Lab - Using OSINT Tools

Objectives

In this lab, you will explore several OSINT tools that are commonly used by pentesters.

- Examine OSINT resources
- Use SpiderFoot
- Investigate Recon-ng
- Find interesting files with Recon-ng

Background / Scenario

When performing information gathering activities, passive reconnaissance uses open, publicly accessible data to guide active reconnaissance efforts and to gather information about the enterprise and employees. In OSINT, it is the data that is open source. OSINT tools may or may not be open source. Some tools are free and open, others require registration to use free versions, and others require a fee for use. OSINT commonly uses data sources that are available to any hacker, so part of the PenTesting effort is to report on sensitive information that is commonly available in order to evaluate vulnerabilities that it may cause. The objectives of OSINT are:

- To determine the digital footprint of the organization.
- Determine what data about the organization is available to cyber criminals.

Required Resources

- Kali VM customized for Ethical Hacker course
- Internet access

Instructions

Part 1: Examine OSINT Resources

Step 1: Access the OSINT Framework

The OSINT Framework is a useful way to visualize the OSINT tools and resources that are available. Unfortunately, it has become a bit outdated, with some resources no longer available. It is still valuable to help you understand the types of tools available and their uses. In many cases, the links are still good.

a. Go to the OSINT Framework site at https://osintframework.com/.
b. You will see a vertical tree-like structure that consists of categories of OSINT tools and resources that can be reached from the framework. Click **Username** at the top of the tree. You will then see two subcategories appear. Click each to open the resource trees for each subcategory. Note in the upper-right corner of the page is a legend that identifies the type of resource.

c. Under **Username Search Engines**, click **"WhatsMyName(T)"**.
d. The link takes you to a Git repository for the WhatsMyName project. In the **README.md** content for the tool, the various sites that implement WhatsMyName are listed. Feel free to explore these, but we will click the first link https://whatsmyname.app/ to visit a free website that implements WhatsMyName.

   The parent organization for the site, https://www.osintcombine.com/, has several interesting free tools available.

e. In the search box, type in a few usernames, each on a separate line. Use your own usernames or others that you find. Try searching the internet for **common username wordlist** for other potential search terms. You can filter the results based on the category filters, but for now, just click the green magnifying glass button to start the search.

   In a pentest, you would use another tool, such as **SpiderFoot** (below) to find usernames in email addresses that are associated with a company or domain.

f. Investigate the results. You can open the links to the accounts either from the green rectangles or the table of results.
g. WhatsMyName provides a very flexible report of the results. The results table can be sorted by column, and you can export the results as CSV or PDF for reporting purposes. In addition, you can easily filter by username and search within the results. Finally, you get links for the profile pages for the users at many different sites.

What is the value of doing username searches and account enumeration?

Answer Area

<!----><!---->

Step 2: Investigate SMART - Start Me Aggregated Resource Tool.

The start.me web service is a popular bookmark manager and productivity tool. The people at My OSINT Training (MOT) have set up a search system that finds all OSINT-related links that people have bookmarked and shared on start.me. There are many. You can enter OSINT-relevant search terms to find links to related resources.

a. Go to https://smart.myosint.training/.
b. In the search box, enter the term **usernames**. You will see a list of username-related OSINT tools that other people have found.
c. Open some of the links to review the resources. Be careful however, these websites come from public sources. Some may be malicious.
d. Choose some of the categories that you saw in the OSINT Framework and see what links appear.
e. Use this site to search for OSINT tools and resources to help you in your pentesting work.

Part 2: Use SpiderFoot

SpiderFoot is an automated OSINT scanner. It is included with Kali. SpiderFoot queries over 1000 open-information sources and presents the results in an easy-to-use GUI. SpiderFoot can also be run from a console. SpiderFoot seeds its scan with one of the following:

- Domain names
- IP addresses

- Subnet addresses
- Autonomous System Numbers (ASN)
- Email addresses
- Phone numbers
- Personal names

SpiderFoot offers the option of choosing scans based on use case, required data, and by SpiderFoot module. The use cases are:

- All – Get every possible piece of information about the target. This use case can take a very long time to complete.
- Footprint – Understand the target's network perimeter, associated identities and other information that is yielded by extensive web crawling and search engine use.
- Investigate – This is or targets that you suspect of malicious behavior. Footprinting, blacklist lookups, and other sources that report on malicious sites will be returned.
- Passive – This type of scan is used if it is undesirable for the target to suspect that it is being scanned. This is a form of passive OSINT.

Step 1: Start and run SpiderFoot.

In a terminal, enter the following command:

```
┌──(kali㉿Kali)-[~]
└─$ spiderfoot -l 127.0.0.1:5001
```

The command should run without errors. Open a browser and enter the IP address and port for the SpiderFoot GUI. You will see the SpiderFoot interface appear. If this is the first time that SpiderFoot has been opened in this VM, you will see the Scans screen. This screen displays a list of all the scans recently run. In this example it is empty.

Step 2: Explore SpiderFoot.

a. Before we get started, look at the scanners that SpiderFoot uses to build its reports. Go to the **Settings** tab.
b. The first two entries in the menu at the left are concerned with the operation of SpiderFoot. The entries below this are for the scanners that SpiderFoot uses. There are over 200 of them. Click the scanners to see their SpiderFoot module name, details about the scanner, and settings that can be made, if any. Complete the table below with some examples. The Scanner name is in the settings menu. The module name appears in the details for the scanner. All SpiderFoot modules are referred to as sfp_[module name].

   Hint: scanners with a lock next to them indicate an API key is necessary. Further information regarding the key requirements is provided in the details for the scanner. Click the "?" icon next to the API Settings option.

   Hint: You can interact with SpiderFoot from the terminal too. You can display all the modules that are available in SpiderFoot and pipe the output to a text file. Enter **spiderfoot –h** to view the command line options.

c. The **grep** command can then be used to search the file for keywords. This will not provide information about API requirements, but it will help you to make sense of the list of available modules.

```
┌──(kali㉿Kali)-[~]
```

```
└─$ spiderfoot -M | grep [search term]
```

d. Using the **grep** command and the GUI, complete the table below.

Note: Answers will vary. Some modules do the same thing or do multiple things.

Information Type	Scanner/Module Name	API key requred? Free?	Comments
Possible accounts associated with a domain	Account Finder sfp_accounts	No, N/A	Over 200 sites like eBay, Redditt, slash dot
Links that are associated with the target	Answer Area <!----><!---->	Answer Area <!----><!---->	Answer Area <!----><!---->
Email addresses associated with the target	Answer Area <!----><!---->	Answer Area <!----><!---->	Answer Area <!----><!---->
Domains and URLs that are associated with the target	Answer Area <!----><!---->	Answer Area <!----><!---->	Answer Area <!----><!---->
Geolocation information	Answer Area <!----><!---->	Answer Area <!----><!---->	Answer Area <!----><!---->
Data breach information	Answer Area <!----><!---->	Answer Area <!----><!---->	Answer Area <!----><!---->

Step 3: Run a SpiderFoot Scan for a Domain.

a. Click the **New Scan** tab in the GUI.
b. Enter a name for the scan and select a target. In this case, we will use **h4cker.org**.
c. You will scan by use case. Note that you can also scan by the type of information required or by selecting the individual scanner modules that you would like to use. By executing narrower scans, you can learn more about the modules and information that can be gathered.
d. Select the scan use case as **Footprint**.

Note: The **All** use case scan may use active scanning. Unless you have permission to scan the target, you should avoid this setting. To be completely safe, the Passive use case should avoid any problems with unauthorized scanning.

e. Click the **Run Scan Now** button.
f. You should see a bar graph appear. The scan statistics will start to increment, and new bars will appear in the graph as new results are obtained. Mouse over the bars for a summary of the findings for that data type.
g. SpiderFoot scans are very detailed and can take a very long time. Give this scan at least 30 minutes so that there is a nice collection of information. To get the most details, a scan could take hours. While the scan is running, you can browse the results.

Step 4: Investigate Scan Results.

a. Go back to the scan results, by clicking the **Scans** tab. You will see a table with the currently running scan and any previous scans displayed.
b. Click the black square in the right-most column of the scans table to stop the scan. Some information is not available until the scan is aborted or completed.
c. Click the name of the scan in the table to return to the scan view. You will be taken to the **Browse** tab. Each row in the table represents data found by the various modules. Some modules contribute to multiple types of data.
d. Investigate the results.

Step 5: Register API Keys (optional).

API keys will enhance the functionality of SpiderFoot. Some of these API keys require free registration. The pentesting tools that are available are constantly evolving. Some tools or services that were once free and open can become fee-based over time.

Note: Some APIs may limit your results after you have reached a prescribed number of uses.

a. Go to the **Settings** tab.
b. Find the four modules in the table below. Open the page for the module and complete the table including the type of information that module searches for. For each module in the table, click the ? next to the API option. Follow the instructions to get API keys for the four modules.

Module	Type of Information	Your API Key, etc
Builtwith	Answer Area <!----><!---->	Answer Area <!----><!---->
Hunter.io	Answer Area <!----><!---->	Answer Area <!----><!---->
Onion.link	Answer Area <!----><!---->	Answer Area <!----><!---->
IntelligenceX	Answer Area <!----><!---->	Answer Area <!----><!---->

c. Enter the API keys in the settings for each module. Be sure to save your changes.
d. Click **New Scan**. Go to the By Module tab. Select only the modules for which you have added API keys. All other modules should be unchecked.
e. Enter the target as **h4cker.org** and click **Start Scan**. Feel free to scan other domains but be sure to observe the terms and conditions of this course.

Step 6: Analyze Results of API Modules Scans.

a. This scan should not take very long.

b. Browse the scan to look at the results. Pay attention to the **Source Module** column. You should see some of the scanners that you configured with API keys.
c. Go to the Leak Site URL type in the table of results.

Which module contributed to this table?

Answer Area

<!----><!---->

d. Double-click several of the entries in the table Data Element column, right click, and select open in new tab.

What do you see?

Answer Area

<!----><!---->

We are just scratching the surface with what this tool can do. Try other searches and see what you can find.

Part 3: Investigate Recon-ng

Recon-ng is an OSINT framework that is similar to the Metasploit exploitation framework or the Social-Engineering Tooklit (SET). If consists of a series of modules that can be run in their own workspaces. The modules can be configured to run with option settings that are specific to the module. This simplifies running Recon-ng at the command line because options for the modules are independently set within the workspace. When you run the module, it uses these settings to perform its searches.

As the name suggests, Recon-ng is used to perform a wide range of reconnaissance activities on different settings that you provide. Some modules are available with the Kali installation and others are available for download and installation in the Recon-ng modules marketplace.

Step 1: Create a workspace.

Recon-ng has auto complete. Press the tab button to complete commands and command options. Use the tab key twice to list the available commands and options at different places in the command line. This is very handy.

a. To run Recon-ng, open a new terminal window and enter **recon-ng**. You can also start the program by going to the Kali tools menu, searching for the app, and clicking the icon.
b. Note that the terminal prompt changes to indicate that you are working within the Recon-ng framework. Enter **help** to get a sense of the commands that are available.
c. Recon-ng uses workspaces to isolate investigations from one another. Workspaces can be created for different parts of a test or different customers for example. Type **workspaces help** to view options for the workspaces command.

How can you display the available workspaces?

Answer Area

<!----><!---->

How can you remove a workspace?

Answer Area

<!----><!---->

d. Create a workspace named **test** by entering **workspaces create** followed by the workspace name. Note that the prompt has changed to indicate that you are in this workspace.
e. Type **help** to see the commands that are available within workspaces.

What command will exit the workspace and return to the main Recon-ng prompt?

Answer Area

<!----><!---->

Step 2: Investigate modules.

Recon-ng is a modular framework. Modules are Python programs with different functions. They are stored in an external marketplace that permits developers to create their own modules and contribute them for use by others.

Return to the Recon-ng prompt. Enter the **modules search** command. This will display the currently installed modules.

How many modules are currently available to you?

Answer Area

<!----><!---->

Step 3: Investigate the module marketplace.

Recon-ng will not function without modules. In this step, we will install modules from the Recon-ng marketplace. The module marketplace is a GitHub public repository. Search the web for **recon-ng-marketplace** to view the repository. Explore the folders to learn more about the modules.

a. In the terminal, view help for the **marketplace** command. Use the **search** option to list all the modules that are currently available.

```
[recon-ng][default] > marketplace search
```

b. Note that the modules are organized by their category and type. This appears as a path prepended to the name of the module. You can filter the output by adding a search term to the marketplace search command. Try a few different search terms that are related to OSINT information to get a sense of the modules that are available.

The module tables have columns for **D** and **K**. Search for shodan modules. What are the requirements for these modules?

Answer Area

<!----><!---->

c. To learn more about individual modules, use the **marketplace info** command followed by the full name of the module, including its category and type. It is easier to select the name of the module and copy and paste it into the command line.

Step 4: Install a new module.

Recon-ng accesses modules from the Github repository and downloads them to Kali when they are installed.

a. Search the marketplace modules using **bing** as a search term. Locate a module that requires no dependencies or API keys.

Which module did you find?

Answer Area

<!----><!---->

b. View information for this module.
c. To install the module, copy the full name, including the path, to the clipboard.
d. Enter the **marketplace install** command followed by the full name of the module.

```
[recon-ng][default] > marketplace install recon/domains-hosts/bing_domain_web
```

e. After installation, enter the **modules search** command to verify that the new module is now available.
f. Repeat the process to install the **hackertarget** module.

Step 5: Run the new modules

a. Create a new workspace. Name it as you wish.
b. To start working with a module, it must be initialized. Enter **modules load hackertarget** to begin working with the module. Note that the prompt changes to reflect the loaded module.
c. Each module is its own environment. The developers of recon-ng have taken care to keep the framework consistent, so the same commands are available for each module. However, the options can vary. Type **info** at the module prompt to view important details about the module.

What information is available for this module?

Answer Area

<!----><!---->

What is the only option for this module?

Answer Area

<!----><!---->

d. Instead of passing options at the command line, in Recon-ng you set the options and then enter a simple command to execute the module. Use the **options set source** command to set the only option for this module. Complete the command by specifying the target as **hackxor.net**.
e. Verify the option setting with the **info** command.
f. Type **run** to execute the module.
g. Inspect the output of the command. The output is stored in a database so you can refer to it later. The data that is stored is specific to the workplace in which it was gathered.
h. Enter the **dashboard** command. This queries the Recon-ng database and provides a summary of the information that has been gathered. It is specific to this workspace.

What is the Recon-ng data label for the subdomains that have been listed? How many were discovered?

Answer Area

<!----><!---->

i. The **show** command displays the data for specific categories. Enter the **show hosts** command to display the list of hosts that were discovered.
j. Now repeat the process with the **bing** module. Compare the results with the **hackertarget** module.

How many subdomains did the module find? How does this compare to the **hackertarget** module?

Answer Area

<!----><!---->

Step 6: Investigate the web interface.

Recon-ng has a web interface that simplifies and improves viewing results that are stored in Recon-ng databases. It also allows easy export of the results tables for reporting purposes.

a. Open a new terminal.
b. Enter the **recon-web** command to start the Recon-ng server process. Note the command output.
c. In a new browser tab, access the webpage using the URL information provided in the output.
d. The web interface shows data from the default workspace when first opened. Click the orange workspace name at the top of the page to display data from different workspaces.

Part 4: Find Interesting Files with Recon-ng

In this part of the lab, we will install and use another plugin.

Step 1: Install another module.

a. Search the marketplace for a module that will discover interesting files in a domain. The plugin that you use should have no dependencies or key requirements.

Which module did you find?

Answer Area

<!----><!---->

b. Install and load the plugin.

Step 2: Run the new module.

a. Set the source option to **hackxor.net** or another location of your choice. (Please comply with the terms of the course when choosing a domain.) The h4cker.org website is interesting also.
b. Run the command. This module creates a .csv file in the recon-ng/data folder.
c. Locate the file and view the contents. Some of these files can be downloaded or viewed using the URLs in the command output.

Step 3: Experiment with different modules and targets.

Investigate other modules. Use the commands that we have learned so far to download, configure options, run, and view results for the module and different targets. It also provides an easy way to export results to various text formats for use in reports.

Reflection Questions

You have experienced the use of several OSINT tools and resources in this lab. There are many more. The best way to learn how to use OSINT in your pentesting practice is to experiment until you find tools and approaches that work for you.

1. What do you think about the recon-ng workspaces feature? How could you use it?

Answer Area

<!----><!---->

2. Recon-ng uses a modular framework architecture. Do modules simplify using the Recon-ng tool? If so, how?

Answer Area

<!----><!---->