Module 2: Footprinting and Reconnaissance

j by Durgesh Thakare

Footprinting

Footprinting and reconnaissance are the initial steps taken to gather information about a target system or network. During this phase, an attacker tries to collect as much information as possible about the target, such as IP addresses, network topology, operating systems, services running on open ports, and other system details. This information is used to identify vulnerabilities and weaknesses in the system that can be exploited later.

Active Footprinting

Active information gathering involves direct interaction with the target system or network. This means sending requests or queries to the target with the intention of receiving responses that reveal information about the system.

Passive Footprinting

Passive information gathering involves collecting data about the target system or network without direct interaction. It does not involve sending requests or queries to the target.

Information obtain in Footprinting

Information obtained through passive footprinting can be used to identify potential vulnerabilities in the target's system. This can include outdated software versions, weak passwords, or misconfigured settings. By gathering this information, an attacker can develop a more targeted and effective attack strategy. It is important for organizations to be aware of the potential risks associated with passive footprinting and take proactive measures to protect their sensitive information.

1 Organization information

Employee details

Telephone numbers

Branch and location details

Background of the organization → Web technologies

News articles, press releases, and related documents

2 Network information

→ Domain and subdomains

Network blocks Network topology, trusted routers, and firewalls

IP addresses of the reachable systems

Whois records → DNS records

3 System information

Web server OS → Location of web servers

Publicly available email addresses

Usernames and passwords

Footprinting Through Search Engines

Footprinting through search engines, also known as "search engine footprinting," is a passive information-gathering technique used to collect publicly available data and information about a target organization, its employees, and its online presence. This process involves leveraging popular search engines like Google, Bing, and specialized search engines to uncover details about the target. Here are the steps

1. Use Search Engines:

- You can use popular search engines like Google, Bing, and specialized search engines like Shodan to uncover information. Here are some techniques and examples:
- Basic Search Queries:

Example:

site:example.com

the target domain "example.com."

• Use basic search queries like "site:example.com" to find all indexed pages associated with

Below is a brief description of what every Google search operator does.

Search Operators for Google:

I've grouped them into three categories:

• Working – Works as intended. • Unreliable - Not officially deprecated by Google, but results are hit-and-miss.

- **Working**

What it does

Example

Search operator

		ZAMPIC
« »	Search for results that mention a word or phrase.	<u>"steve jobs"</u>
OR	Search for results related to X or Y.	jobs OR gates
I	Same as OR:	jobs gates
AND	Search for results related to X and Y.	<u>jobs AND gates</u>
_	Search for results that don't mention a word or phrase.	j <u>obs -apple</u>
*	Wildcard matching any word or phrase.	steve * apple
()	Group multiple searches.	(ipad OR iphone) apple
define:	Search for the definition of a word or phrase.	<u>define:entrepreneur</u>
cache:	Find the most recent cache of a webpage.	cache:apple.com
filetype:	Search for particular types of files (e.g., PDF).	apple filetype:pdf
ext:	Same as filetype:	apple ext:pdf
site:	Search for results from a particular website.	<u>site:apple.com</u>
related:	Search for sites related to a given domain.	<u>related:apple.com</u>
intitle:	Search for pages with a particular word in the title tag.	<u>intitle:apple</u>
allintitle:	Search for pages with multiple words in the title tag.	<u>allintitle:apple iphone</u>
inurl:	Search for pages with a particular word in the URL.	<u>inurl:apple</u>
allinurl:	Search for pages with multiple words in the URL.	allinurl:apple iphone
intext:	Search for pages with a particular word in their content.	intext:apple iphone
allintext:	Search for pages with multiple words in their content.	allintext:apple iphone
weather:	Search for the weather in a location.	weather:san francisco
stocks:	Search for stock information for a ticker.	stocks:aapl
map:	Force Google to show map results.	map:silicon valley
movie:	Search for information about a movie.	movie:steve jobs
in	Convert one unit to another.	<u>\$329 in GBP</u>
source:	Search for results from a particular source in Google News.	<u>apple source:the_verge</u>
before:	Search for results from before a particular date.	<u>apple before:2007-06-29</u>
after:	Search for results from after	<u>apple after:2007-06-29</u>

inanchor:

#..#

Unreliable

Search operator

allinanchor:	Search for pages with backlinks containing multiple words in their anchor text.	allinanchor:apple iphone		
AROUND(X)	Search for pages with two words or phrases within X words of one another.	apple AROUND(4) iphone		
loc:	Find results from a given area.	loc:"san francisco" apple		
location:	Find news from a certain location in Google News.	location:"san francisco" apple		
daterange:	Search for results from a particular date range.	<u>daterange:11278-13278</u>		
GOOGLE HACKING DATABSE GHDB stands for "Google Hacking Database." It is not an official Google product but rather a project that compiles various search queries, known as "Google dorks," used to discover potentially sensitive information and vulnerabilities by leveraging Google's powerful search engine. The GHDB contains a collection of Google dorks that can be used for information				
gathering, security assessments, and penetration testing.Purpose: The GHDB is used by security professionals, ethical hackers, and penetration testers to				

a particular date.

What it does

numbers.

Search within a range of

Search for pages with

backlinks containing

specific anchor text.

iphone case \$50..\$60

inanchor:apple

Example

server details, and more. 3. Categories: The GHDB categorizes Google dorks into various sections, such as "Web Server Detection," "Vulnerable Files," "Error Messages," "Footholds," "Filetype," and more. Each category

online platforms.

🛖 www.exploit-db.com

MW mattw.io

topic keywords.

Shodan

software version.

mattw.io

OffSec's Exploit Database Archive

The GHDB is an index of search queries (we call them dorks) used to find publicly available information, intended for pentesters and security researchers. **Gathering Information from Video Search Engines**

Video search engines are Internet-based search engines that crawl the web for video content.

on their own web servers or parse video content that is hosted externally. The video content

videos allow attackers to search for video content based on the format type and duration.

obtained from video search engines is of high value, as it can be used for gathering information

about the target. Video search engines such as YouTube, Google videos, Yahoo videos, and Bing

These video search engines either provide the functionality of uploading and hosting video content

identify vulnerabilities, exposed data, and misconfigurations on websites, servers, and other

2. **Google Dorks:** Google dorks are specialized search queries that can uncover information not

including login pages, directories with directory listings enabled, error messages disclosing

The GHDB serves as a valuable resource for security professionals to help uncover vulnerabilities and

strengthen the security of web applications, servers, and networks. However, it should be used with

the utmost discretion and only within the bounds of the law and ethical standards.

typically accessible through standard searches. These queries often reveal sensitive information,

Shodan is a specialized search engine and security tool that is often referred to as "the search

engine for hackers." It is designed to scan and index the internet for information about connected

devices and systems. Unlike traditional search engines, Shodan focuses on collecting data about

the online infrastructure of devices, including servers, routers, cameras, industrial control systems,

Created to discover youtube videos based on location and later adapted to search videos by specific channels and

and other Internet of Things (IoT) devices. Here are some key features and uses of Shodan: 1. Device and System Discovery: Shodan scans the internet to discover and index devices and systems. It collects information about open ports, banners, and services running on these devices. 2. Search and Query: Users can search Shodan's database using specialized search queries, similar

to Google dorks, to find specific types of devices, vulnerabilities, or exposed services. For

3. Banner Information: Shodan collects banner information from services running on devices, which can provide details about the software and version, server headers, and other potentially valuable data.

4. **Vulnerability Detection:** Shodan can be used to search for devices that may be vulnerable to

example, users can search for webcams with default passwords or servers running a specific

the physical locations of devices on the internet. 6. Internet of Things (IoT): Shodan is particularly useful for identifying and researching IoT devices that may have security weaknesses.

7. **Security Research and Ethical Hacking:** Security professionals, ethical hackers, and researchers

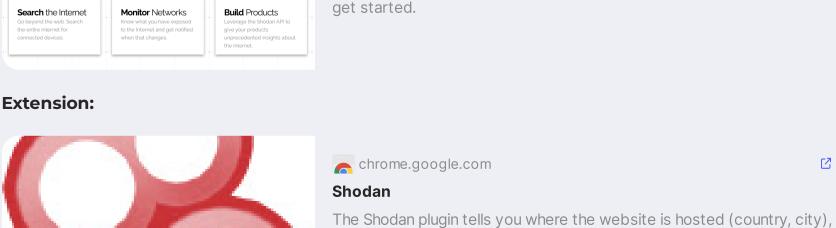
use Shodan to identify security weaknesses, report vulnerabilities, and assess the overall security

specific security issues, such as devices running outdated software or using default credentials.

5. **Geolocation:** Shodan provides geolocation data for devices and systems, helping users pinpoint

- posture of the internet. 8. Monitoring and Alerts: Shodan offers a subscription service called Shodan Monitor, which allows users to set up alerts for specific devices or services. Users can receive notifications when
- website: // USAGE Shodan **?** Shodan \Box More than 4 million users

Shodan







Mozilla Explains: AI

new, potentially vulnerable devices are discovered.

website is hosted (country, city), who owns the IP and what other...

Made with Gamma

Download Shodan for Firefox. The Shodan plugin tells you where the

Search engine of Internet-connected devices. Create a free account to

who owns the IP and what other services/ ports are open.

Network based information gadhering

Network-Specific Information Gathering:

Information gathering in a network-based context, often used in ethical hacking and penetration testing, involves collecting detailed data about a target network to understand its structure, devices, vulnerabilities, and potential attack vectors. Here's a deeper look into network-based information gathering:

Who is lookup A "whois" lookup is a common tool used in network-based information gathering. It allows users to retrieve registration information about a domain or IP address, including details such as the owner, contact information, and registration date. This information can be valuable in assessing the potential risks and vulnerabilities associated with a target network.

Ip-lookup Another useful tool for network-based information gathering is an "IP-lookup." This tool allows users to gather detailed information about an IP address, such as the geolocation, internet service provider (ISP), and potentially associated domains. Conducting an IP-lookup can provide valuable insights into the network's physical location and its potential connections to other systems or entities.

Reverse ip lookup A reverse IP lookup is a technique used to identify the domains that are hosted on a particular IP address. This can be helpful in discovering other websites or systems that are associated with the target network. By conducting a reverse IP lookup, an attacker can gain a better understanding of the network's infrastructure and potentially identify additional vulnerabilities or attack vectors.

Domain lookup A domain lookup is a technique used to gather information about a specific domain name. This can include details such as the registrar, registration date, and expiration date of the domain. By conducting a domain lookup, an attacker can gain insights into the ownership and history of the target domain, helping them assess potential risks and vulnerabilities in the network.

Dns lookup A DNS lookup is a technique used to gather information about the DNS records associated with a domain. This can include details such as the IP addresses linked to the domain, mail server information, and other DNS records like TXT or MX records. Conducting a DNS lookup can provide valuable information about the target domain's infrastructure and configuration, helping an attacker in identifying potential weak points or misconfigurations that could be exploited.

Browser Extension

https://chrome.google.com/webstore/detail/shodan/jjalcfnidlmpjhdfepjhjbhnhkbgleap

https://chrome.google.com/webstore/detail/whatruns/cmkdbmfndkfgebldhnkbfhlneefdaaip

https://chrome.google.com/webstore/detail/wappalyzer-technology-pro/gppongmhjkpfnbhagpmjfkannfbllamg

Websites

Sub Domain Finder

https://dnsdumpster.com/

https://pentest-tools.com/information-gathering/find-subdomains-of-domain#

https://spyse.com/

https://subdomainfinder.c99.nl/

https://www.nmmapper.com/

Server Information

https://www.yougetsignal.com --> reverse IP DOMAIN CHEKUP http://reverseip.domaintools.com/https://whois.net/

Informational Websites

https://whois.icann.org/en https://mxtoolbox.com/

Check Website History

https://archive.org/

Target-Based Information Gathering

Target-based or person-based information gathering is an essential aspect of ethical hacking and penetration testing that focuses on collecting information about specific individuals, organizations, or entities. This type of information gathering helps ethical hackers and security professionals understand their targets, identify potential vulnerabilities, and plan security assessments or penetration tests effectively. Here's a deeper look at target-based and person-based information gathering:

Organization Profiling:

- **Company Website Analysis:** Analyzing the target organization's website to gather information about its products, services, mission, history, and key personnel.
- **News and Press Releases:** Reviewing news articles and press releases related to the organization to understand its recent activities, achievements, and challenges.
- **Financial Reports:** Analyzing financial reports, if available, to gain insights into the organization's financial health and performance.
- **Social Media Accounts:** Exploring the organization's social media profiles for updates and announcements.

Contact Information:

- **Email Address Enumeration:** Collecting known email addresses associated with the individual.
- **Phone Number Identification:** Discovering phone numbers, if available, from public records or social media profiles.
- **Publicly Available Documents:** Searching for publicly available documents containing contact information.

Associations and Relationships:

- Friends and Colleagues: Identifying an individual's personal and professional relationships.
- Affiliations: Finding out an individual's affiliations with organizations or group

Public Records and Online Archives:

- Public Records: Accessing public records such as property records, legal filings, and government documents.
- o **Internet Archive Searches:** Searching the Internet Archive (Wayback Machine) for historical data related to an individual or their online presence.



OSINT Framework

OSINT Framework, also known as "OSINT Framework" (OSINT stands for Open-Source Intelligence), is a comprehensive online resource that provides a curated list of open-source intelligence tools, resources, and references. It is designed to assist individuals and professionals in conducting open-source intelligence activities, which involve gathering information from publicly available sources to gain insights and intelligence.

The OSINT Framework offers a well-organized and categorized collection of tools and websites that can be used for various aspects of open-source intelligence. This includes tools for data collection, information analysis, digital forensics, network reconnaissance, social media investigation, and more. Users can explore and access a wide range of resources and tools for conducting investigations, research, and intelligence gathering.

- 1. **Data Collection Tools:** Tools that help collect data from publicly available sources, such as search engines, social media, news, and websites.
- 2. **Information Analysis:** Tools for processing and analyzing the data collected, including data analysis software and tools for linguistic analysis.
- 3. **Digital Forensics:** Resources for digital forensics investigations, including forensic analysis tools and reference materials.
- 4. **Network Reconnaissance:** Tools and resources for network scanning and reconnaissance.
- 5. **Social Media Investigation:** Tools and techniques for investigating social media profiles and activities.
- 6. **Dark Web Research:** Resources for researching and monitoring activities on the dark web and hidden online communities.
- 7. **Geolocation and Mapping:** Tools for geolocation, mapping, and tracking activities.
- 8. **Phone Number Investigation:** Resources for investigating phone numbers, including reverse phone lookup services.
- 9. **Email Investigation:** Tools for email analysis and email tracking.
- 10. Username Investigation: Resources for investigating usernames and online identities.
- 11. **Miscellaneous:** A category for other relevant resources and tools that may not fit into the previous categories.

osintframework.com

OSINT Framework

(T) - Indicates a link to a tool that must be installed and run locally (D) - Google Dork, for more information: Google Hacking (R) - Requires registration (M) - Indicates a URL that contains the search term and the URL itself must be...