

Лабораторна робота № 7(3.1.18)

Пошук інформації про організацію

Хід роботи:

Частина 1: Пошук інформації про витоки Email

Онлайн сервіси для перевірки витоків

Список сервісів:

- haveibeenpwned.com
- f-secure.com
- hacknotice.com
- breachdirectory.com
- keepersecurity.com

Крок 1: Перевірка статусу вашої email адреси

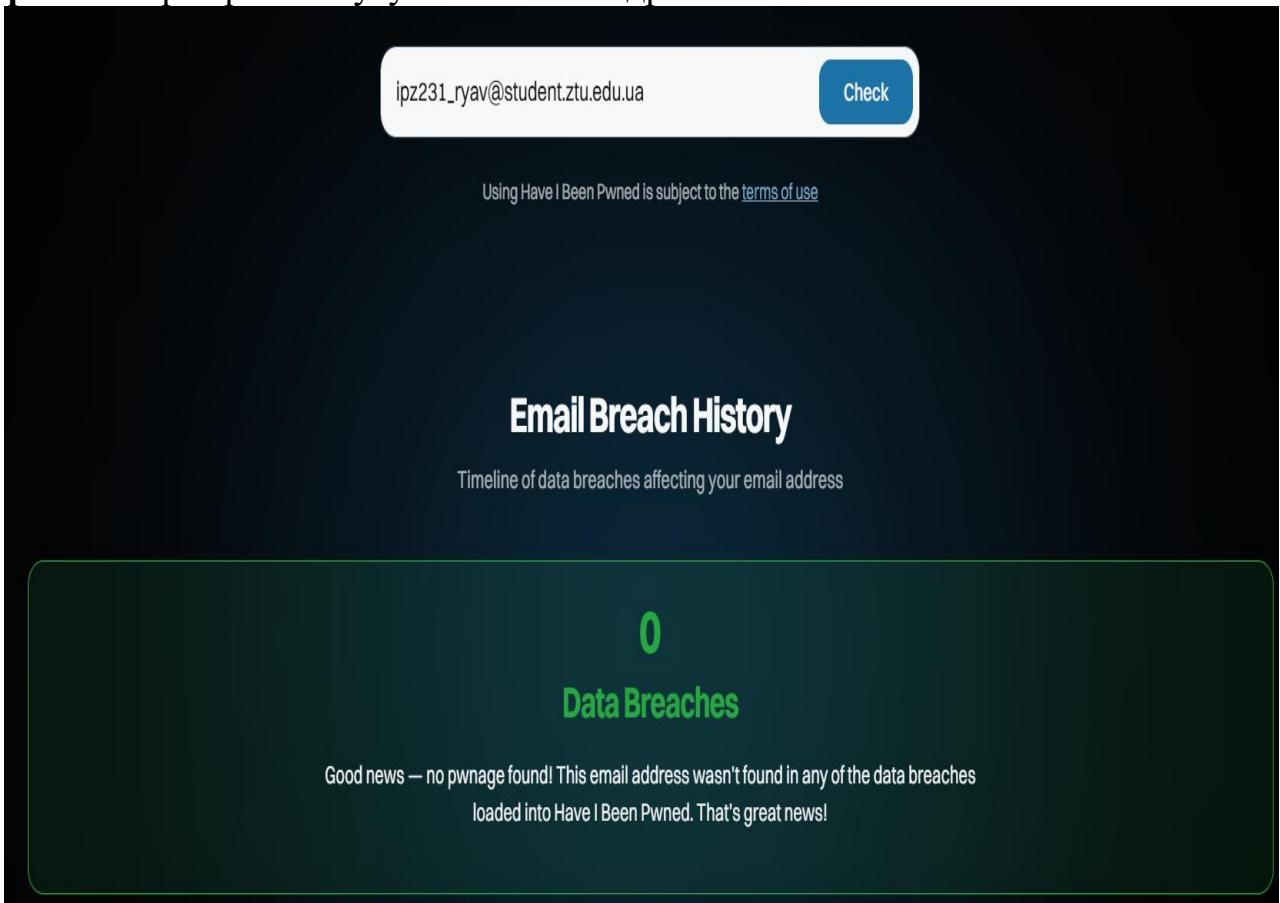


Рис. 1. Перевірка на витоки на сайті <https://haveibeenpwned.com>.

Питання: Чи були ваші email адреси частиною витоку? Якщо так, у яких витоках вони були розкриті?

Відповідь:

Ні, моя пошта не фігурувала в жодному витоку.

Крок 2: Використання інструменту для пошуку email адрес для домену

Змн.	Арк.	№ докум.	Підпис	Дата	ДУ «Житомирська політехніка».23.121.26.000 – Пр7(3.1.18)		
Розроб.	Рижсенко Я.В				Літ.	Арк.	Аркушів
Перевір.	Покотило О.А.					1	12
Керівник					ФІКТ Гр. ІПЗ-23-1[2]		
Н. контр.							
Зав. каф.							

Звіт з лабораторної роботи

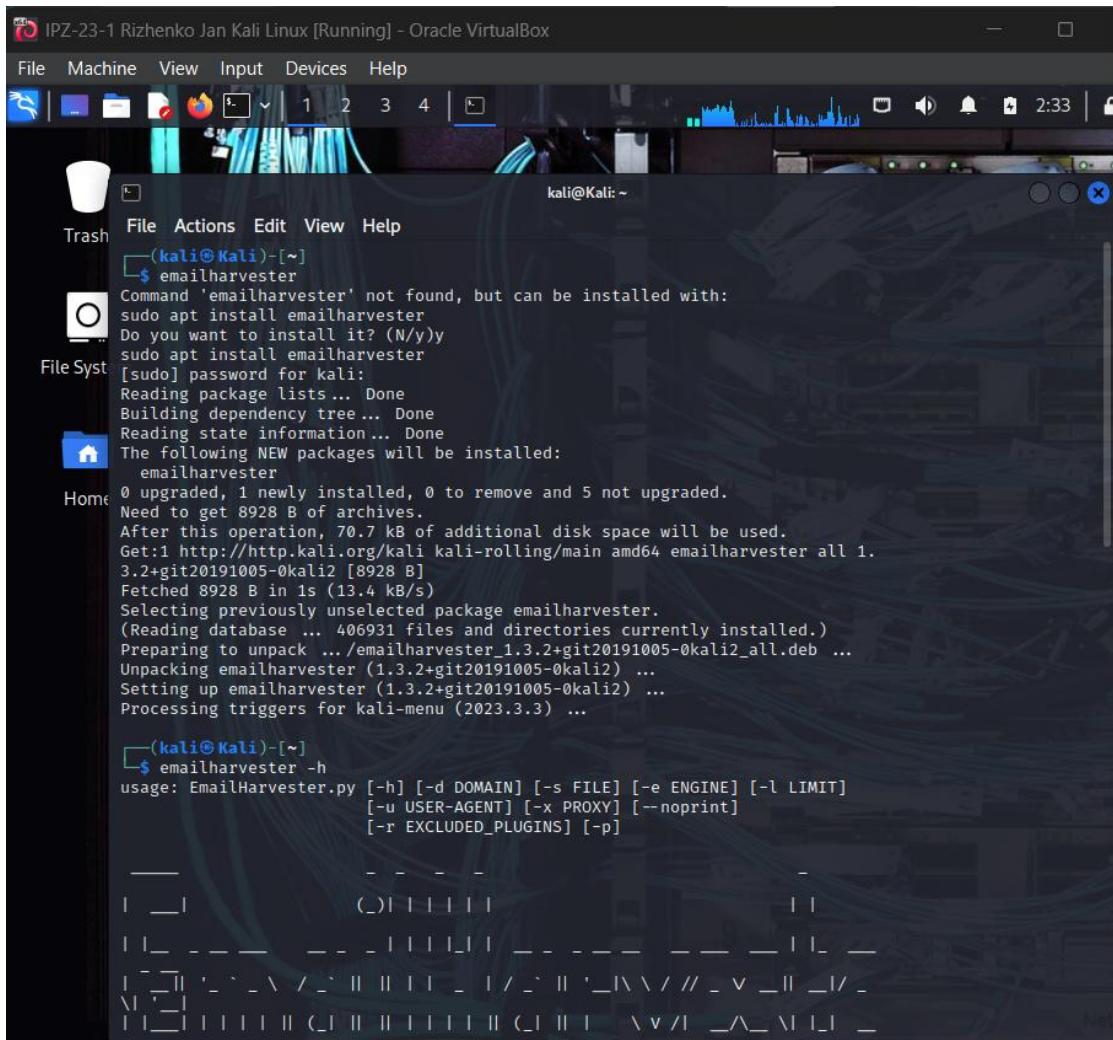


Рис. 2. Робота з emailharvester.

Питання: Що робить опція -d?

Відповідь:

Опція `-d` (або `--domain`) використовується для вказання домену, який потрібно дослідити для пошуку email адрес. EmailHarvester буде шукати всі публічно доступні email адреси, пов'язані з вказаним доменом, використовуючи різні пошукові системи.

Приклад: -d example.com буде шукати всі email адреси, що закінчуються на @example.com

Команди для дослідження доменів:

Пошук email адрес для домену h4cker.org

emailharvester-d-h4cker.org

Пошук з обмеженням результатів

emailharvester -d h4cker.org -l 100

Попуку через всі попукові системи

emailharvester -d h4cker.org -e all

Поиск для hackxor.net

emailharvester -d hackxor.net

Почує для scanme nmap.org

emailharvester -d scanme.nmap.org

		<i>Риженко Я.В</i>				
		<i>Покотило О.А.</i>				
Змн.	Арк.	№ докум.	Підпис	Лата	ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк. 2

Збереження результатів у файл

emailharvester -d h4cker.org -s h4cker_emails

Збереження з повним шляхом

emailharvester -d h4cker.org -s /home/kali/Documents/h4cker_emails

Тільки унікальні результати

emailharvester -d h4cker.org -u -s results

Перевірка знайдених email адрес:

Переглянути створені файли

ls -la /usr/share/emailharvester/

Переглянути текстовий файл

cat /usr/share/emailharvester/h4cker_emails.txt

Переглянути XML файл

cat /usr/share/emailharvester/h4cker_emails.xml

Підрахувати кількість знайдених адрес

cat /usr/share/emailharvester/h4cker_emails.txt | wc -l

Альтернативні інструменти для збору email адрес:

theHarvester (вбудований в Kali)

theHarvester -d h4cker.org -b google

theHarvester з множинними джерелами

theHarvester -d h4cker.org -b all

theHarvester з збереженням результатів

theHarvester -d h4cker.org -b all -f h4cker_harvest

hunter.io через curl (потрібен API ключ)

curl "https://api.hunter.io/v2/domain-search?domain=h4cker.org&api_key=YOUR_API_KEY"

Крок 3: Використання SpiderFoot для дослідження email адрес

Команди для SpiderFoot:

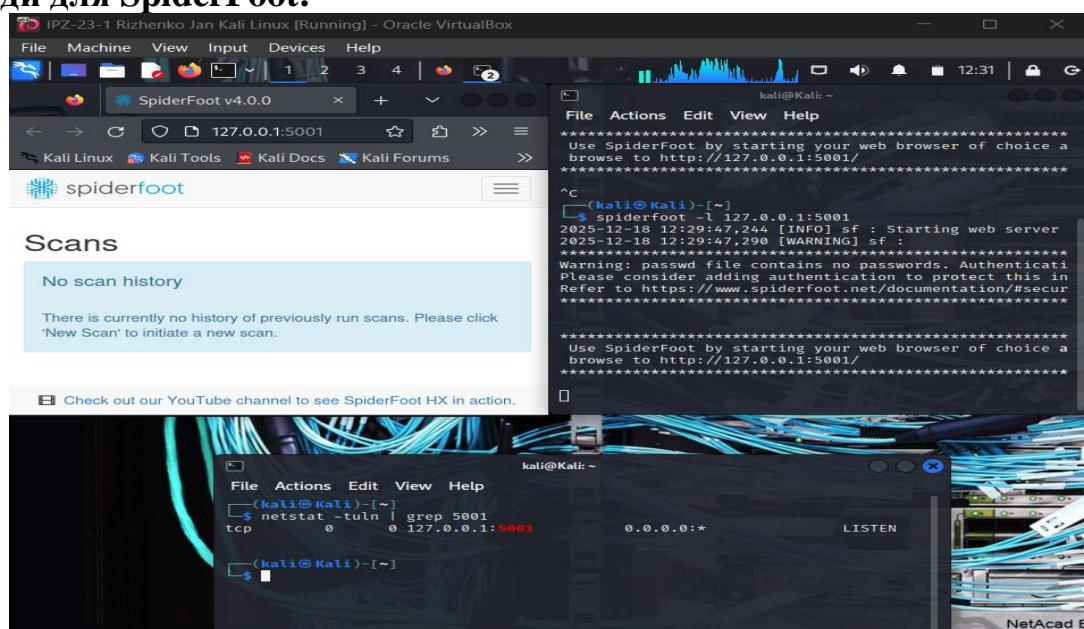


Рис. 3. Запуск та перевірка сервера.

		Рижсенко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				3
Змн.	Арк.	№ докум.	Підпис	Дата		

Модулі для email сканування:

- Ahmia (Dark web search)
- Account Finder (Social media accounts)
- Archive.org (Historical data)
- Bing (Search engine)
- Leak-Lookup (Data breach check)
- CommonCrawl (Web archive)
- Dehashed (Breach database)
- DuckDuckGo (Privacy search)
- EmailCrawlr (Email finder)
- Have I Been Pwned (Breach check)
- PasteBin (Code snippets)
- GitHub (Code repositories)

Корисні модулі для email аналізу:

Have I Been Pwned - Перевірка витоків даних | Так (безкоштовний) |

Account Finder - Пошук облікових записів на 200+ сайтах | Hi |

EmailCrawlr - Пошук email адрес | Так |

Dehashed - База даних витоків | Так (платний) |

Leak-Lookup - Перевірка витоків | Так |

GitHub - Пошук в репозиторіях | Hi (з обмеженнями) |

PasteBin - Пошук в code snippets | Hi |

Archive.org - Історичні дані | Hi |

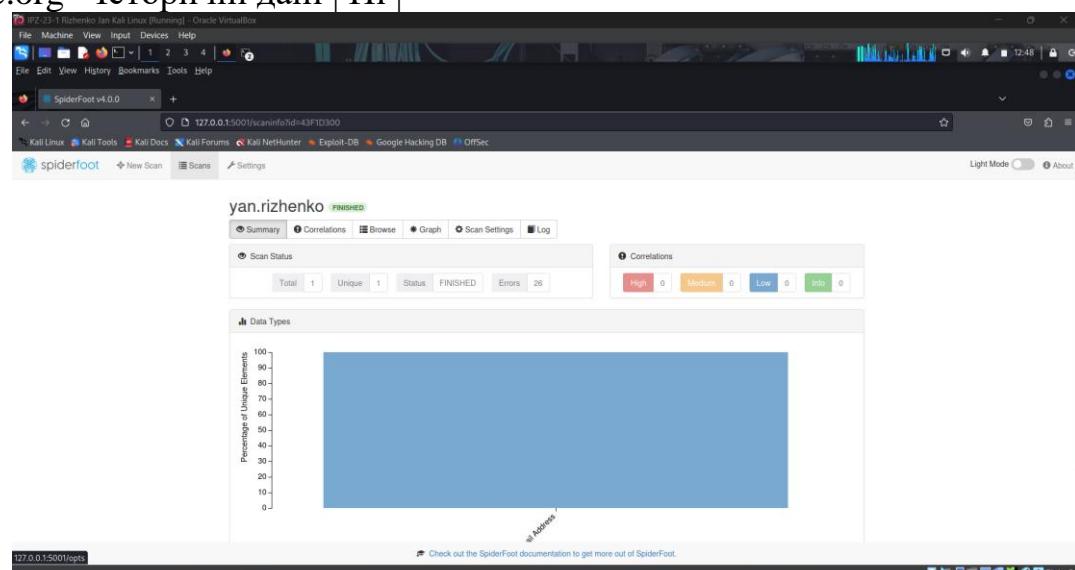


Рис. 4. Результат сканування ЗІ ВСІМА модулями.

Аналіз результатів:

Результати можуть показати:

- Облікові записи в соціальних мережах
- Витоки даних (breaches)
- Згадки в code repositories
- Історичні дані з Archive.org
- Публічні пости на форумах

		Рижсенко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		4

- Інформацію з PasteBin та GitHub

Частина 2: Перегляд метаданих файлів

Крок 1: Встановлення ExifTool

Команди для встановлення:

Оновити список пакетів

sudo apt update

Встановлення через apt

sudo apt install exiftool -y

Перевірити встановлення

exiftool -ver

Переглянути довідку

exiftool -h

Переглянути всі підтримувані теги

exiftool -list

Переглянути підтримувані типи файлів

exiftool -listf

Таблиця підтримуваних форматів файлів:

ТИП	ФОРМАТИ ФАЙЛІВ (РОЗШИРЕННЯ)
Documents	PDF, DOC, DOCX, XLS, XLSX, PPT, PPTX, ODT, ODS, ODP, RTF
Audio	MP3, WAV, AAC, FLAC, OGG, M4A, WMA, APE
Video	MP4, AVI, MOV, MKV, WMV, FLV, MPEG, 3GP, M4V
Graphics	JPG, JPEG, PNG, GIF, TIFF, BMP, PSD, SVG, WEBP, HEIC, RAW, CR2, NEF
Archives	ZIP, RAR, 7Z, TAR, GZ, BZ2, ISO

Крок 2: Використання ExifTool

Базові команди ExifTool:

Переглянути метадані одного файлу

exiftool image.jpg

Переглянути метадані всіх файлів у директорії

exiftool /path/to/directory/

Переглянути метадані рекурсивно (всі піддиректорії)

exiftool -r /path/to/directory/

Переглянути конкретний тег

exiftool -GPS* image.jpg

Експортувати в CSV

exiftool -csv image.jpg > metadata.csv

Експортувати директорію в CSV

exiftool -csv /path/to/directory/ > all_metadata.csv

Експортувати в HTML

exiftool -h image.jpg > metadata.html

Видалити всі метадані з файлу

exiftool -all= image.jpg

Видалити метадані зі збереженням оригіналу

exiftool -all= -o clean_image.jpg image.jpg

		Рижсенко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		5

Пошук файлів через Google Dorks (GHDB):

Пошук PDF файлів

site:example.com filetype:pdf

Пошук документів Word

site:example.com filetype:doc OR filetype:docx

Пошук презентацій

site:example.com filetype:ppt OR filetype:pptx

Пошук таблиць Excel

site:example.com filetype:xls OR filetype:xlsx

Пошук зображень

site:example.com filetype:jpg OR filetype/png

Пошук з конфіденціальною інформацією

site:example.com "confidential" filetype:pdf

site:example.com "internal use only" filetype:doc

Пошук резюме

intitle:"resume" OR intitle:"CV" filetype:pdf

Пошук фінансових документів

site:example.com "budget" OR "financial" filetype:xls

Завантаження файлів:

Використання wget для завантаження

wget https://example.com/document.pdf

Завантаження в конкретну директорію

wget -P ~/Downloads/ https://example.com/document.pdf

Завантаження множинних файлів з файлу

wget -i urls.txt

Використання curl

curl -O https://example.com/document.pdf

Детальний аналіз метаданих:

Аналіз одного файлу

exiftool document.pdf

Пошук конкретних тегів (автор)

exiftool -Author document.pdf

Пошук дати створення

exiftool -CreateDate -ModifyDate document.pdf

Пошук GPS координат в зображеннях

exiftool -GPS* photo.jpg

Пошук інформації про камеру

exiftool -Make -Model -LensModel photo.jpg

Пошук програмного забезпечення

exiftool -Software -Creator document.pdf

Аналіз всіх файлів у директорії

exiftool .

Експорт результатів у CSV для аналізу

exiftool -csv . > metadata_analysis.csv

		Рижсенко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		6

Відкрити CSV у LibreOffice Calc

libreoffice --calc metadata_analysis.csv

Приклади корисних метаданих:

Типові метадані в різних типах файлів:

PDF документи:

- Author (Автор)
- Creator (Програма створення)
- Producer (PDF генератор)
- CreateDate (Дата створення)
- ModDate (Дата модифікації)
- Title (Заголовок)
- Subject (Тема)
- Keywords (Ключові слова)

Зображення (JPEG, PNG):

- Make (Виробник камери)
- Model (Модель камери)
- Software (Програма обробки)
- DateTime (Дата та час)
- GPS coordinates (GPS координати)
- ExifImageWidth/Height (Розміри)
- ISO, Aperture, ShutterSpeed (Налаштування камери)

Microsoft Office документи:

- Author (Автор)
- LastModifiedBy (Останній редактор)
- Company (Компанія)
- Manager (Менеджер)
- CreateDate (Дата створення)
- ModifyDate (Дата модифікації)
- RevisionNumber (Номер ревізії)
- TotalEditTime (Загальний час редагування)

Аудіо файли:

- Artist (Виконавець)
- Album (Альбом)
- Year (Рік)
- Genre (Жанр)
- Encoder (Кодувальник)
- Duration (Тривалість)

Відео файли:

- Duration (Тривалість)
- VideoCodec (Відео кодек)
- AudioCodec (Аудіо кодек)
- FrameRate (Частота кадрів)
- DateTimeOriginal (Дата запису)

		Риженко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		7

Питання: Чи знайшли ви будь-яку інформацію, яка може бути корисною для етичних хакерів?

Відповідь:

Так, в метаданих файлів можна знайти багато корисної інформації:

1. Інформація про персонал:

- Імена авторів: Розкривають імена співробітників
- Email адреси: У деяких файлах зберігається email
- Імена користувачів: Системні імена користувачів (наприклад, "john.smith")
- Посади: Можуть міститися в полі "Manager" або "Title"

Приклад:

Author: John Smith

LastModifiedBy: j.smith

Company: Acme Corporation

Manager: Sarah Johnson

2. Технічна інформація:

- Програмне забезпечення: Версії програм (Microsoft Word 2016, Adobe Photoshop CS6)
- Операційна система: Може вказувати на ОС (Windows 10, macOS)
- Версії бібліотек: Наприклад, "gd-jpeg v1.0" (вразлива версія PHP GD)
- Мережеві шляхи: Інколи містять внутрішні шляхи ('\\server01\\shared\\docs\\')

Приклад вразливості:

Creator: gd-jpeg v1.0 (using IJG JPEG v62)

Це вказує на стару версію PHP GD library з відомими вразливостями

3. Географічна інформація:

- GPS координати: Точне місце знаходження, де було зроблено фото

- Часові зони: Можуть вказувати на локацію офісу

Приклад:

GPS Position: 37°46'30.0"N 122°25'10.0"W

Координати офісу в Сан-Франциско

4. Організаційна структура:

Назви компаній: У полі "Company"

Відділи: Іноді вказані в метаданих

Проекти: Назви проектів у полі "Subject" або "Keywords"

Внутрішні коди: Номери проектів, референси

5. Часові дані:

Робочі години: Час створення/модифікації може показати робочий графік

Патерни роботи: Частота редагування документів

Давність: Старі файли можуть містити застарілу інформацію

6. Внутрішня інфраструктура:

Імена серверів: \\FILESERVER01\\shared\\

Мережеві шляхи: C:\\Users\\john.smith\\Documents\\

Домени: Internal domain names

Приклад використання для атаки:

Знайдена інформація з метаданих:

		Риженко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		8

Author: j.smith@company.com
Software: Microsoft Office 2010 (unpatched)
Last Modified: 2015-03-15
Network Path: \\FILESERVER01\Finance\Reports\

Можливості для атак:

1. Підібрати пароль для j.smith@company.com
2. Використати експлойти для Office 2010
3. Цілити внутрішній сервер FILESERVER01
4. Провести соціальну інженерію проти John Smith

Інші приклади корисних знахідок:

Камера iPhone з GPS

exiftool photo.jpg

Make: Apple

Model: iPhone 12 Pro

GPS Position: 40.7128° N, 74.0060° W

Date/Time Original: 2023:12:18 14:30:00

Внутрішній документ компанії

exiftool confidential.pdf

Author: Sarah Johnson (IT Manager)

Creator: Microsoft Word 2019

Title: Network Architecture - Internal Use Only

Keywords: firewall, VPN, credentials, backup

Company: Acme Corp

Create Date: 2023:11:15 09:45:23

Producer: Acme-Laptop-SJ01

PHP generated image з вразливістю

exiftool generated.jpg

Software: gd-jpeg v1.0 (using IJG JPEG v62)

Вразлива до CVE-2013-2110 та інших

Дослідження вразливостей:

Пошук вразливостей PHP GD

searchsploit php gd

Пошук вразливостей Microsoft Office 2010

searchsploit microsoft office 2010

Онлайн пошук CVE:

<https://cve.mitre.org>

"PHP GD vulnerability" або "Office 2010 CVE"

Додаткові інструменти та техніки

Додаткові інструменти для metadata аналізу:

FOCA (Fingerprinting Organizations with Collected Archives)

Windows-only, але може працювати через Wine

wine FOCA.exe

Metagoofil (вбудований в Kali)

metagoofil -d example.com -t pdf,doc,xls -l 100 -o results/

		Рижсенко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		9

Metadata Anonymization Toolkit (MAT2)

sudo apt install mat2

mat2 --show document.pdf

mat2 --inplace document.pdf - Видалити метадані

Strings (аналіз бінарних даних)

strings document.pdf | grep -i "author|creator|producer"

pdfinfo (аналіз PDF)

pdfinfo document.pdf

Exiv2 (альтернатива для зображенень)

exiv2 photo.jpg

Автоматизація збору та аналізу:

Скрипт для масового завантаження та аналізу

cat > analyze_metadata.sh << 'EOF'

#!/bin/bash

DOMAIN=\$1

OUTPUT_DIR="metadata_\${DOMAIN}"

mkdir -p \$OUTPUT_DIR

Пошук файлів через Google

echo "[+] Searching for files on \$DOMAIN..."

googler -n 50 "site:\$DOMAIN filetype:pdf OR filetype:docx OR filetype:xlsx" > urls.txt

Завантаження файлів

echo "[+] Downloading files..."

wget -i urls.txt -P \$OUTPUT_DIR/files/

Аналіз метаданих

echo "[+] Analyzing metadata..."

exiftool -csv -r \$OUTPUT_DIR/files/ > \$OUTPUT_DIR/metadata.csv

Вилучення email адрес

echo "[+] Extracting email addresses..."

grep -Eiorh '\b[A-Za-z0-9._%+-]+@[A-Za-z0-9.-]+\.[A-Z|a-z]{2,}\b'

\$OUTPUT_DIR/files/ | sort -u > \$OUTPUT_DIR/emails.txt

Вилучення авторів

echo "[+] Extracting authors..."

exiftool -Author -csv \$OUTPUT_DIR/files/ | sort -u > \$OUTPUT_DIR/authors.txt

echo "[+] Done! Results in \$OUTPUT_DIR/"

EOF

Зробити скрипт виконуваним

chmod +x analyze_metadata.sh

Запустити

./analyze_metadata.sh example.com

Reflection (Рефлексія)

Питання: Як би ви описали цей процес? Чи виявили ви велику кількість інформації? Що це говорить про корпоративний процес розвідки персоналу?

Відповідь:

		Риженко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		10

Опис процесу:

Процес збору інформації про організацію та її персонал через OSINT є **багатошаровим, систематичним та надзвичайно продуктивним**. Він включає:

1. **Пасивний характер:** Вся інформація збирається без прямої взаємодії з цільовою системою, що робить її абсолютно безпечною для атакуючого.
2. **Кумулятивний ефект:** Невелика інформація з різних джерел складається в повну картину організації.
3. **Прогресивне поглиблення:** Кожен знайдений елемент (email, ім'я, домен) веде до нових джерел інформації.

Кількість виявленої інформації:

Так, кількість зібраної інформації **вражаюча і часто шокуюча**:

З витоків email:

- Сотні або тисячі email адрес співробітників
- Паролі (хоч і хешовані, але часто зламані)
- Особисті дані (імена, телефони, дати народження)
- Історія витоків за останні 10+ років

З метаданих файлів:

- Імена співробітників та їх позиції
- Організаційна структура компанії
- Технологічний стек (програмне забезпечення, версії)
- Внутрішня мережева інфраструктура
- GPS координатиофісів
- Робочі патерни та графіки

З публічних джерел:

- Субдомени та внутрішні системи (через crt.sh)
- Соціальні медіа профілі співробітників
- Публікації в GitHub, PasteBin
- Історичні дані з Archive.org
- Технічна документація

Що це говорить про корпоративний процес розвідки:

1. Величезна поверхня атаки:

- Сучасні організації залишають **величезний цифровий слід**
- Кожен співробітник є потенційною точкою входу
- Публічно доступна інформація часто недооцінюється

2. Недостатня обізнаність про безпеку:

- Більшість організацій **не усвідомлюють**, скільки інформації вони розкривають
- Співробітники рідко очищають метадані з документів
- Витоки даних накопичуються роками і залишаються актуальними

3. Складність захисту:

- Неможливо повністю контролювати, яка інформація стає публічною
- Старі дані залишаються в архівах назавжди
- Витоки з третіх сторін (LinkedIn, партнери) неможливо запобігти

		Риженко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		11

4. Асиметрія зусиль:

- Атакуючому потрібно **години** для збору інформації
- Організації потрібні **роки** для побудови культури безпеки
- Автоматизація робить розвідку ще простішою

5. Важливість OSINT в modern pentesting:

- До 70-80% інформації для атаки можна зібрати через OSINT
- Соціальна інженерія стає значно ефективнішою
- Targeted attacks можуть бути підготовлені дуже детально

Практичні висновки:

Для захисників (blue team):

- Регулярно проводити OSINT аудит власної організації
- Навчати співробітників очищати метадані
- Моніорити витоки даних та реагувати швидко
- Мінімізувати цифровий слід організації
- Використовувати privacy-preserving реєстрацію доменів

Для атакуючих (red team/pentesters):

- OSINT є **обов'язковим першим етапом** будь-якого пентесту
- Комбінування різних джерел дає синергетичний ефект
- Автоматизація збору даних економить час
- Метадані часто містять "золоті самородки" інформації
- Соціальні витоки даних (LinkedIn, Facebook) надзвичайно цінні

Цей процес демонструє, що **інформаційна безпека - це не тільки технічний захист**, але й культура обізнаності на всіх рівнях організації. Навіть найсильніші технічні засоби захисту можуть бути обійдені, якщо атакуючий має детальну інформацію про персонал, інфраструктуру та процеси організації, зібрану через абсолютно легальні та пасивні методи OSINT.

Кожна організація повинна регулярно проводити **OSINT аудит самих себе**, щоб зрозуміти, яку інформацію про них може зібрати потенційний атакуючий, і вжити заходів для мінімізації цього ризику.

Висновок

У ході виконання лабораторної роботи було досліджено методи збору інформації про організацію через витоки email адрес та аналіз метаданих файлів. Використано онлайн сервіси для перевірки витоків (haveibeenpwned.com, breachdirectory.com), інструменти для збору email адрес (EmailHarvester, theHarvester, SpiderFoot), та утиліти для аналізу метаданих (ExifTool, Metagoofil). Встановлено, що публічно доступна інформація про організацію є надзвичайно багатою і часто недооціненою з точки зору безпеки. Метадані файлів можуть розкривати імена співробітників, технологічний стек, внутрішню інфраструктуру та навіть GPS координати. Процес OSINT розвідки демонструє, що більшість інформації, необхідної для успішної атаки, може бути зібрана абсолютно пасивними методами без прямої взаємодії з цільовою системою, що підкреслює критичну важливість культури інформаційної безпеки та регулярних OSINT аудитів власної організації.

		Риженко Я.В			ДУ «Житомирська політехніка».23.121.26.000 – Лр7(3.1.18)	Арк.
		Покотило О.А.				
Змн.	Арк.	№ докум.	Підпис	Дата		12