НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»
ФАКУЛЬТЕТ КІБЕРБЕЗПЕКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра кібербезпеки

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«ПРОГНОЗУВАННЯ КІБЕРАТАК ІЗ ЗАСТОСУВАННЯМ OPEN-SOURCE INTELLIGENCE (OSINT)»

Виконав: здобувач 4 курсу
Рівень бакалавр
Галузь знань 12 «Інформаційні технології»,
спеціальність 125 «Кібербезпека»
Ярина Владислав Борисович

Керівник: доктор технічних наук, проф.
Соколов Артем Вікторович

Рецензент: кандидат фізико-математичних наук, доц.
Чепурна Олена Євгенієвна

Одеса – 2025

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»

Факультет **кібербезпеки та інформаційних технологій**

Кафедра **кібербезпеки**

Галузь знань: **12 «Інформаційні технології»**

Спеціальність: **125 «Кібербезпека»**

Назва освітньої програми: **Кібербезпека**

ЗАТВЕРДЖУЮ
Завідувач кафедри кібербезпеки
професор С.А. Горбаченко
_____ «____ »_____ 20__ року

З А В Д А Н Н Я

на кваліфікаційну (бакалаврську) роботу

здобувачу Ярині Владиславу Борисовичу

1. Тема роботи: «Прогнозування кібератак із застосуванням Open-source intelligence (OSINT)»

Керівник роботи: д.т.н, проф. Соколов Артем Вікторович

затверджені наказом НУ ОЮА від «18» квітня 2025 року № 548-6

2. Строк подання здобувачем роботи «19» травня 2025 року

3. Дата видачі завдання «16» вересня 2025 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів бакалаврської роботи	Строк виконання етапів роботи	Примітка
1	Аналіз предметної області та пошук науково-технічної інформації за темою роботи	Вересень-жовтень 2024	
2	Узгодження теми з науковим керівником, формулювання мети завдань дослідження	Листопад 2024	
3	Робота над першим розділом	Листопад-грудень 2024	
4	Робота над другим розділом	Січень-лютий 2025	
5	Робота над третім розділом	Лютий-березень 2025	
6	Уточнення подальшого обсягу роботи та його коригування	Березень-травень 2025	
7	Оформлення звітної частини	Травень 2025	
8	Захист роботи	11.06.2025	

Здобувач _____

(підпис) (прізвище та ініціали)

Керівник роботи _____

(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота на тему «Прогнозування кібератак із застосуванням Open-source intelligence (OSINT)» на здобуття першого (бакалаврського) рівня вищої освіти за спеціальністю 125 «Кібербезпека», освітньою програмою: «Кібербезпека», містить 28 рисунків, 21 літературне джерело за переліком посилань. Робота виконана на 82 сторінки загального тексту і 70 сторінках основного тексту.

У роботі досліджено теоретичні основи прогнозування кібератак та концепцію Open-source intelligence (OSINT). Проаналізовано ключові джерела та інструменти OSINT, методи збору та аналізу даних для виявлення індикаторів підготовки атак, а також етичні й юридичні аспекти. Розглянуто практичне застосування OSINT-методик, зокрема з використанням інструментів theHarvester, Censys, Shodan та OSINT Framework, для прогнозування кібератак та оцінено їх ефективність.

Робота є актуальною для фахівців з кібербезпеки, аналітиків кіберзагроз, спеціалістів з OSINT-розвідки, а також для дослідників у сфері проактивного захисту інформації. Висновки та рекомендації, представлені в роботі, можуть бути використані для розробки та вдосконалення методик прогнозування кібератак, підвищення ефективності систем кіберзахисту та підготовки фахівців.

OSINT, КІБЕРАТАКИ, ПРОГНОЗУВАННЯ, КІБЕРБЕЗПЕКА, ВІДКРИТІ ДЖЕРЕЛА, КІБЕРРОЗВІДКА, АНАЛІЗ ДАНИХ

ANNOTATION

The qualification work on the topic "Forecasting Cyberattacks using Open-source Intelligence (OSINT)" for obtaining the first (bachelor's) level of higher education in the specialty 125 "Cybersecurity," educational program: "Cybersecurity," contains 28 figures, 21 literature sources listed in the references. The work is completed on 82 pages of total text and 70 pages of main text.

The work explores the theoretical foundations of cyberattack forecasting and the concept of Open-source Intelligence (OSINT). Key OSINT sources and tools, data collection and analysis methods for identifying indicators of attack preparation, as well as ethical and legal aspects are analyzed. The practical application of OSINT methodologies, particularly using tools like theHarvester, Censys, Shodan, and OSINT Framework, for forecasting cyberattacks and evaluating their effectiveness is examined.

The work is relevant for cybersecurity specialists, cyber threat analysts, OSINT intelligence specialists, as well as for researchers in the field of proactive information protection. The conclusions and recommendations presented in the work can be used for developing and improving cyberattack forecasting methodologies, enhancing the effectiveness of cybersecurity systems, and training specialists.

OSINT, CYBERATTACKS, FORECASTING, CYBERSECURITY, OPEN SOURCES, CYBER INTELLIGENCE, DATA ANALYSIS

ЗМІСТ

ВСТУП

Актуальність теми. Зростання кількості та складності кібератак зумовлює необхідність переходу від реактивних до проактивних стратегій кіберзахисту. Ключовим елементом таких стратегій є прогнозування потенційних атак, що дозволяє організаціям своєчасно вживати превентивних заходів та мінімізувати збитки. Роль OSINT. Розвідка на основі відкритих джерел (Open-source intelligence, OSINT) є важливим інструментом для прогнозування кібератак. Аналіз публічно доступної інформації (форуми, соціальні мережі, технічні ресурси, витоки даних) дозволяє виявляти індикатори підготовки атак (IoPA), аналізувати діяльність зловмисників та формувати обґрунтовані прогнози щодо майбутніх загроз.

Метою дипломної роботи є дослідження можливостей використання OSINT для прогнозування кібератак та розробка методичних засад аналізу відповідних даних.

Завдання дослідження:

1. Аналіз теоретичних основ прогнозування кібератак та концепції OSINT.
2. Ідентифікація та класифікація релевантних джерел, інструментів та методів аналізу OSINT-даних.
3. Розробка концептуального застосування OSINT для прогнозування атак.
4. Оцінка ефективності, обмежень та етичних аспектів OSINT-прогнозування.

Об'єктом дослідження є процес прогнозування кібератак з використанням даних з відкритих джерел.

Предметом дослідження є методи, інструменти, джерела OSINT та індикатори підготовки кібератак у відкритих даних.

Практичне значення результатів. Результати роботи можуть бути використані фахівцями з кібербезпеки для підвищення ефективності проактивного захисту, формування звітів кіберрозвідки та оптимізації розподілу ресурсів.

1 ТЕОРЕТИЧНІ ОСНОВИ ПРОГНОЗУВАННЯ КІБЕРАТАК ТА OSINT

1.1 Сутність та класифікація кібератак. Життєвий цикл кібератаки

Кібератака визначається як цілеспрямований зловмисний вплив на інформаційні системи, мережі, пристрої чи дані. Виконавцями таких дій можуть бути окремі особи, організовані групи або державні структури. Ключовими характеристиками кібератаки є навмисність дій та застосування технологічних засобів для подолання встановлених механізмів захисту цільової системи. Кінцеві цілі варіюються від несанкціонованого доступу до інформації, її викрадення, модифікації чи знищення, до порушення функціонування систем, фінансового шахрайства, шпигунства або завдання репутаційної шкоди.

Для систематизації розуміння цих загроз використовується класифікація кібератак за різними критеріями [18]. Одним з основних критеріїв є вектор атаки, що описує спосіб її реалізації. До поширених векторів належать розповсюдження шкідливого програмного забезпечення (вірусів, хробаків, троянців, програм-вимагачів), застосування методів фішингу та соціальної інженерії для маніпуляції користувачами, експлуатація програмних чи апаратних вразливостей для отримання доступу, здійснення атак на відмову в обслуговуванні (DoS/DDoS) шляхом перевантаження системи, перехоплення комунікацій через атаки типу "людина посередині" (MitM), а також атаки на веб-додатки, такі як SQL-ін'єкції чи міжсайтовий скриптинг (XSS) [18].

Іншим критерієм класифікації є мета атаки, яка може полягати у викраденні даних різного типу (фінансових, персональних, комерційних), отриманні фінансової вигоди, здійсненні шпигунства, саботажі роботи систем чи завданні шкоди репутації. Також атаки класифікують за об'єктом впливу, яким можуть бути як персональні пристрої, так і корпоративні мережі, державні ресурси, об'єкти критичної інфраструктури чи веб-сервіси. Важливо відзначити, що конкретна кібератака може поєднувати ознаки декількох класифікаційних категорій.

Для розробки ефективних методів протидії, зокрема прогностичних, важливим є розуміння життєвого циклу кібератаки. Існують різні моделі опису цього процесу (напр., Cyber Kill Chain, MITRE ATT&CK Framework), проте узагальнено він представляє послідовність етапів [18]. Початковим етапом є розвідка, під час якої збирається інформація про потенційну ціль з використанням пасивних (включаючи OSINT) та активних методів для ідентифікації слабких місць [7]. Далі відбувається розробка та озброєння, тобто підготовка інструментарію атаки (шкідливого ПЗ, експлойтів). На етапі доставки підготовлені засоби передаються до цільової системи. Після цього здійснюється експлуатація вразливості для отримання початкового доступу, за яким слідує етап встановлення, що передбачає закріплення у системі (наприклад, через бекдори). Наступний крок – управління та контроль, що полягає у встановленні каналу зв'язку з командним сервером зловмисника. Завершальним етапом є виконання дій по досягненню мети, що включає ексфільтрацію даних, шифрування, руйнування систем або подальше розповсюдження в мережі.

Аналіз активності, особливо на ранніх стадіях циклу, таких як розвідка з використанням OSINT, формує основу для прогнозування майбутніх атак. Розуміння послідовності дій зловмисника дозволяє ідентифікувати потенційні індикатори підготовки або здійснення атаки.

1.2 Концепція прогнозування кібератак

Концепція прогнозування кібератак є основою проактивної стратегії кібербезпеки, що спрямована на випередження дій зловмисників шляхом передбачення майбутніх загроз [16]. На відміну від традиційних реактивних підходів, які зосереджені на виявленні та реагуванні на атаки, що вже відбуваються або відбулися, прогнозування має на меті завчасну ідентифікацію підвищених ризиків, ймовірних цілей, потенційних векторів атак або часових вікон їх здійснення. Це дозволяє організаціям пріоритезувати заходи захисту, ефективніше розподіляти

ресурси та адаптувати свою інфраструктуру безпеки до очікуваних загроз. Кінцевою метою є не стільки точне передбачення кожної окремої атаки, скільки створення інформаційного підґрунтя для прийняття обґрунтованих рішень щодо превентивних дій.

Існують різні підходи до прогнозування кібератак, які часто використовуються у комплексі:

1. Підхід на основі індикаторів (Indicator-based): Фокусується на виявленні ознак або індикаторів, що передують атаці (Indicators of Attack - IoA, або Indicators of Potential Attack - IoPA). Це може включати моніторинг активності на етапі розвідки (сканування мереж), виявлення прекурсорів шкідливого ПЗ, аналіз обговорень на спеціалізованих форумах (зокрема, у Dark Web), відстеження продажу експлойтів або даних доступу. Аналіз даних OSINT є невід'ємною частиною цього підходу.

2. Підхід на основі вразливостей (Vulnerability-based): Прогнозування будується на аналізі відомих вразливостей у програмному та апаратному забезпеченні, даних про їх експлуатацію "в дикій природі", наявності публічно доступних експлойтів та темпів впровадження оновлень безпеки (патчів) в організаціях. Висока концентрація не виправлених критичних вразливостей може вказувати на підвищену ймовірність атаки.

3. Підхід на основі зловмисника (Attacker-based): Моделювання поведінки, мотивації, можливостей та типових тактик, технік і процедур (TTPs) відомих кіберзлочинних угруповань або типів зловмисників. Аналіз їхніх попередніх цілей та методів може допомогти спрогнозувати майбутні дії.

4. Підхід на основі даних (Data-driven/ML-based): Використання великих обсягів історичних даних про атаки, мережевий трафік, системні журнали, дані з систем безпеки для побудови моделей (часто з використанням машинного навчання), що виявляють складні патерни, які передують атакам.

В рамках зазначених підходів застосовуються різноманітні методи прогнозування. До них належать статистичний аналіз та аналіз часових рядів для

виявлення тенденцій у частоті та типах атак. Широко використовуються методи машинного навчання: класифікація для розрізнення легітимної та зловмисної активності, регресія для прогнозування обсягів атак чи ймовірності їх виникнення, кластеризація для групування схожих інцидентів або зловмисників, а також методи виявлення аномалій для ідентифікації нетипової поведінки систем чи користувачів. Техніки аналізу даних OSINT, включаючи обробку природної мови (NLP) для аналізу текстової інформації з форумів та соціальних мереж, є важливими для підходу на основі індикаторів. Також застосовується структуроване моделювання загроз (Threat Modeling) та аналіз даних, зібраних за допомогою систем-приманок (Honeypots/Honeynets).

Незважаючи на розвиток технологій, прогнозування кібератак стикається з низкою суттєвих викликів. По-перше, це проблеми з даними: їх часто недостатньо, вони мають низьку якість, містять багато шуму або є неструктурованими. Збір та обмін даними про інциденти ускладнюється питаннями конфіденційності та небажанням організацій розкривати інформацію про атаки. По-друге, динамічність ландшафту загроз: зловмисники постійно змінюють свої TTPs, з'являються нові, раніше невідомі (zero-day) вразливості та вектори атак, що робить історичні дані менш релевантними для прогнозування майбутнього. По-третє, складність та масштаб сучасних ІТ-систем та різноманітність мотивацій зловмисників ускладнюють побудову точних моделей. Існує також проблема розрізнення потенційної загрози (наміру, обговорення) від реальної можливості та рішучості її реалізувати. Нарешті, важливим викликом є досягнення прийнятного балансу між виявленням реальних загроз та мінімізацією хибних спрацьовувань (false positives), які можуть призвести до надмірних витрат ресурсів. Ускладнена атрибуція атак також заважає точному моделюванню поведінки конкретних зловмисників.

Подолання цих викликів вимагає комплексного підходу, що поєднує технологічні рішення, якісний аналіз даних (включаючи OSINT) та експертні знання у сфері кібербезпеки.

1.3 Open-source intelligence (OSINT): визначення, джерела, інструменти та методологія. Цикл OSINT

Open-source intelligence (OSINT) визначається як процес збору, обробки та аналізу інформації з публічно доступних та легально отриманих джерел з метою створення розвідувального продукту, що підтримує прийняття рішень [7, 17]. Ключовим аспектом OSINT є використання саме відкритих джерел, що відрізняє його від інших видів розвідки, які можуть застосовувати таємні або нелегальні методи отримання інформації. В контексті кібербезпеки OSINT набуває все більшого значення як засіб для розуміння ландшафту загроз, моніторингу активності потенційних зловмисників та виявлення індикаторів підготовки атак.

Джерела OSINT охоплюють надзвичайно широкий спектр інформації. Сюди належать ресурси глобальної мережі Інтернет, включаючи загальнодоступні веб-сайти, пошукові системи, соціальні мережі, блоги, форуми (в тому числі спеціалізовані та тіньові сегменти, такі як Dark/Deep Web), публічні репозиторії коду та бази даних. Важливу роль відіграють засоби масової інформації – новинні портали, онлайн та друковані видання, теле- та радіомовлення. Джерелами також виступають офіційні публікації урядів та міжнародних організацій, публічні реєстри, законодавчі акти, звіти. Додатково використовуються комерційні бази даних, корпоративна звітність, наукові публікації, матеріали конференцій, патенти, а також так звана "сіра література" – технічні звіти, дисертації та інші документи, що не пройшли формального комерційного видання. Різноманіття та величезний обсяг цих джерел вимагають систематичного підходу до їх відбору та обробки.

Ефективна робота з великими масивами відкритих даних неможлива без використання спеціалізованих інструментів OSINT. Існує значна кількість програмних засобів, які можна класифікувати за їх функціональним призначенням. До них належать пошукові системи, як загального призначення (Google, Bing), так і

спеціалізовані (наприклад, Shodan для пошуку підключених до Інтернету пристроїв, спеціалізовані пошуковики для Dark Web). Важливу роль відіграють інструменти для агрегації, візуалізації та аналізу зв'язків між даними (наприклад, Maltego). Існують засоби для аналізу соціальних мереж, вилучення метаданих з файлів, автоматизованого збору даних з веб-сайтів (веб-скрейпінг, краулінг), часто реалізовані у вигляді окремих скриптів або бібліотек. Також розроблені комплексні платформи та фреймворки (наприклад, OSINT Framework), що структурують доступ до різноманітних інструментів та ресурсів. Вибір конкретних інструментів залежить від поставлених завдань, типу джерел та необхідної глибини аналізу.

OSINT – це не хаотичний пошук інформації, а структурований процес, що описується методологією розвідки. Основою цієї методології є Цикл OSINT, який представляє собою послідовність взаємопов'язаних етапів:

1) Планування та визначення вимог (Planning and Direction): Чітке формулювання інформаційних потреб, визначення цілей розвідки, ключових питань, обсягу роботи та часових рамок. На цьому етапі визначаються пріоритетні напрямки пошуку.

2) Збір даних (Collection): Систематичний пошук та збір релевантної інформації з ідентифікованих відкритих джерел з використанням обраних інструментів та технік.

3) Обробка (Processing): Перетворення зібраної сирої інформації у форму, придатну для аналізу. Це може включати переклад, структурування даних, фільтрацію шуму, видалення дублікатів, розшифровку.

4) Аналіз та вироблення продукту (Analysis and Production): Інтерпретація обробленої інформації, виявлення закономірностей, взаємозв'язків, оцінка достовірності та значущості даних, формулювання висновків та створення кінцевого розвідувального продукту (звіт, аналітична довідка, прогноз) [7, 17].

5) Розповсюдження та використання (Dissemination and Use): Своєчасна доставка розвідувального продукту кінцевим споживачам (особам, що приймають рішення) у зрозумілій та зручній формі.

Зворотний зв'язок (Feedback): Оцінка якості та корисності наданої інформації споживачами, що дозволяє коригувати та вдосконалювати весь цикл OSINT на наступних ітераціях.

Цей цикл є ітеративним, тобто результати одного етапу можуть вимагати повернення до попередніх для уточнення вимог, пошуку додаткових даних або перегляду аналізу. Суворе дотримання методології забезпечує системність, об'єктивність та ефективність OSINT-діяльності.

1.4 Роль OSINT у виявленні та аналізі кіберзагроз. OSINT та прогнозування

Розвідка на основі відкритих джерел (OSINT) відіграє дедалі важливішу роль у сучасному ландшафті кібербезпеки, слугуючи ключовим інструментом для виявлення та аналізу кіберзагроз. Використання OSINT дозволяє організаціям отримати глибоке розуміння тактик, технік та процедур (TTPs), що застосовуються зловмисниками, шляхом аналізу загальнодоступних звітів про інциденти, технічних описів шкідливого ПЗ, дискусій на спеціалізованих форумах та блогах дослідників безпеки. OSINT-методи використовуються для профілювання відомих кіберзлочинних угрупувань та індивідуальних акторів, допомагаючи зрозуміти їхню мотивацію, можливості, інфраструктуру та ймовірні цілі. Це здійснюється через моніторинг соціальних мереж, тіньових сегментів Інтернету, аналіз витоків даних та публічних заяв.

OSINT є ефективним засобом для відстеження інформації про нові вразливості та їх можливу експлуатацію. Моніторинг баз даних вразливостей (напр., CVE), репозиторіїв коду (GitHub), списків розсилки та форумів дозволяє оперативно дізнаватися про появу нових загроз та публічно доступних експлойтів. Також OSINT використовується для ідентифікації потенційно зловмисної інфраструктури –

командних серверів (C2), фішингових доменів, IP-адрес, асоційованих з атаками, що згадуються у відкритих джерелах. Аналіз публічних даних може надати ранні попередження про підготовку масштабних кіберкампаній, наприклад, через виявлення обговорень планованих атак, продажу несанкціонованого доступу до мереж організацій або викрадених баз даних.

Зв'язок OSINT з прогнозуванням кібератак є прямим та фундаментальним. Інформація, отримана за допомогою OSINT, становить значну частину вихідних даних, необхідних для функціонування прогностичних моделей та підходів, розглянутих у пункті 1.2.

1) OSINT як джерело індикаторів: Моніторинг активності зловмисників на етапі розвідки, аналіз дискусій на форумах, відстеження продажу експлойтів та інші OSINT-активності безпосередньо забезпечують виявлення індикаторів підготовки атаки (IoPA), що є основою для індикаторного підходу до прогнозування.

2) Контекстуалізація вразливостей: Дані OSINT про активну експлуатацію конкретних вразливостей, наявність інструментів для їх використання та обговорення їх зловмисниками доповнюють технічну інформацію про вразливості, дозволяючи більш точно оцінити ризики та прогнозувати атаки в рамках підходу на основі вразливостей.

3) Інформація для профілювання акторів: Зібрані за допомогою OSINT відомості про TTPs, мотивацію та цілі конкретних кіберзлочинних груп слугують вхідними даними для моделей прогнозування, що базуються на аналізі поведінки зловмисників.

4) Дані для моделей машинного навчання: Масштабний збір та аналіз OSINT-даних (наприклад, частота згадувань певних технологій чи організацій у контексті загроз, сентимент-аналіз дискусій) може забезпечити необхідні обсяги даних для тренування моделей машинного навчання, спрямованих на виявлення складних прогностичних патернів.

Особливо важливим є те, що OSINT дозволяє отримати уявлення про активність на ранніх етапах життєвого циклу кібератаки, зокрема на етапі розвідки, який є критично важливим для успішного прогнозування. Виявлення ознак підготовки до атаки надає захисникам цінний часовий ресурс для вжиття превентивних заходів.

Дані OSINT можуть містити шум, дезінформацію або бути неповними. Тому для ефективного прогнозування часто необхідна їх кореляція з даними з інших джерел розвідки, включаючи внутрішню телеметрію систем безпеки організації. Проте, OSINT залишається незамінним компонентом сучасної кіберрозвідки та необхідною передумовою для побудови дієвих систем прогнозування кібератак.

2 АНАЛІЗ ДЖЕРЕЛ ТА МЕТОДІВ OSINT ДЛЯ ПРОГНОЗУВАННЯ КІБЕРАТАК

2.1 Огляд ключових джерел OSINT для виявлення індикаторів підготовки атак.

Ефективність прогнозування кібератак значною мірою залежить від здатності ідентифікувати та аналізувати інформацію з релевантних джерел відкритих даних. Хоча потенційних джерел OSINT існує безліч, для виявлення індикаторів підготовки атак (IoPA) особливу увагу слід приділити тим ресурсам, де з високою ймовірністю може з'являтися інформація про наміри, можливості або активність зловмисників. Розглянемо ключові категорії таких джерел:

1) Форуми та спільноти Dark/Deep Web: Цей сегмент Інтернету, недоступний для стандартних пошукових систем, є важливим осередком активності кіберзлочинців. На спеціалізованих форумах, торгових майданчиках та чат-групах відбувається обговорення та продаж експлойтів (включаючи zero-day), шкідливого програмного забезпечення, викрадених баз даних, облікових записів, даних платіжних карток та доступу до скомпрометованих мереж. Моніторинг цих ресурсів може надати прямі індикатори підготовки конкретних атак, появи нових інструментів чи вразливостей, а також інформацію про цілі та мотивацію зловмисників. Однак збір даних з Dark/Deep Web пов'язаний з технічними труднощами, ризиками безпеки, мовними бар'єрами та проблемами верифікації інформації.

2) Соціальні мережі (Twitter, Telegram, LinkedIn тощо): Публічні платформи активно використовуються різними суб'єктами кіберпростору. Хактівістські групи часто анонсують свої кампанії або звітують про атаки. Дослідники безпеки обговорюють нові загрози та вразливості. Співробітники компаній можуть ненавмисно розголошувати чутливу інформацію, корисну для розвідки зловмисників. Моніторинг ключових слів, хештегів та профілів, пов'язаних з кібербезпекою, хакерськими угрупуваннями чи потенційними цілями, може виявити ранні ознаки загроз або зміни у тактиках зловмисників.

3) Сайти для обміну текстом та кодом (Pastebin, GitHub Gist тощо): Ці платформи часто використовуються для тимчасового розміщення фрагментів коду, конфігураційних файлів, списків облікових даних або інших текстових даних. Зловмисники можуть використовувати їх для обміну інформацією, публікації викрадених даних або навіть розміщення частин шкідливого коду. Автоматизований моніторинг цих ресурсів за певними ключовими словами чи патернами (напр., імена компаній, IP-адреси, типові формати облікових даних) може виявити витоки або іншу релевантну активність.

4) Публічні репозиторії коду (GitHub, GitLab тощо): Ці ресурси є джерелом інформації про розробку програмного забезпечення, включаючи інструменти для кібербезпеки (як захисні, так і наступальні). Аналіз репозиторіїв може виявити публікацію нових експлойтів (Proof-of-Concept, PoC), інструментів для автоматизації атак, витоки вихідного коду з вразливостями або вбудованими обліковими даними. Також тут можна відстежувати активність відомих дослідників або розробників шкідливого ПЗ.

5) Технічні блоги, дослідницькі сайти та списки розсилки: Публікації дослідників безпеки, звіти компаній, що займаються кіберрозвідкою (Threat Intelligence), та обговорення у спеціалізованих спільнотах є цінним джерелом детальної інформації про нові TTPs, аналіз шкідливих кампаній та технічні деталі вразливостей. Зловмисники активно відстежують ці ресурси для пошуку нових методів атак, тому публікація інформації про нову техніку експлуатації може бути сильним індикатором її майбутнього масового застосування.

6) Бази даних вразливостей (CVE, NVD, Exploit-DB тощо): Самі по собі ці бази даних фіксують вже відомі вразливості. Однак їх моніторинг на предмет появи нових записів, особливо з високим рівнем критичності (високий бал CVSS), наявністю публічного експлойта або інформації про активне використання "в дикій природі", є важливим для прогнозування. Кореляція цієї інформації з даними з інших джерел (напр., обговорення на форумах) підвищує точність прогнозу.

7) Витоки даних та ресурси моніторингу витоків: Інформація про скомпрометовані облікові записи (пари логін/пароль, email), що з'являється у відкритому доступі в результаті витоків, є прямим індикатором підвищеного ризику атак типу credential stuffing або фішингових кампаній проти відповідних користувачів чи організацій.

8) Новинні ресурси та публічні звіти: Повідомлення у ЗМІ та офіційні звіти державних установ (напр., CERT) або компаній про великі кіберінциденти, нові типи атак, геополітичну напруженість, що може впливати на кіберпростір, надають важливий контекст для розуміння загального ландшафту загроз та прогнозування атак на певні сектори чи регіони.

Аналіз інформації з цих джерел вимагає комбінованого підходу, що поєднує автоматизовані засоби моніторингу та збору даних з кваліфікованим людським аналізом для верифікації, інтерпретації та оцінки значущості виявлених індикаторів. Жодне джерело окремо не є достатнім, і лише кореляція даних з різних ресурсів дозволяє сформувати найбільш повну картину та підвищити точність прогнозування.

2.2 Інструментарій для збору та аналізу OSINT-даних

Ефективний збір та аналіз інформації з різноманітних відкритих джерел у великих обсягах неможливий без застосування спеціалізованого інструментарію. Сучасний арсенал OSINT-фахівця включає широкий спектр програмних засобів – від графічних платформ для аналізу зв'язків до консольних утиліт та власних скриптів. Вибір інструментів залежить від конкретних завдань, типу джерел та необхідного рівня автоматизації.

Графічні платформи, такі як Maltego, надають потужні можливості для візуалізації та аналізу зв'язків між різними сутностями (домени, IP-адреси, email, профілі в соцмережах, організації тощо). Maltego використовує так звані "трансформації" (Transforms) для автоматичного збору даних з різних джерел (DNS-

записи, WHOIS, соціальні мережі, Shodan та ін.) та побудови графів взаємозв'язків. Це дозволяє виявляти приховані патерни, інфраструктуру зловмисників або зв'язки між різними індикаторами компрометації.

Спеціалізовані пошукові системи, як Shodan, фокусуються на індексації пристроїв, підключених до Інтернету (серверів, маршрутизаторів, IoT-пристроїв, систем промислової автоматизації) [2, 13]. Shodan дозволяє шукати пристрої за банерами сервісів, геолокацією, портами, програмним забезпеченням та відомими вразливостями. Для аналізу кіберзагроз Shodan може використовуватися для ідентифікації потенційно вразливих систем, виявлення командних серверів (C2) або моніторингу інфраструктури певних організацій чи країн, що може вказувати на підготовку атаки.

Для автоматизації процесу розвідки на початкових етапах часто використовуються консольні інструменти та фреймворки. Recon-ng є потужним фреймворком, написаним на Python, що надає модульну структуру для проведення розвідки. Він включає модулі для роботи з різними джерелами даних (пошукові системи, Shodan, VirusTotal, GitHub тощо) та виконання різноманітних завдань – від пошуку email-адрес та субдоменів до аналізу репозиторіїв коду. The Harvester – ще одна популярна утиліта для збору email-адрес, імен користувачів, субдоменів та відкритих портів, пов'язаних з певним доменом, використовуючи різні публічні джерела (Google, Bing, PGP-сервери, LinkedIn та ін.) [3].

Інформаційні ресурси, такі як OSINT Framework, не є інструментами у прямому сенсі, але надають структурований каталог посилань на сотні різноманітних OSINT-інструментів та джерел даних, згрупованих за категоріями (імена користувачів, email-адреси, IP-адреси, соціальні мережі, Dark Web тощо) [20]. Це цінний ресурс для пошуку відповідного інструменту під конкретне завдання.

Окрім готових інструментів, важливу роль відіграє можливість написання власних скриптів для веб-скрейпінгу або взаємодії з програмними інтерфейсами (API). Веб-скрейпінг дозволяє автоматично видобувати дані з веб-сайтів, які не

надають зручного API. Наприклад, за допомогою бібліотек Python, таких як requests (для надсилання HTTP-запитів) та BeautifulSoup (для парсингу HTML/XML), можна написати скрипт для моніторингу певного форуму чи блогу на предмет появи ключових слів.

Багато сервісів та платформ (VirusTotal, Shodan, Twitter, GitHub та ін.) надають програмні інтерфейси (API), які дозволяють програмно отримувати дані у структурованому форматі (зазвичай JSON). Використання API є більш надійним та ефективним способом отримання даних порівняно зі скрейпінгом.

Таким чином, ефективна OSINT-діяльність для прогнозування кібератак вимагає гнучкого підходу до вибору інструментарію, поєднуючи використання готових платформ, спеціалізованих утиліт та, за необхідності, розробку власних засобів для збору та первинної обробки даних з відкритих джерел.

2.3 Методи аналізу OSINT-даних для прогнозування

Зібрані за допомогою інструментів OSINT дані, хоч і потенційно цінні, є лише сировиною. Для перетворення їх на дієві розвідувальні продукти, здатні підтримати прогнозування кібератак, необхідно застосовувати відповідні методи аналізу. Ці методи варіюються від базових технік фільтрації до складних алгоритмів машинного навчання.

Аналіз ключових слів (Keyword Analysis): Один з найпростіших, але фундаментальних методів. Він полягає у пошуку та підрахунку частоти згадувань певних термінів, назв організацій, технологій, імен користувачів, ідентифікаторів вразливостей (CVE) або інших специфічних маркерів у великих масивах текстових даних (напр., з форумів, новин, соціальних мереж). Різке збільшення частоти згадувань певної вразливості або компанії у негативному контексті на тіньових форумах може бути раннім індикатором підготовки атаки.

Сентимент-аналіз (Sentiment Analysis): Метод, спрямований на визначення емоційного забарвлення або тональності тексту (позитивна, негативна, нейтральна). В контексті OSINT для кібербезпеки, сентимент-аналіз може допомогти оцінити ставлення до певної технології, компанії або вразливості у дискусіях на форумах чи в соціальних мережах. Наростання негативного сентименту щодо потенційної цілі може сигналізувати про зростання інтересу зловмисників.

Аналіз соціальних мереж (Social Network Analysis - SNA): Застосовується для візуалізації та аналізу структури взаємозв'язків між учасниками спільнот (наприклад, користувачами форумів, розробниками на GitHub). SNA дозволяє ідентифікувати ключових, найбільш впливових акторів (висока центральність), виявляти щільні групи (кластери) спілкування, відстежувати потоки інформації та розуміти динаміку взаємодії всередині спільноти. Аналіз зв'язків між відомими зловмисниками та іншими учасниками може допомогти виявити нових акторів або спільні інтереси, що потенційно вказують на підготовку спільних дій.

Виявлення аномалій (Anomaly Detection): Методи, спрямовані на пошук даних або подій, що суттєво відрізняються від очікуваної або нормальної поведінки. В OSINT це може бути раптовий сплеск активності певного користувача, нетипова частота публікацій на певну тему, поява дискусій про зазвичай не обговорювану технологію або організацію, різка зміна сентименту. Виявлення таких аномалій може сигналізувати про початок розвідки, координацію дій або інші нетипові процеси, що передують атаці.

Кореляційний аналіз: Важливий метод, що полягає у пошуку та оцінці взаємозв'язків між різними подіями або даними з різних джерел OSINT. Наприклад, виявлення кореляції між публікацією нового експлойта на GitHub, його обговоренням на Dark Web форумі та збільшенням сканувань відповідного порту (за даними Shodan) значно підвищує ймовірність підготовки атаки з використанням цієї вразливості. Кореляція дозволяє об'єднати розрізнені фрагменти інформації в єдину, більш значущу картину.

Базові моделі машинного навчання (Machine Learning - ML): Зі зростанням обсягів OSINT-даних, все частіше застосовуються методи ML.

1. Класифікація: Моделі можуть навчатися автоматично класифікувати тексти (наприклад, розрізняти пости про продаж доступу від звичайних обговорень), ідентифікувати типи зловмисників за їхньою онлайн-активністю.

2. Кластеризація: Дозволяє автоматично групувати схожі загрози, TTPs, зловмисників або інциденти на основі зібраних даних, виявляючи неочевидні зв'язки.

3. Аналіз часових рядів: Моделі можуть аналізувати історичні дані (наприклад, частоту згадувань CVE) для виявлення трендів та прогнозування майбутніх значень.

Впровадження ML вимагає значних обсягів якісних даних для навчання та експертизи у галузі Data Science, однак може суттєво підвищити ефективність та швидкість аналізу OSINT-даних для прогнозування.

На практиці, найбільш ефективним є комбіноване застосування декількох методів аналізу. Результати автоматизованого аналізу мають обов'язково перевірятися та інтерпретуватися кваліфікованим аналітиком, який може врахувати контекст, оцінити достовірність джерел та зробити остаточні висновки щодо потенційних загроз.

2.4 Етичні та юридичні аспекти використання OSINT

Open-source intelligence (OSINT) за визначенням оперує даними з відкритих та публічно доступних джерел, його застосування, особливо у сфері кібербезпеки та прогнозування загроз, не є вільним від суттєвих етичних дилем та юридичних обмежень. Нехтування цими аспектами може призвести не тільки до порушення законодавства, але й до репутаційних втрат та підриву довіри до OSINT-діяльності.

Етичні міркування при роботі з OSINT включають:

1. Конфіденційність та приватність: Навіть публічно доступна інформація може бути чутливою, особливо коли вона стосується приватних осіб. Збір, агрегація та аналіз таких даних, навіть з відкритих джерел (напр., профілі в соціальних мережах, дані з витоків), можуть порушувати обґрунтовані очікування приватності осіб. Необхідно ретельно оцінювати необхідність збору персональних даних та потенційний вплив на приватність.

2. Точність та інтерпретація: Інформація у відкритих джерелах не завжди є достовірною, вона може бути застарілою, неповною або навіть навмисно сфальсифікованою (дезінформація). Аналітик має критично оцінювати надійність джерел та уникати упереджених інтерпретацій, які можуть призвести до хибних висновків та необґрунтованих звинувачень.

3. Потенційна шкода: Необережне поводження зі зібраною OSINT-інформацією або її витік може завдати шкоди особам чи організаціям, дані про яких було зібрано. Крім того, публікація певних OSINT-знахідок (наприклад, деталей про вразливості чи методи атак) може бути використана іншими зловмисниками.

4. Пропорційність та необхідність: Збір даних має бути пропорційним до поставленої мети. Слід уникати надмірного збору інформації ("принцип пилососа"), зосереджуючись лише на даних, що є дійсно необхідними для аналізу та прогнозування конкретних загроз.

Юридична база використання OSINT є комплексною та залежить від юрисдикції, типу даних та методів їх збору. Ключовими аспектами є:

1. Законодавство про захист персональних даних: В Україні основним нормативно-правовим актом є Закон України "Про захист персональних даних" [1]. При обробці будь-якої інформації, що дозволяє ідентифікувати фізичну особу (ПІБ, email, IP-адреса, фото тощо), необхідно дотримуватись принципів законності, визначеності мети, мінімізації даних, точності, обмеження зберігання та забезпечення безпеки. Міжнародним орієнтиром у цій сфері є Загальний регламент про захист

даних (GDPR) Європейського Союзу, принципи якого впливають на глобальні практики обробки даних [6].

2. Авторське право та інтелектуальна власність: Автоматизований збір (скрейпінг) та використання контенту з веб-сайтів може порушувати авторські права власників контенту.

3. Умови використання ресурсів (Terms of Service): Багато веб-сайтів та онлайн-сервісів мають умови використання, які можуть забороняти або обмежувати автоматизований збір даних. Порушення цих умов може мати юридичні наслідки (хоча б у вигляді блокування доступу).

4. Законодавство про доступ до комп'ютерної інформації: Хоча OSINT використовує відкриті джерела, надмірно агресивні методи збору (напр., інтенсивне сканування, що межує з DoS-атакою) або спроби обійти базові механізми контролю доступу можуть бути розцінені як неправомірні дії.

Щодо нормативних документів та стандартів, хоча не існує єдиного міжнародного стандарту, що регулює виключно етику чи юридичні аспекти OSINT, релевантними є стандарти у сфері інформаційної безпеки та управління приватністю. Зокрема, серія стандартів ISO/IEC 27000 надає комплексну основу для систем управління інформаційною безпекою (СУІБ):

1. ISO/IEC 27001 (в Україні прийнятий як ДСТУ ISO/IEC 27001) встановлює вимоги до СУІБ, включаючи управління ризиками та захист інформаційних активів (в т.ч. зібраних OSINT-даних).

2. ISO/IEC 27002 (прийнятий як ДСТУ ISO/IEC 27002) надає каталог заходів контролю інформаційної безпеки, які можуть застосовуватися для захисту процесів OSINT.

3. ISO/IEC 27701 (прийнятий як ДСТУ ISO/IEC 27701) є розширенням ISO/IEC 27001 та ISO/IEC 27002 для управління інформацією про приватність (PIMS) і безпосередньо стосується обробки персональних даних, що є критично важливим для OSINT.

4. ISO/IEC 27005 (прийнятий як ДСТУ ISO/IEC 27005) стосується управління ризиками інформаційної безпеки, що включає оцінку ризиків, пов'язаних з OSINT-діяльністю [17].

Також загальні принципи управління ризиками, викладені в ISO 31000 (ДСТУ ISO 31000), можуть застосовуватися для оцінки та мінімізації етичних та юридичних ризиків OSINT.

Таким чином, відповідальне використання OSINT вимагає від організацій та фахівців розробки чітких внутрішніх політик та процедур, які б враховували етичні норми та вимоги чинного законодавства, зокрема Закону України "Про захист персональних даних", а також спиралися на найкращі практики, викладені у міжнародних стандартах серії ISO/IEC 27000. Необхідним є постійне навчання персоналу та забезпечення належного контролю за процесами збору та аналізу даних.

3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ OSINT: МЕТОДИКИ, ЕФЕКТИВНІСТЬ ТА ЕТИЧНІ АСПЕКТИ

3.1 Постановка задачі дослідження: Оцінка можливостей OSINT для виявлення та реагування на кіберзагрози

Метою даного у є дослідження та демонстрація практичних аспектів застосування Open-source intelligence (OSINT) у сфері кібербезпеки, зокрема для виявлення потенційних кіберзагроз та формування стратегій реагування на них. З огляду на зростаючу кількість та складність кіберінцидентів, всебічний аналіз можливостей OSINT набуває особливої актуальності. Дане дослідження спрямоване на систематизацію підходів до проведення OSINT-досліджень, аналіз ефективності конкретних інструментів та розгляд прикладів їх практичного застосування.

Об'єктом дослідження є процес виявлення кіберзагроз та реагування на них з використанням інформації з відкритих джерел. Це включає аналіз методик, інструментів та стратегій, що дозволяють своєчасно ідентифікувати потенційні загрози та мінімізувати їхній вплив.

Цільовий контекст дослідження передбачає розгляд широкого спектру кіберзагроз, які можуть бути виявлені за допомогою OSINT, включаючи:

1. Витоки конфіденційних даних (персональних, корпоративних).
2. Виявлення вразливостей у програмному забезпеченні та мережевій інфраструктурі.
3. Розслідування фішингових кампаній та інших форм соціальної інженерії.
4. Моніторинг активності зловмисників та кіберзлочинних угрупувань.
5. Аналіз інфраструктури, що використовується для проведення атак.

Обґрунтування вибору даного об'єкта дослідження:

1. Актуальність та значущість: OSINT стає все більш важливим інструментом у контексті зростаючих кіберзагроз. Здатність ефективно

використовувати відкриті джерела для виявлення та реагування на загрози є критично важливою для організацій будь-якого масштабу.

2. Доступність даних: Основна перевага OSINT полягає у використанні публічно доступної інформації, що робить його привабливим методом для дослідження без необхідності проникнення в закриті системи.

3. Практична цінність: Результати дослідження можуть бути корисними для фахівців з кібербезпеки, аналітиків загроз, а також для розробки рекомендацій щодо покращення методів використання OSINT.

4. Комплексність: Дослідження охоплює як теоретичні аспекти (методологія), так і практичне застосування (інструменти, приклади), а також етичні та правові питання.

Часовим горизонтом для аналізу прикладів та ефективності інструментів буде обрано актуальні та нещодавні випадки (за останні кілька років), що дозволить продемонструвати сучасні можливості та обмеження OSINT. Моніторинг у рамках дослідження може бути короткостроковим, але методики, що розглядаються, передбачають можливість довгострокового та безперервного застосування.

Очікуваним результатом практичного дослідження є:

1. Систематизація методики проведення OSINT-досліджень для виявлення кіберзагроз.

2. Аналіз ефективності використання конкретних OSINT-інструментів (наприклад, Shodan, Maltego, TheHarvester, SpiderFoot, Recon-ng, FOCA) на реальних або модельованих прикладах.

3. Формулювання стратегій реагування на виявлені за допомогою OSINT загрози.

4. Оцінка можливостей та обмежень OSINT у сучасному ландшафті кібербезпеки.

5. Розгляд правових та етичних аспектів використання OSINT, включаючи законодавчі обмеження та питання конфіденційності.

3.2 Вибір джерел та інструментів OSINT для практичного дослідження

Для проведення всебічного практичного дослідження можливостей OSINT у виявленні та реагуванні на кіберзагрози, необхідно обрати репрезентативний набір джерел інформації та інструментів. Вибір має враховувати їхню ефективність, доступність та релевантність до типових завдань OSINT-аналітика.

Вибір ключових OSINT-джерел: Відповідно до загальноприйнятої практики та прикладів, найпоширенішими, ключовими категоріями джерел для OSINT-дослідження є:

1. Публічні веб-ресурси та пошукові системи: Загальнодоступні веб-сайти, новинні портали, блоги, офіційні сайти організацій та урядових установ. Пошукові системи (Google, Bing, DuckDuckGo) є первинним інструментом для доступу до цієї інформації.

2. Соціальні медіа: Платформи як Facebook, Twitter, LinkedIn, Instagram, Telegram, Reddit. Є багатим джерелом інформації про осіб, організації, події, настрої та обговорення, релевантні для кібербезпеки.

3. Спеціалізовані бази даних та сервіси:

1) Бази даних вразливостей (CVE, NVD).

2) Сервіси моніторингу витоків даних (напр., Have I Been Pwned).

3) Реєстри доменних імен та IP-адрес (WHOIS-сервіси).

4. Dark/Deep Web форуми та торгові майданчики: Для виявлення інформації про продаж експлойтів, викрадених даних, інструментів для атак.

5. Технічні блоги та звіти компаній з кібербезпеки: Для отримання детальної інформації про нові загрози, TTPs та методи захисту.

6. Публічні репозиторії коду (GitHub, GitLab): Для аналізу вихідного коду на предмет вразливостей або інструментів, що можуть бути використані у зловмисних цілях.

7. Метадані файлів: Для вилучення прихованої інформації з документів, зображень тощо.

Вибір інструментів OSINT та аналіз їх переваг і недоліків: На основі огляду інструментів, наведеного у теоретичних ах , та їхньої доцільності для вирішення типових OSINT-задач, для практичного дослідження обрано наступні інструменти. Їх вибір також корелює з основними найпоширенішими проблемами .

Maltego є програмним інструментом, призначеним для візуалізації даних та аналізу зв'язків, який знаходить широке застосування у розслідуваннях у сфері кібербезпеки та цифровій криміналістиці [21]. Ключовою функціональністю платформи є можливість агрегації та графічного представлення взаємозв'язків між різноманітними інформаційними об'єктами. Це досягається шляхом збору даних з численних відкритих джерел, що включають інформацію про доменні імена, IP-адреси, профілі в соціальних мережах, а також інші релевантні онлайн-ресурси. Таким чином, Maltego функціонує як інструментарій для розвідки на основі відкритих джерел (OSINT).

Програмне забезпечення Maltego забезпечує користувачам можливість проведення інтерактивного аналізу даних, використовуючи розширені можливості візуалізації для ефективного дослідження існуючих зв'язків між об'єктами. Цей інструмент використовується для проведення онлайн-досліджень, спрямованих на виявлення та аналіз взаємозв'язків між інформаційними фрагментами, отриманими з різних інтернет-джерел. Платформа дозволяє виявляти приховані або неочевидні зв'язки між фізичними особами, організаціями, а також агрегувати загальнодоступну інформацію, що стосується об'єктів дослідження.

Maltego позиціонується як комплексний фреймворк для OSINT, який дозволяє систематизувати процеси збору, візуалізації та аналізу інформації. Важливою особливістю є здатність генерувати графічні представлення (графи), що наочно демонструють зв'язки між різними типами сутностей. До таких сутностей належать особи, компанії, організації, інфраструктурні елементи, геолокаційні дані та інші

об'єкти, що становлять інтерес у рамках конкретного розслідування чи аналітичного завдання.

Maltego є інструментом, що надає можливості для збору, структурування та візуального аналізу інформації з відкритих джерел, що робить його корисним компонентом у процесах цифрової розвідки та розслідувань. Його здатність представляти складні взаємозв'язки у графічному форматі сприяє більш глибокому розумінню досліджуваних ситуацій та виявленню неочевидних патернів.

Переваги та можливості Maltego

1. Візуалізація даних: Maltego дозволяє створювати графічні схеми, що показують зв'язки між об'єктами, що значно спрощує процес аналізу великих обсягів даних.

2. Інтеграція: Підтримує інтеграцію з різними джерелами даних та іншими інструментами, що розширює його функціональні можливості.

3. Кастомізація: Користувачі можуть налаштовувати інструмент під свої потреби, створюючи власні трансформації для збору специфічних даних.

4. Використовується для розслідувань у сфері кібербезпеки.

5. Дозволяє аналізувати посилання.

6. Може виявляти зв'язки між людьми та компаніями.

7. Знаходить загальнодоступну інформацію.

8. Може бути використаний для збору інформації з соціальних мереж, таких як LinkedIn, Facebook, Twitter, і допомагає розкривати зв'язки між користувачами.

9. Використовується для розслідування фішингових атак. Дозволяє виявити зв'язки між підозрілими листами та раніше відомими фішинговими кампаніями. Інструмент дозволяє створити схеми, що показали джерела цих атак і допомогли ідентифікувати зловмисників.

10. Може бути використаний для пошуку вразливостей електронних ресурсів.

11. Застосовується для аналізу зв'язків між різними компонентами системи та виявлення потенційних вразливостей.

Недоліки Maltego

1. Висока вартість: Ліцензія на повну версію Maltego може бути досить дорогою, що може бути перешкодою для малих компаній або індивідуальних користувачів.

2. Складність в навчанні: Інструмент має досить складний інтерфейс, що може вимагати значного часу для освоєння.

Shodan є пошуковою системою, спеціалізованою на виявленні та індексації пристроїв, підключених до мережі Інтернет [2, 13]. Її функціональність базується на безперервному скануванні глобального адресного простору з метою ідентифікації різноманітних мережевих вузлів, включаючи компоненти Інтернету речей (IoT), сервери, маршрутизатори, промислові системи управління (ICS) та інші пристрої, що мають мережевий інтерфейс.

Функціонал Shodan надає можливість ідентифікувати не лише типи підключених пристроїв, але й збирати детальну інформацію про них. До таких даних належать відомості про відкриті порти, версії програмного забезпечення, що використовуються, конфігураційні параметри, дані SSL/TLS сертифікатів та інші технічні характеристики. Особлива увага приділяється збору інформації про потенційно вразливі пристрої та сервіси, такі як веб-камери з дефолтними налаштуваннями або незахищені промислові контролери.

Застосування Shodan знаходить своє відображення у сфері кібербезпеки, зокрема для виявлення вразливих систем, аналізу поверхні атаки, ідентифікації неправильно налаштованих пристроїв та потенційних витоків конфіденційної інформації. Інструмент дозволяє проводити оцінку загального рівня безпеки досліджуваних мережевих активів шляхом аналізу публічно доступних даних про їхню конфігурацію та поведінку.

Shodan класифікується як інструмент для пасивного збору інформації та розвідки на основі відкритих джерел (OSINT). Користувачі взаємодіють з попередньо зібраною та проіндексованою базою даних, що мінімізує пряму взаємодію з цільовими

системами на етапі початкового збору даних. Це дозволяє проводити аналіз вразливостей та збір розвідувальної інформації з меншим ризиком виявлення. Таким чином, Shodan надає доступ до великого обсягу даних про глобальну інтернет-інфраструктуру, що є цінним ресурсом для аналітичних та дослідницьких завдань у сфері інформаційної безпеки.

Переваги та можливості Shodan:

1. Сканування мережі: Shodan дозволяє сканувати мережі на наявність підключених пристроїв.
2. Виявлення вразливостей: Інструмент допомагає ідентифікувати вразливі пристрої, що можуть бути потенційними цілями для атак. Він також виявляє відкриті порти та інші конфігурації.
3. Моніторинг мережі: Shodan забезпечує можливість моніторингу стану мережі та виявлення змін у конфігурації пристроїв.
4. Надає інформацію про різні типи підключених пристроїв, включаючи промислові системи та веб-камери.
5. Дозволяє знаходити конфіденційну інформацію, витоки даних та вразливості.
6. Демонструє ефективність у швидкому скануванні мережі та наданні точної інформації про конфігурацію пристроїв.
7. Використовується для пасивного сканування та виявлення вразливостей.
8. Наводиться як приклад успішного використання для виявлення вразливих пристроїв у мережі фінансової установи, а також для виявлення вразливості відеокамери.
9. Може бути використаний для захисту від атак, наприклад, DDoS-атак.
10. Допомагає встановити джерело загроз, наприклад, визначити IP-адресу відправника та провайдера.

Недоліки Shodan:

1. Обмеженість даних: Shodan збирає дані тільки з відкритих портів. Це означає, що він може не виявити всі потенційні загрози в мережі.

2. Необхідність технічних знань: Для ефективного використання Shodan користувачі повинні мати досить високий рівень технічних знань та розуміння мережевих технологій.

Платформа Censys являє собою спеціалізовану пошукову систему, призначену для моніторингу та комплексного аналізу стану глобальної мережі Інтернет [10]. Основним завданням Censys є систематичне сканування адресного простору Інтернету для ідентифікації активних пристроїв та сервісів. Процес сканування охоплює широкий спектр мережевих вузлів, включаючи веб-сервери, маршрутизатори, пристрої Інтернету речей (IoT), компоненти промислових систем управління (ICS/SCADA) та інші. Для реалізації цих завдань платформа використовує спеціалізовані інструменти, такі як ZMap для широкомасштабного виявлення відкритих портів та ZGrab2 для збору деталізованої інформації про сервіси, що функціонують на виявлених портах.

Агреговані дані надають дослідникам можливість аналізувати великий масив технічної інформації. Цей масив включає дані про відкриті порти, детальну інформацію про SSL/TLS сертифікати (емітент, термін дії, криптографічні алгоритми), а також конфігураційні параметри пристроїв, зокрема банери сервісів, що часто розкривають версії програмного забезпечення, HTTP-заголовки та іншу релевантну метадані. Вся зібрана інформація індексується та стає доступною для аналізу через веб-інтерфейс та програмний інтерфейс додатків (API), що забезпечує можливість виконання складних пошукових запитів на основі специфічних технічних атрибутів.

Застосування Censys має значну цінність у контексті виявлення потенційних вразливостей та ідентифікації активних сервісів на мережевих вузлах. Фахівці в галузі кібербезпеки використовують платформу для ідентифікації систем із застарілим

програмним забезпеченням, помилковими конфігураціями (наприклад, сервіси зі стандартними обліковими даними, незахищені бази даних) та для оцінки стану криптографічних параметрів. Хоча дані, надані Censys, можуть бути використані зловмисниками для розвідки та вибору цілей, для фахівців з інформаційної безпеки та тестувальників на проникнення цей інструмент є важливим засобом для проактивного виявлення ризиків, аналізу поверхні атаки та контролю відповідності встановленим політикам безпеки.

Censys функціонує як інструмент розвідки на основі відкритих джерел (OSINT), оскільки агреговані ним дані збираються з публічно доступних сегментів мережі Інтернет. З перспективи кінцевого користувача, взаємодія з платформою кваліфікується як пасивний збір інформації, оскільки запити спрямовуються до попередньо зібраної бази даних Censys, а не безпосередньо до цільових систем. Сама ж платформа здійснює активне, хоча й розподілене, сканування для наповнення цієї бази. Такий підхід дозволяє проводити дослідження, мінімізуючи ризик виявлення з боку об'єктів аналізу. Важливою особливістю є обробка значних обсягів даних та збереження історичних зрізів, що уможливлює аналіз динаміки змін у конфігураціях та стані безпеки інтернет-ресурсів у часі.

Переваги та можливості Censys:

1. Глобальне сканування Інтернету: Censys регулярно сканує Інтернет, збираючи дані про підключені пристрої та їх конфігурації.

2. Аналіз безпеки: Інструмент дозволяє проводити аналіз безпеки мережевих пристроїв, виявляти вразливості та відстежувати їх усунення. Він допомагає виявити відкриті порти та сервіси, які можуть бути використані зловмисниками для атак, дозволяючи своєчасно закрити або захистити ці вразливі місця.

3. Пошук даних: Censys надає можливість пошуку інформації за різними параметрами, що дозволяє знаходити конкретні пристрої чи конфігурації.

4. Виявлення підозрілої активності: Використовується для виявлення підозрілих з'єднань з зовнішніми IP-адресами, що можуть свідчити про зловмисні дії. Також допомагає виявити шкідливе програмне забезпечення.

5. Використання у реальних сценаріях: Застосовувався фінансовою установою для виявлення вразливих IoT-пристроїв, які могли бути використані для DDoS-атак. IT-компанія використовувала Censys для виявлення шкідливого програмного забезпечення на своїх серверах завдяки ідентифікації підозрілих з'єднань.

6. Аналіз сертифікатів SSL/TLS: Дозволяє аналізувати сертифікати SSL/TLS.

7. Ефективність: Демонструє ефективність у виявленні аномалій та шкідливого програмного забезпечення.

Недоліки Censys:

1. Висока вартість преміум-версії: Повна функціональність Censys доступна тільки у платній версії, що може бути дорогим для малих організацій.

2. Відсутність деяких даних: Censys може не охоплювати всі пристрої та мережі, оскільки його сканування залежить від багатьох факторів, включаючи географічне розташування та типи мереж.

TheHarvester є одним з інструментів, який використовується в Open Source Intelligence (OSINT) розвідці [3]. Це інструмент, розроблений на Python. Він широко використовується для розвідки у сфері кібербезпеки та підготовки до тестування на проникнення.

Суть TheHarvester полягає в автоматизованому зборі даних з різних відкритих джерел. Він призначений для збору конкретних типів інформації, таких як електронні адреси, субдомени, хости, імена співробітників та відкриті порти. TheHarvester збирає ці дані з різних загальнодоступних джерел, включаючи пошукові системи, соціальні мережі, DNS-сервери, сервери ключів PGP та базу даних SHODAN. Таким чином, його основна функція зосереджена на автоматизації процесу збору інформації для

швидкого отримання необхідних даних. Він є інструментом, орієнтованим в основному на збір даних, а не на їх розширений аналіз.

Переваги TheHarvester (Можливості):

1. Збирає дані з різних джерел, таких як пошукові системи, соціальні мережі, DNS-сервери.
2. Може збирати дані про домени, електронні адреси, IP-адреси та інші мережеві ресурси.
3. Дозволяє зібрати субдомени, хости, імена співробітників та відкриті порти.
4. Має можливість інтеграції з іншими інструментами та платформами для розширення своїх можливостей.
5. Автоматизує процес збору інформації, що дозволяє швидко отримувати необхідні дані.
6. Використовується для розвідки у сфері кібербезпеки та підготовки до тестування на проникнення.
7. Згадується як інструмент, що може бути використаний для виявлення витоку конфіденційних даних через соціальні медіа та для моніторингу соціальних медіа та форумів, де зловмисники можуть обговорювати плани атак.

Недоліки TheHarvester:

1. Обмеженість джерел: Інструмент обмежується тими джерелами, з яких він може збирати дані, тому деяка інформація може бути пропущена.
2. Відсутність аналізу: Інструмент в основному орієнтований на збір даних і не надає розширених можливостей для їх аналізу, тому отримані дані потребують додаткової обробки.

На рисунку 3.1 показаний графічний інтерфейс програми TheHarvester у середовищі kalilinux , воно дає змогу побачити поточну версію застосунку та основні використання команд.

Рисунок 3.1 – Графічний інтерфейс TheHarvester у середовищі kalilinux

SpiderFoot є програмним інструментом, призначеним для автоматизованого збору та аналізу даних з відкритих джерел (OSINT). Його функціональність дозволяє здійснювати комплексний аналіз потенційних кіберзагроз шляхом агрегації інформації з численних публічно доступних ресурсів. Інструмент часто характеризується як засіб для проведення розширеної "дактилоскопії" цифрових активів, завдяки інтеграції з великою кількістю OSINT-джерел.

Ключовою особливістю SpiderFoot є його здатність автоматично взаємодіяти з понад сотнею різноманітних відкритих джерел даних, збираючи та структурувати релевантну інформацію. До таких джерел можуть належати пошукові системи, соціальні мережі, бази даних доменних імен, репозиторії програмного коду, сервіси аналізу шкідливого ПЗ та інші. Основною метою розробки даного інструменту є суттєве спрощення та автоматизація процесу розвідки та збору інформації.

Застосування SpiderFoot охоплює завдання моніторингу та аналізу даних, отриманих з відкритих джерел. Це включає виявлення потенційних вразливостей, пов'язаних з досліджуваними об'єктами, а також контроль за можливими витоками конфіденційної інформації. Інструмент надає можливість візуалізації зібраних даних та зв'язків між ними, що сприяє більш ефективному аналізу та прийняттю рішень у контексті інформаційної безпеки.

SpiderFoot є багатофункціональною платформою для OSINT, що автоматизує збір та первинну обробку даних з великої кількості джерел. Його використання дозволяє фахівцям з кібербезпеки отримувати більш повне уявлення про досліджувані об'єкти та потенційні вектори атак, що є важливим елементом у процесах розвідки загроз та управління ризиками.

Переваги та можливості SpiderFoot:

1. Автоматизований збір даних: SpiderFoot автоматизує збір даних з багатьох джерел, що робить процес швидшим та ефективнішим.
2. Широке охоплення джерел: Він запитує інформацію з понад 100 відкритих джерел.
3. Модульна структура: Інструмент підтримує модулі для різних видів даних, що робить його дуже гнучким і настроюваним під конкретні потреби.
4. Комплексний аналіз: Дозволяє проводити комплексний аналіз зібраної інформації для виявлення та реагування на кіберзагрози.
5. Виявлення вразливостей: Використовується для сканування мереж та ідентифікації критичних вразливостей, як у прикладі з державною установою.
6. Моніторинг витоків даних: Допомагає виявляти витік конфіденційних даних, зокрема через соціальні медіа, та швидко збирати інформацію про джерело витоку.
7. Виявлення інформації про активи: Збирає інформацію про IP-адреси, доменні імена, веб-сервери, адреси електронної пошти тощо.
8. Підтримка візуалізації: Надає можливість візуалізації знайдених результатів.
9. Проактивне виявлення загроз: Сприяє своєчасному виявленню потенційних загроз та запобіганню інцидентам.
10. Ефективність у реальних випадках: Демонструє високу ефективність у різних сценаріях реагування на кіберзагрози.

Недоліки SpiderFoot:

1. Складність налаштування: Для ефективного використання може вимагатися складне налаштування та конфігурація, що потребує технічних

2. Проблеми з точністю даних: Деякі модулі можуть надавати неточні або неповні дані, що може впливати на результати розвідки.

SpiderFoot є потужним, але складним в налаштуванні інструментом OSINT, який автоматизує збір даних з великої кількості джерел для цілей кібербезпеки, але вимагає уваги до якості отриманої інформації.

На рисунку 3.2 показаний графічний інтерфейс програми SpiderFoot у середовищі командної строки kali linux , яка дає змогу ознайомитися з основними командами.

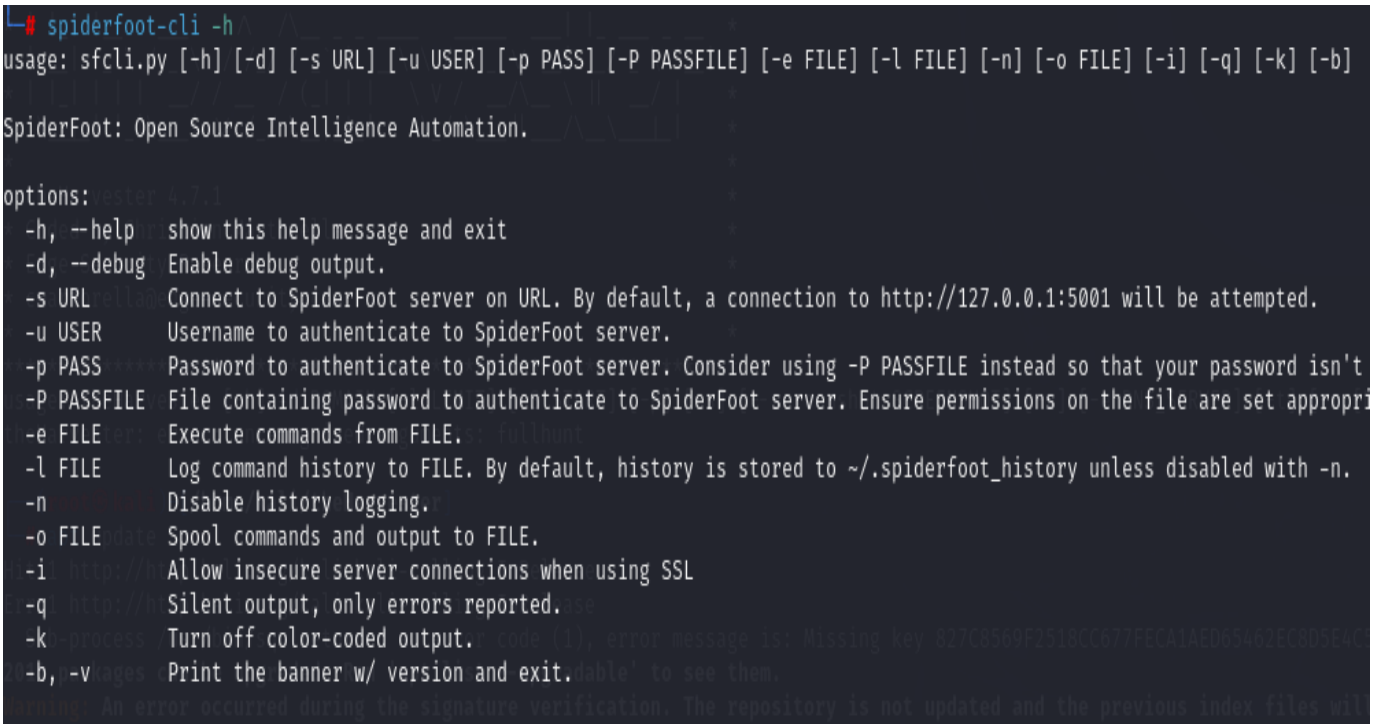

Рисунок 3.2 – Інтерфейс SpiderFoot у середдовищі kali linux

Програмний комплекс FOCA (Fingerprinting Organizations with Collected Archives) є спеціалізованим інструментом, призначеним для автоматизованого

вилучення, аналізу та візуалізації метаданих з різноманітних електронних документів. Його ключова функція полягає у виявленні прихованої інформації, яка ненавмисно залишається у файлах під час їх створення чи редагування. FOCA підтримує обробку широкого спектра файлових форматів, зокрема документів Microsoft Office, Open Office, PDF, а також графічних файлів та інших поширених типів, що дозволяє проводити комплексний аналіз цифрових активів.

Застосування FOCA в процесі обробки та аналізу зібраних даних дозволяє отримати значний обсяг цінних відомостей. Інструмент здійснює глибокий аналіз метаданих, виявляючи таку інформацію, як імена користувачів, назви комп'ютерів, версії операційних систем та програмного забезпечення, що використовувалися для роботи з документом, шляхи до файлів на сервері або локальному диску, дати створення та останньої модифікації, інформацію про принтери та, в окремих випадках, геолокаційні дані. Ця деталізована інформація, часто невидима для звичайного користувача, може бути критично важливою для ідентифікації потенційних загроз та вразливостей. Наприклад, виявлення застарілих версій програмного забезпечення може свідчити про наявність відомих експлойтів, а інформація про користувачів та структуру каталогів – розкривати внутрішню інфраструктуру та технологічний стек організації.

Окрім аналізу локально збережених файлів, FOCA має функціонал для пошуку документів, що належать певній організації, у публічному інтернет-просторі за допомогою інтеграції з пошуковими системами. Після виявлення та завантаження таких документів інструмент проводить їх комплексний аналіз метаданих. Цей процес дозволяє сформувати уявлення про цифрові "відбитки" організації, виявити потенційні вектори атаки або ненавмисні витоки конфіденційної інформації. Таким чином, FOCA слугує потужним засобом для рекогносцировки та збору розвідувальних даних, а результати його роботи є важливою основою для подальшого аналізу безпеки та розробки заходів щодо мінімізації виявлених ризиків [4].

Переваги та можливості FOCA:

1. Аналіз метаданих: FOCA дозволяє аналізувати метадані в документах для виявлення прихованої інформації.

2. Підтримка різних форматів: Інструмент підтримує роботу з багатьма форматами документів, включаючи PDF, DOC, XLS та інші.

3. Виявлення прихованих вразливостей: Ефективний у виявленні прихованих вразливостей у документах.

4. Використання у реальних випадках: Був успішно використаний юридичною фірмою для виявлення конфіденційної інформації в документах та вжиття заходів для запобігання витокам даних. Також допоміг виявити внутрішнього зловмисника та упровадити нові політики безпеки після аналізу метаданих викрадених файлів.

Недоліки FOCA:

1. Обмежена функціональність: FOCA спеціалізується лише на аналізі метаданих, тому його можливості обмежені порівняно з іншими інструментами OSINT.

2. Необхідність додаткової обробки даних: Отримані дані потребують додаткової обробки для використання у комплексних розслідуваннях.

Recon-ng є модульною платформою для розвідки на основі відкритих джерел (OSINT), призначеною для автоматизованого збору та первинної систематизації даних [5]. Його ключова функціональність реалізована через систему незалежних модулів, кожен з яких орієнтований на взаємодію з конкретними джерелами інформації або застосування специфічних технік збору. Така архітектура дозволяє користувачеві гнучко конфігурувати інструмент для вирішення конкретних завдань розвідки шляхом вибору та активації релевантних модулів.

Основне призначення Recon-ng полягає в автоматизації процесу збору даних з різноманітних відкритих джерел, серед яких пошукові системи, соціальні медіа, DNS-

сервери, репозиторії коду та публічні бази даних. Інструмент забезпечує вилучення інформації, її структурування у внутрішній базі даних для подальшого аналізу та кореляції. Важливою функцією є можливість генерації звітів про зібрані дані у різних форматах, що полегшує їх подальшу обробку та використання для оцінки цифрового сліду, виявлення потенційних загроз та прийняття рішень в рамках завдань інформаційної безпеки.

Переваги (Можливості) Recon-ng:

1. Модульна структура: Має модульну архітектуру, що дозволяє користувачам додавати різні модулі для збору та аналізу даних залежно від конкретних потреб. Ця гнучкість робить його дуже універсальним інструментом для розвідки.

2. Автоматизація збору даних: Автоматизує процес збору даних з різноманітних джерел, таких як пошукові системи, соціальні медіа, DNS-сервери та інші.

3. Звітність: Дозволяє генерувати детальні звіти про зібрані дані, що є корисним для подальшого аналізу та прийняття рішень.

4. Ефективність у реагуванні на загрози: Використання модулів для збору даних з відкритих джерел дозволяє зібрати інформацію про зловмисників та їхні зв'язки з іншими подібними атаками, що допомагає організації зміцнити систему безпеки та запобігти майбутнім атакам. Це підкреслює ефективність Recon-ng у проактивному моніторингу та аналізі загроз.

Недоліки Recon-ng:

1. Потреба в налаштуванні: Для максимальної ефективності інструменту потрібні початкові налаштування, що можуть вимагати певних технічних знань.

2. Обмежена база даних: Як і інші інструменти, Recon-ng залежить від доступності та якості даних у відкритих джерелах, що може обмежувати його можливості у певних випадках.

На рисунку 3.3 показаний графічний інтерфейс програми Recon-ng у середовищі командної строки kalilinux , воно дає змогу побачити поточну версію застосунку.

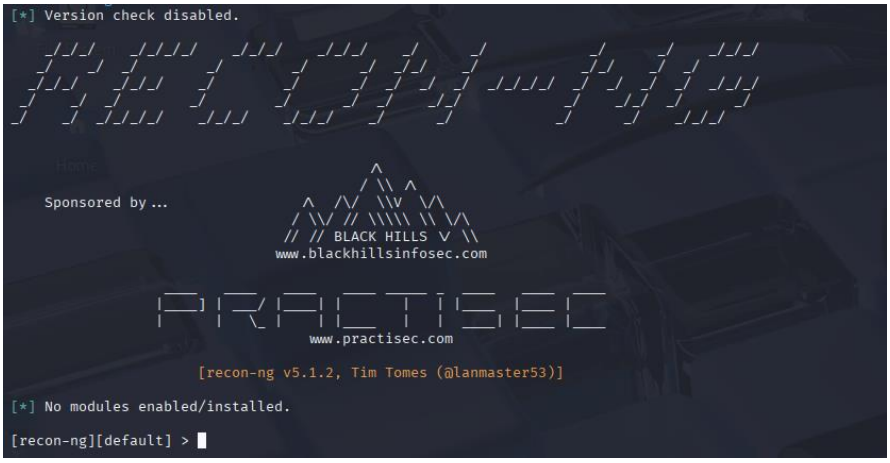

Рисунок 3.3 – Графічний інтерфейст Recon-ng у середовищі kali linux

OSINT Framework є веб-сайтом, створеним ентузіастом Джастіном Нордінормом, який слугує визначним ресурсом для слідчих OSINT для вивчення інструментів. Він є одним з найпопулярніших інструментів у цій сфері [20].

OSINT Framework надає зручний перегляд та доступ до різних інструментів та ресурсів, які можуть бути корисними в процесі OSINT розвідки. Він містить велику кількість категорій інструментів та джерел інформації, що охоплюють різні аспекти OSINT. До цих категорій належать соціальні мережі, пошукові системи, блоги, форуми, географічні ресурси, медіа-платформи, інструменти аналізу та інше. Кожна категорія містить посилання на відповідні інструменти, ресурси та сервіси для отримання інформації.

Фреймворк також надає корисні посилання на спеціалізовані інструменти для пошуку за електронною поштою, IP-адресами, телефонними номерами, хеш-значеннями та іншими важливими елементами OSINT розвідки. Він також містить посилання на інструменти для виявлення зображень, аналізу тексту, моніторингу соціальних мереж та багато іншого.

Основні функції OSINT Framework включають:

1. Централізований доступ: він збирає різні інструменти та ресурси в одному місці, що дозволяє ефективно організувати та керувати процесом OSINT розвідки.

2. Оновлення та розширення: OSINT Framework постійно оновлюється новими інструментами та ресурсами відповідно до змінюваних потреб OSINT розвідки. Це дозволяє користувачам отримувати доступ до найновіших інструментів та методик.

3. Систематизація інформації: він розподіляє інструменти та ресурси за категоріями, що допомагає користувачам орієнтуватися та швидко знаходити потрібні інструменти.

4. Ефективність: використання OSINT Framework дозволяє заощаджувати час та зусилля, оскільки користувачам не потрібно витрачати час на пошук окремих інструментів та ресурсів. Вони можуть швидко знайти все необхідне на одному веб-сайті.

На рисунку 3.4 показаний графічний інтерфейс інструменту OSINT Framework , що дає змогу ознайомитися з його основними функціональними розгалуженнями для пошуку потрібних сервісів та інструментів.

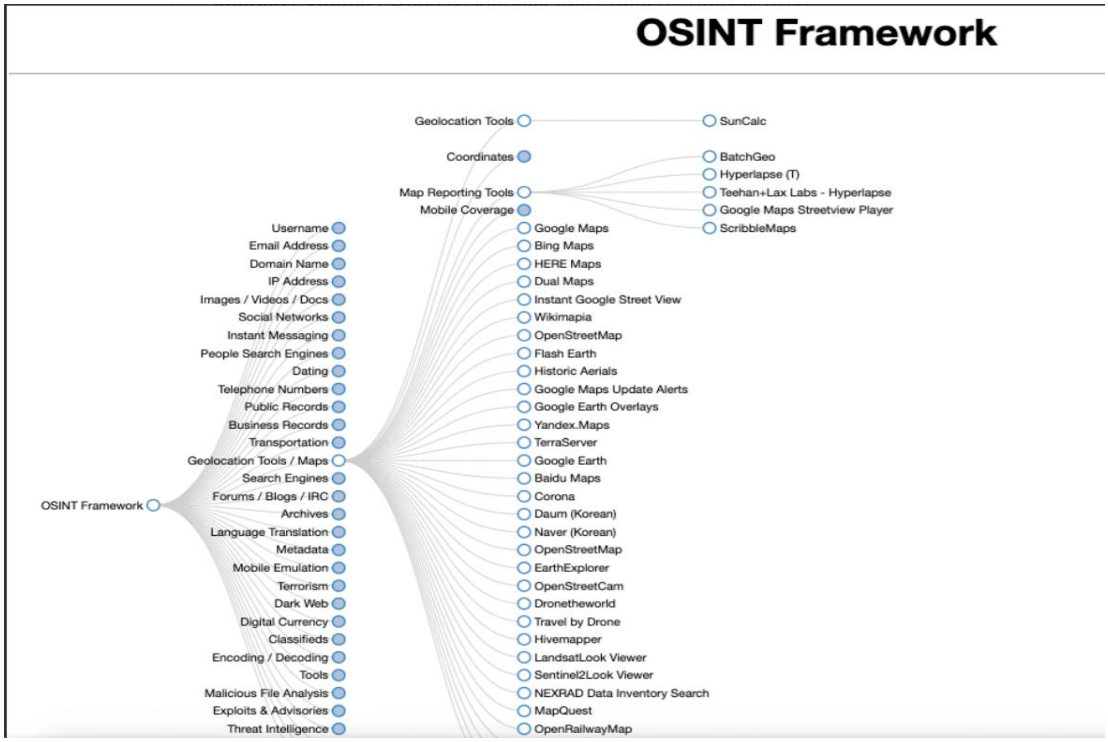

Рисунок 3.4 – Сторінка OSINT Framework

Дані українських державних установ за 2025 рік окреслюють пріоритетні напрямки для OSINT-прогнозування кібератак. Департамент кіберполіції фіксує домінування фішингу, онлайн-шахрайства та атак на вебсайти й бази даних. CERT-UA підтверджує значне зростання інцидентів (на 69,8%), з переважанням розповсюдження шкідливого ПЗ, фішингу та компрометацій систем, спрямованих на викрадення чи знищення даних [14, 16]. Ця статистика визначає ключові індикатори (нові фішингові домени, обговорення вразливостей, поява специфічного ПЗ), моніторинг яких через OSINT є критичним для превентивного виявлення ознак підготовки атак.

Переваги:

1) Обґрунтованість та релевантність: Дані від офіційних джерел, таких як Кіберполіція та CERT-UA, відображають реальну картину загроз у конкретному регіоні/країні, дозволяючи сфокусувати OSINT-зусилля на найбільш актуальних векторах атак.

2) Визначення пріоритетних індикаторів: Статистика допомагає ідентифікувати типові інструменти, тактики та цілі зловмисників, що, своєю чергою, дозволяє визначити конкретні індикатори для моніторингу через OSINT (наприклад, типи шкідливого ПЗ, фішингові шаблони).

3) Підвищення ефективності OSINT-розвідки: Замість широкого та нецільового пошуку, аналітики можуть концентруватися на зборі інформації, пов'язаної з уже відомими та поширеними загрозами, збільшуючи шанси на виявлення підготовчих дій.

Недоліки:

1) Реактивний характер даних: Статистика відображає вже реалізовані інциденти, тому прогнозування на її основі може мати обмеження щодо виявлення абсолютно нових або нетипових загроз, які ще не потрапили у звіти.

2) Затримка в отриманні інформації: Офіційні звіти публікуються з певною періодичністю, тому найсвіжіші тенденції можуть не бути відображені оперативно, що критично для динамічного середовища кіберзагроз.

3) Узагальненість даних: Публічні звіти часто надають узагальнену інформацію без глибокої технічної деталізації, яка може бути необхідна для тонкого налаштування OSINT-інструментів та специфічного пошуку.

4) Ризик "тунельного бачення": Надмірна концентрація на відомих загрозах, підтверджених статистикою, може призвести до ігнорування менш очевидних, але потенційно небезпечних векторів атак, які не потрапили до офіційних зведень.

У таблиці 3.1 наведена узагальнена порівняльна характеристика розглянутих раніше інструментів

Таблиця 3.2.1 – Переваги та недоліки обраних OSINT-інструментів

Інструмент	Переваги	Недоліки
Maltego	Потужна візуалізація зв'язків, велика кількість трансформацій, інтеграція, кастомізація.	Висока вартість повної версії, складність в освоєнні.

Shodan / Censys	Виявлення підключених пристроїв, відкритих портів, вразливостей. Регулярне сканування Інтернету (Censys).	Обмеженість даних (Shodan), необхідність технічних знань, вартість преміум-версій, неповнота охоплення (Censys).
TheHarvester	Автоматизація збору email, субдоменів; інтеграція з різними джерелами.	Обмеженість джерел, орієнтація на збір, а не аналіз.
SpiderFoot	Автоматизований збір з багатьох джерел, модульна структура, гнучкість.	Складність налаштування, можлива неточність деяких модулів.
FOCA	Ефективний аналіз метаданих документів.	Обмежена функціональність, дані потребують додаткової обробки.
Recon-ng	Модульність, автоматизація збору даних з веб-джерел, генерація звітів.	Потреба в налаштуванні, залежність від якості відкритих даних.
OSINT Framework	Зображує загальну картину інцидентів , та структурує інструменти	Вимагає навичок роботи за OSINT даними, часозатратність.
Укр. держ. ресурси	Авторитетність, релевантність для України, оперативність.	Може бракувати глибоких технічних деталей.

Вибір конкретного набору інструментів для кожного етапу OSINT-дослідження буде залежати від специфіки завдання. Пріоритет надаватиметься комплексному використанню декількох інструментів для перехресної перевірки та збагачення даних, а також аналізу інформації з офіційних українських джерел для забезпечення контекстуальної релевантності.

3.3 Процес збору даних: Опис методики, приклад зібраних даних

Етап збору даних передбачає безпосереднє отримання інформації з раніше обраних джерел із залученням спеціалізованих OSINT-інструментів. До основних процедур належить моніторинг соціальних медіа та форумів, де потенційно обговорюються кіберзагрози, з використанням автоматизованих засобів збору даних. Процес сканування мережевих пристроїв реалізується за допомогою таких інструментів, як Shodan та Censys, що дозволяє виявляти підключені до глобальної

мережі пристрої та їхні потенційні уразливості. Агрегація відомостей про домени та IP-адреси здійснюється із застосуванням theHarvester, який забезпечує збір даних про доменні імена, електронні пошти, IP-адреси та інші мережеві атрибути. Для аналізу метаданих, що містяться в документах і можуть нести цінну інформацію про потенційні загрози, використовується інструмент FOCA.

Методологія проведення OSINT-досліджень являє собою багатогранний процес, що охоплює планування, збір, обробку й аналіз даних, а також підготовку звітів та реагування на виявлені загрози. Застосування сучасного OSINT-інструментарію, такого як Maltego, Shodan, Censys, theHarvester, SpiderFoot, FOCA та Recon-ng, забезпечує ефективне виявлення кіберзагроз та оперативне реагування на них, що сприяє підтримці надійного захисту інформаційних систем.

Приклад збору даних за допомогою сервісу Shodan . У цьому сервісі вже є приклади команд які можна виконати . Одна з таких є шпаргалка з фільтрами Shodan наразі сканує майже 4000 портів в Інтернеті. Ось кілька найпоширеніших фільтрів пошуку.

Основні фільтри Shodan:

1. `country:"UA"`: Пошук пристроїв в Україні. Це базовий фільтр для географічного обмеження дослідження.
2. `city:"Kyiv"`: Пошук пристроїв у конкретному місті (наприклад, Києві).
3. `org:"[Назва організації]"`: Пошук пристроїв, що належать конкретній організації (наприклад, інтернет-провайдеру або компанії).
4. `asn:[AS_номер]`: Пошук за номером автономної системи (ASN), що дозволяє ідентифікувати пристрої в мережі конкретного провайдера.
5. `port:"[номер_порту]"`: Пошук пристроїв з відкритим конкретним портом (наприклад, port:"22" для SSH, port:"3389" для RDP, port:"80" для HTTP).
6. `product:"[назва_продукту]"`: Пошук пристроїв з конкретним програмним забезпеченням (наприклад, product:"Apache httpd", product:"Microsoft IIS", product:"OpenSSH").

7. `version:"[версія_продукту]"`: Пошук конкретної версії програмного забезпечення (часто використовується разом з product).

8. `os:"[операційна_система]"`: Пошук пристроїв з певною операційною системою (наприклад, os:"Windows Server 2012").

9. `hostname:"[шаблон_імені_хоста]"`: Пошук за частиною імені хоста (наприклад, hostname:".gov.ua" для державних сайтів).

10. `vuln:[CVE-ID]`: Пошук пристроїв, для яких Shodan ідентифікував відому вразливість за її CVE-ідентифікатором (наприклад, vuln:CVE-2021-44228 для Log4Shell).

11. `has_screenshot:true`: Пошук пристроїв, для яких Shodan зробив скріншот (часто це веб-інтерфейси, RDP, VNC).

12. `http.title:"[заголовок_сторінки]"`: Пошук веб-сторінок за їх заголовком.

13. `ssl:"[фрагмент_сертифікату]"`: Пошук за інформацією в SSL-сертифікаті (наприклад, назва організації в сертифікаті).

14. `tag:[тег]`: Пошук за тегами, які Shodan присвоює пристроям (наприклад, tag:ics для промислових систем управління, tag:database для баз даних).

На рисунку 3.5 наведено приклад використання запитів у ресурсі Shodan.

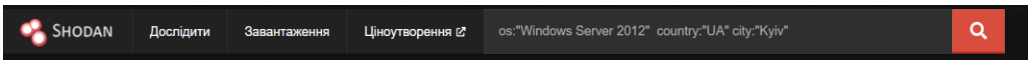

Рисунок 3.5 – Приклад використання комбінування фільтрів

Використані фільтри :

1. `os:"Windows Server 2012"`
2. `country:"UA"`
3. `city:"Kyiv"`

Результат виконання пошуку наведено на Рисунку 3.6.

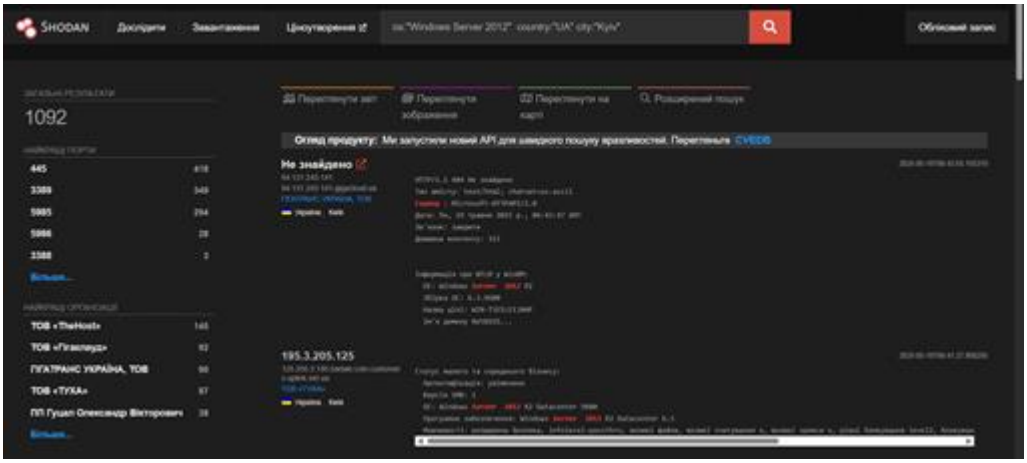

Рисунок 3.6 – Результат використання комбінування фільтрів

На рисунку 3.7 представлено загальну інформацію про хост з IP-адресою 195.3.205.125, зібрану за допомогою сервісу Shodan, включаючи його домен, країну (Україна), місто (Київ), організацію (ТОВ «ТУХА») та операційну систему (Windows Server 2012 R2). Також відображено перелік відкритих портів, зокрема 135, 445 та 3389, що вказують на активні мережеві сервіси.

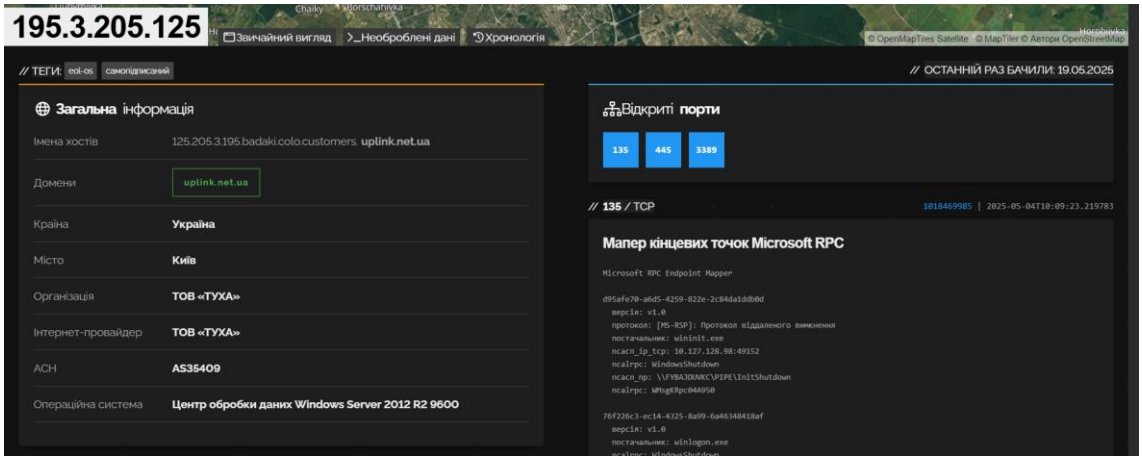

Рисунок 3.7 – Результат інформації за хостом 195.3.205.125

На рисунку 3.8 показано детальну інформацію про сервіс, що працює на порту 445/TCP хоста 195.3.205.125, зокрема стан SMB (автентифікація увімкнена) та використання застарілої версії протоколу SMBv1. Також підтверджується, що

операційна система хоста – Windows Server 2012 R2 Datacenter, що вказує на потенційні вразливості, пов'язані з цим протоколом та ОС.

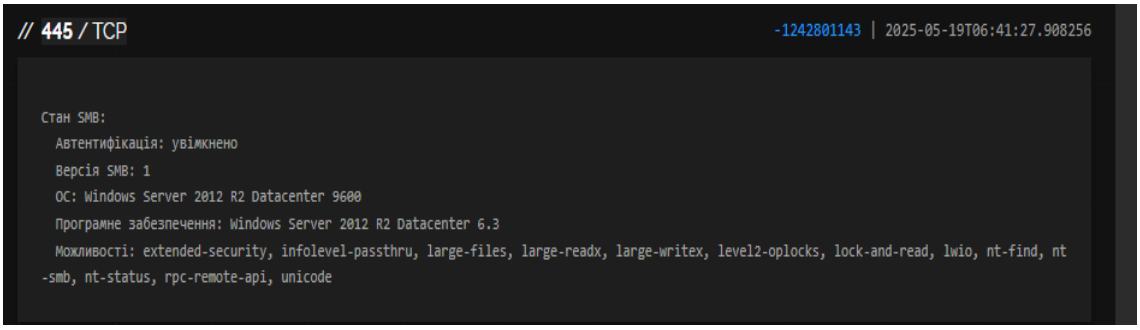

Рисунок 3.8 – Результат інформації 445/TCP за хостом 195.3.205.125

На рисунку 3.9 показана інформація про відкритий порт 3389/TCP, що відповідає протоколу віддаленого робочого столу (RDP) на хості 195.3.205.125. Зазначено, що операційна система, асоційована з RDP, є Windows 8.1/Windows Server 2012 R2, а також наведено ідентифікатори NetBIOS та DNS, що можуть бути використані для подальшої розвідки.

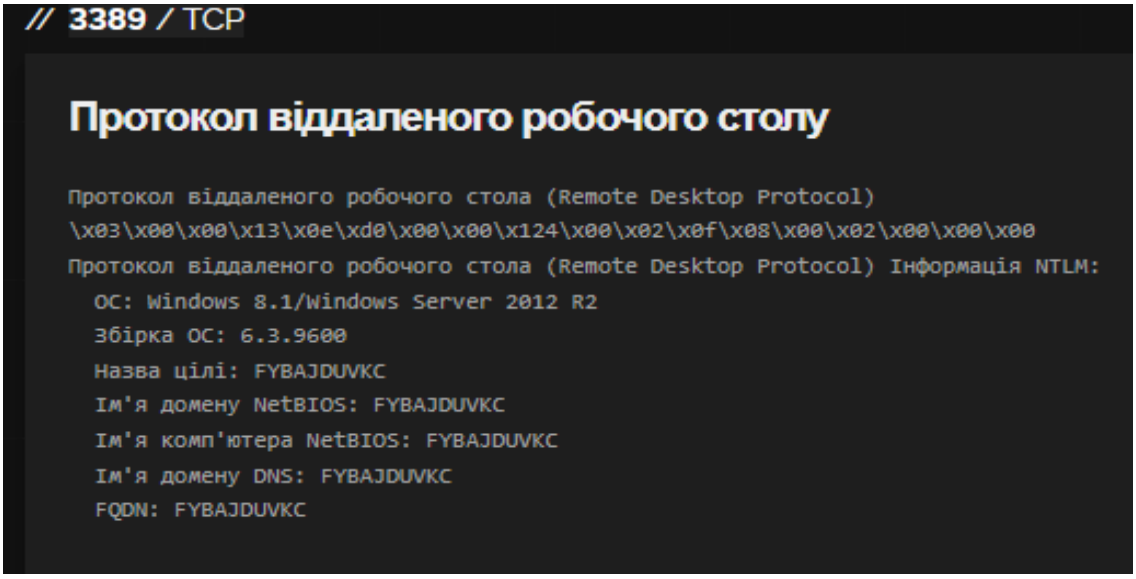

Рисунок 3.9 – Результат інформації 3389/TCP за хостом 195.3.205.125

На рисунку 3.10-3.11 показано деталі SSL-сертифікату, включаючи його версію, серійний номер, алгоритм підпису (sha256WithRSAEncryption), видавця та тему (CN=FYBAJDUVKC), термін дії (з 28 грудня 2024 р. до 29 червня 2025 р.), а також розширення X509v3, що вказують на використання ключа для автентифікації веб-сервера TLS та шифрування ключа/даних.

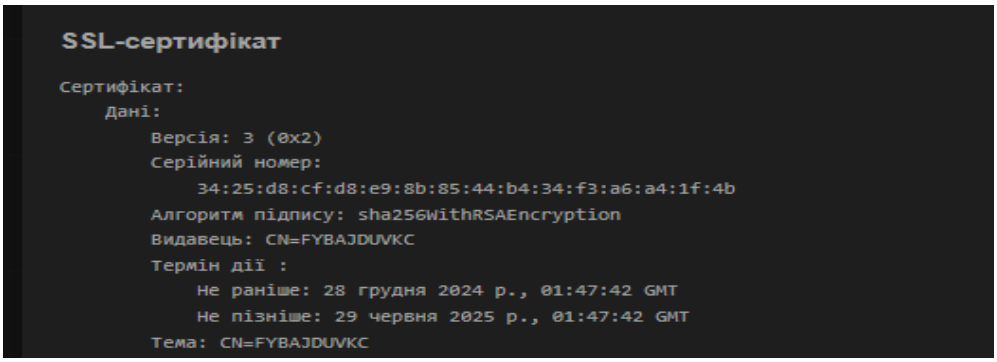

Рисунок 3.10 – зображення інформації про сертифікати SSL

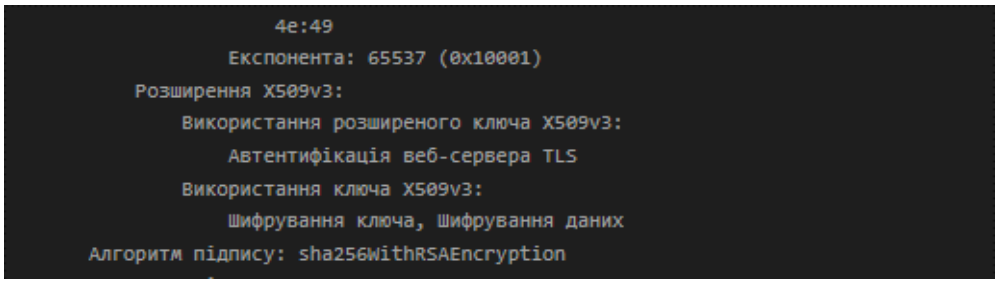

Рисунок 3.11– зображення інформації про сертифікати SSL

У таблиці 3.2 представлено загальні ідентифікаційні відомості про досліджуваний хост, зокрема його IP-адресу, доменні атрибути, географічну приналежність, організаційну приналежність та характеристики операційної системи.

Таблиця 3.2 – Загальна інформація хоста 195.3.205.125

Параметр	Значення
IP-адреса	195.3.205.125
Ім'я хоста	125.205.3.195.badaki.colo.customers.uplink.net.ua
Домени	uplink.net.ua
Країна	Україна
Місто	Київ
Організація	ТОВ «ТУХА»
Інтернет-провайдер	ТОВ «ТУХА» (ASN: AS35409)
Операційна система	Windows Server 2012 R2 Datacenter Build 9600
Останній раз скановано	2025.05.19

Таблиця 3.3 містить результати аналізу відкритих портів та активних мережевих сервісів хоста, що становлять потенційні точки доступу для несанкціонованого втручання.

Таблиця 3.3 – Відкриті порти та сервіси хоста 195.3.205.125

Параметр	Значення
Порт 135/TCP	MS RPC - Microsoft RPC Endpoint Mapper. Служби: wininit.exe, winlogon.exe. RPC інтерфейси: InitShutdown, WMsgKRpc...
Порт 445/TCP	Microsoft-DS (SMB v1). Автентифікація увімкнена. Можливості: large-files, rpc-remote-api тощо.
Порт 3389/TCP	RDP. ОС: Windows Server 2012 R2. NetBIOS: FYBAJDUVKC. Сертифікат: самопідписаний, SHA256, 2048 біт.

У таблиці 3.4 наведено виявлені вразливості, пов’язані з конфігурацією програмного забезпечення та сервісів хоста, що створюють передумови для здійснення кіберзагроз.

Таблиця 3.4 – Аналіз вразливостей хоста 195.3.205.125

Параметр	Значення
Windows Server 2012 R2	Критична вразливість через відсутність підтримки після 10.10.2023.
SMBv1 (порт 445)	Критична вразливість. Піддається атакам типу EternalBlue (WannaCry, NotPetya).
Відкритий RDP (порт 3389)	Високий ризик Brute-force, експлуатації RDP-вразливостей.
Самопідписаний сертифікат	Можливість MitM-атаки через відсутність довіри до сертифіката.
Порт 135 відкритий	Розкриває інформацію про RPC-сервіси. Може бути використано для рекогносцировки.

Таблиця 3.5 містить узагальнення основних кіберризиків, виявлених у процесі аналізу, а також сформульовані рекомендації щодо підвищення рівня інформаційної безпеки хоста.

Таблиця 3.5 – Загальні ризики та рекомендації хоста 195.3.205.125

Параметр	Значення
Критичний ризик	ОС без підтримки, SMBv1, відкритий RDP створюють надзвичайну вразливість.
Рекомендація 1	Негайно відключити SMBv1. Увімкнути SMBv2/SMBv3.
Рекомендація 2	Заблокувати порт 445 з Інтернету. Для доступу використовувати VPN.
Рекомендація 3	Оновити ОС або придбати ESU. Ізолювати сервер.
Рекомендація 4	Захистити RDP: VPN, сильні паролі, NLA, IP-фільтрація.
Рекомендація 5	Заблокувати порт 135 з Інтернету.
Рекомендація 6	Регулярні аудити безпеки та сканування.

Результати, отримані шляхом OSINT-розвідки хоста 195.3.205.125, надають критично важливі дані, що дозволяють з високою ймовірністю прогнозувати потенційні вектори кібератак та оцінювати загальний рівень ризику.

Зібрана OSINT-інформація вказує на те, що хост 195.3.205.125 є високопріоритетною та легкою ціллю для широкого спектра кібератак. Комбінація застарілої ОС, активного SMBv1 та відкритого RDP створює ідеальні умови для успішної компрометації з боку як автоматизованих інструментів (ботнетів, сканерів вразливостей), так і цілеспрямованих зловмисників.

Приклад збору даних за допомогою сервісу theHarvester , робота з цим ресурсом виконується у режимі командного рядка .

На операційних системах: Linux (особливо дистрибутиви для пентестингу, як Kali Linux, Parrot OS, де він часто вже встановлений або легко встановлюється) , macOS , Windows (з встановленим Python та необхідними залежностями, можливо, через WSL - Windows Subsystem for Linux для кращої сумісності)

Основні опції (options):

Довідка:

1) `-h, --help`: Показати це довідкове повідомлення та вийти.

Цільовий об'єкт:

2) `-d, --domain DOMAIN`: (Обов'язково) Назва компанії або домен для пошуку.

Контроль результатів:

3) `-l, --limit LIMIT`: Обмежити кількість результатів пошуку. За замовчуванням: 500.

4) `-S, --start START`: Почати з результату номер X. За замовчуванням: 0.

Джерела даних та інтеграції:

5) `-b, --source SOURCE`: Джерела даних для використання. Можливі значення: anubis, baidu, bevigil, binaryedge, bing, bingapi, brave, bufferoverun, censys, certspotter, criminalip, crtsh, duckduckgo, fullhunt, github-code, hackertarget, hunter,

hunterhow, intelx, netlas, onyphe, otx, pentesttools, projectdiscovery, rapiddns, rocketreach, securityTrails, sitedossier, subdomaincenter, subdomainfinderc99, threatminer, tomba, urlscan, virustotal, yahoo, whoisxml, zoomeye, venacus.

6) `-s, --shodan`: Використовувати Shodan для запитів до виявлених хостів.

Робота з DNS:

7) `-v, --virtual-host`: Перевірити ім'я хоста через DNS-резолюцію та шукати віртуальні хости.

8) `-e, --dns-server DNS_SERVER`: DNS-сервер для використання при пошуку.

9) `-r, --dns-resolve` [DNS_RESOLVE]: Виконати DNS-резолюцію для субдоменів, використовуючи список резолверів або передані резолвери. За замовчуванням: False (вимкнено).

10) `-n, --dns-lookup`: Увімкнути пошук через DNS-сервер. За замовчуванням: False (вимкнено).

11) `-c, --dns-brute`: Виконати DNS брутфорс (перебір за словником) для домену.

Додаткові функції:

12) `-p, --proxies`: Використовувати проксі для запитів. Проксі вказуються у файлі proxies.yaml.

13) `--screenshot SCREENSHOT`: Зробити скріншоти для знайдених доменів; вказати каталог для збереження скріншотів: --screenshot output_directory.

14) `-t, --take-over`: Перевірити на можливість захоплення субдоменів (takeovers).

Збереження результатів:

15) `-f, --filename FILENAME`: Зберегти результати у файл формату XML та JSON.

Також у theHarvester є джерела які потребують API пакетів і ті що працюють без них.

Для дослідження було обрано домен microsoft.com та onua.edu.ua і виконано наступну команду `theHarvester -d microsoft.com -l 20 -b securityTrails` та отримано результат на рисунку 3.12

Рисунок 3.11 – результат виконання команди `theHarvester -d microsoft.com -l 20 -b securityTrails`

З рисунку видно , що команда `-b securityTrails` не працює без налаштування API key , щоб його додати треба виконати наступну команду у режимі суперкористувача `sudo nano api-keys.yaml` та у відповідному полі ввести свій API key так як це показано на рисунку 3.3.5 та зберегти відредагований файл. Після цього виконання команди буде налаштовано належним чином та буде виконуватися аналіз із заданим джерелом для обраного домену.

На рисунку 3.12 показано метод додавання ключа API.

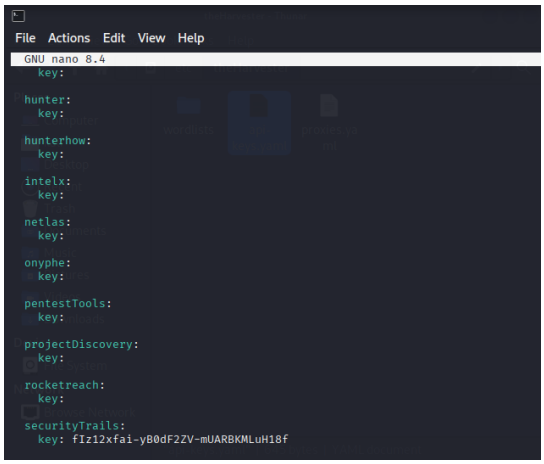

Рисунок 3.12 – додавання API key

На рисунку 3.13 показаний результат виконання тієї ж команди при налаштованому API ключі.

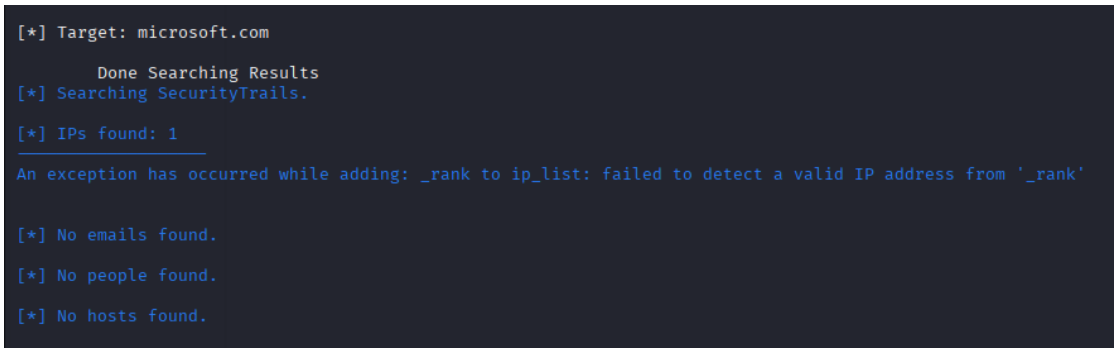

Рисунок 3.13 – результат виконання команди `theHarvester -d microsoft.com -l 20 -b securityTrails` з налштованим API key

Для отримання більш повної інформації треба використати більшу кількість джерел як зображено на рисунку 3.14

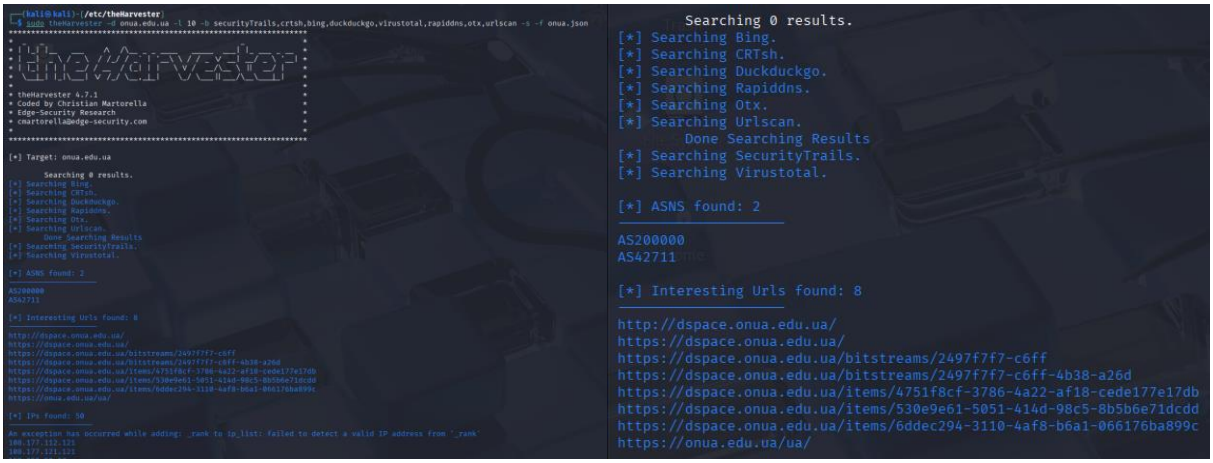

Рисунок 3.14 – Результат виконання команди `theHarvester -d onua.edu.ua -l 10 -b securityTrails , crtsh , bing, duckduckgo, virustotal, rapiddns,otx,urlscan -s -f onua.json`

Команда дає можливість зберігати отриманий результат у форматі .json , що значно полегшує подальший аналіз та дає можлтвість створення структурований аналіз отриманих даних з onua.json:

У таблиці 3.6 наведений аналіз отриманих даних за допомогою theHarvester, що демонструє категорії зібраної інформації (автономні системи, email-адреси, субдомени/хости, цікаві URL, IP-адреси) та їх корисність для маршрутизації трафіку, виявлення цілей для фішингу, оцінки поверхні атаки, пошуку вразливостей веб-додатків та ідентифікації серверної інфраструктури.

Таблиця 3.6 – Аналіз отриманих даних за домомогою theHarvester

Категорія	Дані	Аналіз / Корисність
Автономні системи (ASNs)	AS200000, AS42711	Маршрутизація трафіку. Допомагає ідентифікувати інфраструктуру, пов’язані IP-діапазони.
Email-адреси	chancellery@onua.edu.ua, international@onua.edu.ua, press@onua.edu.ua, vstup@onua.edu.ua	Фішинг, соц. інженерія. Перевірка на витоки. Структура організації.
Субдомени / хости	14congress.onua.edu.ua, agroecozem.onua.edu.ua, mail.onua.edu.ua тощо	Поверхня атаки, ідентифікація сервісів, старі/тестові активи, зовнішні сервіси.
Цікаві URL	http://dspace.onua.edu.ua/, https://onua.edu.ua/ua/ тощо	Веб-додатки, пошук вразливостей, метадані в документах, структура сайту.

Категорія	Дані	Аналіз / Корисність
IP-адреси	108.177.112.121, 159.253.20.18, 194.110.129.33 тощо	Ідентифікація серверів, сканування портів, аналіз геолокації та провайдерів.

У таблиці 3.7 наведена оцінка отриманої інформації, що ілюструє напрями використання зібраних OSINT-даних, такі як прогнозування кібератак та оцінка ризиків (визначення поверхні атаки, ідентифікація слабких місць, пріоритезація захисту), тестування на проникнення, конкурентна розвідка та моніторинг безпеки.

Таблиця 3.7 – Оцінка отриманої інформації

Напрям використання	Опис користі
Прогнозування кібератак та оцінка ризиків	Визначення поверхні атаки: Велика кількість субдоменів та IP-адрес вказує на численні потенційні точки входу.
	Ідентифікація слабких місць: Виявлення застарілого ПЗ, відкритих портів та email-адрес, які можуть бути використані у фішингу.
	Пріоритезація захисту: Визначення критичних активів для концентрації зусиль на їх захисті.
Тестування на проникнення (Pentesting)	Зібрані дані служать стартовою точкою для подальших дій: сканування портів, пошук вразливостей, соціальна інженерія.
Конкурентна розвідка та інфраструктура	Дозволяє оцінити технології, структуру та зовнішні залежності організації (наприклад, хостинг чи поштові сервіси).
Моніторинг безпеки	Регулярний збір подібної інформації дає змогу виявляти нові або неправильно налаштовані активи.

Приклад збору даних за допомогою сервісу Censys . Платформа Censys надає аналітичні дані в режимі реального часу, яким команди безпеки можуть довіряти для швидшого виявлення загроз, впевненого визначення пріоритетів ризиків та пришвидшення розслідувань. Для дослідження було обрано домен omr.gov.ua та результат пошуку за цим доменом було проілюстровани на рисунку 3.15

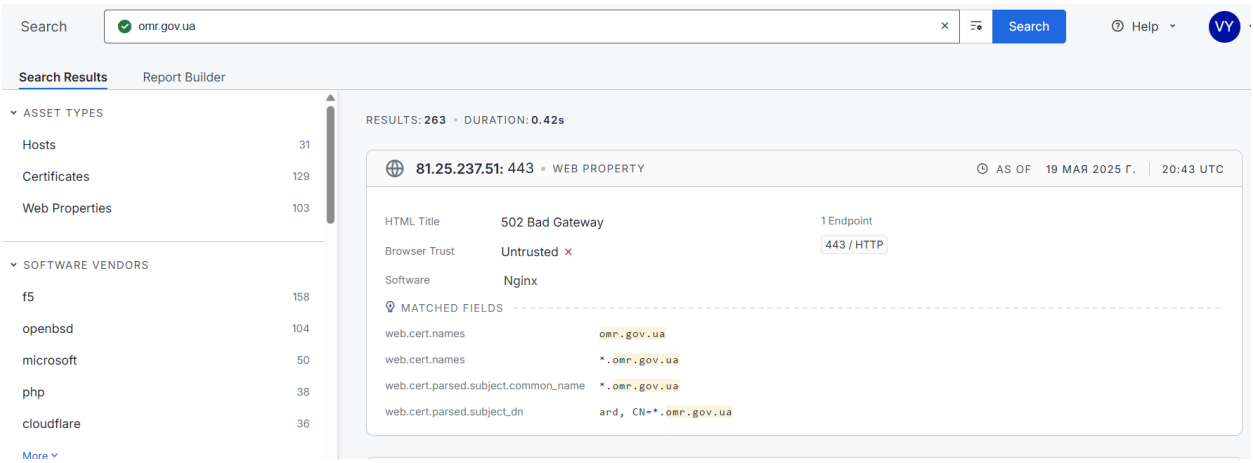

Рисунок 3.15 – Результат запиту omr.gov.ua

На рисунку 3.16 показан приклад використання фільтрів для пошуку веб серверів apache у сервісі Censys

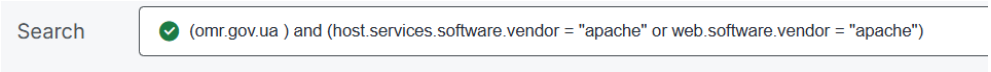

Рисунок 3.16 – Команда веб серверів apache

Ця команда показує усі веб сервери apache приєднані до досліджуваного домену omr.gov.ua , результат виконання команди наведений на рисунку 3.3.12

На рисунку 3.17 показан результат використання фільтрів для пошуку веб серверів apache у сервісі Censys

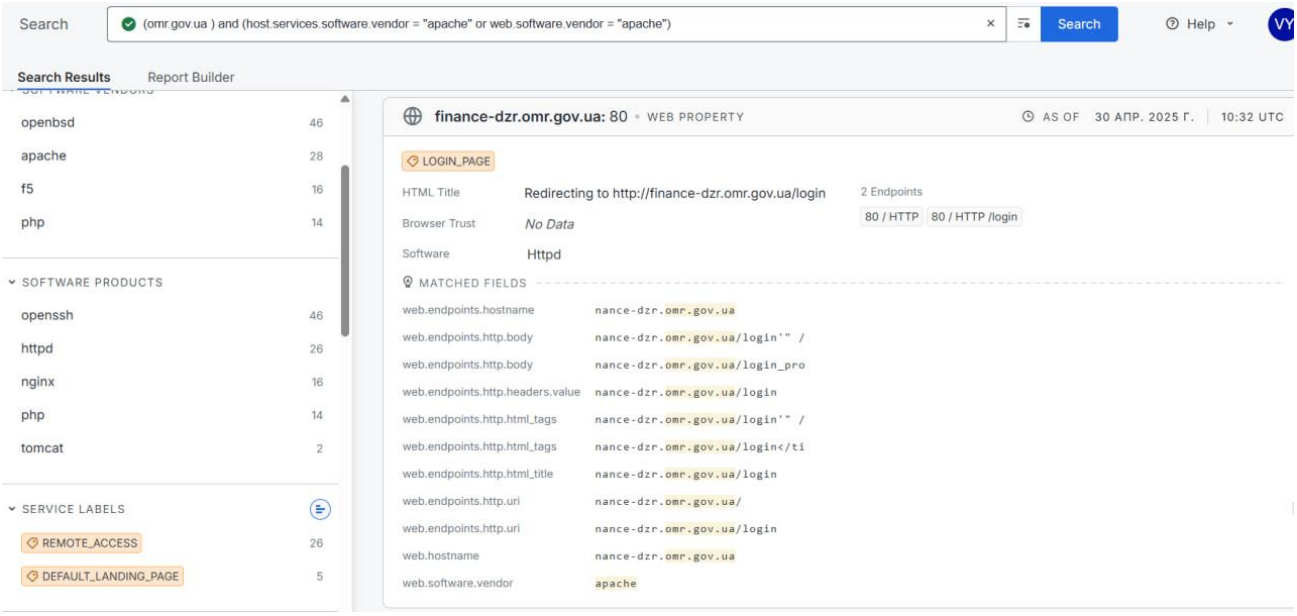

Рисунок 3.17 – Результат виконання команди з пошуку веб серверів

На рисунку 3.18 показана більш детальна інформація за вказаним веб-сервером, що демонструє відповіді HTTP 80/TCP для URL http://finance-dzr.omr.gov.ua/ (статус 302 Found, редирект на сторінку логіну) та http://finance-dzr.omr.gov.ua/login (статус 200 OK), що підтверджує наявність та доступність сторінки входу через незахищений протокол.

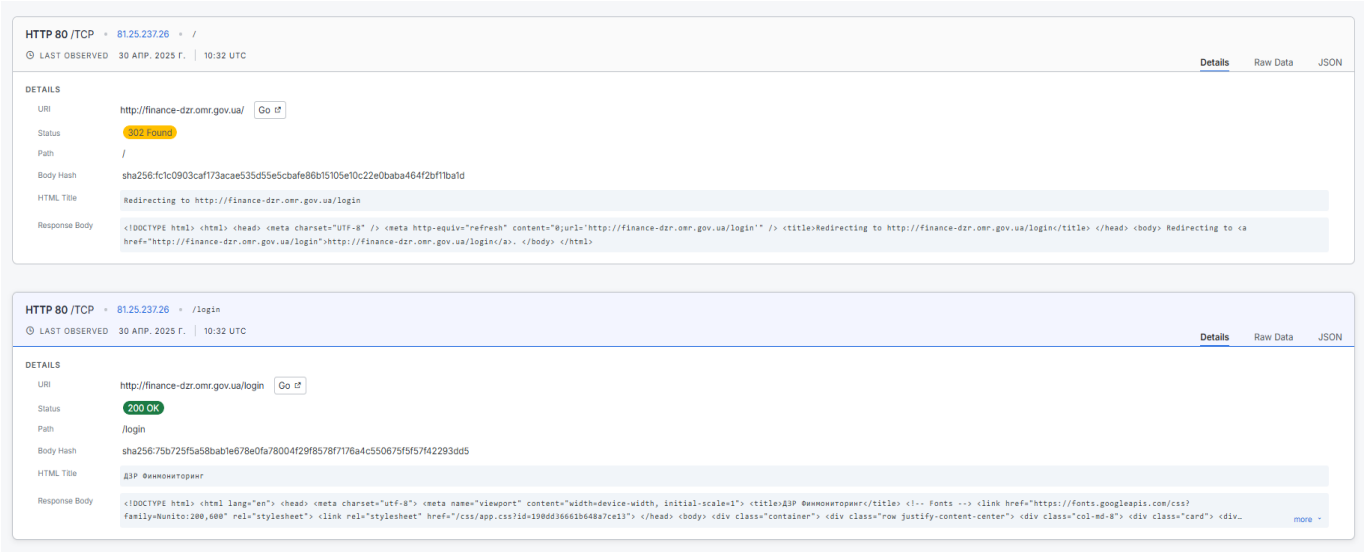

Рисунок 3.18 – Більш детальна інформація за вказаним веб сервером

На рисунку одразу видно , що субдомен finance-dzr.omr.gov.ua є високопріоритетною ціллю через наявність сторінки входу до, ймовірно, фінансової системи. Найбільш критичною проблемою є використання HTTP для сторінки входу, що робить облікові дані надзвичайно вразливими до перехоплення.

Також можливо подивитись сертифікати документ, який містить інформацію про власника доменного імені і дані про терміни реєстрації домену. Реальний власник домену, вказаний в базі WHOIS

На рисунку 3.19 показано команду пошуку (omr.gov.ua) and cert.fingerprint_sha256:* в інтерфейсі Censys, спрямовану на виявлення всіх SSL/TLS сертифікатів, пов'язаних з доменом omr.gov.ua.

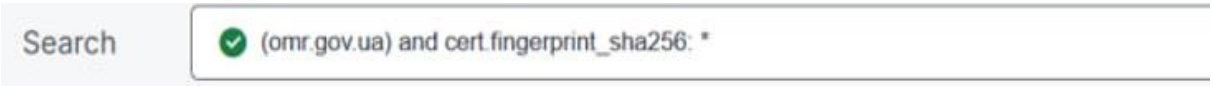

Рисунок 3.19 – Команда для відображення сертифікатів домену

На рисунку 3.20 представлено результати пошуку сертифікатів для домену omr.gov.ua в Censys, де видно інформацію про сертифікати для субдоменів citizen.omr.gov.ua (з позначкою "expired") та mail.dpsp.omr.gov.ua.

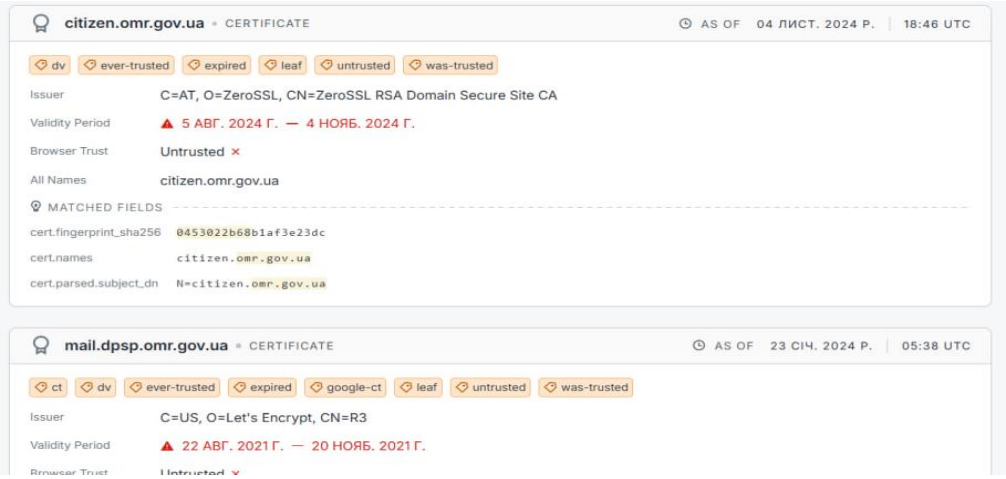

Рисунок 3.20 – Результат команди відображення сертифікатів домену

Даний рисунок демонструє розширену інформацію про SSL/TLS сертифікат для субдомену citizen.omr.gov.ua, отриману з Censys, що вказує на його статус "expired" (термін дії закінчився), "leaf" (кінцевий сертифікат), "untrusted" (недовірений) та "was-trusted" (колись був довіреним), що свідчить про низку проблем з його валідністю.

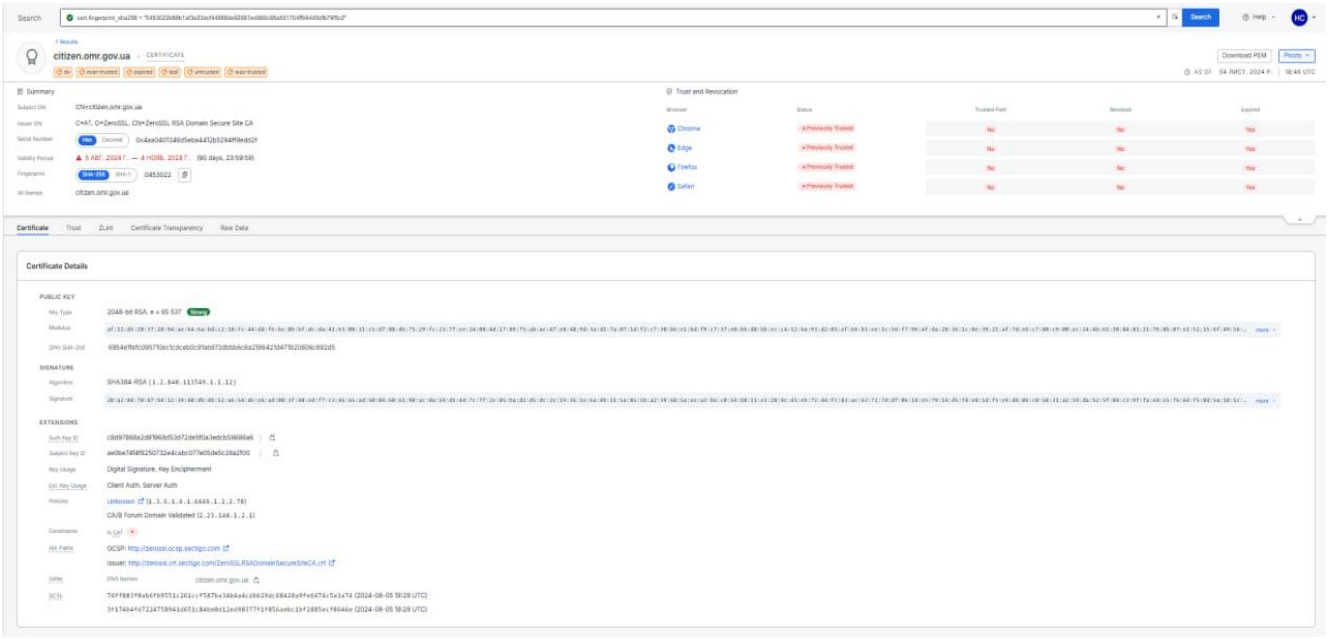

Рисунок 3.21 – Результат більш детальної інформації про сертифікат

Сертифікат для citizen.omr.gov.ua, який має низку проблем, що роблять його недовіреним для браузерів. Censys чітко маркує його як expired (термін дії закінчився) , leaf (кінцевий сертифікат) , untrusted (недовірений) та was-trusted (колись був довіреним).

OSINT Framework зосереджений на зборі інформації з безкоштовних інструментів або ресурсів . Він представлений у вигляді ментальної карти з розгалуженнями до кожної гілки і початком в OSINT Framework . Приклад можливого розгалуження: OSINT Framework → Email Address → Email Verification → MailScrap . Цей інструмент корисний своєю кількістю зазначених ресурсів , що значно пришвидшує роботу , приклад пошуку за розгалуженям зображено на рисунку 3.22 .

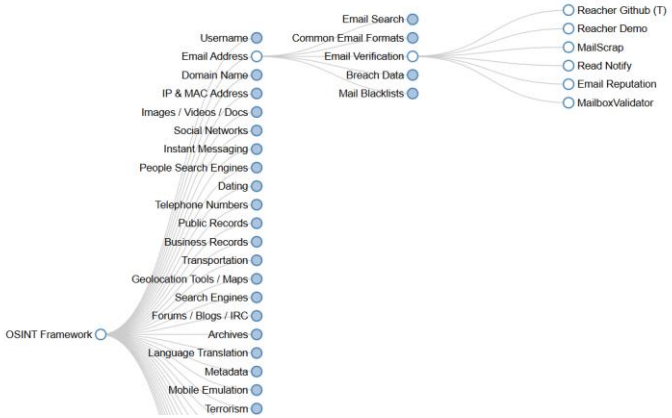

Рисунок 3.22 – Приклад розгалуження до MailScrap

Приклад використання сервісу з запропанованих на рисунку 3.23 , для ддослідження був обраний сервіс перевірки поштових скриньок MailScrap та сумнівна поштова скринька jeff.dasovich@enron.com .

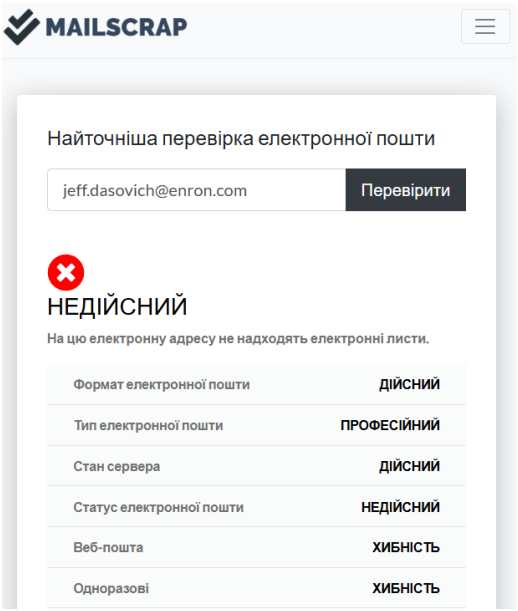

Рисунок 3.23 – Результат перевірки поштової скриньки

Також цей сервіс надзвичайно зручний при пошуку інформації у соціальним мережах послідовна розгалуженість зображену на рисунку 3.24 OSINT Framework → Social Networks → Instagram → Instagram Profile Analyzer

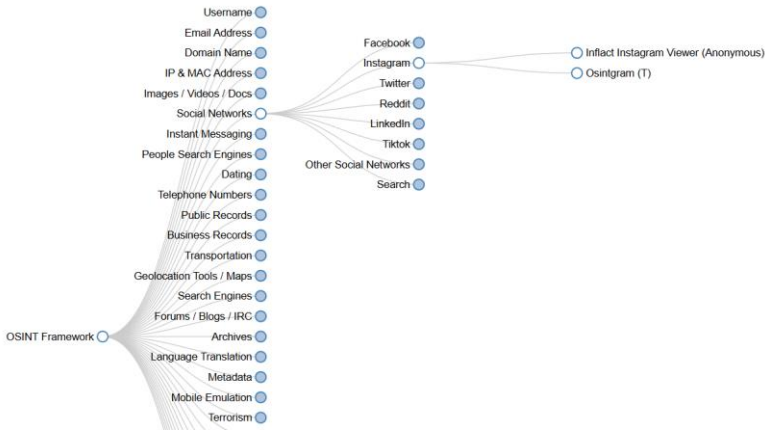

Рисунок 3.24 – Розгалуження до Instagram

На рисунку 3.25 зображено використання сервісу для аналізу профілю Instagram Profile Analyzer, що демонструє результати аналізу профілю користувача "jlindsley" (Joe Lindsley), включаючи показники залученості (Engagement), кількість підписників, публікацій, середню кількість лайків та коментарів, а також глобальний та регіональний рейтинг.

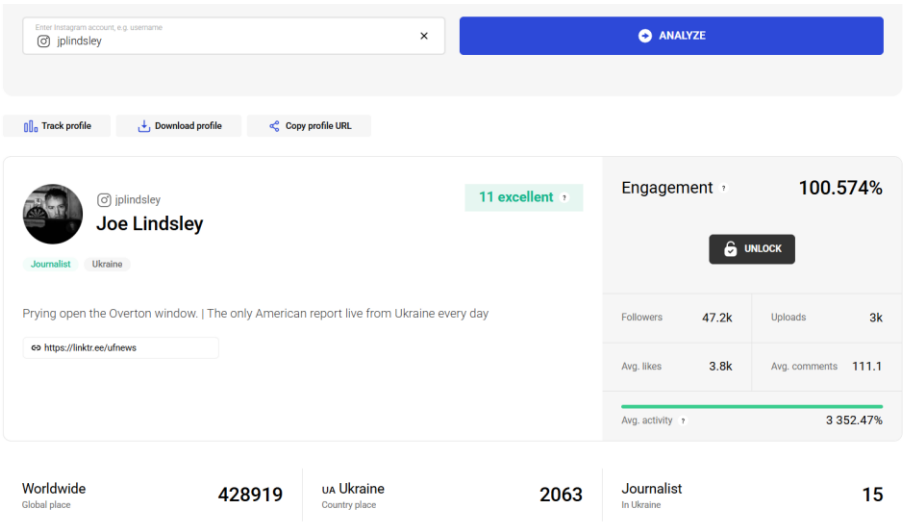

Рисунок 3.25 – Використання сервісу для аналізу профілю Instagram Profile Analyzer

На рисунку 3.26 представлена інфорграфіка, що відображає аналіз активності публікацій в Instagram-профілі, включаючи частоту публікацій за днями тижня, найпопулярніший час для постів (вівторок, 12:00), топові хештеги та ключові слова в описах, а також результати семантичного аналізу настрою контенту.

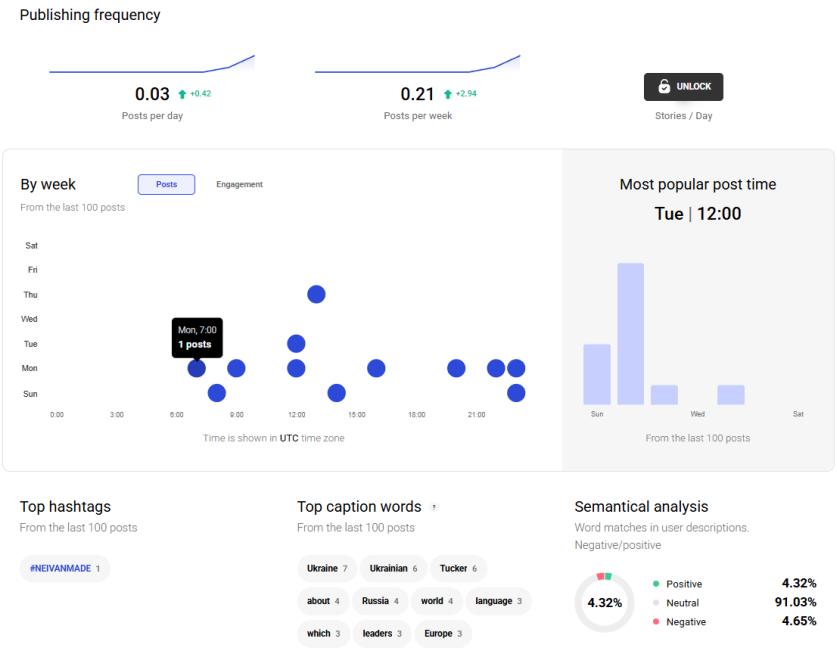

Рисунок 3.26 – Інфорграфіка частоти викладення постів

Рисунок 3.27 ілюструє деревоподібну структуру OSINT Framework, що показує розгалуження від загальних категорій пошуку до конкретних інструментів, зокрема шлях до інструментів для перевірки телефонних номерів (наприклад, Telephone Numbers → International → Whocalld, True Caller).

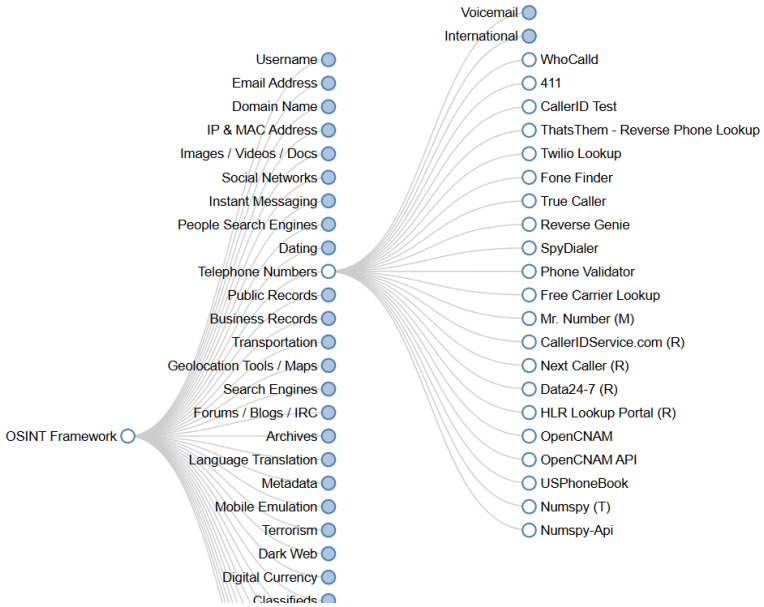

Рисунок 3.27– Розгалуження для перевірки номеру

На рисунку 3.28 зображено результат пошуку за номером у сервісі GetContact , де можна побачити як рейтинг довіри так і теги інших користувачів

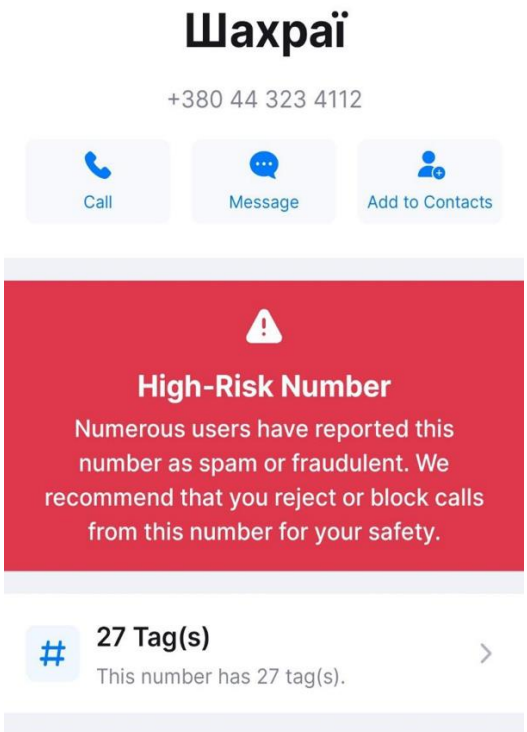

Рисунок 3.28 – Результат перевірки за номером

Ресурс OSINT Framework найкраще підходить для швидкого реагування та простий у використанні . Найбільшу користь він може принести коли персонал будь-якого підприємства буде ознайомлений і матиме навички його використання , що значно зменшить ризик фішинговивих атак та телефонних шахрайств .

3.4 Аналіз зібраних даних: Як отримані дані можуть бути використані для прогнозування кібератак

Після завершення етапу збору даних з відкритих джерел, наступним критично важливим кроком є їх аналіз з метою виявлення значущих патернів, індикаторів компрометації (IoC) або індикаторів підготовки атак (IoPA), які можуть слугувати основою для прогнозування кіберзагроз.

Описана методологія збору даних за допомогою OSINT-інструментів є фундаментальною для прогнозування кібератак. Моніторинг соціальних медіа та форумів дозволяє виявляти ранні ознаки обговорення потенційних цілей або нових векторів атак, що може свідчити про підготовку кампаній. Сканування мережевих пристроїв інструментами Shodan та Censys надає інформацію про конфігурації та версії програмного забезпечення, що дозволяє ідентифікувати системи з відомими вразливостями, які з високою ймовірністю стануть цілями для експлуатації. Хоча дані з theHarvester будуть проаналізовані окремо, агрегація відомостей про домени та IP-адреси є ключовою для розуміння цифрового сліду організації та виявлення її публічно доступних активів. Аналіз метаданих документів за допомогою FOCA може розкрити внутрішню інформацію (імена користувачів, шляхи до файлів, версії ПЗ), яка, потрапивши до зловмисників, значно полегшить їм підготовку цільових атак, таких як фішинг або спроби несанкціонованого доступу. Застосування комплексного OSINT-інструментарію, включаючи Maltego, SpiderFoot та Recon-ng, забезпечує багатогранний збір даних, що дозволяє виявити слабкі місця та потенційні точки входу, тим самим прогнозуючи можливі напрямки атак.

Наведений приклад використання Shodan для дослідження хоста 195.3.205.125 яскраво демонструє, як конкретні фільтри допомагають ідентифікувати критично вразливі системи. Виявлення операційної системи Windows Server 2012 R2, підтримка якої припинена, є прямим індикатором високого ризику, оскільки нові вразливості не виправлятимуться. Це дозволяє прогнозувати, що даний хост стане ціллю для атак, що експлуатують відомі або нові (0-day для цієї ОС) вразливості. Відкриті порти, такі як 135/TCP (MS RPC), 445/TCP (SMBv1) та 3389/TCP (RDP), вказують на конкретні сервіси, які часто стають векторами атак. Зокрема, наявність SMBv1 робить систему вразливою до атак типу EternalBlue (WannaCry, NotPetya). Прогнозується висока ймовірність компрометації через цей протокол. Відкритий RDP з самопідписаним сертифікатом збільшує ризик атак перебору паролів та "людина посередині". Це дозволяє прогнозувати спроби несанкціонованого віддаленого доступу та перехоплення облікових даних. Загалом, комбінація застарілої ОС, активного SMBv1 та відкритого RDP робить хост легкою ціллю, дозволяючи з високою ймовірністю прогнозувати успішну компрометацію. Рекомендації, наведені в Таблиці 3.1 , спрямовані на мінімізацію саме цих прогнозованих загроз.

Використання OSINT Framework для навігації по безкоштовних інструментах та ресурсах є ефективним методом систематизації процесу збору інформації. Приклад розгалуження до MailScrap для перевірки email-адрес, як-от jeff.dasovich@enron.com, ілюструє можливість верифікації контактних даних. Якщо така перевірка показує, що адреса дійсна і належить до професійного типу, але, наприклад, сервер не відповідає або статус електронної пошти "недійсний", це може мати різне значення. Для прогнозування атак, це має значення при навчанні персоналу коричтуватися елементами цього сервісу .

Instagram Profile Analyzer Аналіз профілів у соціальних мережах, таких як Instagram, за допомогою спеціалізованих інструментів, може надати цінну інформацію про особу або організацію. Дані про частоту публікацій, популярний час для постів, топові хештеги та слова в описах, а також семантичний аналіз

(позитивний/нейтральний/негативний настрій) можуть бути використані для створення психологічного портрету, виявлення інтересів, зв'язків та навіть рутини. Для прогнозування атак ця інформація може бути використана для підготовки більш персоніфікованих фішингових листів або повідомлень у соціальній інженерії. Наприклад, знаючи інтереси або нещодавні події з життя цілі, зловмисник може створити більш правдоподібну легенду. Аналіз активності може також вказати на періоди, коли особа менш пильна або більш схильна до необережних дій в Інтернеті.

Перевірка телефонних номерів через сервіси типу Whocalld дозволяє отримати інформацію про оператора зв'язку та країну. У контексті прогнозування кібератак, якщо номер телефону пов'язаний з корпоративним акаунтом або використовується для двофакторної автентифікації, знання оператора може бути використано для більш цілеспрямованих атак SMS-фішингу (смішингу) або спроб перехоплення SIM-карти (SIM swapping). Якщо виявлено, що номер належить до певного регіону або оператора, який відомий своєю вразливістю до певних типів шахрайства, це підвищує ризик відповідних атак.

Платформа Censys надає більш глибокий технічний аналіз публічно доступних хостів та сервісів, особливо в частині конфігурації SSL/TLS сертифікатів та деталей роботи мережевих протоколів. Ця інформація є надзвичайно цінною для прогнозування атак, заснованих на слабкостях шифрування та автентифікації.

Інструмент theHarvester є ключовим на початкових етапах OSINT-розвідки, агрегуючи дані про домени, IP-адреси, email та інші мережеві атрибути з публічних джерел (пошукові системи, SecurityTrails, CRTsh, VirusTotal, OTX, UrlScan). Зібрана інформація критична для прогнозування кібератак. Наприклад , численні субдомени (onua.edu.ua: 14congress.onua.edu.ua , mail.onua.edu.ua) розширюють поверхню атаки, адже кожен може містити вразливості, що скануються зловмисниками. Email-адреси (chancellery@onua.edu.ua, vstup@onua.edu.ua) стають цілями для фішингу та соціальної інженерії. IP-адреси та ASN допомагають картувати інфраструктуру для глибокого сканування (Nmap, Shodan/Censys) портів та сервісів, особливо пов'язаних

з критичними субдоменами. URL на кшталт dspace.onua.edu.ua можуть вказувати на вразливі веб-додатки. Збереження результатів у JSON полегшує автоматизований аналіз. Гнучкість theHarvester дозволяє використовувати джерела як з API-ключами (SecurityTrails), так і без них.

Дослідження домену omr.gov.ua через Censys (рис. 3.3.9-3.3.14) виявляє статус SSL/TLS сертифікатів та конфігурації ПЗ. Наприклад, прострочений сертифікат для citizen.omr.gov.ua (рис. 3.3.13, 3.3.14) вказує на ризик MitM-атак та перехоплення даних, тоді як актуальний (напр., для osvita-omr.gov.ua) свідчить про кращий захист. Аналіз деталей сертифікату (видавець, термін дії, алгоритми) може виявити додаткові вразливості, як-от слабкі алгоритми шифрування. Ідентифікація ПЗ, наприклад, веб-серверів Apache (рис. 3.3.9, 3.3.10), та їх версій дозволяє перевіряти відомі CVE, роблячи застарілі системи цілями для експлойтів. Критичним є виявлення сторінок входу (LOGIN_PAGE для finance-dzr.omr.gov.ua, рис. 3.3.10) через незахищений HTTP (рис. 3.3.11), що підвищує ризик перехоплення облікових даних . Таким чином, Censys дозволяє ідентифікувати активи та глибоко аналізувати їх конфігурацію безпеки для виявлення слабких місць і прогнозування атак.

ВИСНОВКИ

У кваліфікаційній роботі було проведено комплексне дослідження можливостей застосування методів та інструментів Open-source intelligence (OSINT) для прогнозування кібератак. Метою роботи було вивчення теоретичних засад, аналіз практичного інструментарію та методик OSINT, а також демонстрація їх ефективності для виявлення потенційних загроз та формування обґрунтованих прогнозів щодо ймовірних векторів та цілей кібератак.

Для досягнення поставленої мети було вирішено наступні завдання:

Проаналізовано теоретичні основи прогнозування кібератак та концептуальні засади OSINT. Розглянуто класифікацію кібератак, етапи їх життєвого циклу та існуючі підходи до прогнозування. Детально вивчено сутність OSINT, його ключові джерела (відкриті дані з Інтернету, соціальні мережі, технічні форуми, бази даних вразливостей), інструментарій, методологію та ітераційний цикл OSINT-розвідки. Встановлено фундаментальну роль OSINT у процесах виявлення, аналізу кіберзагроз та його невід'ємний зв'язок з проактивним прогнозуванням атак [7, 17].

Ідентифіковано та класифіковано релевантні джерела, інструменти та методи аналізу OSINT-даних, що є ключовими для виявлення індикаторів підготовки атак (IoPA). Оглянуто такі джерела, як форуми Dark/Deep Web, соціальні мережі, публічні репозиторії коду, бази даних вразливостей та витоків. Досліджено функціональні можливості та обмеження основного інструментарію (Maltego, Shodan, Censys, theHarvester, SpiderFoot, Recon-ng, FOCA, OSINT Framework). Розглянуто методи аналізу OSINT-даних, зокрема аналіз ключових слів, сентимент-аналіз, кореляційний аналіз та потенціал застосування базових моделей машинного навчання. Особливу увагу приділено етичним та юридичним аспектам використання OSINT.

Продемонстровано практичне застосування OSINT-методик для прогнозування кібератак на основі зібраних даних. На конкретних прикладах дослідження доменів onua.edu.ua, omr.gov.ua та microsoft.com було показано процес збору інформації за

допомогою інструментів theHarvester, Censys, Shodan та OSINT Framework. В результаті було виявлено значну кількість субдоменів, IP-адрес, адрес електронної пошти та інших мережевих атрибутів, що суттєво розширюють поверхню потенційної атаки. Аналіз отриманих даних, зокрема стану SSL/TLS сертифікатів, версій програмного забезпечення (наприклад, веб-серверів Apache) та конфігурацій мережевих сервісів (наприклад, відкриті порти SMBv1, RDP, використання незахищеного протоколу HTTP для сторінок входу), дозволив ідентифікувати конкретні вразливості. На основі цих вразливостей та зібраних індикаторів сформовано прогнози щодо ймовірних векторів атак, таких як експлуатація відомих слабкостей застарілого ПЗ, фішингові кампанії, атаки типу "людина посередині" та компрометація облікових даних.

Оцінено ефективність, обмеження та етичні аспекти застосування OSINT для прогнозування кібератак. Практичні приклади підтвердили, що зібрані OSINT-дані є критично важливими для формування обґрунтованих прогнозів. Наприклад, виявлення застарілої операційної системи Windows Server 2012 R2 та активного протоколу SMBv1 на хості 195.3.205.125 дозволяє прогнозувати високий ризик компрометації через атаки типу EternalBlue. Прострочений SSL-сертифікат для citizen.omr.gov.ua є індикатором, що дозволяє прогнозувати можливі MitM-атаки. Водночас було виявлено обмеження, такі як залежність від API-ключів для деяких джерел (наприклад, SecurityTrails), потенційні помилки в роботі інструментів та нагальна необхідність кореляції даних з різних джерел для підвищення точності та надійності прогнозів.

Поставлена в кваліфікаційній роботі мета була досягнута. Доведено, що Open-source intelligence є потужним та доступним інструментом для прогнозування кібератак, що дозволяє на ранніх етапах ідентифікувати потенційні загрози, оцінювати поверхню атаки та виявляти слабкі місця в IT-інфраструктурі. Ефективне використання OSINT-методик та інструментів (theHarvester, Censys, Shodan, OSINT Framework), особливо у поєднанні з аналізом статистичних даних (наприклад, звітів

CERT-UA та Кіберполіції), надає фахівцям з кібербезпеки цінну інформаційну основу для прийняття проактивних захисних заходів та обґрунтованого розподілу ресурсів.

Практичне значення отриманих результатів полягає у можливості їх використання фахівцями з кібербезпеки для вдосконалення існуючих та розробки нових методик прогнозування кібератак, підвищення загального рівня захищеності організацій та підготовки кваліфікованих кадрів у сфері кіберрозвідки.

Перспективними напрямками подальших досліджень є розробка автоматизованих систем для інтелектуальної кореляції OSINT-даних з різних джерел, глибше застосування методів машинного навчання для виявлення складних та неочевидних патернів підготовки атак, а також інтеграція OSINT-прогнозів з сучасними системами управління інформаційною безпекою (SIEM/SOAR) для автоматизації та підвищення ефективності реагування на інциденти.

ПЕРЕЛІК ПОСИЛАНЬ

1. *Перелік документів, що регулюють питання кібербезпеки України.* URL: https://ligabezinfo.org/legislation/ (дата звернення: 10.05.2025).
2. Ko Y. S., Ra I.-K., Kim C.-S. A Study on IP Exposure Notification System for IoT Devices Using IP Search Engine Shodan. *International Journal of Multimedia and Ubiquitous Engineering.* 2015. Vol. 10, no. 12. P. 61–66. URL: https://doi.org/10.14257/ijmue.2015.10.12.07.
3. laramies. *theHarvester: E-mails, subdomains and names Harvester* - OSINT. GitHub. URL: https://github.com/laramies/theHarvester (дата звернення: 14.05.2025).
4. ElevenPaths. *FOCA - Metadata and Information Gathering Tool* : [Електронний ресурс]. 2023. URL: https://www.elevenpaths.com/technology/foca (дата звернення: 11.05.2025).
5. lanmaster53. *Recon-ng - Web Reconnaissance Framework* : [Електронний ресурс]. 2023. GitHub. URL: https://github.com/lanmaster53/recon-ng (дата звернення: 17.05.2025).
6. European Parliament and Council. *General Data Protection Regulation* (GDPR).
7. Bazzell M. *Open Source Intelligence Techniques: Resources for Searching and Analyzing.* 6th ed. 2018. 58 p.
8. Rumsey S. *How to Find Information*: *A guide for researchers*. SE. England: Open University Press, 2008. 223 p.
9. *How to find fishing letter*. Binance. URL: https://www.binance.com/uk-UA/support/faq/detail/360020817051 (дата звернення: 03.05.2025).
10. *Censys home*. URL: https://platform.censys.io/home (дата звернення: 05.05.2025).
11. Задерейко О.В., Логінова Н.І., Толокнов А.А. *Комп'ютерні мережі* : навчальний посібник [Електронний ресурс]. Одеса, 2022. 249 с. URL: https://hdl.handle.net/11300/19423 (дата звернення: 01.05.2025).

12. vtu10547. *Attacked dataset*. Kaggle. URL: https://www.kaggle.com/datasets/vtu10547/attacked/data (дата звернення: 09.05.2025).
13. Shodan (*вебсайт*). Матеріал з Вікіпедії – вільної енциклопедії. URL: https://uk.wikipedia.org/wiki/Shodan (дата звернення: 07.05.2025).
14. *CERT-UA*: Про нас. URL: https://cert.gov.ua/about-us (дата звернення: 02.05.2025).
15. *WHOIS lookup for onua.edu.ua*. DomainTools. URL: https://whois.domaintools.com/onua.edu.ua (дата звернення: 17.05.2025).
16. *Стратегія кібербезпеки України: проект*. РНБО України. URL: https://www.rnbo.gov.ua/files/2021/20strategiikyberbezpekiUkr.pdf (дата звернення: 17.05.2025).
17. ДСТУ 8845:2023. *Кібербезпека. Процеси розвідки. Розвідка на основі відкритих джерел (OSINT). Вимоги та настанови*. [Чинний від 2024-01-01]. Київ, 2023.
18. ДСТУ 8760:2022. *Кібербезпека. Загальні положення. Класифікація та термінологія кібератак*. [Чинний від 2023-01-01]. Київ, 2022.
19. *SecurityTrails*. URL: https://securitytrails.com/app/auth/login?return=/app/account/credentials (дата звернення: 17.05.2025).
20. *OSINT Framework*. URL: https://www.osintframework.com/ (дата звернення: 17.05.2025).
21. *Maltego*. URL: https://app.maltego.com/ (дата звернення: 15.05.2025).

ДОДАТОК А. Приклад скрейпера на Python

```
import requests
from bs4 import BeautifulSoup

# URL цільового ресурсу (приклад)
url = 'https://example-security-blog.com/new-threats'
# Ключові слова для пошуку
keywords = ['vulnerability', 'exploit', 'malware_campaign']

try:
    response = requests.get(url, timeout=10)
    response.raise_for_status() # Перевірка на помилки HTTP
    soup = BeautifulSoup(response.text, 'html.parser')

    # Пошук ключових слів у тексті сторінки (приклад пошуку в тегах <p>)
    page_text = ' '.join(p.get_text().lower() for p in
soup.find_all('p'))

    found_keywords = [kw for kw in keywords if kw in page_text]
    if found_keywords:

        print(f"Знайдено ключові слова на {url}: {',
'.join(found_keywords)}")

    else:
        print(f"Ключових слів на {url} не знайдено.")

except requests.exceptions.RequestException as e:
    print(f"Помилка при доступі до {url}: {e}")

except Exception as e:
    print(f"Виникла помилка: {e}")
```

ДОДАТОК Б. Приклад використання гіпотетичного API для перевірки IP-адреси (Python)

```
import requests
import json

# Гіпотетичний URL API та ключ
api_url = 'https://api.threatintel-provider.com/v1/ip-check'
api_key = 'YOUR_API_KEY'
ip_to_check = '8.8.8.8' # IP для перевірки
headers = {'Authorization': f'Bearer {api_key}'}
params = {'ip': ip_to_check}
try:

    response = requests.get(api_url, headers=headers, params=params,
timeout=10)

    response.raise_for_status()
    data = response.json() # Отримання даних у форматі JSON

    # Приклад обробки відповіді
    if data.get('is_malicious'):
        print(f"IP-адреса {ip_to_check} позначена як зловмисна.")
        print(f"Деталі: {data.get('details', 'N/A')}")
    else:
        print(f"IP-адреса {ip_to_check} не позначена як зловмисна.")

except requests.exceptions.RequestException as e:
    print(f"Помилка запиту до API: {e}")

except json.JSONDecodeError:
    print("Помилка декодування відповіді JSON від API.")

except Exception as e:
    print(f"Виникла помилка: {e}")
```

ДОДАТОК В. Концептуальний приклад підрахунку ключових слів (Python)

```
 def count_keywords(text_data, keywords):
    """Підраховує входження ключових слів у тексті."""
    text_lower = text_data.lower() # Переведення тексту в нижній
регістр
    keyword_counts = {kw: 0 for kw in keywords}

    for kw in keywords:
        # Простий підрахунок входжень (можна вдосконалити)
        keyword_counts[kw] = text_lower.count(kw.lower())

    return keyword_counts

# Приклад використання:
texts_from_forum = [
    "Discussion about new CVE-2025-XXXX exploit.",
    "Anyone selling access to ExampleCorp?",
    "Looking for partners for a new campaign targetting ExampleCorp.",
    "How to use the CVE-2025-XXXX PoC?"
]
target_keywords = ["CVE-2025-XXXX", "ExampleCorp", "exploit"]
total_counts = {kw: 0 for kw in target_keywords}

for text in texts_from_forum:
    counts = count_keywords(text, target_keywords)
    for kw, count in counts.items():
        total_counts[kw] += count

print("Загальна кількість знайдених ключових слів:", total_counts)
# Вивід: Загальна кількість знайдених ключових слів: {'CVE-2025-XXXX':
2, 'ExampleCorp': 2, 'exploit': 1}
```

ДОДАТОК Г . Концептуальний приклад сентимент-аналізу (Python, використовуючи простий словник)

```
   def simple_sentiment(text, positive_words, negative_words):
    """Проста оцінка тональності на основі словників."""
    text_lower = text.lower()
    score = 0

    for word in positive_words:
        score += text_lower.count(word)
    for word in negative_words:
        score -= text_lower.count(word)

    if score > 0:
        return "Positive"
    elif score < 0:
        return "Negative"
    else:
        return "Neutral"

# Приклад використання:
positive = ["good", "stable", "secure", "patched"]
negative = ["vulnerable", "exploit", "down", "attack", "compromised"]

post = "The new patch fixed the vulnerable component, but the exploit is
still discussed."

sentiment = simple_sentiment(post, positive, negative)
print(f"Тональність посту: {sentiment}")
# Вивід: Тональність посту: Negative
```