Reconnaissance Methodology

Reconnaissance (also known as footprinting) refers to the preparatory phase where an attacker seeks to gather as much information as possible about a target of evaluation prior to launching an attack. An essential aspect of footprinting is identifying the level of risk associated with the organization's publicly accessible information. Footprinting, the first step in ethical hacking, refers to the process of collecting information about a target network and its environment. Using footprinting, you can find a number of opportunities to penetrate and assess the target organization's network.

After you complete the footprinting process in a methodological manner, you will obtain the blueprint of the security profile of the target organization. Here, the term "blueprint" refers to the unique system profile of the target organization acquired by footprinting. There is no single methodology for footprinting, as information can be traced in a number of ways. However, the activity is important, as you need to gather all the crucial information about the target organization before beginning the hacking phase. For this reason, footprinting needs to be carried out in an organized manner. The information gathered in this step helps in uncovering vulnerabilities existing in the target network and in identifying different ways of exploiting these vulnerabilities.

Types of Footprinting/Reconnaissance

Footprinting can be categorized into passive footprinting and active footprinting.

- **Passive Footprinting**

Passive footprinting involves gathering information about the target without direct interaction. It is mainly useful when the information gathering activities are not to be detected by the target. Performing passive footprinting is technically difficult, as active traffic is not sent to the target organization from a host or anonymous hosts or services over the Internet. We can only collect archived and stored information about the target using search engines, social networking sites, and so on.

- Open-source Intelligence (OSINT) gathering
- Proprietary databases and paid services
- Sharing intelligence with partner organizations or industry groups

- **ActiveFootprinting**

Active footprinting involves gathering information about the target with direct interaction. In active footprinting, the target may recognize the ongoing information gathering process, as we overtly interact with the target network. Active footprinting requires more preparation than passive footprinting, as it may leave traces that may alert the target organization.

It involves:

- DNS interrogation
- Social engineering

- Network/port scanning
- User and service enumeration

➢ Information Obtained in Footprinting

The major objectives of footprinting include collecting the network information, system information, and organizational information of the target. By conducting footprinting across different network levels, you can gain information such as network blocks, specific IP addresses, employee details, and so on. Such information can help attackers in gaining access to sensitive data or performing various attacks on the target network.

Organization Information:

The information about an organization is available from its website. In addition, you can query the target's domain name against the Whois database and obtain valuable information.

The information collected includes:

- Employee details (employee names, contact addresses, designations, and work experience)
- Addresses and mobile/telephone numbers
- Branch and location details
- Partners of the organization
- Web links to other company-related sites o Background of the organization
- Web technologies
- News articles, press releases, and related documents
- Legal documents related to the organization
- Patents and trademarks related to the organization

Attackers can access organizational information and use such information to identify key personnel and launch social engineering attacks to extract sensitive data about the entity.

Network Information:

You can gather network information by performing Whois database analysis, trace routing, and so on.

- The information collected includes:
- Domain and sub-domains
- Network blocks
- Network topology, trusted routers, and firewalls
- IP addresses of the reachable systems
- Whois records
- DNS records and related information

System Information:

You can gather system information by performing network footprinting, DNS footprinting, website footprinting, email footprinting, and so on.

The information collected includes:

- Web server OS
- Location of web servers
- Publicly available email addresses

Footprinting Threats

The following are assorted threats made possible through footprinting:

Social Engineering: Without using any intrusion methods, hackers directly and indirectly collect information through persuasion and other means. Hackers gather crucial information from willing employees who are unaware of the hackers' intent.

System and Network Attacks: Footprinting enables an attacker to perform system and network attacks. Thus, attackers can gather information related to the target organization's system configuration, the operating system running on the machine, and so on. Using this information, attackers can find vulnerabilities in the target system and then exploit such vulnerabilities. They can then take control of a target system or the entire network.

Information Leakage: Information leakage poses a threat to any organization. If sensitive information of an entity falls into the hands of attackers, they can mount an attack based on the information or alternatively use it for monetary benefit.

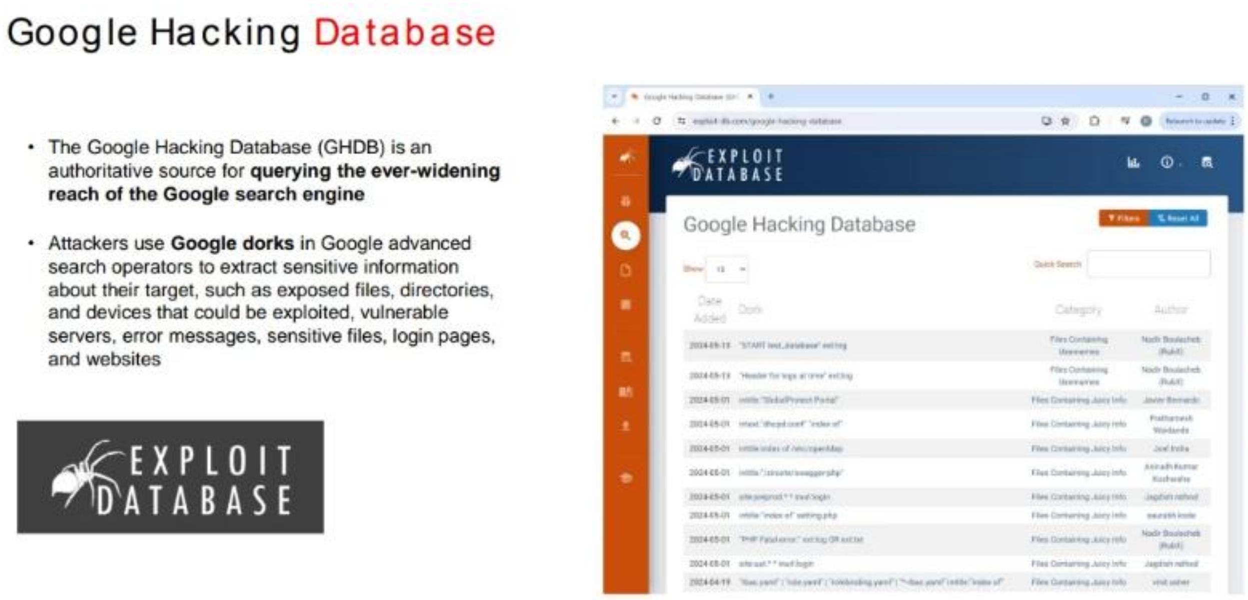
Privacy Loss: Through footprinting, hackers can access the systems and networks of the organization and even escalate the privileges up to admin levels, resulting in the loss of privacy for the organization as a whole and for its individual personnel.

Corporate Espionage: Corporate espionage is a central threat to organizations, as competitors often aim to attempt to acquire sensitive data through footprinting. Through this approach, competitors can launch similar products in the market, alter prices, and generally undermine the market position of a target organization.

Business Loss: Footprinting can have a major effect on organizations such as online businesses and other e-commerce websites as well as banking and finance-related businesses. Billions of dollars are lost every year due to malicious attacks by hackers.

What can a Hacker Do with Google Hacking?

An attacker can create complex search-engine queries to filter large amounts of search results to obtain information related to computer security. The attacker can use Google operators to locate specific strings of text within search results. Thus, the attacker can not only detect websites and web servers that are vulnerable to exploitation but also locate private and sensitive information about the target. Once a vulnerable site is identified, attackers attempt to launch various possible attacks, such as buffer overflow and SQL injection, which compromise information security. Examples of sensitive information on public servers that an attacker can extract with the help of Google Hacking Database (GHDB) queries include: ▪ Error messages that contain sensitive information ▪ Files containing passwords ▪ Sensitive directories ▪ Pages containing logon portals ▪ Pages containing network or vulnerability data, such as IDS, firewall logs, and configurations ▪ Advisories and server vulnerabilities ▪ Software version information ▪ Web application source code ▪ Connected IoT devices and their control panels, if unprotected ▪ Hidden web pages such as intranet and VPN services.



Google Hacking Database

- The Google Hacking Database (GHDB) is an authoritative source for **querying the ever-widening reach of the Google search engine**
- Attackers use **Google dorks** in Google advanced search operators to extract sensitive information about their target, such as exposed files, directories, and devices that could be exploited, vulnerable servers, error messages, sensitive files, login pages, and websites

Google Hacking Database

Date Added	Dork	Category	Author
2024-05-13	"START test_database" ext:log	Files Containing Usernames	Nadir Boulacheb (RubX)
2024-05-13	"Header for logs at time" ext:log	Files Containing Usernames	Nadir Boulacheb (RubX)
2024-05-01	intitle:"GlobalProtect Portal"	Files Containing Juicy Info	Javier Bernardo
2024-05-01	intext:"dhcpd.conf" "index of"	Files Containing Juicy Info	Prathamesh Wadarde
2024-05-01	intitle:index of /etc/openldap	Files Containing Juicy Info	Joel Indra
2024-05-01	intitle:"/zinurls/swagger.php"	Files Containing Juicy Info	Aniruddh Kumar Kushwaha
2024-05-01	site:prod.* * inurl:login	Files Containing Juicy Info	Jagdish rathod
2024-05-01	intitle:"index of" setting.php	Files Containing Juicy Info	saurabh kode
2024-05-01	"PHP Fatal error:" ext:log OR ext:txt	Files Containing Juicy Info	Nadir Boulacheb (RubX)
2024-05-01	site:uat.* * inurl:login	Files Containing Juicy Info	Jagdish rathod
2024-04-19	"docs.yaml" | "site.yaml" | "bitbinding.yaml" | "*-docs.yaml" intitle:"index of"	Files Containing Juicy Info	vinit usher

https://www.exploit-db.com/google-hacking-database

Google Hacking Database Source: https://www.exploit-db.com/google-hacking-database

The GHDB is a subset of the Exploit-DB and focuses on using Google search queries (often referred to as "Google Dorks") to find sensitive information inadvertently exposed on the web.

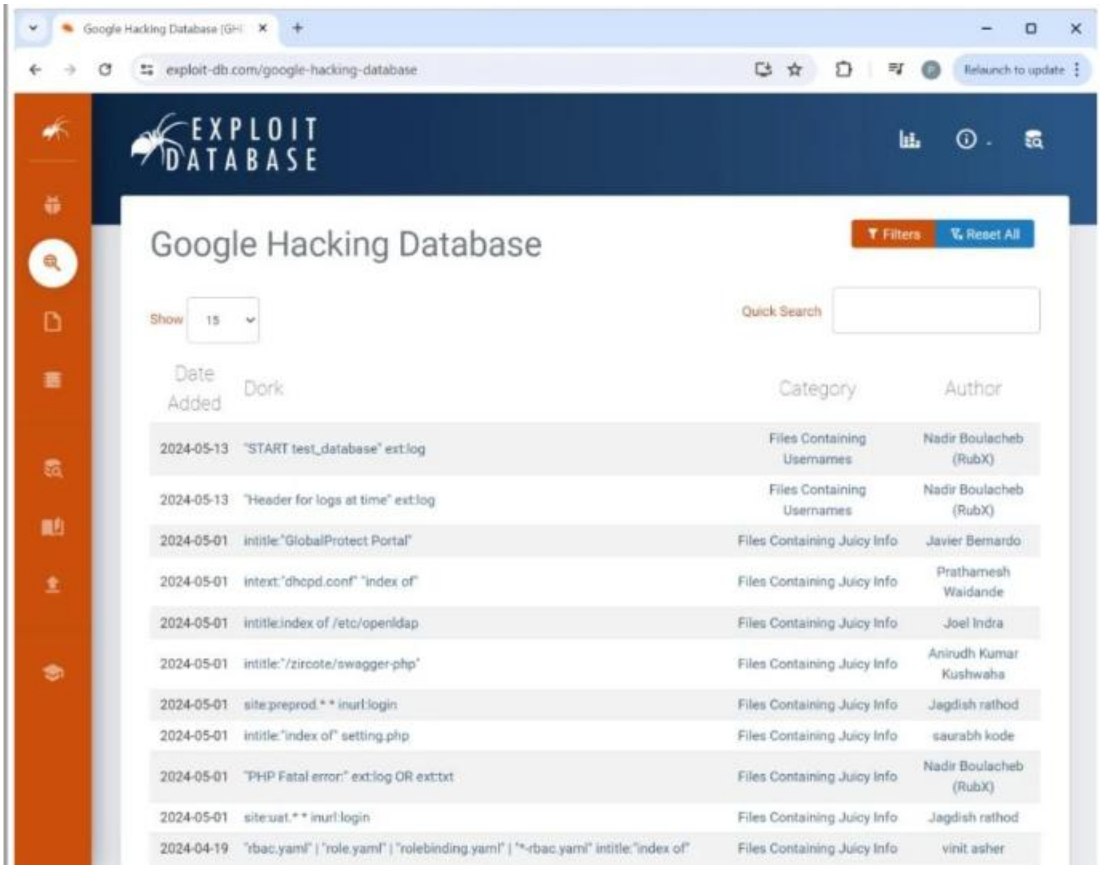
These queries exploit advanced Google search operators to uncover the following:

- Sensitive files: Such as configuration files, database dumps, and log files that may contain usernames, passwords, or other confidential data.
- Exposed directories: Open directories on web servers that might contain sensitive information.
- Error messages: Web server or application error messages that may reveal server configurations or vulnerabilities.
- Vulnerable devices: Identifying specific types of devices or software versions known to have vulnerabilities

Google Hacking Database Categories:

- Footholds
- Files Containing Usernames
- Sensitive Directories
- Web Server Detection
- Vulnerable Files
- Vulnerable Servers
- Error Messages
- Files Containing Juicy Info

- Files Containing Passwords
- Sensitive Online Shopping Info
- Network or Vulnerability Data
- Pages Containing Login Portals
- Various Online Devices
- Advisories and Vulnerabilities


Google Hacking Database (GH
exploit-db.com/google-hacking-database
Relaunch to update
EXPLOIT DATABASE
Filters
Reset All

Google Hacking Database

Show 15

Quick Search

Date Added	Dork	Category	Author
2024-05-13	"START test_database" ext:log	Files Containing Usernames	Nadir Boulacheb (RubX)
2024-05-13	"Header for logs at time" ext:log	Files Containing Usernames	Nadir Boulacheb (RubX)
2024-05-01	intitle:"GlobalProtect Portal"	Files Containing Juicy Info	Javier Bernardo
2024-05-01	intext:"dhcpd.conf" "index of"	Files Containing Juicy Info	Prathamesh Waidande
2024-05-01	intitle:index of /etc/openldap	Files Containing Juicy Info	Joel Indra
2024-05-01	intitle:"/zircote/swagger-php"	Files Containing Juicy Info	Anirudh Kumar Kushwaha
2024-05-01	site:preprod.* * inurl:login	Files Containing Juicy Info	Jagdish rathod
2024-05-01	intitle:"index of" setting.php	Files Containing Juicy Info	saurabh kode
2024-05-01	"PHP Fatal error:" ext:log OR ext:txt	Files Containing Juicy Info	Nadir Boulacheb (RubX)
2024-05-01	site:uat.* * inurl:login	Files Containing Juicy Info	Jagdish rathod
2024-04-19	"rbac.yaml" | "role.yaml" | "rolebinding.yaml" | "*-rbac.yaml" intitle:"index of"	Files Containing Juicy Info	vinit asher

Attackers can leverage the GHDB in various ways to identify and exploit vulnerabilities:

- Reconnaissance: Attackers use GHDB queries to gather information about potential targets, including exposed files, directories, and devices that could be exploited.
- Exploiting Misconfigurations: By identifying sensitive information exposed through misconfigured web servers or services, attackers can exploit these misconfigurations to gain unauthorized access.
- Finding Vulnerable Systems: Using GHDB, attackers can locate systems running outdated or vulnerable software versions, providing a starting point for further exploitation.
- Credential Harvesting: Sensitive information found through GHDB queries can include usernames and passwords, which attackers can use for credential stuffing or brute force attacks.
- Identifying Open Ports and Services: Some GHDB queries can reveal open ports and services on a network, giving attackers a map of potential entry points.

Attackers can also use SearchSploit, which is a command-line search tool for Exploit-DB that allows taking a copy of the Exploit database for remote use. It allows attackers to perform detailed offline searches through their locally checked-out copy of the repository. This capability is particularly useful for security assessments of segregated or air-gapped networks without Internet access.

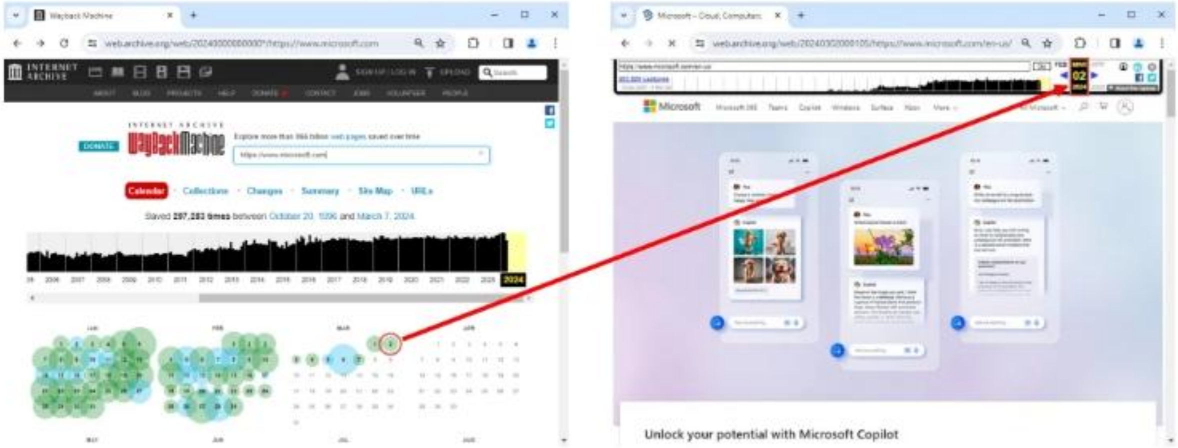
Footprinting through Internet Research Services

Internet research services such as people search services can provide sensitive information about the target. People search services, alerting services, financial services, and job sites provide information about a target such as infrastructure details, physical location, and employee details. Using this information, an attacker may build a hacking strategy to break into the target organization's network and carry out other types of advanced system attacks.

This section aims to familiarize you with finding the target company's top-level domains, subdomains, and geographical location, performing people search on people search services, gathering information from job sites, financial services, third-party data repositories, performing dark web footprinting, gathering competitive intelligence, etc.

Extracting Website Information from https://archive.org

- Internet Archive's Wayback Machine allows one to visit **archived versions of websites**


Wayback Machine
INTERNET ARCHIVE
DONATE
WaybackMachine
Calendar · Collections · Changes · Summary · Site Map · URLs
Saved 297,283 times between October 20, 1996 and March 7, 2024.
2024
Microsoft
Unlock your potential with Microsoft Copilot

- Attackers can use tools such as **Photon** to retrieve archived URLs of the target website from archive.org

Extracting Website Information from https://archive.org

Source: https://archive.org

Archive is an Internet Archive Wayback Machine that explores archived versions of websites. Such exploration allows an attacker to gather information on an organization's web pages since its creation. As the website https://archive.org keeps track of web pages from the time of their creation, an attacker can retrieve even information removed from the target website, such as web pages, audio files, video files, images, text, and software programs. Attackers use this information to perform phishing and other types of web application attacks on the target organization.

Footprinting through Social Networking Sites

While footprinting through social networking sites may seem similar to footprinting through social engineering (which is discussed in greater detail later), there are some differences between the two methods. In footprinting through social engineering, the attacker tricks people into revealing information, whereas in footprinting through social networking sites, the attacker gathers information available on those sites. Attackers can even use social networking sites as a medium to perform social engineering attacks.

This section explains the type of information one can collect from social networking sites and how it can be obtained. It aims to familiarize you with locating information from social media sites using various online services and resources.

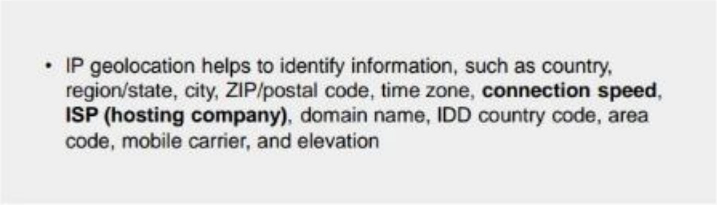
People Search on Social Networking Sites

Searching for a particular person on a social networking website is fairly easy. Social networking services are online services, platforms, or sites that focus on facilitating the building of social networks or social relations among people. These websites contain information provided by users in their profiles. They help relate people directly or indirectly to each other through various fields, such as common interests, work locations, and education. Social networking sites allow people to share information quickly, as they can update their personal details in real time. Such sites allow users to update facts about upcoming or current events, recent announcements and invitations, and so on. Social networking sites are a great platform for finding people and their related information. Many social networking sites allow visitors to search for people without registering on the site; this makes people searching on social networking sites an easy and anonymous task. A user can search for a person using the name, email, or address. Some sites allow users to check whether an account is active, which then provides information on the status of the person being searched.

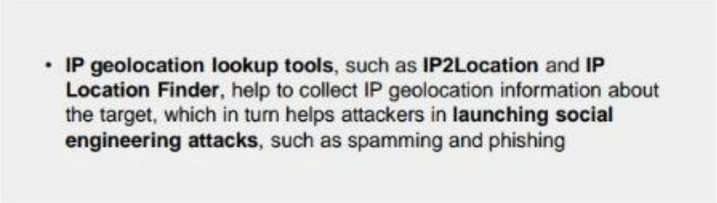
Social networking sites such as Facebook, Twitter, LinkedIn, and Instagram allow you to find people by name, keyword, company, school, friends, colleagues, and the people living around them. Searching for people on these sites returns personal information such as name, position, organization name, current location, and educational qualifications. In addition, you can also find professional information such as company or business, current location, phone number, email ID, photos, videos and so on. Social networking sites such as Twitter are used to share advice, news, concerns, opinions, rumors, and facts.

Whois Footprinting

Gathering network-related information such as “Whois” information about the target organization is important when planning an attack. In this section, we will discuss Whois footprinting, which helps in gathering domain information such as information regarding the owner of an organization, its registrar, registration details, its name server, and contact information. Whois footprinting focuses on how to perform a Whois lookup, analyze the Whois lookup results, and find IP geolocation information, as well as the tools used to gather Whois information.



- IP geolocation helps to identify information, such as country, region/state, city, ZIP/postal code, time zone, **connection speed**, **ISP (hosting company)**, domain name, IDD country code, area code, mobile carrier, and elevation



- **IP geolocation lookup tools**, such as **IP2Location** and **IP Location Finder**, help to collect IP geolocation information about the target, which in turn helps attackers in **launching social engineering attacks**, such as spamming and phishing

Finding IP Geolocation Information

IP geolocation helps to obtain information regarding a target such as its country, region/state, city, latitude and longitude of its city, ZIP/postal code, time zone, connection speed, ISP (hosting company), domain name, IDD country code, area code, weather station code and name, mobile carrier, and elevation.

Using the information obtained from IP geolocation, an attacker may attempt to gather more information about a target with the help of social engineering, surveillance, and non-technical attacks such as dumpster diving, hoaxing, or acting as a technical expert. With the help of the information obtained, an attacker can also set up a compromised web server near the victim's location, and if the exact location of the victim is detected, the attacker can perform malicious activities and infect the victim with malware designed for that specific area or gain unauthorized access to the target device or attempt to launch an attack on the target device.

IP geolocation lookup tools such as IP2Location, IP Location Finder, and IP Address Geographical Location Finder help to collect IP geolocation information about the target, which enables attackers to launch social engineering attacks such as spamming and phishing.

IP Geolocation Lookup Tools

IP2Location Source: https://www.ip2location.com

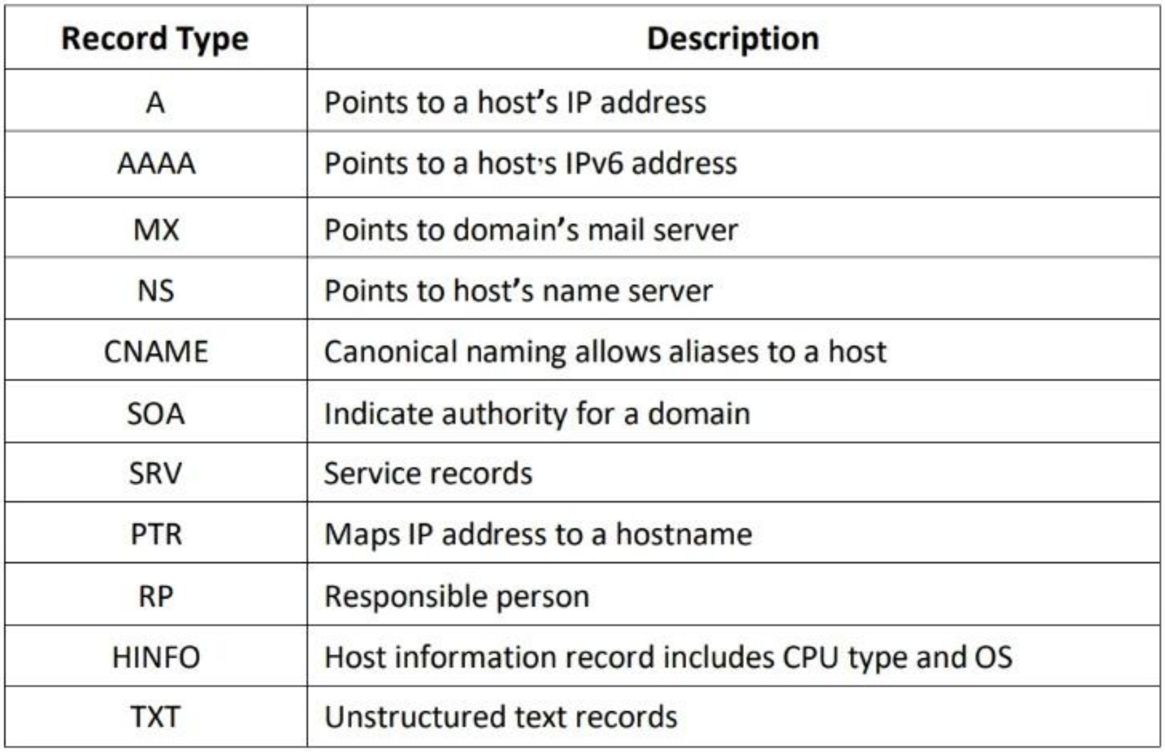
Attackers use IP2Location tool to identify a visitor's geographical location, i.e., country, region, city, latitude and longitude of city, ZIP code, time zone, connection speed, ISP, domain name, IDD country code, area code, weather station code and name, mobile carrier, elevation, and usage type information using a proprietary IP address lookup database and technology.

DNS Footprinting

After collecting the Whois records of the target, the next phase of the footprinting methodology is Domain Name System (DNS) footprinting. Attackers perform DNS footprinting to gather information about DNS servers, DNS records, and the types of servers used by the target organization. This information helps attackers identify the hosts connected in the target network and further exploit the target organization. This section describes how to extract DNS information and perform reverse DNS lookups using various DNS interrogation tools.

Extracting DNS Information

DNS footprinting reveals information about DNS zone data. DNS zone data include DNS domain names, computer names, IP addresses, and much more information about a network. An attacker uses DNS information to determine key hosts in the network and then performs social engineering attacks to gather even more information. DNS footprinting helps in determining the following records about the target DNS:



Record Type	Description
A	Points to a host’s IP address
AAAA	Points to a host's IPv6 address
MX	Points to domain’s mail server
NS	Points to host’s name server
CNAME	Canonical naming allows aliases to a host
SOA	Indicate authority for a domain
SRV	Service records
PTR	Maps IP address to a hostname
RP	Responsible person
HINFO	Host information record includes CPU type and OS
TXT	Unstructured text records

Network and Email Footprinting

The next step after retrieving DNS information is to gather network-related information and track email communications. This section describes the method to locate the network range, traceroute analysis, and traceroute tools. It also describes how to track email communications, how to collect information from email headers, and email tracking tools.

Locate the Network Range To perform network footprinting, one needs to gather basic and important information about the target organization, such as what the organization does, who works there, and what type of work it does. The answers to these questions provide information that helps identify the internal structure of the target network. After gathering the information, an attacker can determine the network range of the target system. Detailed information regarding IP allocation and the nature of allocation is available with the appropriate regional registry database. An attacker can also determine the subnet mask of the domain and trace the route between the system and target system. Widely used traceroute tools include NetScanTools Pro and PingPlotter.

Obtaining private IP addresses can be useful to attackers. The Internet Assigned Numbers Authority (IANA) has reserved the following three blocks of IP address space for private internets: 10.0.0.0–10.255.255.255 (10/8 prefix), 172.16.0.0–172.31.255.255 (172.16/12 prefix), and 192.168.0.0–192.168.255.255 (192.168/16 prefix).

Using the network range, an attacker can obtain information about how the network is structured and which machines in the network are alive. The network range also helps identify the network topology, access control device, and OS used in the target network. To find the network range of the target network, one must enter the server IP address (gathered in Whois footprinting) in the ARIN Whois database search tool.

A user can also visit the ARIN website (https://www.arin.net/about/welcome/region) and enter the server IP into the SEARCH Site or Whois text box. This yields the network range of the target network. Improperly set-up DNS servers offer attackers a good chance of obtaining a list of internal machines in the network. Additionally, if an attacker traces a route to a machine, it could be possible to obtain the internal IP address of the gateway, which can be useful.

Collecting Information Using Eavesdropping, Shoulder Surfing, Dumpster Diving, and Impersonation

Eavesdropping, shoulder surfing, dumpster diving, and impersonation are social engineering techniques widely used to collect information from people.

- **Eavesdropping**
  Eavesdropping is the act of intercepting communication in any form, such as audio, video, or text, without the consent of the communicating parties. It also includes reading confidential messages from communication media such as instant messaging or fax transmissions. The attacker can gain information by tapping phone conversations or intercepting audio, video, or written communications.

- **Shoulder Surfing**
  Shoulder surfing is a technique whereby attackers secretly observe the target to gain critical information. In the shoulder surfing technique, an attacker stands behind the victim and secretly observes the victim’s activities on the computer, such as keystrokes while entering usernames, passwords, and so on. The technique is effective in gaining passwords, personal identification numbers, security codes, account numbers, credit card information, and similar data. Attackers can easily perform shoulder surfing in a crowded place, as it is relatively easy to stand behind and watch the victim without his or her knowledge

- **Dumpster Diving**
  This uncouth technique, also known as trashing, involves the attacker rummaging for information in garbage bins. The attacker may gain vital information such as phone bills, contact information, financial information, operations-related information, printouts of source codes, printouts of sensitive information, and so on from the target company’s trash bins, printer waste bins, sticky notes at users’ desks, and so on. The attacker may also gather account information from ATM trash bins.

- **Impersonation**
  Impersonation is a technique whereby an attacker pretends to be a legitimate or authorized person. Attackers perform impersonation attacks personally or use phones or other communication media to mislead targets and trick them into revealing information. The attacker might impersonate a courier/delivery person, janitor, businessman, client, technician, or he/she may pretend to be a visitor. Using this technique, an attacker gathers sensitive information by scanning terminals for passwords, searching important documents on desks, rummaging bins, and so on. The attacker may even try to overhear confidential conversations and “shoulder surf” to obtain sensitive information.

Footprinting Countermeasures

Thus far, we have discussed the importance of footprinting, various methods to perform footprinting, and tools that help in its execution. Now, we discuss footprinting countermeasures, i.e., the measures or actions taken to prevent or offset information disclosure.

Some of the footprinting countermeasures are as follows:

- Restrict the employees’ access to social networking sites from the organization’s network.
- Configure web servers to avoid information leakage.
- Educate employees to use pseudonyms on blogs, groups, and forums.
- Do not reveal critical information in press releases, annual reports, product catalogs, etc.
- Limit the amount of information published on a website or the Internet.
- Use footprinting techniques to discover and remove any sensitive information that is publicly available.
- Prevent search engines from caching a web page and use anonymous registration services.
- Develop and enforce security policies such as information security and password policies to regulate the information that employees can reveal to third parties.
- Implement multi-factor authentication mechanisms to enhance the security of the organization's systems and resources.
- Set apart internal and external DNS or use split DNS, and restrict zone transfer to authorized servers.