мету впливати на політичну ситуацію в країні або навіть на міжнародному рівні, формуючи віртуальну політичну реальність.

Розгляд соціально-політичних аспектів дозволяє не лише аналізувати технічні аспекти атак, а й розуміти їхній соціальний контекст. Врахування цих аспектів сприяє встановленню мотивацій атак та ідентифікації потенційних ініціаторів, що є ключовим елементом у процесі атрибуції в кіберпросторі.

3.ІНДИКАТОРИ КОМПРОМЕТАЦІЇ

Індикатор компрометації (IoC) у сфері комп'ютерної безпеки - це спостережений у мережі або на конкретному пристрої об'єкт (чи активність), який з великою ймовірністю вказує на несанкціонований доступ до системи (тобто її компрометацію). Такі індикатори використовуються для виявлення шкідливої активності на ранніх етапах, а також для запобігання відомим загрозам IoC може приймати форму різних об'єктів чи дій, таких як аномальні мережеві запити, зміни в системних файлах, невідомі процеси або надзвичайний трафік. Виявлення цих індикаторів може служити сигналом для оперативних заходів з безпеки, щоб вчасно реагувати на потенційну загрозу та мінімізувати можливі наслідки. Однією з ключових функцій IoC є раннє виявлення атак, навіть до того, як вони спричинить суттєвий збиток. Це дозволяє безпековим командам ефективно реагувати на інциденти та приймати необхідні заходи для забезпечення цілісності та безпеки системи.

Індикатори компрометації (іноді відомі як індикатори компромісу або IOC – Indicators of Compromise) є ознаками або слідами, які свідчать про те, що комп'ютерна система, мережа або пристрій були скомпрометовані або порушені. Ці індикатори можуть бути використані для виявлення атак та надають цінну інформацію для подальшого розслідування та відновлення системи.

Індикатори компрометації можуть мати різноманітну природу та охоплювати різні аспекти комп'ютерної безпеки. Деякі з найпоширеніших типів індикаторів компрометації включають:

— файли та хеш-суми. Індикатори, пов'язані з певними файлами або хеш-сумами файлів, можуть вказувати на наявність шкідливого програмного забезпечення або компрометацію. Це можуть бути відомі вразливі файли,

зловмисні виконувані файли, файлові розширення або конкретні хеш-суми, які відповідають відомим шкідливим програмам;

— мережевий трафік. Аналіз мережевого трафіку може розкрити індикатори компрометації, такі як підозрілі підключення до віддалених серверів, використання незвичних протоколів або незвичайна активність на певних портах. Нестандартний мережевий трафік може свідчити про присутність зловмисного програмного забезпечення або несанкціонований доступ;

— підозрілі процеси та поведінка. Індикатори, пов'язані з підозрілими процесами або незвичайною поведінкою системи, можуть свідчити про компрометацію. Це можуть бути незнайомі процеси, незвичайні зміни в системних налаштуваннях, виконання підозрілих команд або збільшення привілеїв користувача без належної авторизації;

— виявлення аномалій. Виявлення аномальних дій або змін у системі, які не збігаються з типовими шаблонами або налаштуваннями, може вказувати на компрометацію. Це можуть бути незвичайні зміни в реєстрі, системні виклики, незвичайні підключення до системи або незвичайна активність користувача;

— журнали подій. Аналіз журналів подій системи або програм може розкрити індикатори компрометації, такі як невдалі спроби входу, помилки аутентифікації, несподівана активність або зміни в системних налаштуваннях;

— сигнатури шкідливого програмного забезпечення. Індикатори, що використовуються для виявлення відомих сигнатур шкідливого програмного забезпечення, таких як відомі вразливості, хеш-суми файлів або сигнатури виконуваних модулів;

— доменні імена та URL-адреси. Підозрілі доменні імена або URL-адреси можуть свідчити про використання зловмисниками для ведення фішингових атак, розповсюдження шкідливого програмного забезпечення або взаємодії зі злонамірeними серверами;

— автономні системи та IP-адреси. Аналіз автономних систем (AS) та IP-адрес може вказувати на підозрілі джерела атаки або присутність відомих зловмисних мереж;
— інформація про сертифікати. Неспівпадання або недійсні сертифікати можуть свідчити про підробку або незаконне використання сертифікатів для здійснення атак;
— зміни в системних файлів та конфігурацій. Виявлення змін в системних файлів, конфігураційних файлах або реєстрі може свідчити про компрометацію або незаконний доступ до системи;
— зловмисні підписи та ключі. Виявлення підозрілих або відомих зловмисних підписів або ключів може допомогти виявити зловмисників та їх діяльність;
— підписи атак та вразливостей. Використання підписів атак та вразливостей дозволяє виявити певні типи атак або використання відомих вразливостей;
— аналіз вмісту пакетів. Виявлення підозрілого вмісту пакетів мережевого трафіку може допомогти виявити активність зловмисників, таку як витік конфіденційної інформації, надання команд або розповсюдження шкідливого програмного забезпечення;
— аналіз електронної пошти. Підозрілі електронні листи, включаючи фішингові повідомлення, можуть свідчити про спроби соціального інженерингу або розповсюдження шкідливого програмного забезпечення;
— індикатори аномальної активності користувачів. Аналіз активності користувачів може виявити підозрілі дії, такі як спроби несправедливого отримання доступу, невдалі спроби входу, зміна прав доступу або незвичайна активність в системі;
— індикатори зловмисного програмного забезпечення на кінцевих пристроях. Виявлення ознак зловмисного програмного забезпечення на

кінцевих пристроях, таких як аномальне використання ресурсів, незвичайні мережеві з'єднання або виконання небезпечних дій, може свідчити про компрометацію;

— індикатори соціального інженерингу. Виявлення ознак соціального інженерингу, таких як небезпечні посилання, вимоги про передачу конфіденційної інформації або спроби використання маніпуляційного впливу на людей, може допомогти виявити атаки, спрямовані на отримання нелегального доступу до системи;

— індикатори аномальних аудиторських журналів. Аналіз аудиторських журналів може розкрити незвичайні записи або аномальні події, які можуть свідчити про компрометацію або відхилення від звичайної поведінки;

— індикатори фізичної компрометації. Фізичні індикатори, такі як пошкодження або зміни на фізичних пристроях, відсутність або зміна міток, можуть свідчити про фізичну компрометацію системи або пристроїв;

— індикатори вторгнень. Виявлення підозрілих або невідомих вторгнень, які можуть бути виявлені за допомогою систем виявлення вторгнень (IDS) або систем виявлення аномалій (ADS), може допомогти виявити активність зловмисників у системі;

— індикатори мережевої активності. Моніторинг мережевої активності може виявити підозрілі пакети, незвичайний обсяг трафіку, використання нестандартних протоколів або незвичайну мережеву поведінку, що може свідчити про атаку або незаконну активність;

— індикатори фішингу. Виявлення ознак фішингових атак, таких як підозрілі посилання, запити конфіденційних даних або підроблені веб-сайти, може допомогти виявити спроби шахрайства та втягнення користувачів у пастку;

— індикатори зміни поведінки. Виявлення зміни поведінки системи, програм або користувачів може свідчити про компрометацію. Наприклад, незвичайна активність, збільшене використання ресурсів або незвичайні спроби доступу можуть бути ознаками атаки;
— індикатори вразливостей. Виявлення вразливостей в системі, програмах або пристроях може свідчити про можливі точки входу для зловмисників. Незвичайна активність або експлуатація вразливостей можуть бути ознаками атаки;
— індикатори зловживання привілеїв. Виявлення незвичайної або неправомірної активності з використанням підвищених привілеїв адміністратора або привілеїв користувачів може свідчити про незаконний доступ до системи або зловживання правами доступу.

Це лише кілька прикладів індикаторів компрометації, а їхній перелік може бути значно більшим та розширюватися залежно від контексту та особливостей атаки. Виявлення та аналіз цих індикаторів є важливою складовою процесу виявлення та відповіді на кібератаки.

4 МЕТОДИКА ДОСЛІДЖЕННЯ

4.1 Середа розробки

За середу розробки було обрано середовище “Anaconda”. Anaconda - це платформа, яка має надзвичайну популярність серед мільйонів користувачів у всьому світі. Вона є вільно розповсюджуваним дистрибутивом програмних продуктів, спеціалізованим на наукових обчисленнях. Anaconda надає зручний спосіб управління пакетами та розгортанням, спрощуючи рутинні завдання для науковців і аналітиків даних.

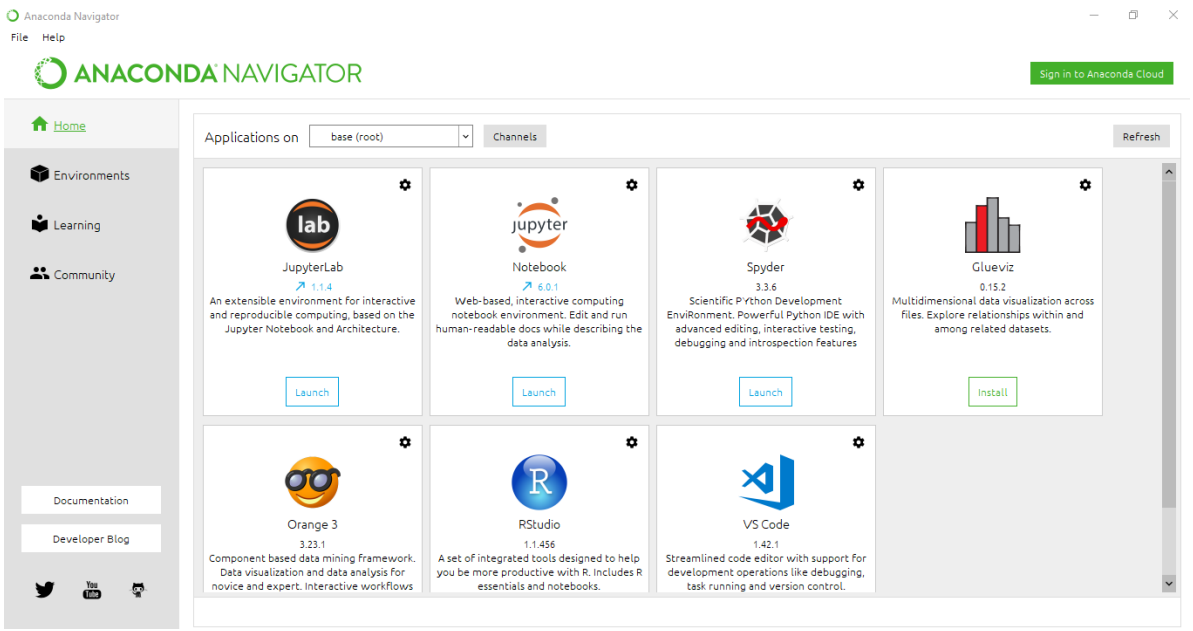


Рисунок 1 – головне меню Anaconda Navigator

Цей потужний дистрибутив включає більше 1500 популярних пакетів наукових даних, що забезпечують широкі можливості для розвитку проектів у сферах обробки даних, статистики, машинного навчання та багатьох інших. Серед них варто відзначити такі інструменти, як NumPy для роботи з

масивами та науковими обчисленнями, SciPy для високорівневих математичних функцій і Ggplot2 для створення привабливих графіків.

Однак Anaconda - це не просто набір інструментів. Вона також поєднується з потужним інтерактивним обчислювальним середовищем - Jupyter Notebook, що надає зручну платформу для створення і виконання "живих" документів, які можуть містити код, графіки, математику та багато іншого. Завдяки йому ви можете ефективно працювати з даними, виконувати аналіз та демонструвати результати у привабливому та зрозумілому форматі.

За допомогою Jupyter Notebook можливе підключення до різних ядер, що дозволяє вам програмувати на різних мовах програмування. Крім того, ви можете легко конвертувати ваші блокноти в різні формати, такі як HTML, LaTeX або PDF, що дозволяє вам зручно ділитися результатами своєї роботи з колегами або публікувати їх в Інтернеті.

Зараз важко уявити наукове співтовариство без Anaconda і Jupyter Notebook. Ці інструменти забезпечують швидку та зручну розробку, дозволяючи науковцям і дослідникам фокусуватися на важливих аспектах своєї роботи і робити значний внесок у світ науки та даних.

Таким чином, використання Anaconda та Jupyter Notebook в сфері наукових обчислень стало справжньою революцією, що забезпечує широкі можливості для дослідників і дослідницьких груп у всьому світі. Завдяки їм збільшується продуктивність, спрощується робота з даними і створюється нова якість в аналізі та візуалізації інформації.

4.2 Досліджувані хакерські групи

APT28, також відома як Fancy Bear, є однією з найвідоміших і активних кіберзагроз в світі. Це група хакерів, яка вважається спонсорованою російською державою інформаційною або розвідувальною службою. APT28 відома своїми цілеспрямованими кібератаками на різні цілі, зокрема урядові

установи, політичні організації, оборонні підприємства, технологічні компанії та міжнародні організації.

Група використовує різноманітні методи та техніки для здійснення своїх кібератак. Деякі з найвідоміших і характерних методів, що використовуються APT28, включають:

— соціальний інжиніринг. APT28 використовує фішингові атаки, спам-кампанії та спеціально спроектовані електронні листи, щоб впровадити шкідливі програми в цільові системи. Вони намагаються використовувати вразливості в людському факторі та переконати користувачів у відкритті шкідливих вкладень або посилань;

— атаки на веб-інфраструктуру. APT28 атакує веб-сервери та системи керування вмістом (CMS) з метою отримання незаконного доступу, внесення змін до веб-сайтів або встановлення шкідливого програмного забезпечення для витоку інформації;

— використання вразливостей. Група активно використовує вразливості в операційних системах, програмах та програмному забезпеченні для отримання несанкціонованого доступу до систем та виконання атак;

— використання шпигунського програмного забезпечення. APT28 створює та використовує шпигунське програмне забезпечення (spyware) для збору розумової власності, конфіденційної інформації та розвідки на комп'ютерах та мережах жертв;

— злам аккаунтів та розподіл вірусів. APT28 використовує методи злому аккаунтів, включаючи викрадення облікових записів та розподіл вірусів, щоб отримати доступ до систем та розповсюджувати шкідливе програмне забезпечення.

APT28 є досить складною та стійкою групою хакерів, яка постійно розвиває свої методи та техніки. Їхні атаки вимагають високого рівня кібербезпеки та заходів захисту для ефективного протистояння їм.

APT29, також відома як Cozy Bear або The Dukes, є ще однією з відомих кіберзагроз, пов'язаних з російською розвідкою. Ця група займається комп'ютерними вторгненнями та кібершпигунством з метою отримання доступу до конфіденційної інформації з різних секторів, включаючи урядові установи, дипломатичні організації, високотехнологічні компанії та дослідницькі організації.

APT29 відома своїми продуманими та ретельно планованими атаками. Ось кілька розповсюджених методів, що використовуються групою APT29:

— соціальний інжиніринг. APT29 використовує соціальний інжиніринг техніки, такі як фішингові атаки, спам-кампанії та вбудовані віруси в електронних листах, щоб впровадити шкідливі програми на комп'ютери жертв;
— атаки на інфраструктуру. Група атакує мережеву інфраструктуру, використовуючи різні техніки, включаючи витік інформації, атаки на сервери електронної пошти та DNS-зловживання, щоб отримати незаконний доступ та зловживати цільовими системами;
— використання вразливостей. APT29 використовує вразливості в програмному забезпеченні та операційних системах, включаючи відомі та невідомі уразливості, для отримання доступу до систем та розповсюдження шкідливого програмного забезпечення;
— викрадення облікових записів. APT29 здійснює атаки на облікові записи користувачів, включаючи злом електронної пошти та соціальних мереж, для отримання несанкціонованого доступу до конфіденційної інформації та перехоплення комунікацій;
— використання розширених персистентних загроз (APT). APT29 використовує складні та продумані APT, такі як програми-шпигуни, троянці та інші загрози, які дозволяють їм залишатися в системах-жертвах протягом тривалого часу і непомітно виконувати шпигунські операції.

APT29 є досить спритною та високотехнологічною кіберзагрозою, яка вимагає постійного моніторингу, заходів кібербезпеки та протидії для ефективного захисту від їхніх атак.

Sandworm (пісочний черв'як) є групою хакерів, пов'язаною з російськими розвідувальними службами, яка відома своєю активністю в кіберпросторі. Ця група вперше набула широкої популярності в 2014 році після своїх атак на Україну, зокрема на енергетичний сектор та урядові установи.

Sandworm відомий своїми складними та координованими кібератаками з використанням різноманітних інструментів і методів. Ось кілька розповсюджених атак, пов'язаних з групою Sandworm:

— використання вразливостей. Sandworm використовує вразливості в операційних системах та програмах для здійснення атак. Наприклад, вони використовували вразливість у програмі Microsoft Office для впровадження шкідливого коду та виконання атак на цільові системи;
— фішинг та соціальний інжиніринг. Група Sandworm використовує фішингові електронні листи та соціальний інжиніринг для отримання несанкціонованого доступу до систем. Вони можуть використовувати підроблені електронні листи, викликати довіру та переконати користувачів у відкритті шкідливих вкладень або переході на компрометовані веб-сторінки;
— розподіл шкідливого програмного забезпечення. Sandworm використовує різні методи розповсюдження шкідливого програмного забезпечення, включаючи використання торрент-мереж, зараження USB-носіїв та розподіл інфікованих файлів через інтернет;
— кібершпигунство та розвідка. Група Sandworm відома своїми спробами отримання конфіденційної інформації та проведення кібершпигунських операцій. Вони можуть проникати в мережі та системи цілей для отримання важливої інформації та виконання розвідувальних дій.

Sandworm вважається впливовою та небезпечною групою хакерів, яка продовжує активно діяти в кіберпросторі. Їх атаки вимагають постійного вдосконалення кібербезпекових заходів та протидії для забезпечення захисту від їхньої діяльності.

Gamaredon, також відомий як ПГУ, є назвою кіберзагрози, пов'язаної з групою хакерів, яка визначається своїми зв'язками з Росією. Ця група активна вже кілька років і спеціалізується на здійсненні кібератак на урядові установи, оборонні структури, енергетичні компанії та інші сектори.

Основні характеристики та методи, пов'язані з групою Gamaredon, включають наступне:

— фішингові атаки. Група Gamaredon використовує фішингові електронні листи, які містять шкідливі вкладення або посилання на компрометовані веб-сторінки. Це дозволяє їм здійснювати атаки на комп'ютери та мережі жертв, зловживаючи довірою користувачів;

— використання вразливостей. Група Gamaredon активно використовує вразливості в програмному забезпеченні, операційних системах та інших компонентах для отримання несанкціонованого доступу. Вони шукають та використовують раніше невідомі уразливості, а також використовують відомі вразливості, для отримання контролю над цільовими системами;

— розповсюдження шкідливого програмного забезпечення. Gamaredon використовує різноманітні методи розповсюдження шкідливих програм, включаючи електронну пошту, соціальні мережі, форуми та інші канали. Вони можуть використовувати троянські програми, віруси та інші типи шкідливого коду для зламу систем і розподілу шпигунського програмного забезпечення;

— викрадення даних та розвідка; Gamaredon зацікавлений у здобутті конфіденційної інформації. Вони можуть використовувати шпигунське

програмне забезпечення для перехоплення комунікацій, отримання доступу до важливих даних та виконання розвідувальних дій.

Група Gamaredon відома своєю активністю в кіберпросторі, особливо на пострадянському просторі. їх атаки можуть мати серйозні наслідки для жертв, включаючи втрату конфіденційної інформації, порушення безпеки та пошкодження репутації.

InvisiMole (також відомий як USB Thief) - це назва кіберзагрози, яка була виявлена у 2018 році. Ця загроза відноситься до групи атак, пов'язаних з кібершпигунством. InvisiMole є високо спеціалізованою і складною атакою, спрямованою на шпигунство і збір конфіденційної інформації.

Основні характеристики та методи, пов'язані з InvisiMole, включають наступне:

— постійний доступ. InvisiMole розроблений для довготривалого і непомітного проникнення в систему. Він встановлює свої складні компоненти на заражені системи, що дозволяє зловмисникам мати постійний доступ до цілі та контролювати її;
— шифрування та стеганографія. InvisiMole використовує різні методи шифрування для захисту своїх складних компонентів та забезпечення непомітності. Він також використовує стеганографію - техніку приховування інформації в інших носіях, таких як зображення або звукові файли, для ускладнення виявлення та аналізу;
— збір інформації. InvisiMole здатний збирати різноманітну конфіденційну інформацію зі заражених систем, включаючи вміст файлів, паролі, логи клавіатури та інші дані. Він також може реалізувати засоби нагляду та перехоплення комунікацій, включаючи веб-камеру та мікрофон;
— розповсюдження. InvisiMole може розповсюджуватися через заражені USB-носії, шляхом інфікування портативних пристроїв та мереж. Він

також може використовувати соціальний інжиніринг та фішингові методи для залучення жертв до виконання шкідливих дій.

InvisiMole є складною і небезпечною загрозою, спрямованою на цілеспрямоване шпигунство та збір інформації. Виявлення та протидія цій загрозі вимагають високого рівня кібербезпеки та використання відповідних заходів захисту.

GhostWriter (також відомий як UNC1151) є назвою кіберзагрози, пов'язаної з групою хакерів. Ця група відома своєю спроможністю проводити напади на веб-сайти та блоги з метою розповсюдження фальшивої або перекрученої інформації.

Основні характеристики та методи, пов'язані з GhostWriter, включають наступне:

— маніпуляція інформацією. GhostWriter спеціалізується на створенні та поширенні фальшивих або перекручених новинних матеріалів. Вони можуть створювати вигадані історії, перекручувати факти або використовувати маніпулятивні заголовки з метою впливу на громадську думку або впливу на політичні процеси;

— злам веб-сайтів. Група використовує різні методи для зламу веб-сайтів та отримання несанкціонованого доступу до них. Вони можуть використовувати вразливості в програмному забезпеченні, слабкі паролі адміністраторів, а також соціальний інжиніринг для отримання доступу до панелей управління веб-сайтів;

— розповсюдження дезінформації. GhostWriter використовує різні канали для розповсюдження своєї дезінформації. Це може включати використання соціальних мереж, поштових розсилок, форумів та інших онлайн-платформ. їхні цілі полягають у поширенні вигаданих історій та впливу на громадську думку;

— інформаційна війна. GhostWriter може використовувати свої навички та ресурси для ведення інформаційної війни. Вони можуть спрямовувати

свою дезінформацію на конкретні цільові аудиторії з метою впливу на їхні погляди, настрої та рішення.

Група GhostWriter використовує кіберпростір для поширення фальшивої або перекрученої інформації з метою впливу на громадську думку та політичні процеси. Їх дії можуть мати серйозні наслідки, такі як порушення довіри громадськості, погіршення взаємин між країнами та загроза національній безпеці.

Snake (також відомий як Turla або Uroburos) є назвою продвинутої кіберзагрози, пов'язаної з висококваліфікованою хакерською групою. Ця група відома своїми спритними та довготривалими кібератаками на різноманітні цілі, включаючи урядові організації, військові установи, дипломатичні місії та промислові компанії.

Основні характеристики та методи, пов'язані з Snake, включають наступне:

— розширена персистентність. Snake володіє високим рівнем персистентності та непомітності. Він встановлює свої компоненти на заражених системах, що дозволяє зловмисникам зберігати стійкий доступ та контроль над цільовими системами на тривалий час;

— засоби розпізнавання та обхід захисту. Snake використовує різні техніки для розпізнавання та обходу захисних механізмів, таких як антивіруси та фаєрволи. Він може захоплювати легітимні процеси та використовувати криптографію для ухилення від виявлення та аналізу;

— адаптивність та розвиток. Snake постійно розвивається та адаптується до нових захисних заходів. Він використовує нові техніки та інструменти, включаючи використання нульових днів та невідомих вразливостей, для проведення своїх кібератак;

— шпигунство та збір інформації. Snake спеціалізується на шпигунстві та зборі конфіденційної інформації. Він може перехоплювати комунікації, включаючи електронну пошту та файли, збирати логи клавіатури, витягувати паролі та інші важливі дані з заражених систем.

Snake є висококваліфікованою та небезпечною кіберзагрозою, яка вимагає ретельного моніторингу та заходів кібербезпеки для виявлення, захисту та протидії її кібератакам.

Dragonfly (також відомий як Energetic Bear) є назвою кіберзагрози, пов'язаної з хакерською групою, яка спеціалізується на кібершпигунстві та кібератаках на енергетичні компанії та інфраструктуру. Ця група стала відомою своїми нападами на критичну енергетичну інфраструктуру в різних країнах.

Основні характеристики та методи, пов'язані з Dragonfly, включають наступне:

— цілеспрямованість. Dragonfly зосереджує свої кібератаки на енергетичних компаніях та інфраструктурі. Їхні цілі полягають у здобутті конфіденційної інформації, викраденні даних та зламі систем, пов'язаних з енергетикою;
— соціальний інжиніринг. Dragonfly використовує соціальний інжиніринг, щоб впровадитися в цільові системи. Вони можуть відправляти фішингові електронні листи або використовувати інші методи, щоб зламати користувачів та отримати доступ до систем;
— використання вразливостей. Dragonfly використовує вразливості в програмному забезпеченні та мережевих протоколах для злому цільових систем. Вони можуть використовувати вразливості в операційних системах, програмах та інших компонентах, щоб отримати несанкціонований доступ;
— довільний доступ та шпигунство. Dragonfly прагне отримати довільний доступ до систем інфраструктури енергетики. Це дозволяє їм переглядати, копіювати та збирати конфіденційну інформацію, включаючи дані про енергетичні об'єкти та системи управління.

Група Dragonfly є серйозною кіберзагрозою для енергетичних компаній та інфраструктури. Їхні кібератаки можуть мати серйозні наслідки,

включаючи порушення роботи електромереж, виток конфіденційної інформації та загрозу енергетичній безпеці.

4.3 Методи машинного навчання

Машинне навчання (ML) є одним із найважливіших напрямів розвитку кібербезпеки. Воно використовується для вирішення різноманітних завдань, у тому числі для виявлення та ідентифікації кібератак.

В загалом існують два основних напрямки в машинному навчанні: це навчання за прецедентами, відоме також як індуктивне навчання, і дедуктивне навчання. До останнього відносять експертні системи, проте на сьогодні термін "машинне навчання" часто використовується як синонім для "навчання за прецедентами". Індуктивне навчання є актуальним напрямком, в той час як експертні системи зазнають кризи. Важкість узгодження баз знань експертних систем з реляційною моделлю даних ускладнює їх ефективне використання в промислових системах баз даних.

Навчання за прецедентами можна розділити на три основних типи: контрольоване навчання, також відоме як навчання з вчителем, неконтрольоване навчання або навчання без вчителя, і навчання з підкріпленням.

Окрім цього, розробляються й інші методи навчання, такі як активне, багатозадачне, різноманітне, трансферне тощо. Особливий успіх останнім часом має "глибоке навчання", яке успішно поєднує в собі алгоритми навчання з вчителем і без вчителя.

Контрольоване навчання застосовується в ситуаціях, де величезний обсяг даних, наприклад, тисячі зображень домашніх тварин, позначені ярликами, щоб вказати, чи це кішка, чи собака. У цьому випадку "вчитель" (людина) надає позначки, і машина, опираючись на ці дані, сама вивчає ознаки, які відрізняють кішок від собак. Даний метод навчання робить можливим швидке адаптування отриманого алгоритму для нових завдань,

наприклад, розпізнавання інших видів тварин. Таким чином, нейромережу, яка вивчила розпізнавати кішок і собак, можна легко переналаштувати для обробки результатів комп'ютерної томографії або інших завдань, де потрібно виявлення складних ознак.

Існує значна надлишковість даних без маркерів, в порівнянні з маркованими даними, такими як зображення без підписів, аудіозаписи без коментарів та тексти без анотацій. Завдання машини у неконтрольованому навчанні полягає в тому, щоб виявити взаємозв'язки між окремими даними, визначити закономірності, виявити шаблони або описати їхню структуру, здійснити класифікацію даних.

Неконтрольоване навчання застосовується в різних областях, таких як рекомендаційні системи, де, наприклад, на основі аналізу попередніх покупок, пропонуються товари, які можуть зацікавити покупця. Також воно використовується в інших сценаріях, таких як портал YouTube, який, враховуючи переглянуті відеокліпи, рекомендує подібний контент. Крім того, у пошукових системах, наприклад, Google, результати пошуку можуть відрізнятися для кожного користувача, враховуючи їх історію пошуків.

Навчання з підкріпленням є варіацією контрольованого навчання, проте вчителем у цьому випадку виступає саме "середовище". У цьому контексті машина (або агент) не має попередньої інформації про середовище, але може виконувати в ньому дії. Середовище реагує на ці дії, надаючи агенту дані, які дозволяють йому реагувати та вчитися. Фактично агент і середовище створюють систему з зворотнім зв'язком.

Застосування даного підходу для розв'язання складніших завдань, ніж навчання з учителем або без нього. Наприклад, воно використовується в системах навігації для роботів, які навчаються уникати зіткнень з перешкодами, отримуючи досвід та зворотний зв'язок при кожному зіткненні. Також навчання з підкріпленням застосовується в логістиці, плануванні завдань та в навчанні машин логічним іграм, таким як покер, нарди чи го.

У сучасному парадигмі машинного навчання різноманітні технології та алгоритми використовуються для аналізу та опрацювання даних. При цьому деякі з них, такі як дискримінантний аналіз та байєсівські класифікатори, базуються на традиційних математичних методах. Проте протягом останніх десятиліть, особливо з кінця XX століття, особлива увага приділяється штучним нейронним мережам (ANN).

Штучні нейронні мережі — це системи, що складаються з взаємодіючих штучних нейронів, які функціонують на основі простих процесорів. Вони відображають спробу моделювання роботи людського мозку, де нейрони взаємодіють між собою, обробляючи та передаючи імпульси.

Зокрема, ANN використовуються для вирішення різноманітних завдань, від розпізнавання образів до прогнозування тенденцій. Важливим кроком у розвитку ANN став "Метод зворотного поширення помилки", що значно покращив їхню здатність до навчання.

ANN може бути організована у вигляді шарів, де перший шар — вхідні нейрони, отримують дані, а кожен наступний шар — це приховані та вихідні нейрони. Приховані рівні можуть виконувати складні логічні перетворення, дозволяючи нейромережі виявляти взаємозв'язки у данних.

Глибоке навчання, яке використовується у контексті ANN, передбачає використання нейронних мереж із кількома прихованими рівнями. Ці рівні, чергуючись, можуть виявляти взаємозв'язки не лише на рівні окремих елементів, але й між взаємозв'язками.

Глибоке навчання призводить до високої якості у вирішенні завдань різної складності. Зокрема, його успішне впровадження у систему "Перекладач" від Google дозволило великою мірою покращити якість машинного перекладу між англійською та французькою мовами. Такий успіх демонструє потенціал глибокого навчання у покращенні різноманітних аспектів технологічного ландшафту.

Атрибуція кібератак — це процес визначення відповідальних за інцидент. Він включає в себе ідентифікацію зловмисників, їхніх мотивацій, інструментів та методів використання.

Машинне навчання може бути використано для вдосконалення процесів атрибуції кібератак у декількох аспектах:

— аналіз логів та поведінкової аналітики. Алгоритми ML можуть використовуватися для виявлення аномалій у системах та користувачах, які можуть вказувати на можливі кібератаки. Наприклад, ML може бути використано для виявлення аномалій у тому, як користувачі роблять вхід в систему або як вони використовують мережу;

— кластеризація та класифікація. Алгоритми ML можуть використовуватися для групування атак за подібністю та визначення їхніх характеристик. Наприклад, ML може бути використано для групування атак за методом, який вони використовують, або за типом даних, які вони атакують;

— прогнозування ризиків. Алгоритми ML можуть використовуватися для прогнозування ризиків кібератак. Наприклад, ML може бути використано для прогнозування того, які системи найбільш вразливі до атак, або які користувачі найбільш схильні до атак;

— глибоке навчання. Глибоке навчання є типом ML, який може аналізувати різноманітні дані, такі як тексти та зображення. Це дозволяє глибокому навчанню виявляти складні патерни, які можуть бути недоступні для традиційних методів ML.

Крім того, ML може використовуватися для інтеграції з Threat Intelligence, автоматизації процесу атрибуції та розробки систем інтелектуальної захисту. Перелічені техніки можуть допомогти організаціям ефективно реагувати на нові та раніше невідомі загрози.

Threat Intelligence — це інформація про поточні загрози та їхні джерела. ML може використовуватися для інтеграції з Threat Intelligence, щоб

забезпечити організації більш повним і актуальним уявленням про поточні загрози. Це може допомогти організаціям краще захиститися від атак.

Автоматизація процесу атрибуції. ML може використовуватися для автоматизації процесу атрибуції, щоб полегшити організаціям визначення відповідальних за кібератаки. Це може допомогти організаціям швидше реагувати на атаки та запобігти повторним атакам.

ML може використовуватися для розробки систем інтелектуальної захисту, які можуть самостійно виявляти та реагувати на кібератаки. Ці системи можуть допомогти організаціям забезпечити більш ефективний захист від кібератак.

У цілому, машинне навчання є потужним інструментом, який може допомогти організаціям покращити їхню кібербезпеку. Воно може використовуватися для виявлення, ідентифікації та реагування на кібератаки.

5 ПРОГРАМНА РЕАЛІЗАЦІЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ

5.1 Збір даних

У цьому розділі коду дані збираються з двох джерел:

— MITRE ATT&CK — це стандартна таксономія для опису кібер атак. Вона містить інформацію про різні тактики, техніки та процедури, які використовуються хакерами.

— VirusTotal — це онлайн-сервіс, який аналізує файли та веб-сайти на наявність шкідливого коду.

```
mitre_data = pd.read_json("mitre_attack.json")
virustotal_data = pd.read_csv("virustotal.csv")
```

Лістинг 5.1 – Отримання даних з фреймворків

Дані з MITRE ATT&CK та VirusTotal можна отримані за допомогою API.

```
def get_mitre_data():
    url = "https://attack.mitre.org/techniques/enterprise/"
    response = requests.get(url)
    if response.status_code == 200:
        soup = BeautifulSoup(response.text, "html.parser")
        parsed_data = {}
        return parsed_data
    else:
        print(f"Failed to fetch MITRE data. Status code: {response.status_code}")
        return None

def get_virustotal_data(api_key, resource):
    url = f"https://www.virustotal.com/api/v3/files/{resource}"
    headers = {"x-apikey": api_key}
    response = requests.get(url, headers=headers)

    if response.status_code == 200:
        parsed_data = {}
        return parsed_data
    else:
        print(f"Failed to fetch VirusTotal data. Status code: {response.status_code}")
        return None
```

Лістинг 5.2 – Підключення за API ключем

5.2 Очищення даних

Дані, які ми отримали з двох джерел, можуть містити помилки або шум. Перед тим, як використовувати їх для навчання моделі машинного навчання, їх необхідно очистити.

Для очищення даних з MITRE ATT&CK ми можемо видалити дублікати та записи, які містять неповні або невірні дані.

Для очищення даних з VirusTotal ми можемо видалити записи, які не містять інформації про шкідливий код.

```
mitre_data.drop_duplicates(inplace=True)
virustotal_data = virustotal_data.loc[virustotal_data["malicious"] == 1]
```

Лістинг 5.3 – Очищення даних

5.3 Підготовка даних та навчання моделі

Дані для машинного навчання часто вимагають додаткової підготовки. Наприклад, ми можемо перетворити їх у формат, який підходить для конкретної моделі машинного навчання, яку ми збираємося використовувати.

Для підготовки даних з MITRE ATT&CK ми можемо створити вектори характеристик для кожної атаки. Вектори характеристик можуть містити інформацію про такі характеристики атаки, як тактика, техніка, процедура, використовувані інструменти та цілі.

Для підготовки даних з VirusTotal ми можемо створити дескриптори для кожного файлу або веб-сайту. Дескриптори можуть містити інформацію про такі характеристики файлу або веб-сайту, як розмір, метадані та сигнатури шкідливого коду.

```
mitre_features = mitre_data[["tactic", "technique", "procedure", "tools", "targets"]].to_numpy()
virustotal_features = virustotal_data[["size", "metadata", "signatures"]].to_numpy()

model = LogisticRegression()
model.fit(np.concatenate([mitre_features, virustotal_features], axis=1), mitre_data["group"])
```

Лістинг 5.4 – Підготовки та навчання моделі

Після того, як очищено та підготовлено дані, розпочинається навчання моделі машинного навчання. Для цього використовуються різні алгоритми машинного навчання, такі як логістична регресія, підтримка векторних машин або нейронні мережі.

Для навчання моделі використано метод навчання з учителем. Це означає, що ми будемо мати набір даних, який містить як нові дані атаки, так і інформацію про те, до якої групи хакерів відноситься кожна атака.

5.4 Виконуюча частина

У цьому розділі коду дані з нового звіту про атаку порівнюються з даними з минулих звітів і Excel таблиці.

Для порівняння даних з минулих звітів у коді використовується функція compare_data(). Ця функція отримує список хешів, файлів, посилань, ір тощо з нового звіту та порівнює їх з даними з минулих звітів. Якщо дані збігаються, вони додаються до списку знайдених даних.

Для порівняння даних з Excel таблиці код використовує функцію compare_data_with_excel(). Ця функція отримує список технік з Excel таблиці та порівнює їх з техніками, які були знайдені в новому звіті. Якщо техніки збігаються, вони додаються до списку знайдених технік.

На основі знайдених даних код визначає можливі групи хакерів. Для цього код використовує функцію get_groups(). Ця функція отримує список технік і повертає список груп хакерів, які використовують ці техніки.

писок хеш-сумок файлів або веб-сайтів, які були знайдені в новому звіті. Хеш-сумки можна використовувати для виявлення повторних атак.

```python
def compare_data(new_data):
    new_hashes = new_data["hashes"]
    new_files = new_data["files"]
    new_links = new_data["links"]
    new_ips = new_data["ips"]

    matched_hashes = []
    matched_files = []
    matched_links = []
    matched_ips = []
    for hash in new_hashes:
        if hash in virustotal_data["sha256"]:
            matched_hashes.append(hash)
    for file in new_files:
        if file in virustotal_data["filename"]:
            matched_files.append(file)
    for link in new_links:
        if link in virustotal_data["url"]:
            matched_links.append(link)
    for ip in new_ips:
        if ip in virustotal_data["ip"]:
            matched_ips.append(ip)

    matched_techniques = []
    for hash in matched_hashes:
        matched_techniques.extend(virustotal_data.loc[virustotal_data["sha256"] == hash]["technique"].tolist())
    for file in matched_files:
        matched_techniques.extend(virustotal_data.loc[virustotal_data["filename"] == file]["technique"].tolist())
    for link in matched_links:
        matched_techniques.extend(virustotal_data.loc[virustotal_data["url"] == link]["technique"].tolist())
    for ip in matched_ips:
        matched_techniques.extend(virustotal_data.loc[virustotal_data["ip"] == ip]["technique"].tolist())

    return matched_hashes, matched_files, matched_links, matched_ips, matched_techniques
```

Лістинг 5.5 – Отримання даних та порівняння з минулими звітами

За результатом аналізу отриманих даних нової кібератаки отримаємо звіт у вигляді списку що містить у собі:

— файли. Список файлів, які були знайдені в новому звіті. Файлові імена можна використовувати для виявлення повторних атак.

— посилання. Їх можна використовувати для виявлення повторних атак.

— IP-адреси. Можна використовувати для виявлення повторних атак.

— техніки. Використовуються для визначення можливих груп хакерів.

— можливі групи хакерів. Список груп хакерів, які, ймовірно, відповідають за атаку. Цей список може бути використаний для подальшого розслідування атаки.

```
def attribute_attack(new_data, excel_file):
    matched_hashes, matched_files, matched_links, matched_ips, matched_techniques = compare_data(new_data)


    excel_matched_techniques = compare_data_with_excel(new_data, excel_file)

    all_matched_techniques = matched_techniques + excel_matched_techniques

    possible_groups = []
    for technique in all_matched_techniques:
        for group in mitre_data.loc[mitre_data["technique"] == technique]["group"].unique():
            if group not in possible_groups:
                possible_groups.append(group)

    print("Порівняння даних з минулими звітами:")
    print("Знайдені хеш-сумки:", matched_hashes)
    print("Знайдені файли:", matched_files)
    print("Знайдені посилання:", matched_links)
    print("Знайдені IP-адреси:", matched_ips)
    print("Знайдені техніки:", all_matched_techniques)
    print("Можливі групи хакерів:", possible_groups)
```

Лістинг 5.6 – Атрибуції кібератаки

5.5 Результати тесту моделі

При тестуванні даної моделі було створено датасет з наступними характеристиками:

— розмір. датасет містить понад 2 тисяч записів;

— різноманітність. Датасет включає записи про кібератаки, проведені різними групами зловмисників у різних країнах;

— якість. Дані в датасеті були ретельно очищені та перевірені на точність.

Датасет є важливим ресурсом для розробки моделей атрибуції кібератак. Він надає широкий спектр даних, які можуть бути використані для навчання моделей, які можуть точно ідентифікувати осіб або групи, відповідальні за кібератаки.

Дані що містяться у датасеті:

— тип атаки;

— використані методи атаки;

— метадані атак;

— використані інструменти та тактики;

— тип зловмисників;

— дати та часи атаки.

Для розрахунку метрики класифікатора, таких як точність, чутливість, специфічність, точність та оцінка, використовуються наступні формули:
«У наступні вирази підставлено числові значення

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad 5.1$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \quad 5.2$$

$$\text{Specificity} = \frac{TN}{TN + FP}, \quad 5.3$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad 5.4$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}, \quad 5.5$$

де символи мають наступні значення: True Positive (TP) = 265, False Positive (FP) = 46, True Negative (TN) = 112, False Negative (FN) = 27».

Таблиця 5.1 – Показники роботи класифікатора

Показник	Згорткова нейронна мережа
Точність (Accuracy)	0.838
Чутливість (Sensitivity)	0.907
Специфічність (Specificity)	0.709
Точність передбачення (Precision)	0.852
F1-оцінка	0.879

Згідно з числовими метриками, отриманими для згорткової нейронної мережі, можна визначити, що модель є ефективною в класифікації з високим показником точності (83.8%). Вона також виявила високу чутливість (90.7%), що свідчить про ефективність виявлення позитивних елементів.

Специфічність становить 70.9%, існує можливість її покращення. Точність передбачення висока (85.2%), що означає правильне визначення більшості позитивних припущень. F1-оцінка складає 87.9%, вказуючи на збалансованість між чутливістю та точністю. У даному випадку згорткова нейронна мережа проявила себе як ефективна модель класифікації, здатна досягати високих показників якості моделі. Однак важливо враховувати контекст та вимоги задачі для повного розуміння висновків та можливості подальшої оптимізації моделі.

ВИСНОВКИ

Модель машинного навчання для атрибуції кібератак є потужним підходом до виявлення та аналізу кібератак. Цей метод використовує індикатори компрометації, які є ознаками або вказівниками на наявність атаки, і використовує машинне навчання для автоматизованої обробки та аналізу цих індикаторів.

Застосування моделі дозволяє автоматизувати процес виявлення та аналізу загроз, зменшує час реакції на кібератаки та поліпшує точність виявлення. У моделі можливе використання різних алгоритмів та методів машинного навчання для навчання на великих обсягах даних та виявлення складних патернів та зв'язків між індикаторами компрометації.

В подальшому, завдяки глибинному машинному навчанню використання штучного інтелекту в методі атрибуції кібератак можна покращити ефективність і точність виявлення загроз, а також забезпечити більш швидку реакцію на атаки. Відповідно, цей підхід може допомогти організаціям забезпечити більш високий рівень кібербезпеки та захисту своїх інформаційних ресурсів.

Проте, важливо враховувати, що розроблена модель атрибуції не є бездоганним. Він може потребувати великих обсягів даних для навчання та налаштування моделей, а також постійного оновлення та вдосконалення для виявлення нових та еволюціонуючих загроз. Також важливо враховувати проблеми приватності та етики, пов'язані з обробкою та аналізом великих обсягів інформації.

Машинне навчання відіграє ключову роль у вдосконаленні процесів атрибуції кібератак. Воно може використовуватися для виявлення аномалій у системах та користувачах, групування атак за подібністю, прогнозування ризиків та виявлення складних патернів. Крім того, ML може використовуватися для інтеграції з Threat Intelligence, автоматизації процесу

атрибуції та розробки систем інтелектуальної захисту. Ці техніки можуть допомогти організаціям ефективно реагувати на нові та раніше невідомі загрози.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Фленов М.В., «Web-сервер глазами хакера»: 2-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2009. – 320с.
2. Peter Szor., «Art of Computer Virus Research and Defense», Addison-Wesley Professional (February 3, 2005).
3. Jon Erickson, «Hacking: The Art of Exploitation» (ISBN 1-59327-007-0).
4. MITRE ATT&CK [Електронний ресурс] / Режим доступу до ресурсу: https://attack.mitre.org
5. Lee, R. M., & Hutchins, E. M. (2013). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Computers & Security, 34, 10-17.
6. Kassner, M. (2016). Cyber attribution: What it is and why it matters. TechRepublic.
7. Mell, P., & Grance, T. (2012). The NIST Definition of Cloud Computing. National Institute of Standards and Technology.
8. ЗВІТ ПРО РОБОТУ - SCPC [Електронний ресурс] / Режим доступу до ресурсу: https://scpc.gov.ua/api/docs/19b0a96e-8c31-44bf-863e-cd3e0b651f20/19b0a96e-8c31-44bf-863e-cd3e0b651f20.pdf
9. Радіоелектроніка та молодь у XXI столітті : матеріали 27-го Міжнар. молодіж. форуму, 10-12 трав. 2023 р.: зб. матеріалів форуму. Т. 3. Конференція "Інформаційні радіотехнології та технічних захист інформації" / М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. – Харків : ХНУРЕ, 2023. – 329 с.
10. Machine Learning, ML [Електронний ресурс] / Режим доступу до ресурсу: https://www.it.ua/knowledge-base/technology-innovation/machine-learning
11. VirusTotal [Електронний ресурс] / Режим доступу до ресурсу: https://www.virustotal.com/gui/home/upload

12. CERT-UA - Урядова команда реагування на комп’ютерні надзвичайні події України, яка функціонує в складі Державної служби спеціального зв’язку та захисту інформації України. [Електронний ресурс] / Режим доступу до ресурсу: https://cert.gov.ua/