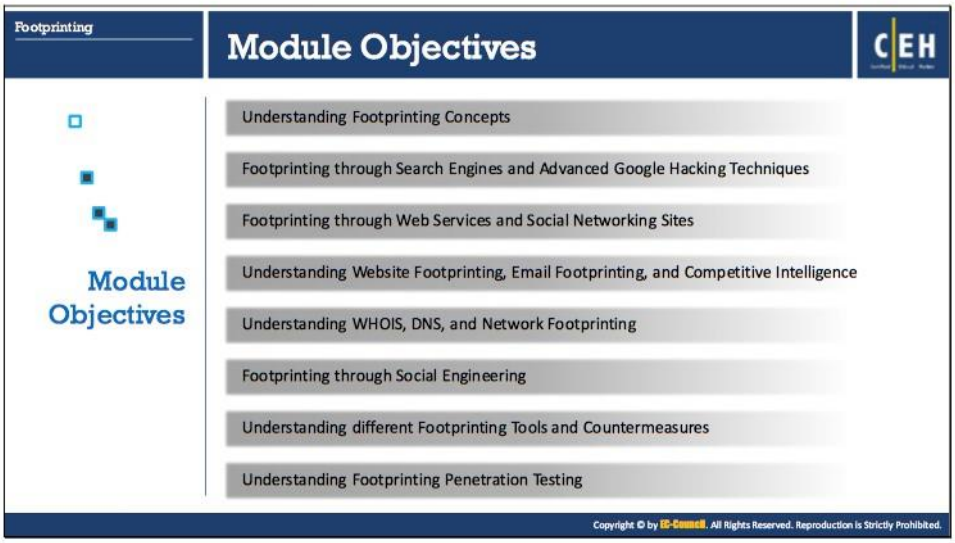
C|EH™
Certified Ethical Hacker

Module 02

Footprinting and Reconnaissance

This page is intentionally left blank.



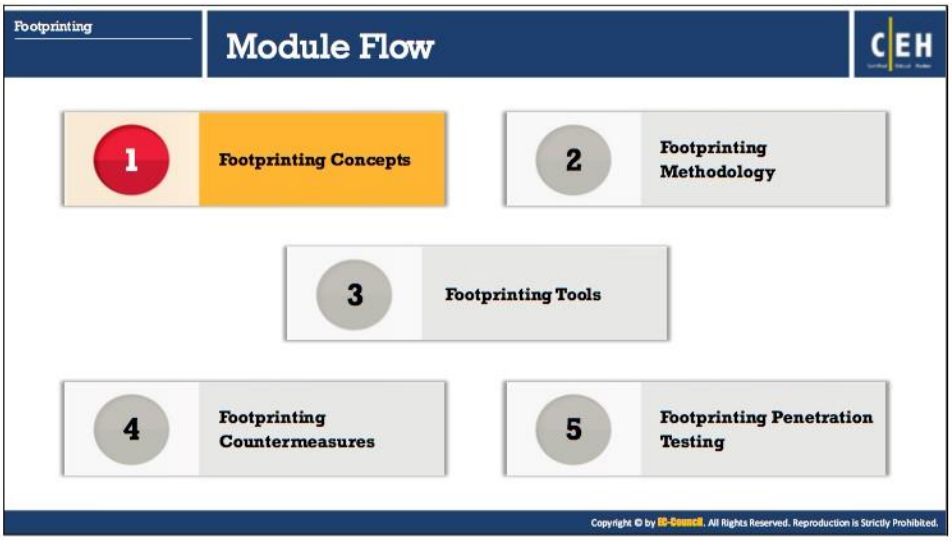
Module Objectives

Footprinting is a first step in the evaluation of the security posture of the target organization's IT infrastructure. Through footprinting and reconnaissance, one can gather maximum information about a computer system or a network and about any devices connected to that network. In other words, footprinting provides security profile blueprint for an organization, and should be undertaken in a methodological manner.

This module starts with an introduction to footprinting concepts and provides insight into footprinting methodology. Later the module discusses footprinting tools and countermeasures. The module ends with an overview of penetration ('pen') testing steps that an ethical hacker should follow to perform the security assessment of a target.

At the end of this module, you will be able to:

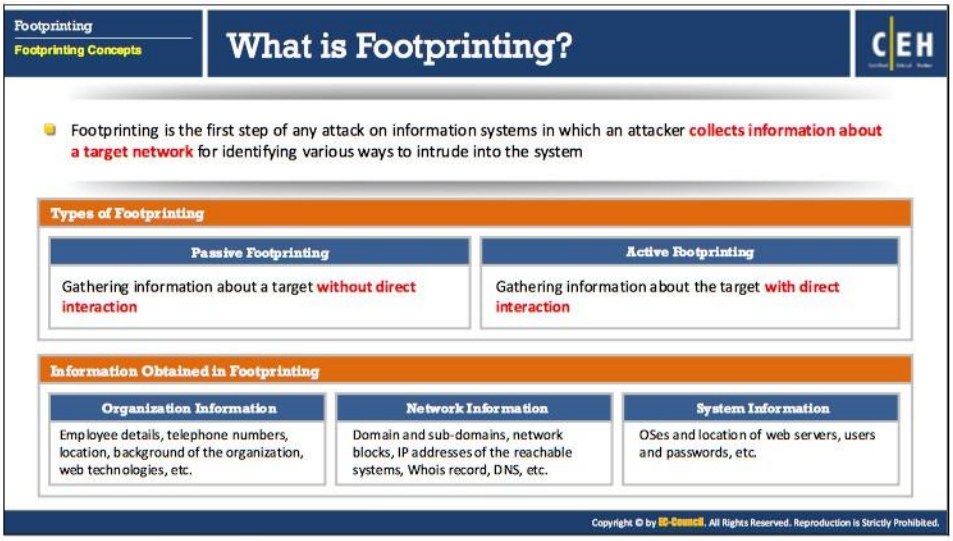
- Describe footprinting concepts
- Perform footprinting through search engines and advanced google hacking techniques
- Perform footprinting through web services and social networking sites
- Perform website footprinting, email footprinting, and competitive intelligence
- Perform Whois, DNS, and network footprinting
- Perform footprinting through social engineering
- Use different footprinting tools
- Apply footprinting best practice
- Perform footprinting penetration testing



Footprinting Concepts

Ethical hacking is legal in nature and conducted in order to evaluate the security of a target organization's IT infrastructure with their consent. Footprinting, where an attacker tries to gather information about a target, is the first step in ethical hacking.

The footprinting concepts section aims to get you familiarized with footprinting, why it is necessary, and its objectives.



What is Footprinting?

An essential aspect of footprinting 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.

Types of Footprinting

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 a type of footprinting that is mainly useful when there is a requirement that 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 from anonymous hosts or services over the Internet. We can

only collect the archived and stored information about the target using search engines, social networking sites, and so on.

Passive footprinting techniques include:

- Finding information through search engines
- Finding the Top-level Domains (TLDs) and sub-domains of a target through web services
- Collecting location information on the target through web services
- Performing people search using social networking sites and people search services
- Gathering financial information about the target through financial services
- Gathering infrastructure details of the target organization through job sites
- Monitoring target using alert services
- Gathering information using groups, forums, and blogs
- Determining the operating systems in use by the target organization
- Extracting information about the target using Internet archives
- Performing competitive intelligence
- Monitoring website traffic of the target
- Tracking the online reputation of the target
- Collecting information through social engineering on social networking sites

- **Active Footprinting**

  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 techniques include:

  - Querying published name servers of the target
  - Extracting metadata of published documents and files
  - Gathering website information using web spidering and mirroring tools
  - Gathering information through email tracking
  - Performing Whois lookup
  - Extracting DNS information
  - Performing traceroute analysis
  - Performing social engineering

Information Obtained in Footprinting

The major objectives of footprinting include collecting the network information, system information, and the 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.

- **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
  - IP addresses of the reachable systems
  - Whois record
  - 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:

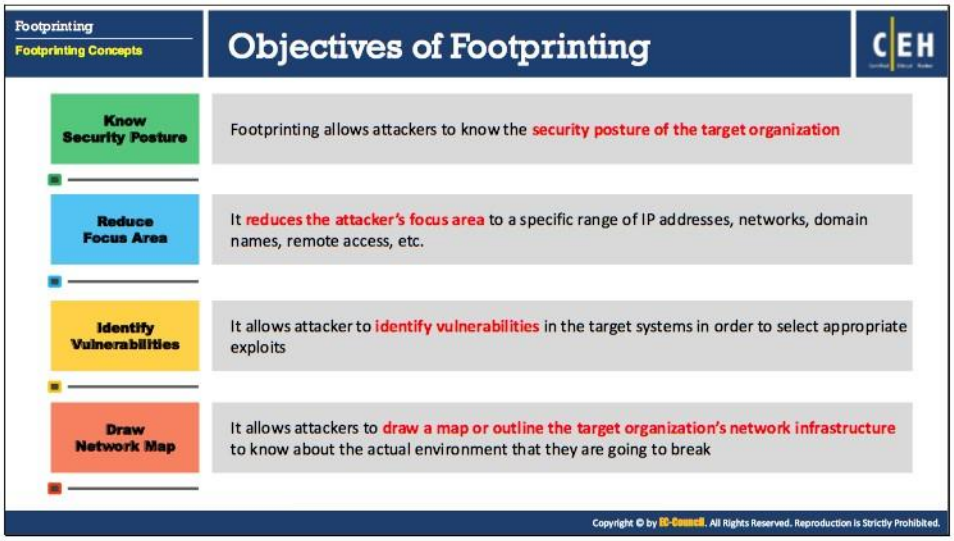
  - Web server OSes
  - Location of web servers
  - Users and passwords and so on.

- **Organization Information**: Such 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, designation, and work experience)
  - Address and mobile/telephone numbers
  - Location details
  - Background of the organization
  - Web technologies
  - News articles, press releases, and related documents

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



Objectives of Footprinting

For attackers to build a hacking strategy, they need to gather information about the target organization's network. They then use such information to locate the easiest way to break through the organization's security perimeter. As mentioned previously, footprinting methodology makes it easy to gather information about the target organization; this plays a vital role in the hacking process.

Footprinting helps to:

- **Know Security Posture**: Performing footprinting on the target organization gives the complete profile of the organization's security posture. Hackers can then analyze the report to identify loopholes in the security posture of the target organization and then build a hacking plan accordingly.
- **Reduce Focus Area**: By using a combination of tools and techniques, attackers can take an unknown entity (for example, XYZ Organization) and reduce it to a specific range of domain names, network blocks, and individual IP addresses of systems directly connected to the Internet, as well as many other details pertaining to its security posture.
- **Identify Vulnerabilities**: A detailed footprint provides maximum information about the target organization. Attackers can build their own information database about security weaknesses of the target organization. Such a database can then help in identifying the weakest chain in the link of the organization's security perimeter.
- **Draw Network Map**: Combining footprinting techniques with tools such as Tracert allows the attacker to create diagrammatic representations of the target organization's network

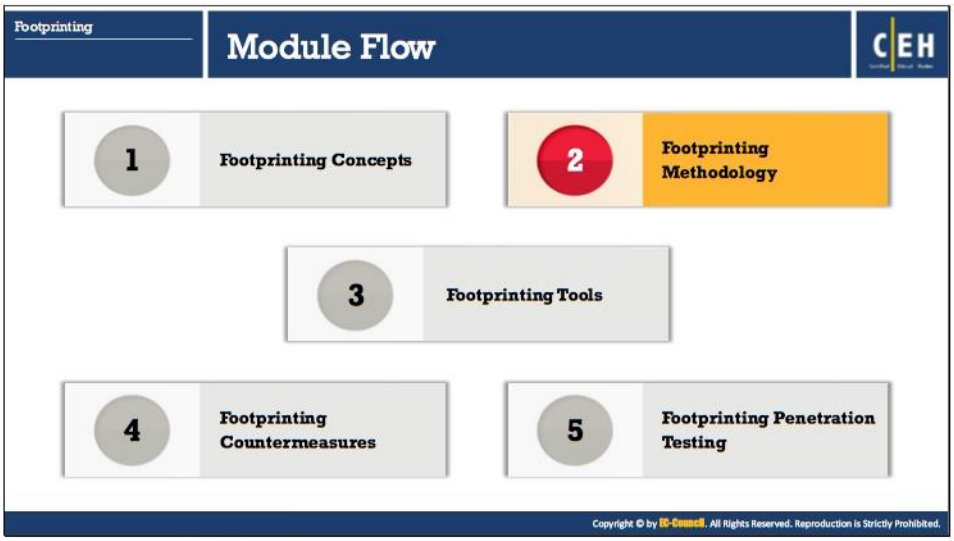
presence. A network map will depict the rogue's understanding of the target's Internet footprint. These network diagrams can guide the attacker in performing an attack.

Footprinting Threats

Attackers perform footprinting as the first step of any attack on information systems. In this phase, attackers attempt to collect valuable system-level information such as account details, operating system and other software versions, server names, database schema details, and so on that will be useful in the hacking process.

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 helps an attacker to perform system and network attacks. Through such, attackers can gather information related to the target organization's system configuration, operating system running on the machine, and so on. Using this information, rogues are able to find vulnerabilities in the target system and then exploit those vulnerabilities. Attackers can then take control over 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, those persons can mount an attack plan based on the information, or alternatively use it for monetary benefit.
- **Privacy Loss**: With the help of footprinting, hackers are able to 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 its individual personnel.
- **Corporate Espionage**: Corporate espionage is central threats to organizations, as competitors often aim to attempt to secure sensitive data through footprinting. Vis this approach, competitors are able to launch similar products in the market, alter prices, and generally adversely affect the market position of a target organization.
- **Business Loss**: Footprinting can have a major effect on organizations such as online businesses and other ecommerce websites, banking and financial related businesses. Billions of dollars are lost every year due to malicious attacks by hackers.



Footprinting Methodology

Now that you are familiar with footprinting concepts and potential threats, we will discuss its methodology. Footprinting methodology is a procedure for collecting information about a target organization from all available sources. It involves gathering information about a target organization such as URLs, locations, establishment details, number of employees, the specific range of domain names, contact information, and other related information. Attackers collect this information from publicly accessible sources such as search engines, social networking sites, Whois databases and so on. The footprinting methodology section discusses the common techniques used to collect information about the target organization from different sources.

Footprinting techniques:

- Footprinting through search engines
- Footprinting through web services
- Footprinting through social networking sites
- Website footprinting
- Email footprinting
- Competitive intelligence
- Whois footprinting
- DNS footprinting
- Network footprinting
- Footprinting through social engineering

Footprinting through Search Engines

Search engines are the main information sources to locate key information about a target organization. Search engines play a major role in extracting critical details about a target from the Internet. It returns a list of Search Engine Results Pages ('SERPs'). Many search engines can extract target organization information such as technology platforms, employee details, login pages, intranet portals, contact information and so on. The information helps attacker in performing social engineering and other types of advanced system attacks.

A Google search could reveal submissions to forums by security personnel that disclose brands of firewalls or antivirus software in use at the target. Attackers sometimes discover even the network diagrams, which enable them to launch an attack.

For example, consider an organization, perhaps Microsoft. Type **Microsoft** in the **Search** box of a search engine and press **Enter**; this will display the results containing information about Microsoft. Browsing the results often provides critical information such as physical location, contact address, the services offered, number of employees, and so on, that may prove to be a valuable source for hacking.

Examples of major search engines include Google, Yahoo, Bing, Ask, Aol, Baidu, and DuckDuckGo.

Attackers can use advanced search operators available with these search engines and create complex queries to find, filter, and sort specific information regarding the target. Search engines are also used to find other sources of publicly accessible information resources. For example, you can type "Top Job Portals" to find major job portals that provide critical information about the target organization.

As an ethical hacker, if you find any deleted pages/information about your company in SERPs or search engine cache, then seek authority to request that the search engine removes the page/information from their indexed cache.

Footprint Using Advanced Google Hacking Techniques

Google hacking refers to use of advanced Google search operators for creating complex search queries in order to extract sensitive or hidden information. The accessed information is then used by attackers to find vulnerable targets. Footprinting using advanced Google hacking techniques gathers information by Google hacking, a hacking technique to locate specific strings of text within search results using an advanced operator in the Google search engine.

Advanced Google Hacking refers to the art of creating complex search engine queries. Queries can retrieve valuable data about a target company from the Google search results. Through Google Hacking, an attacker tries to find websites that are vulnerable to exploitation. Attackers can use the Google Hacking Database ('GHDB'), a database of queries, to identify sensitive data. Google operators help in finding required text and avoiding irrelevant data. Using advanced Google operators, attackers can locate specific strings of text such as specific versions of vulnerable web applications. When a query without advanced search operators is specified, Google traces for the search terms in any part of the webpage that includes the title, text, URL and so on. In order to confine a search, Google offers advanced search operators. These search operators help to narrow down the search query and get the most relevant and accurate output.

The syntax to use an advanced search operator: operator**: search_term**

Note: Do not enter any spaces between the operator and the query.

Some of the popular Google advanced search operators include:

Source: *http://www.googleguide.com*

- **site**: This operator restricts search results to the specified site or domain.

  For example, the [games site: www.certifiedhacker.com] query gives information on games from the certifiedhacker site.

- **allinurl**: This operator restricts results to only those pages containing all the query terms specified in the URL.

  For example, the [allinurl: google career] query returns only pages containing the words "google" and "career" in the URL.

- **inurl**: This operator restricts the results to only those pages containing the word specified in the URL.

  For example, the [inurl: copy site:www.google.com] query returns only pages in Google site in which the URL has the word "copy."

- **allintitle**: This operator restricts results to only those pages containing all the query terms specified in the title.

  For example, the [allintitle: detect malware] query returns only pages containing the words "detect" and "malware" in the title.

- **intitle**: This operator restricts results to only those pages containing the specified term in the title.

  For example, the [malware detection intitle:help] query returns only pages that have the term "help" in the title, and "malware" and "detection" terms anywhere within the page.

- **inanchor**: This operator restricts results to only those pages containing the query terms specified in the anchor text on links to the page.

  For example, the [Anti-virus inanchor:Norton] query returns only pages with anchor text on links to the pages containing the word "Norton" and the page containing the word "Anti-virus."

- **allinanchor**: This operator restricts results to only those pages containing all query terms specified in the anchor text on links to the page.

  For example, the [allinanchor: best cloud service provider] query returns only pages in which the anchor text on links to the pages contain the words "best," "cloud," "service," and "provider."

- **cache**: This operator displays Google's cached version of a web page, instead of the current version of the web page.

  For example, [cache:www.eff.org] will show Google's cached version of the Electronic Frontier Foundation home page.

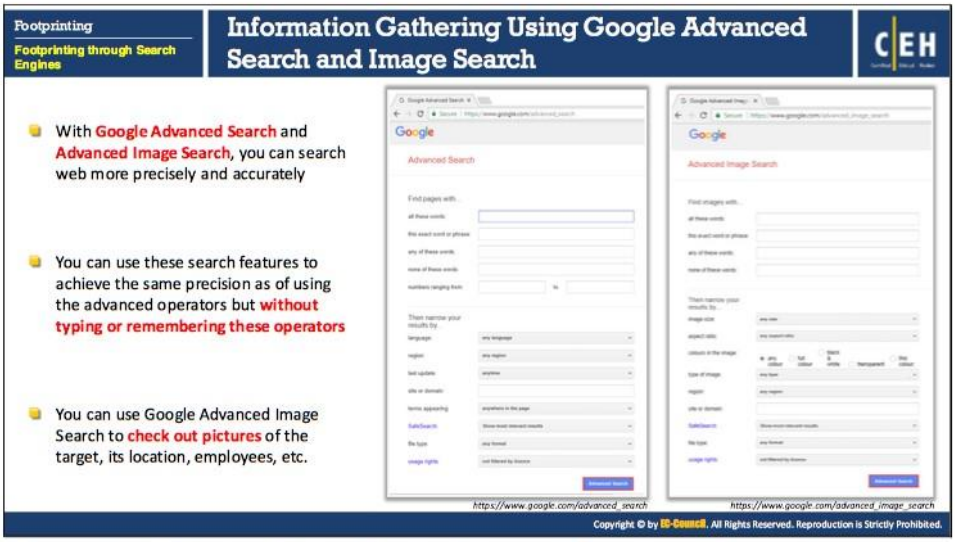
- **link**: This operator searches websites or pages that contain links to the specified website or page.

  For example, [link:www.googleguide.com] finds pages that point to Google Guide's home page.

  Note: According to Google's documentation, "you cannot combine a link: search with a regular keyword search."

Also note that when you combine link: with another advanced operator, Google may not return all the pages that match.

- **related**: This operator displays websites that are similar or related to the URL specified.
  For example, [related:www.microsoft.com] provides the Google search engine results page with websites similar to microsoft.com.
- **info**: This operator finds information for the specified web page.
  For example, [info:gothotel.com] provides information about the national hotel directory GotHotel.com home page.
- **location:** This operator finds information for a specific location.
  For Example, [location: 4 seasons restaurant] will give you results based around the term 4 seasons restaurant.
- **Filetype:** This operator allows you to search your results based on its file extension.
  For Example, [jasmine:jpg] will provide jpg files based on jasmine.



Information Gathering Using Google Advanced Search and Image Search

An attacker cannot always gather information easily from an information-rich site using only a normal search box. A complicated search involves a number of interrelated conditions.

Google's Advanced search feature helps an attacker to perform complex web searching. With **Google Advanced Search** and **Advanced Image Search**, one can search web more precisely and accurately. You can use these search features to achieve same precision as of using the advanced operators but without typing or remembering the operators. Using Google's Advanced search option, you can find sites that may link back to the target organization's website. This helps to extract information such as partners, vendors, clients, and other affiliations of the target website. You can use Google Advanced Image Search to check out pictures of the target, its location, employees and so on.

To perform an advanced search in Google, click **Settings** at the bottom-right of the **Google** home page and then choose **Advanced search** in the menu or directly type *https://www.google.com/advanced_search* in the address bar. Advanced search allows you to specify any number of criteria that the search must match, as this pattern builds on the search box pattern by adding more search options. To do this, you choose a field. Then enter the string you want to search for in the field's text box, and click on the **Advanced Search** button. By default, various values are joined together with "and" (meaning all of them need to match) except for sets, blocks and formats, which are joined together with "or" (meaning any of them can match).

To perform an advanced image search in Google, type *https://www.google.com/advanced_image_search* in the address bar. Advanced image search allows you to tweak your image search in a number of ways. You can search based on image

color, domain, file type, size, keyword, and so on. To do this, you choose a field. Then enter the string you want to search for in the field's text box, and click on the **Advanced Search** button.

What can a Hacker do with Google Hacking?

An attacker can create complex search engine queries in order to filter large amounts of search results to obtain information related to computer security. The hacker uses Google operators that help to locate such specific strings of text within the search results. Doing so, an attacker is able to detect websites and web servers that are vulnerable to exploitation, as well as locate private, sensitive information about others, such as credit card numbers, social security numbers, passwords and so on. Once a vulnerable site is identified, attackers try to launch various possible attacks such as buffer overflows, SQL Injection, among others that compromise information security.

Examples of sensitive information left 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 firewall logs
- Advisories and server vulnerabilities
- Software version information
- Web application source code

Example: Use Google Advance Operator syntax `[intitle:intranet inurl:intranet +intext:"human resources"]` to find sensitive information about a target organization and its employees. Attackers use the gathered information to perform social engineering attacks.

The screenshot below shows a Google search engine results page displaying the results for the query mentioned before.

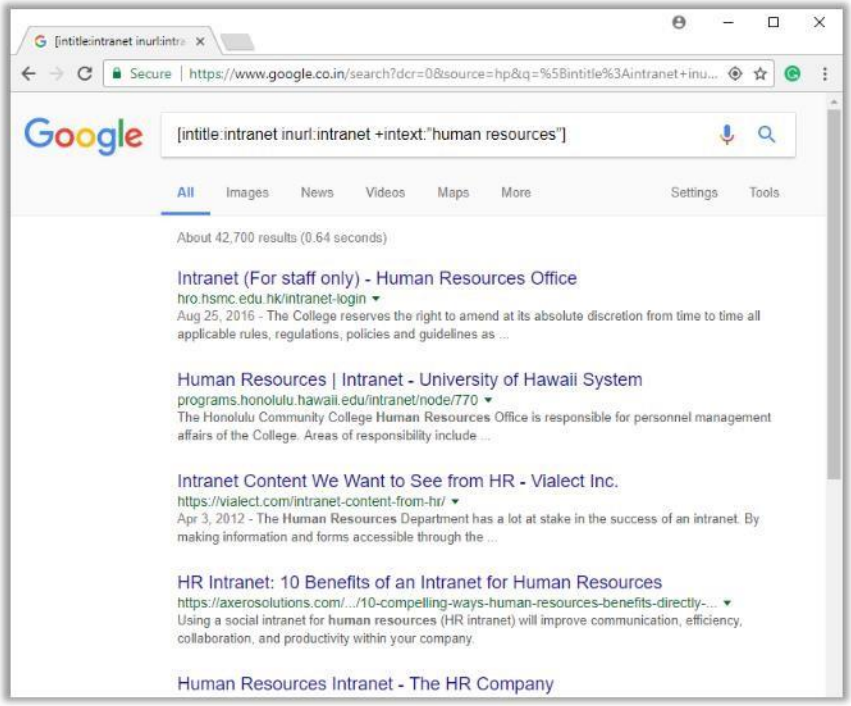


FIGURE 2.1: Search engine showing results for given Google Advance Operator syntax

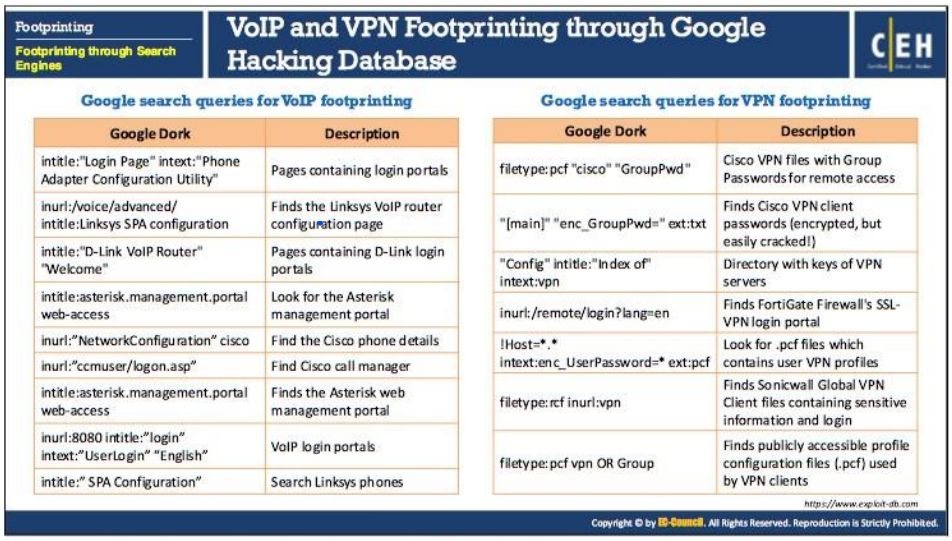
Google Hacking Database

Source: *https://www.exploit-db.com*

The Google Hacking Database ('GHDB') is an authoritative source for querying the ever-widening reach of the Google search engine. In the GHDB, you will find search terms for files containing usernames, vulnerable servers, and even files containing passwords. The Exploit Database is a Common Vulnerabilities and Exposures (CVE) compliant archive of public exploits and corresponding vulnerable software, developed for use by penetration testers and vulnerability researchers.

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



VoIP and VPN Footprinting through Google Hacking Database

Google hacking involves the implementation of advanced operators in the Google search engine to match for the specific strings of text within the search result. These advanced operators help refine searches to expose sensitive information, vulnerabilities, and passwords. You can use these google hacking operators or Google dorks for footprinting VoIP and VPN networks. You can extract information such as pages containing login portals, VoIP login portals, directory with keys of VPN servers, and so on.

The following tables list some of the google hacking operators or google dorks to obtain specific information related to VoIP and VPN footprinting respectively.

Google search queries for VoIP footprinting

Google Dork	Description
intitle:"Login Page" intext:"Phone Adapter Configuration Utility"	Pages containing login portals
inurl:/voice/advanced/ intitle:Linksys SPA configuration	Finds the Linksys VoIP router configuration page
intitle:"D-Link VoIP Router" "Welcome"	Pages containing D-Link login portals
intitle:asterisk.management.portal web-access	Look for the Asterisk management portal
inurl:"NetworkConfiguration" cisco	Find the Cisco phone details
inurl:"ccmuser/logon.asp"	Find Cisco call manager

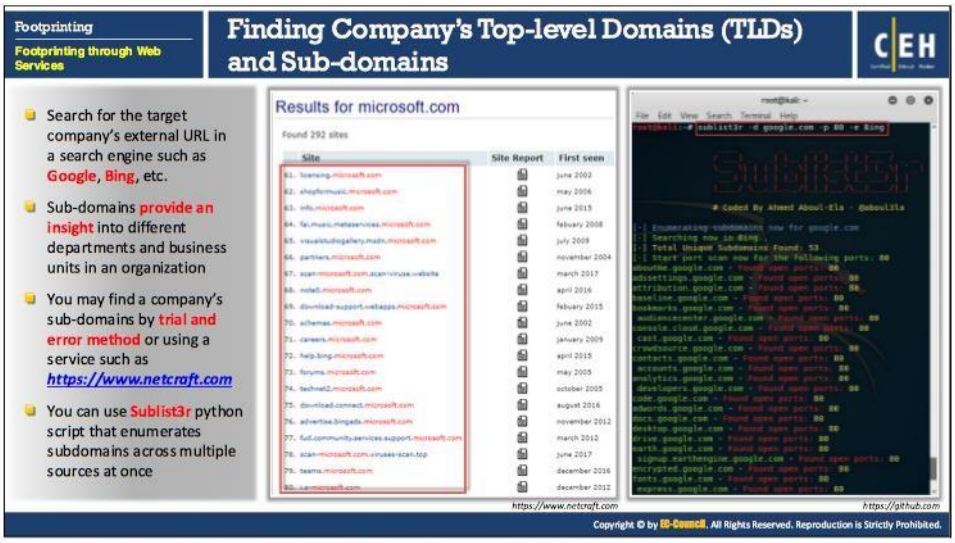
intitle:asterisk.management.portal web-access	Finds the Asterisk web management portal
inurl:8080 intitle:"login" intext:"UserLogin" "English"	VoIP login portals
intitle:" SPA Configuration"	Search Linksys phones

TABLE 2.1: Google search queries for VoIP footprinting

Google search queries for VPN footprinting

Google Dork	Description
filetype:pcf "cisco" "GroupPwd"	Cisco VPN files with Group Passwords for remote access
"[main]" "enc_GroupPwd=" ext:txt	Finds Cisco VPN client passwords (encrypted, but easily cracked!)
"Config" intitle:"Index of" intext:vpn	Directory with keys of VPN servers
inurl:/remote/login?lang=en	Finds FortiGate Firewall's SSL-VPN login portal
!Host=*.* intext:enc_UserPassword=* ext:pcf	Look for .pcf files which contains user VPN profiles
filetype:rcf inurl:vpn	Finds Sonicwall Global VPN Client files containing sensitive information and login
filetype:pcf vpn OR Group	Finds publicly accessible profile configuration files (.pcf) used by VPN clients

TABLE 2.2: Google search queries for VPN footprinting



Footprinting through Web Services

Web services such as people search services can provide sensitive information about the target. Internet archives may also provide sensitive information that has been removed from the World Wide Web ('WWW'). Social networking sites, people search services, alerting services, financial services and job sites provide information about a target such as infrastructure details, physical location, and employee details. Moreover, groups, forums, and blogs can help attackers in gathering sensitive information about a target such as public network information, system information, and personal information. Using this information, an attacker may build a hacking strategy to break into the target organization's network and may carry out other types of advanced system attacks.

Finding Company's Top-level Domains (TLDs) and Sub-domains

A company's top-level domains ('TLDs') and sub-domains can provide a lot of useful information to an attacker. A public website is designed to show the presence of an organization on the Internet. It is available for free access and is accessible by anyone. It is designed to attract customers and partners. It may contain information such as organizational history, services and products, and contact information. The target organization's external URL can be located with the help of search engines such as Google, Bing among others.

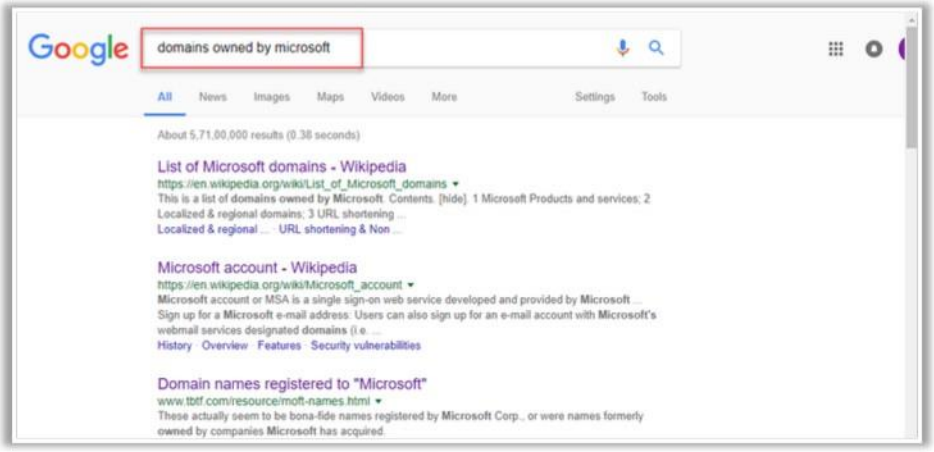


FIGURE 2.2: Google Search Engine showing results for given syntax

The sub-domain is available to only a few people. These persons may be employees of an organization or members of a department. Sub-domains provide an insight into different departments and business units in an organization. Access restrictions can be applied based on the IP address, domain or subnet, username, and password. The sub-domain helps to access the private functions of an organization. Most organizations use common formats for sub-domains. Therefore, a hacker who knows the external URL of a company can often discover the sub-domain through trial and error, or by using a service such as Netcraft.

Tool to Search Company's Sub-domains

- **Netcraft**

  Source: *https://www.netcraft.com*

  Netcraft provides internet security services including anti-fraud and anti-phishing services, application testing and PCI scanning. They also analyze the market share of web servers, operating systems, hosting providers and SSL certificate authorities and other parameters of the internet.

- **Sublist3r**

  Source: *https://github.com*

  Sublist3r is a python script designed to enumerate subdomains of websites using OSINT. It enables you to enumerate subdomains across multiple sources at once. It helps penetration testers and bug hunters collect and gather subdomains for the domain they are targeting. It enumerates subdomains using many search engines such as Google, Yahoo, Bing, Baidu, and Ask. It also enumerates subdomains using Netcraft, Virustotal,

ThreatCrowd, DNSdumpster, and ReverseDNS. It has integrated the venerable SubBrute, allowing you to also brute force subdomains using a wordlist.

Syntax:

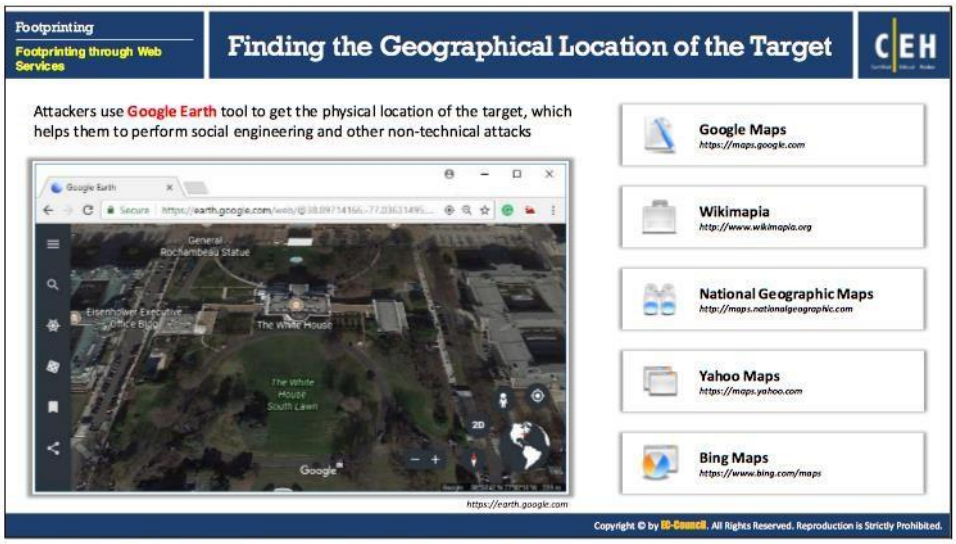
```
sublist3r [-d DOMAIN] [-b BRUTEFORCE] [-p PORTS] [-v VERBOSE][-t
THREADS] [-e ENGINES] [-o OUTPUT]
```

Short Form	Long Form	Description
-d	--domain	Domain name to enumerate subdomains of
-b	--bruteforce	Enable the subbrute bruteforce module
-p	--ports	Scan the found subdomains against specific TCP ports
-v	--verbose	Enable the verbose mode and display results in realtime
-t	--threads	Number of threads to use for subbrute bruteforce
-e	--engines	Specify a comma-separated list of search engines
-o	--output	Save the results to text file
-h	--help	Show the help message and exit

TABLE 2.3: Table showing Sublist3r options with description

Examples:

Search for subdomains of google.com (-d google.com) using the Bing search engine (-e Bing) with port 80 (-p 80)



Finding the Geographical Location of the Target

Information such as the physical location of an organization plays a vital role in the hacking process. Attackers can obtain this information using footprinting. In addition to physical location, a hacker can also collect information such as surrounding public Wi-Fi hotspots that may prove to be a way to break into the target organization's network.

Attackers with the knowledge of a target organization's location may attempt dumpster diving, surveillance, social engineering, and other non-technical attacks to gather more information. Once the attackers discern the location of the target, they can obtain detailed satellite images of the location using various sources available on the Internet such as Google Earth and Google Maps. Attackers can use this information to gain unauthorized access to buildings, wired and wireless networks, systems.

Example: Google Earth (*https://earth.google.com*)

The Google Earth tool allows you to find the exact location of a target. It can even provide access to 3D images that depict most of the populated Earth's surface in high-resolution detail. The detail allows you to view places street view, altitude, and even coordinates.

Tools for Finding the Geographical Location

The tools for geographical location allow you to find and explore most locations on the earth. They provide information such as images of buildings, as well as surroundings, including Wi-Fi networks. Tools such as Google Maps even locate entrances of building, security cameras, and gates. These tools provide interactive maps, outline maps, satellite imagery, and information on how to interact with and create one's own maps. Google Maps, Yahoo Maps, and other tools

provide driving directions, traffic conditions, locate landmarks, give us detailed information about address and contact information.

Attackers may use these tools to find or locate entrances to buildings, security cameras, gates, places to hide, weak spots in perimeter fences, and utility resources like electricity connections, to measure distance between different objects, and so on.

Some of the tools used to find geographical location information include:

- Google Maps (*https://maps.google.com*)
- Wikimapia (*http://www.wikimapia.org*)
- National Geographic Maps (*http://maps.nationalgeographic.com*)
- Yahoo Maps (*https://maps.yahoo.com*)
- Bing Maps (*https://www.bing.com/maps*)

People Search on Social Networking Sites

Searching for a particular person on a social networking website is easier than most would think. 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 that users provide in their profiles. They help to directly or indirectly relate people to each other through various fields such as common interests, work location, and educational communities.

Social networking sites allow people to share information quickly, as they can update their personal detail in real time. Such sites allow updating facts about upcoming or current events, recent announcements and invitations, and so on. Social networking sites are a great platform for searching 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 task. A user can search a person using 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 allow a you to find people by name, keyword, company, school, their 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, a 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. Through people searching on social networking services, an attacker can gather critical information that will be helpful in performing social engineering or other kinds of attacks.

Some of the social networking sites are:

- Facebook (*https://www.facebook.com*)
- LinkedIn (*https://www.linkedin.com*)
- Twitter (*https://twitter.com*)
- Google+ (*https://plus.google.com*)
- YouTube (*https://www.youtube.com*)
- Pinterest (*https://www.pinterest.com*)
- Instagram (*https://www.instagram.com*)
- MySpace (*https://myspace.com*)

People Search on People Search Services

You can use public record websites to find information about email addresses, phone numbers, house addresses, and other information. Many individuals use people search online services to find information about other people. Generally, people search online services provide people's names, addresses, contact details, date of birth, photographs, videos, profession, details about their family and friends, social networking profiles, property information and optional background on criminal checks. Often people search online services may also reveal the type of work an individual does, businesses owned by a person, upcoming projects and operating environment, website and blogs, contact numbers, important dates, company email addresses, cell phone numbers, fax numbers, and personal e-mail addresses. Using this information, a hacker can try to obtain bank details, credit card details, past history, and so on. This information proves to be highly beneficial for attackers to launch attacks. There are many people search online services available that help find people. Examples of such people search services include pipl, and AnyWho.

- **People search service - pipl**

  Source: *https://pipl.com*

  pipl is an online people search tool to find other users through their name, email, username or phone number. It has an Identity Resolution engine that focuses on finding the right person and provides accurate results for people search.

Some of the people search online services include:

- Intelius (*https://www.intelius.com*)
- BeenVerified (*https://www.beenverified.com*)
- Spokeo (*https://www.spokeo.com*)
- AnyWho (*https://www.anywho.com*)
- US Search (*https://ussearch.com*)
- 411 (*http://www.411.com*)
- Veromi (*http://www.veromi.net*)
- PrivateEye (*http://www.privateeye.com*)
- Public Background Checks (*http://www.publicbackgroundchecks.com*)
- Zaba Search (*http://www.zabasearch.com*)
- WebMii (*http://webmii.com*)

Gathering Information from LinkedIn

LinkedIn is a social networking website for professionals. It connects the world's human resources to aid productivity and success. The site contains personal information such as name, position, organization name, current location, educational qualifications, and so on. Information gathered from LinkedIn helps an attacker in performing social engineering or other kinds of attacks.

Attackers can use InSpy utility to gather information from LinkedIn based on job title, company, or email address.

- **InSpy**

  Source: *https://github.com*

  InSpy is a python based LinkedIn enumeration tool. It performs enumeration on LinkedIn and finds people based on job title, company, or email address. InSpy has two functionalities:

  - **TechSpy**: Crawls LinkedIn job listings for technologies used by the provided company. InSpy attempts to identify technologies by matching job descriptions to keywords from a new line delimited file.
  - **EmpSpy**: Crawls LinkedIn for employees working at the provided company. InSpy searches for employees by title and/or departments from a new line delimited file. InSpy may also create emails for the identified employees if the user specifies an email format.

Gathering Information from Financial Services

Attackers who seek access to personal information or financial information often target financial data such as stock quotes and charts, financial news, and portfolios. Financial services such as Google Finance, and Yahoo! Finance, can provide a lot of useful information such as the market value of a company's shares, company profile, competitor details, stock exchange rates, corporate press releases, financial reports along with news and blog search articles about each corporation. The information offered varies from one service to the other. Financial firms rely on web services to perform transactions and grant users access to their accounts. Attackers can obtain sensitive and private information on these firms and by using malware, exploiting software design flaws, breaking authentication mechanisms, service flooding, performing brute force attacks and phishing attacks.

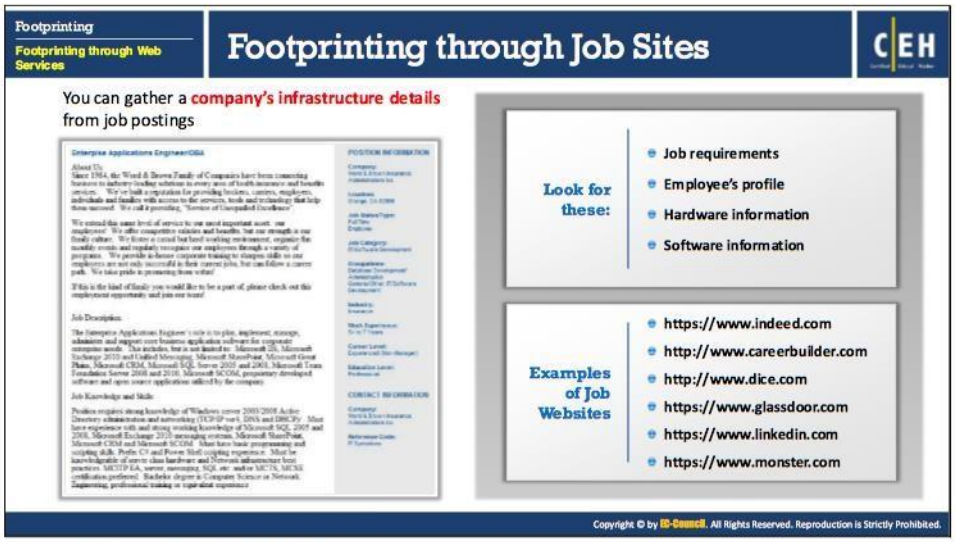
- **Google Finance**

  Source: *https://www.google.com/finance*

  Google finance service features business and enterprise headlines for many corporations, including their financial decisions and major news events. Stock information is available, as are stock price charts that contain marks for major news events and corporate actions. The site also aggregates Google news and Google blog search articles about each corporation.

Some of the financial services include:

- Yahoo! Finance (*https://finance.yahoo.com*)
- TheStreet (*https://www.thestreet.com*)
- MarketWatch (*https://www.marketwatch.com*)

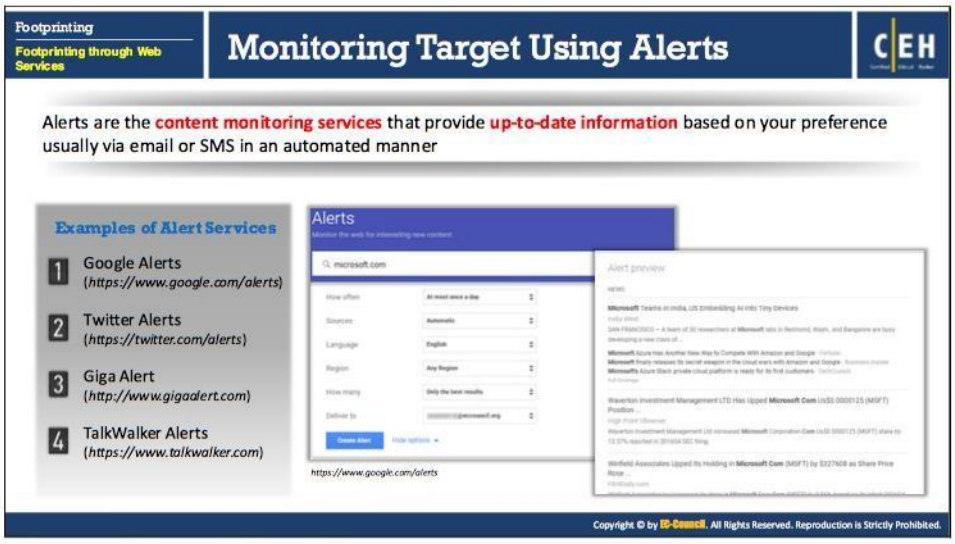


Footprinting through Job Sites

Attackers can gather valuable information about the operating system, software versions, company's infrastructure details, and database schema of an organization, through footprinting job sites using different techniques. Many organizations' websites provide recruiting information on a job posting page that in turn reveals hardware and software information, network-related information, and technologies used by the company (e.g., firewall, Internal server type, OS used, network appliances and so on.). In addition, the website may have a key employee list with email addresses. Such information may prove to be beneficial for an attacker. For example, if an organization advertises a Network Administrator job, it posts the requirements related to that position.

Some of the popular job sites include:

- *https://www.indeed.com*
- *http://www.careerbuilder.com*
- *http://www.dice.com*
- *https://www.glassdoor.com*
- *https://www.linkedin.com*
- *https://www.monster.com*
- *http://www.simplyhired.com*
- *https://www.usajobs.gov*
- *https://www.idealist.org*



Monitoring Target Using Alerts

Alerts are content monitoring services that provide automated up-to-date information based on user preference, usually via email or SMS. To receive alerts, a user must register on the website and provide either an email address or phone number. Online alert services automatically notify users when new content from news, blogs, and discussion groups matches a set of search terms selected by the user. These services provide up-to-date information about competitors and industry. Alerts are sent via email or SMS notifications.

Some of these tools also help to track mentions of the organization's name, member names, website, or any people or projects that are important. Attackers can gather updated information about the target periodically from the alert services and use it for further attacks.

- **Google Alerts**

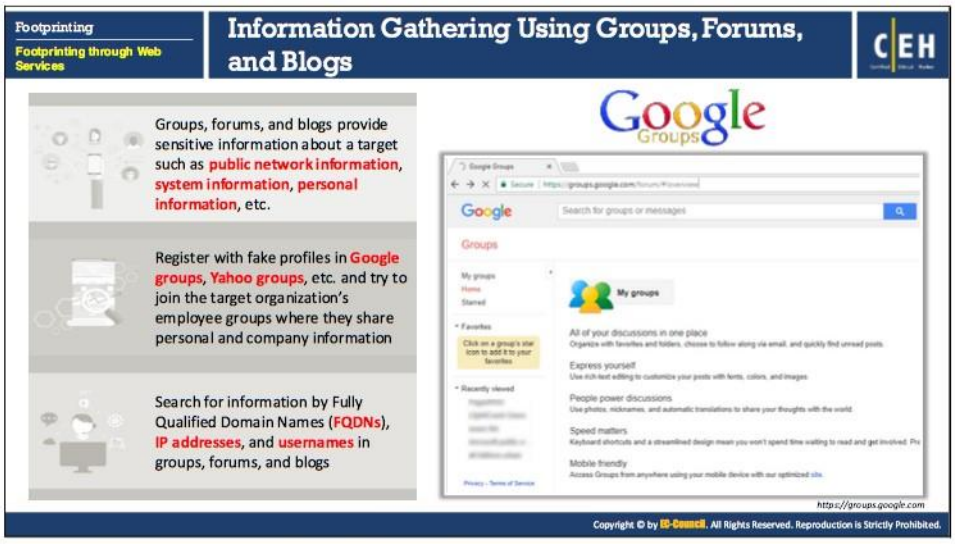
  Source: *https://www.google.com/alerts*

  Google Alerts automatically notifies users when new content from news, web, blogs, video, and/or discussion groups matches a set of search terms selected by the user and stored by the Google Alerts service.

  Google Alerts aids in monitoring a developing news story, keeping current on a competitor or industry, getting the latest on a celebrity or event, and keeping tabs on sports teams.

Some of the more popular alert services include:

- Twitter Alerts (*https://twitter.com/alerts*)
- Giga Alert (*http://www.gigaalert.com*)
- TalkWalker Alerts (*https://www.talkwalker.com*)

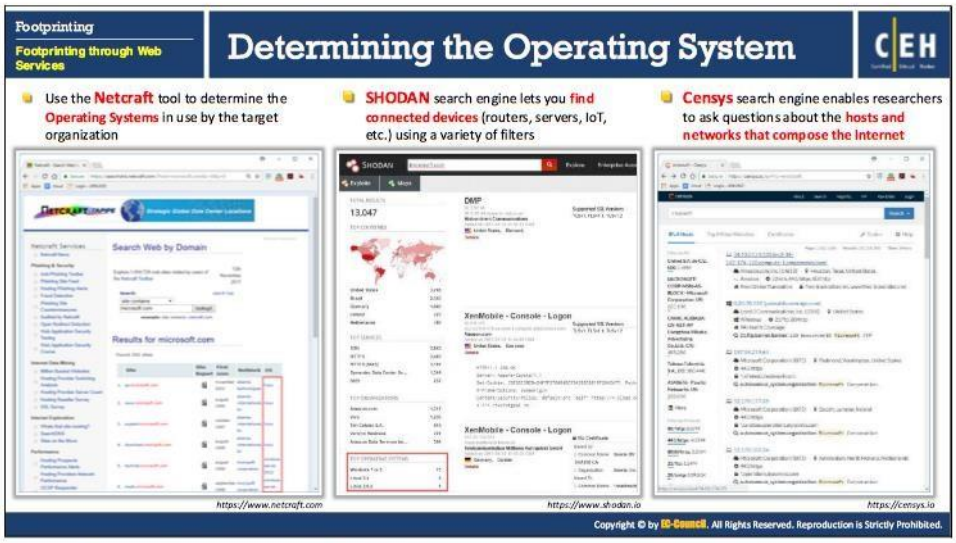


Information Gathering Using Groups, Forums, and Blogs

Many Internet users use blogs, groups, and forums for knowledge sharing purposes. For this reason, attackers often focus on groups, forums, and blogs to find information about a target organization and its people. Organizations generally fail to monitor the exchange of information that employees reveal to other users in forums, blogs, and group discussions. Attackers see this as an advantage and collect sensitive information about the target such as public network information, system information, and employee personal information. Attackers can register with fake profiles in Google groups, Yahoo groups and so on. and try to join the target organization's employee groups, where they can share personal and company information. Attackers can also search for information in groups, forums, and blogs by Fully Qualified Domain Names (FQDNs), IP addresses, and usernames.

Employee information that an attacker can gather from groups, forums, and blogs can include:

- Full name of the employee
- Place of work and residence
- Home telephone, cell number, or office number
- Personal and organizational email address
- Pictures of the employee residence or work location that include identifiable information
- Pictures of employee awards and rewards or upcoming goals



Determining the Operating System

- **Netcraft**

  Source: *https://www.netcraft.com*

  The technique of obtaining information about the target network operating system is OS fingerprinting. Open ***https://www.netcraft.com*** in the browser and type the domain name of the target network in the **What's that site running?** Field. Netcraft displays the sites associated with that domain along with the operating system running at each site.

- **SHODAN Search Engine**

  Source: *https://www.shodan.io*

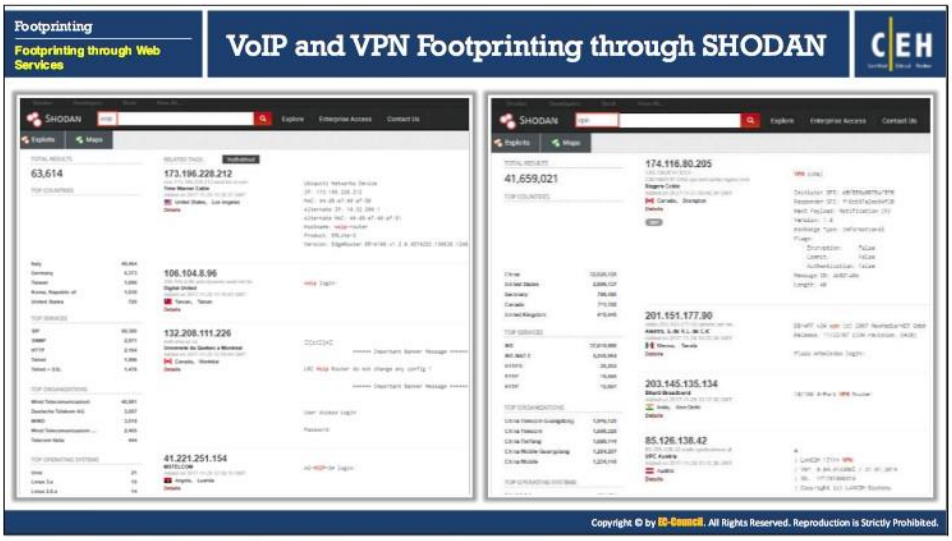
  Shodan is the computer search engine that searches the Internet for connected devices (routers, servers, and IoT.). You can use Shodan to discover which devices are connected to the Internet, where they are located and who is using them. It allows one to keep track of all the devices on the network that are directly accessible from the Internet. It also allows the user to find devices based on city, country, latitude/longitude, hostname, operating system, and IP address. Further it helps the user to search for known vulnerabilities and exploits across Exploit DB, Metasploit, CVE, OSVDB, and Packetstorm with a single interface.

  This information helps attacker to identify potential vulnerabilities and find effective exploits.

- **Censys**

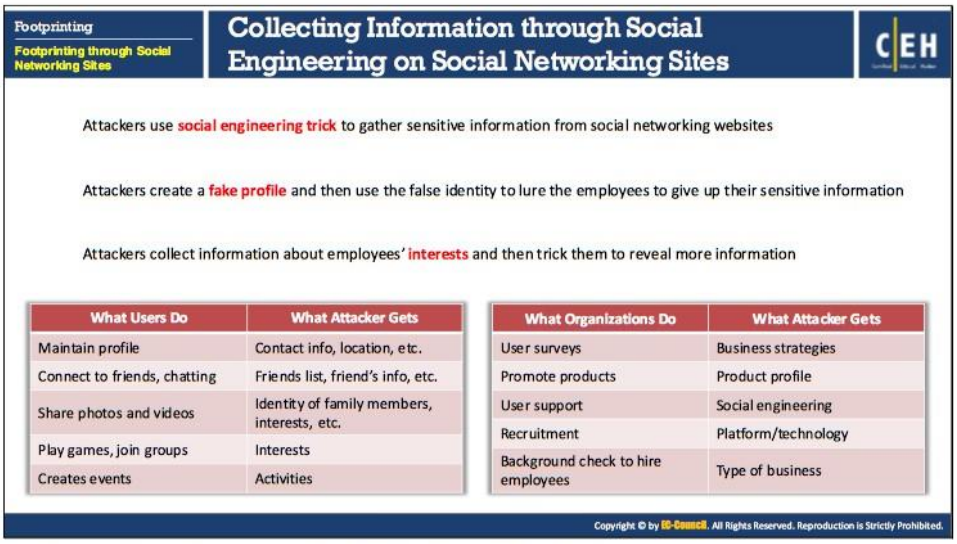
  Source: *https://censys.io*

  Censys is a search engine that enables researchers to ask questions about the hosts and networks that compose the Internet. Censys collects data on hosts and websites through daily ZMap and ZGrab scans of the IPv4 address space, in turn maintaining a database of how hosts and websites are configured. Researchers can interact with this data through a search interface, report builder, and SQL engine.



VoIP and VPN Footprinting through SHODAN

Shodan is a search engine that is helpful for the hacker community to perform footprinting at various levels. It is used to detect the devices and networks with vulnerabilities. A search in Shodan for VoIP and VPN footprinting can deliver various results, which will be helpful to gather VPN and VoIP related information. The following figures show some of the VPN and VoIP footprinting search results through Shodan.



Footprinting through Social Networking Sites

While footprinting through social networking sites may seem similar to footprinting through social engineering (such is discussed in more 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 by means of social engineering, and how it can be done.

Collecting Information through Social Engineering on Social Networking Sites

Social networking sites are the online services, platforms, or other sites that allow people to connect with each other and to build interpersonal relations. The use of social networking sites is increasing rapidly. Examples of such sites include LinkedIn, Facebook, MySpace, Twitter, Pinterest, Google+, Instagram and so on. Each social networking site has its own purpose and features. One site may connect friends, family and so on, while another helps users to share professional profiles. Social networking sites are open to everyone. Attackers may take advantage of this to gather sensitive information from users either by browsing through users' public profiles or by creating a fake profile to pose as a genuine user. On social networking sites, people may post personal information such as date of birth, educational information, employment backgrounds, spouse's names and so on. Organizations often post information such as potential partners, websites, and upcoming news about the company.

For an attacker, social networking sites can be great sources of information about the target person or organization. The attacker can only gather information that is posted by individuals.

There are no barriers for attackers to access the public pages of accounts created on the social networking sites. To obtain more information about the target, attackers may create fake accounts and use social engineering techniques to lure the victim to reveal more information. For example, the attacker can send a friend request to the target person from the fake account; if the victim accepts the request, then the attacker can access even the restricted pages of the target person on that website. Social networking sites often prove to be valuable information resources for attackers.

Information Available on Social Networking Sites

So far, we have discussed *how* an attacker can collect information from social networking sites. Now we will discuss *what* information an attacker can get from social networking sites.

People usually maintain profiles on social networking sites in order to provide basic information about themselves and to help make and keep connections with others. The profile generally contains information such as name, contact information (cell phone number, email address), friends' information, information about family members, their interests, activities and so on. People usually connect with friends and chat with them. Attackers can gather sensitive information through these chats. Social networking sites also allow people to share photos and videos. If the people fail to set privacy settings for their albums, then attackers can see the pictures and videos shared by them. Users may join groups to play games or to share their views and interests. Attackers can collect information about a victim's interests by tracking his or her groups and then can mislead the victim into revealing more information. Users may create events to notify other users about upcoming occasions, from which attackers will come to know user activities.

The activities of users on the social networking sites and the respective information that an attacker can collect is shown in the following table.

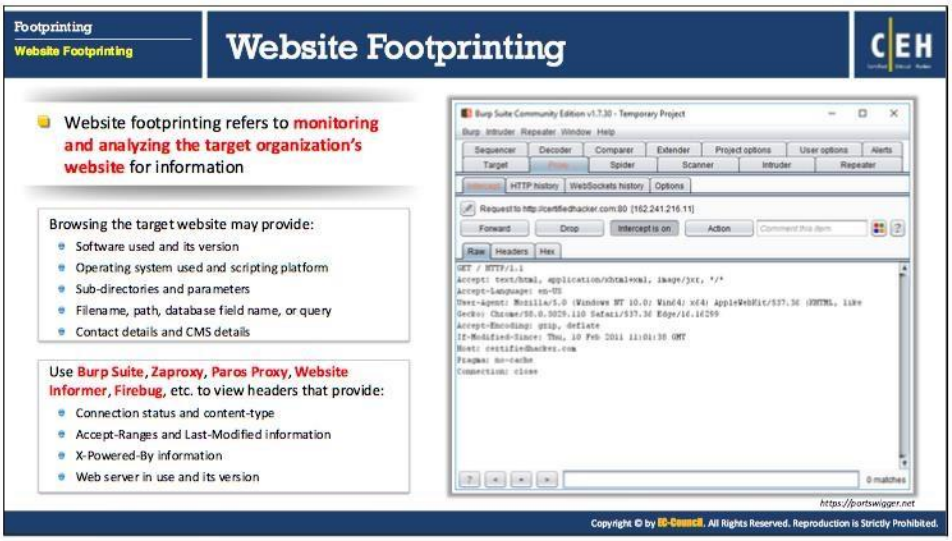
What Users Do	What Attacker Gets
Maintain profile	Contact info, location and related information.
Connect to friends, chatting	Friends list, friend's info and related information.
Share photos and videos	Identity of a family members, interests and related information.
Play games, join groups	Interests
Creates events	Activities

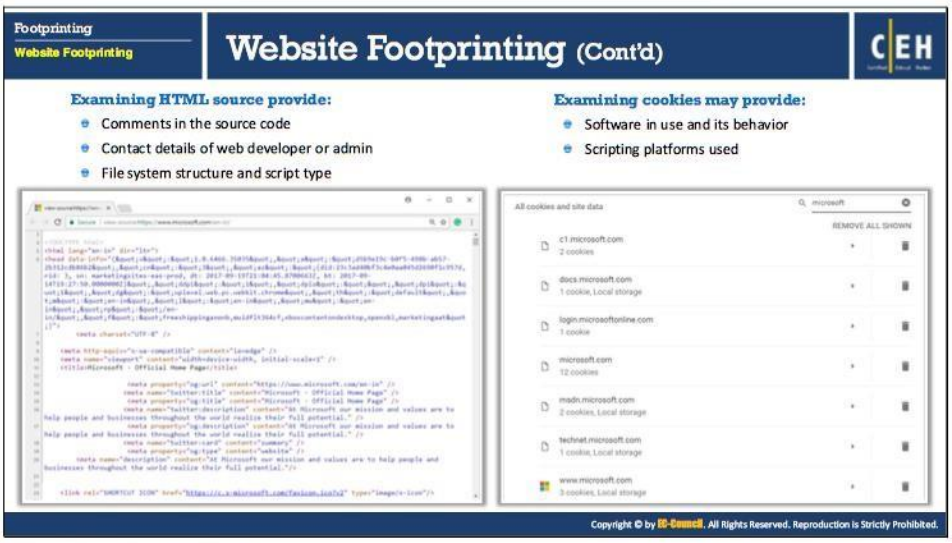
TABLE 2.4: Table showing activities of users on the social networking sites and the respective information

Like individuals, organizations also use social networking sites to connect with people, promote their products, and to gather feedback about their products and services. The activities of an organization on the social networking sites and the respective information that an attacker can collect is shown in below table.

What Organizations Do	What Attacker Gets
User surveys	Business strategies
Promote products	Product profile
User support	Social engineering
Recruitment	Platform/technology information
Background check to hire employees	Type of business

TABLE 2.5: Table showing activities of organization on the social networking sites and the respective information





Website Footprinting

So far, we have discussed footprinting through social networking sites. Here on, we will discuss website footprinting. An organization's website is the first place to get sensitive information such as names and contact details of the leaders of the organization, upcoming project details and so on. This section covers the website footprinting concept, mirroring websites, the tools used for mirroring, and monitoring web updates.

Website footprinting refers to monitoring and analyzing the target organization's website for information. It is possible for an attacker to build a detailed map of a website's structure and architecture without triggering the IDS or without raising any system administrator's suspicions. Attackers use sophisticated footprinting tools or the basic tools that come with the operating system, such as Telnet, or by using a browser.

The Netcraft tool can gather website information such as IP address, registered name and address of the domain owner, domain name, host of the site, and OS details. However, the tool may not give all these details for every site. In such cases, the attacker can browse the target website.

Browsing the target website will typically provide the following information:

- **Software used and its version**: An attacker can easily find the software and version in use on an off-the-shelf software-based website.
- **Operating system used**: Usually the operating system in use can also be determined.
- **Sub-directories and parameters**: Searches can reveal the sub-directories and parameters by making a note of the URLs while browsing the target website.
- **Filename, path, database field name, or query**: The attacker will often carefully analyze anything after a query that looks like a filename, path, database field name, or query to check whether it offers opportunities for SQL injection.
- **Scripting platform**: With the help of script filename extensions such as .php, .asp, or .jsp, one can easily determine the scripting platform that the target website is using.
- **Contact details and CMS details**: The contact pages usually offer details such as names, phone numbers, email addresses, and locations of admin or support people. An attacker can use these details to perform a social engineering attack. CMS software allows URL rewriting in order to disguise the script filename extensions, if the attacker is willing to put in a little more effort to determine the scripting platform.

Use Burp Suite, Zaproxy, Paros Proxy, Website Informer, and Firebug to view headers that provide:

- Connection status and content-type
- Accept-Ranges and Last-Modified information
- X-Powered-By information
- Web server in use and its version

Burp Suite

Source: *https://portswigger.net*

Burp Suite is an integrated platform for performing security testing of web applications. Its various tools work together to support the entire testing process, from initial mapping and analysis of an application's attack surface, through finding and exploiting security vulnerabilities.

Website footprinting can be performed by examining HTML source code and cookies.

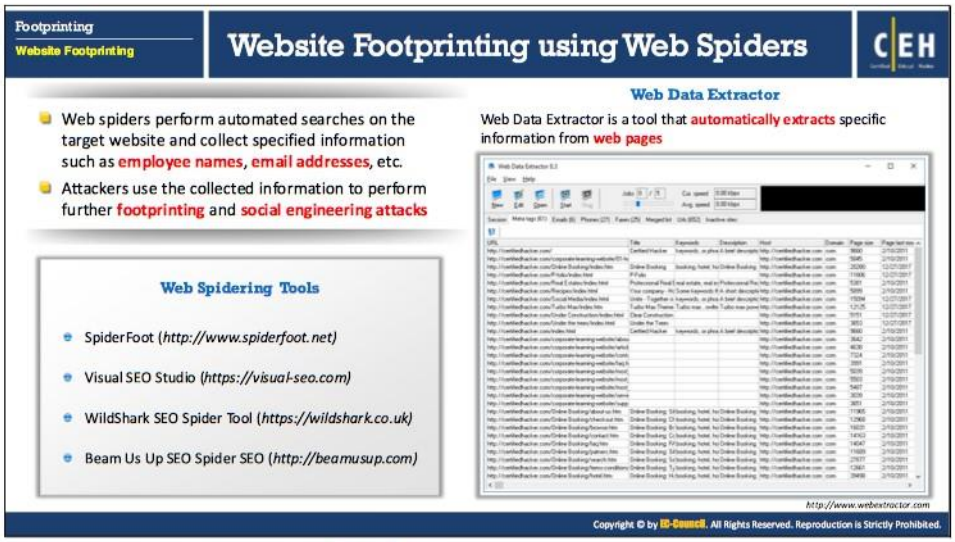
- **Examining the HTML source code**

  Attackers can gather sensitive information by examining the HTML source code, and following the comments that are inserted manually or that the CMS system creates. The comments may provide clues to what is running in the background. This may even provide contact details of the web developer or administrator.

  Observe all the links and image tags, in order to map the file system structure. This will reveal the existence of hidden directories and files. Enter fake data to determine how the script works. It is sometimes possible to edit the source code.

- **Examining Cookies**

  To determine the software running and its behavior, one can examine cookies set by the server. Identify the scripting platforms by observing sessions and other supporting cookies. The information about cookies name, value, domain size can also be extracted.



Website Footprinting using Web Spiders

Web spider (also known as web crawler or web robot) is a program or automated script that browses websites in a methodical manner to collect specified information such as employee names, email addresses and so on. Attackers then use the collected information to perform footprinting and social engineering attacks. Web spidering fails if the target website has the robots.txt file in its root directory, with a listing of directories to prevent crawling.

Web Spidering Tools

Web spidering tools can collect sensitive information from the target website.

- **Web Data Extractor**

  Source: *http://www.webextractor.com*

  Web Data Extractor automatically extracts specific information from web pages. It extracts targeted contact data (email, phone, and fax) from the website, extracts the URL and meta tag (title, description, keyword) for website promotion, searches directory creation, web research and so on. With Web Data Extractor, one can automatically get lists of meta-tags, e-mail addresses, phone and fax numbers and so on. and store them in different formats for future use.

Some of the additional web spidering tools are listed following:

- SpiderFoot (*http://www.spiderfoot.net*)
- Visual SEO Studio (*https://visual-seo.com*)
- WildShark SEO Spider Tool (*https://wildshark.co.uk*)

- Beam Us Up SEO Spider SEO (*http://beamusup.com*)
- Scrapy (*https://scrapy.org*)
- Screaming Frog (*https://www.screamingfrog.co.uk*)
- Xenu (*http://home.snafu.de*)

Mirroring Entire Website

Website mirroring is the process of creating an exact replica or clone of the original website. Users can duplicate the websites by using mirroring tools such as HTTrack Web Site Copier, and NCollector Studio. These tools download a website to a local directory, building recursively all directories, HTML, images, flash, videos, and other files from the web server to another computer.

Website mirroring has the following benefits:

- It is helpful for offline site browsing
- It supports an attacker in spending more time viewing and analyzing the website for vulnerabilities and loop holes
- It assists in finding directory structure and other valuable information from the mirrored copy without multiple requests to the web server

Website Mirroring Tools

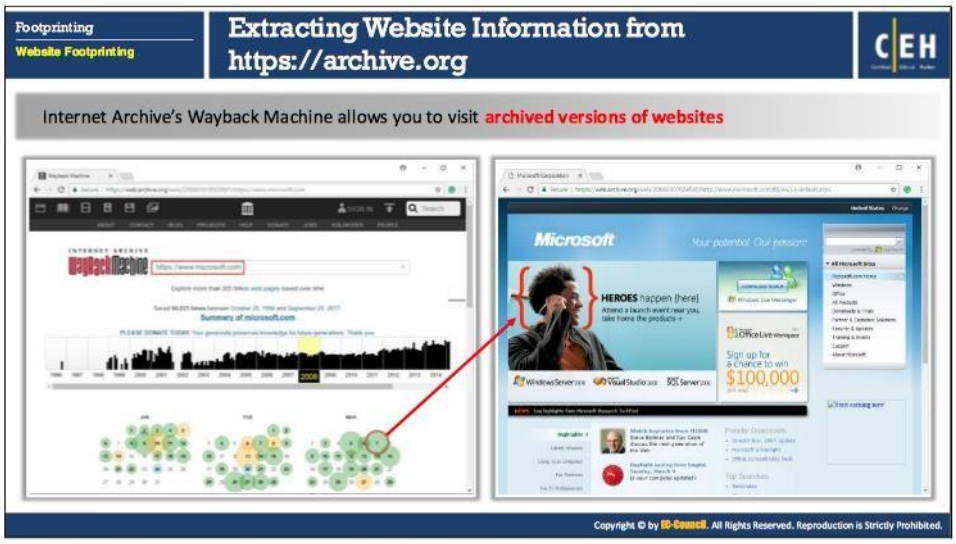
- **HTTrack Web Site Copier**

  Source: *http://www.httrack.com*

  HTTrack is an offline browser utility. It downloads a website from the Internet to a local directory, building recursively all directories, getting HTML, images, and other files from the web server to another computer. HTTrack arranges the original site's relative link-structure. HTTrack can also update an existing mirrored site, and resume interrupted downloads.

Some of the additional website mirroring tools are listed:

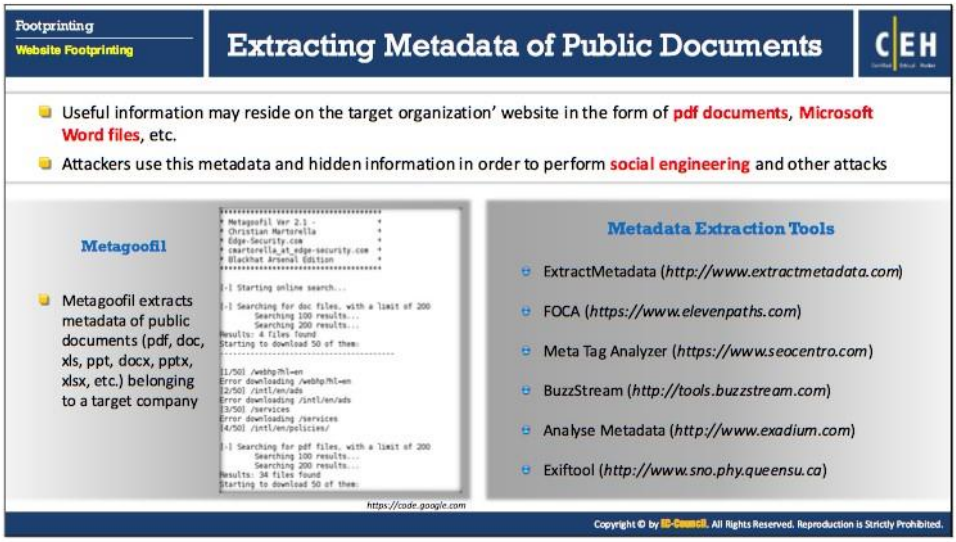
- NCollector Studio (*http://www.calluna-software.com*)
- Teleport Pro (*http://www.tenmax.com*)
- Portable Offline Browser (*http://www.metaproducts.com*)
- Offline Explorer Enterprise (*http://www.metaproducts.com*)
- Website Ripper Copier (*http://www.tensons.com*)
- Gnu Wget (*http://www.gnu.org*)
- Pavuk (*http://pavuk.sourceforge.net*)
- BackStreet Browser (*http://www.spadixbd.com*)
- SurfOffline (*http://www.surfoffline.com*)
- BlackWidow (*https://www.softbytelabs.com*)



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 their creation. As the website ***https://archive.org*** keeps track of web pages from the time of their inception, an attacker can retrieve even information removed from the target website.



Extracting Metadata of Public Documents

Useful information may reside on the target organization' website in the form of pdf documents, Microsoft Word files and other formats. You should be able to extract valuable data, including metadata and hidden information from such documents. It mainly contains hidden information about the public documents that can be analyzed in order to obtained information such as title of the page, description, keywords, creation/modification data and time of the content, usernames and e-mail addresses of employees of the target organization.

An attacker can misuse the information to perform malicious activities on the target organization by Brute-forcing authentication using usernames and e-mail addresses of employees and perform social engineering in order to send malware which can infect target system.

Metadata Extraction Tools

Metadata extraction tools automatically extract critical information that includes username of the clients, operating systems (exploits are OS-specific), email addresses (possibly for social engineering), list of software (version and type) used, list of servers and document date creation/modification, authors of the website and so on.

- **Metagoofil**

  Source: *https://code.google.com*

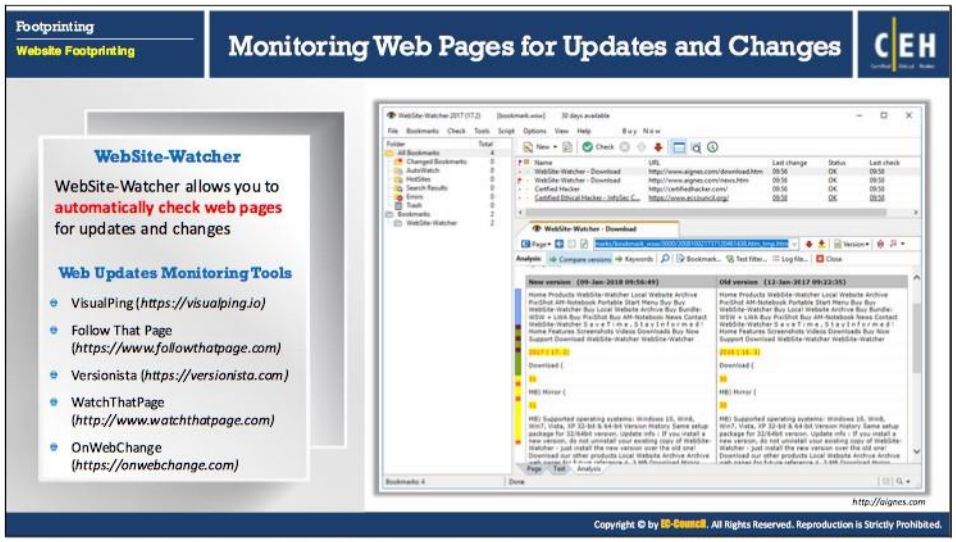
  Metagoofil extracts metadata of public documents (pdf, doc, xls, ppt, docx, pptx, and xlsx) belonging to a target company.

  It performs a Google search to identify and download the documents to local disk and then extracts the metadata with different libraries like Hachoir, PdfMiner and others.

With the results, it generates a report with usernames, software versions and servers or machine names that will help penetration testers in the information gathering phase.

Some of the additional metadata extraction tools are listed below:

- ExtractMetadata (*http://www.extractmetadata.com*)
- FOCA (*https://www.elevenpaths.com*)
- Meta Tag Analyzer (*https://www.seocentro.com*)
- BuzzStream (*http://tools.buzzstream.com*)
- Analyse Metadata (*http://www.exadium.com*)
- Exiftool (*http://www.sno.phy.queensu.ca*)
- Web Data Extractor (*http://www.webextractor.com*)



Monitoring Web Pages for Updates and Changes

Web Updates Monitoring Tools

Web updates monitoring tools are capable of detecting any changes or updates in a particular website and can give notifications or send alerts to the interested users through email or SMS.

- **WebSite-Watcher**

  Source: *http://aignes.com*

  WebSite-Watcher helps to track websites for updates and automatic changes. When an update or change occurs, WebSite-Watcher automatically detects and saves the last two versions onto your disk. It also highlights the changes in-text. It is a useful tool for monitoring sites to gain competitive advantage.

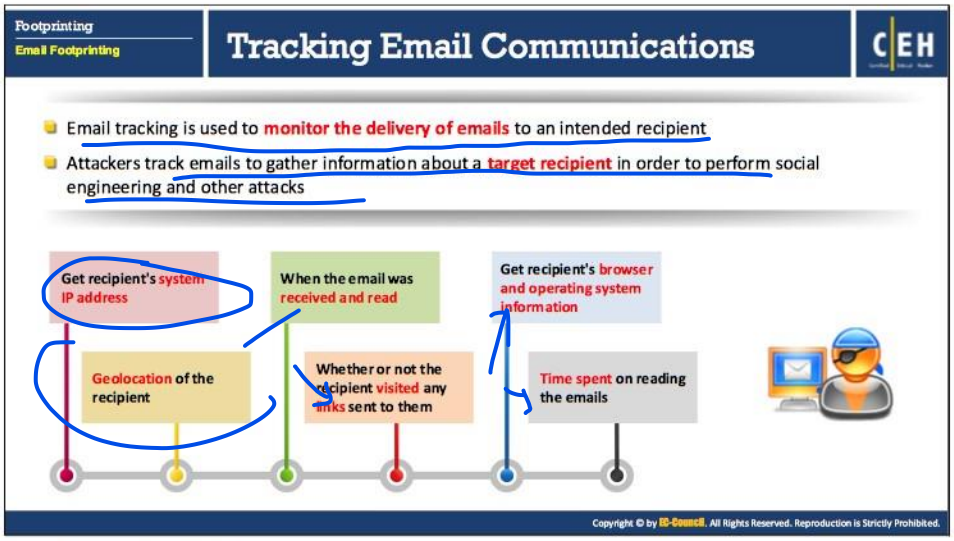
  Features:

  - Frequent manual checking of updates is unnecessary. WebSite-Watcher can automatically detect and notify users about updates
  - Monitors web pages, password protected pages, forums for new postings and replies, RSS feeds, newsgroups and local files
  - It allows you to know what your competitors are doing by scanning their ´websites
  - The site can keep track of new software versions or driver updates
  - It stores images of the modified websites to a disk

Some of the additional website updates monitoring tools are listed below:

- VisualPing (*https://visualping.io*)

- Follow That Page (*https://www.followthatpage.com*)
- Versionista (*https://versionista.com*)
- WatchThatPage (*http://www.watchthatpage.com*)
- OnWebChange (*https://onwebchange.com*)
- InfoMinder (*http://www.infominder.com*)
- Update Scanner (*https://addons.mozilla.org*)
- Check4Change (*https://addons.mozilla.org*)



Email Footprinting

So far, we have discussed footprinting through search engines, footprinting using Google, footprinting through social networking sites, and website footprinting. Now we will discuss email footprinting. This section describes how to track email communications, how to collect information from email headers, and email tracking tools.

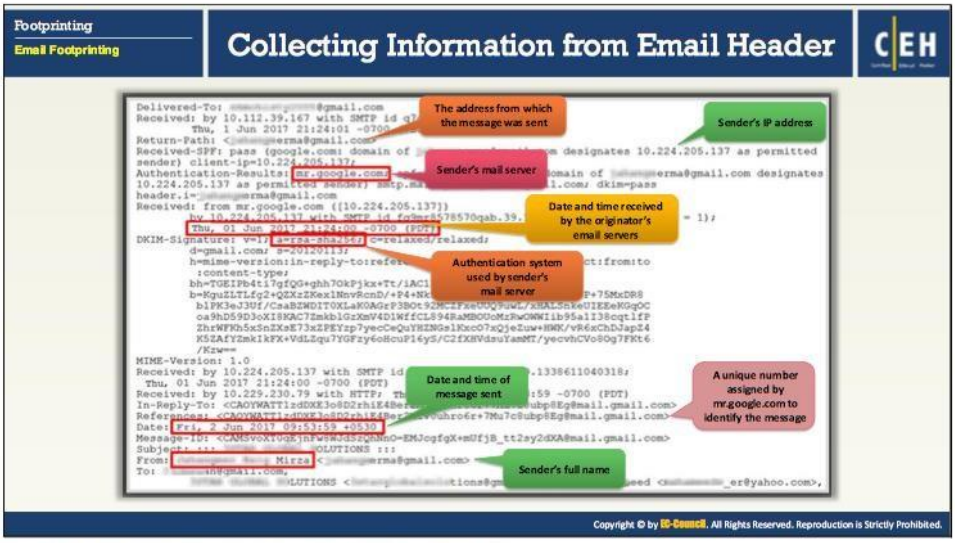
Tracking Email Communications

Email tracking monitors the emails of a particular user. This kind of tracking is possible through digitally time stamped records that reveal the time and date when the target receives and opens a specific email. Email tracking tools allows an attacker to collect information such as IP addresses, mail servers, and service provider involved in sending the mail. Attackers can use this information to build a hacking strategy and to perform social engineering and other attacks. Examples of email tracking tools include eMailTrackerPro, Yesware, ContactMonkey and so on.

Information gathered about the victim using email tracking tools:

- **Recipient's system IP address**: Allows to track the recipients IP address
- **Geolocation**: Estimates and displays the location of the recipient on the map and may even calculate the distance from the attacker's location
- **Email received and Read**: Notifies when the email is received and read by the recipient
- **Read duration**: The duration of time spent by the recipient on reading the mail sent by the sender
- **Proxy detection**: Provides information about the type of server used by the recipient

- **Links**: Checks whether the links sent to the recipient through email have been checked
- **Operating system and Browser information**: Reveals information about the operating system and the browser used by the recipient. The attacker can use this information to find loopholes in that version of operating system and browser, in order to launch further attacks
- **Forward Email:** Determines whether the email sent to the user is forwarded to another person
- **Device Type:** Provides information about the type of device used to open and read the email e.g., desktop computer, mobile device, or laptop.



Collecting Information from Email Header

An email header contains the details of the sender, routing information, date, subject, and recipient. Each is a great source of information for an attacker to launch attacks against the target. The process of viewing the email header varies with different email programs.

Commonly used email programs:

- eM Client
- Mailbird Free
- Claws Mail
- Opera Mail
- Mozilla Thunderbird
- SmarterMail Webmail
- Outlook and Outlook Express
- Eudora
- Entourage
- Netscape Messenger
- MacMail

The email header contains the following information:

- Sender's mail server
- Data and time received by the originator's email servers
- Authentication system used by the sender's mail server
- Data and time of message sent
- A unique number assigned by mr.google.com to identify the message
- Sender's full name
- Senders IP address and address from which the message was sent

The attacker can trace and collect all of this information by performing a detailed analysis of the complete email header.

Email Tracking Tools

Email tracking tools allow an attacker to track an email and extract information such as sender identity, mail server, sender's IP address, location and so on. These tools send notifications automatically when the recipients open the mail and gives status information about whether the email was successfully delivered or not. Attackers use the extracted information to attack the target organization's systems by sending malicious emails.

- **eMailTrackerPro**

  Source: *http://www.emailtrackerpro.com*

  eMailTrackerPro analyzes email headers and reveals information such as sender's geographical location, IP address and so on. It allows an attacker to review the traces later by saving past traces.

The following are a few of the most widely used email tracking tools:

- PoliteMail (*http://www.politemail.com*)
- Yesware (*http://www.yesware.com*)
- ContactMonkey (*https://contactmonkey.com*)
- Zendio (*http://www.zendio.com*)
- ReadNotify (*http://www.readnotify.com*)
- DidTheyReadIt (*http://www.didtheyreadit.com*)
- Trace Email (*http://whatismyipaddress.com*)

- Email Lookup – Free Email Tracker (*http://www.ipaddresslocation.org*)
- Pointofmail (*https://www.pointofmail.com*)
- WhoReadMe (*http://whoreadme.com*)
- GetNotify (*http://www.getnotify.com*)
- G-Lock Analytics (*https://glockanalytics.com*)

Competitive Intelligence

The next phase in footprinting methodology is competitive intelligence.

Competitive intelligence is a process that involves the gathering, analyzing, and distribution of information about products, customers, competitors, and technologies using the Internet. The information that is gathered can help managers and executives of a company make strategic decisions. This section is about competitive intelligence gathering, and sources of valuable information.

Competitive Intelligence Gathering

Competitive intelligence gathering is the process of identifying, gathering, analyzing, verifying, and using information about your competitors from resources such as the Internet. Competitive intelligence means understanding and learning what about other businesses, in order to become as competitive as possible. It is non-interfering and subtle in nature compared to the direct intellectual property theft carried out through hacking or industrial espionage. It concentrates on the external business environment. In this method, professionals gather information ethically and legally instead of gathering it secretly.

Competitive intelligence helps in determining:

- What the competitors are doing
- How competitors are positioning their products and services.

Companies carry out competitive intelligence either by employing people to search for the information, or by utilizing a commercial database service, which can be lower in cost.

Sources of Competitive Intelligence

Competitive Intelligence gathering can be performed either using direct or indirect approach.

- **Direct Approach**

  Direct approach serves as primary sources for competitive intelligence gathering. Direct approach techniques include gathering information from trade shows, social engineering employees and customers and so on.

- **Indirect Approach**

  Through an indirect approach, information about competitors are gathered using online resources. Indirect approach techniques include:

  - Company websites and employment ads
  - Search engines, Internet, and online database
  - Press releases and annual reports
  - Trade journals, conferences, and newspapers
  - Patent and trademarks
  - Product catalogues and retail outlets
  - Analyst and regulatory reports
  - Customer and vendor interviews
  - Agents, distributors, and suppliers
  - Industry-specific blogs and publications
  - Legal databases E.g., LexisNexis
  - Business information databases E.g., Hoover's

Competitive Intelligence - When Did this Company Begin? How Did it Develop?

Gathering competitor documents and records helps to improve productivity and profitability that in turn stimulates the growth of the company. It helps in determining answers to the following:

- **When did it begin?**

  Through competitive intelligence, companies can collect the history of a particular company, such as its establishment date. Sometimes, they gather crucial information that is not often available to others.

- **How did it develop?**

  What are the various strategies the company uses? Development intelligence can include advertisement strategies, customer relationship management and so on.

- **Who leads it?**

  This information helps a company learn about the competitor's decision makers.

- **Where is it located?**

  Competitive intelligence also includes the location of the company and information related to various branches and their operations.

Attackers can use the information gathered through competitive intelligence to build a hacking strategy.

Information Resource Sites

Information resource sites that help to gain competitive intelligence include:

- **EDGAR Database**

  Source: *https://www.sec.gov/edgar.shtml*

  EDGAR, the Electronic Data Gathering, Analysis, and Retrieval system, performs automated collection, validation, indexing, acceptance, and forwarding of submissions by companies and others who are required by law to file with the U.S. Securities and Exchange Commission ('SEC'). Its primary purpose is to increase the efficiency and fairness of the securities market for the benefit of investors, corporations, and the economy by accelerating the receipt, acceptance, dissemination, and analysis of time-sensitive corporate information filed with the agency.

- **Hoovers**

  Source: *http://www.hoovers.com/about-us.html*

  Hoovers is a business research company that provides complete details about companies and industries all over the world. Hoovers provides patented business-related information through the Internet, data feeds, wireless devices, and co-branding agreements with other online services. It gives complete information about the organizations, industries, and people that drive the economy.

- **LexisNexis**

  Source: *https://www.lexisnexis.com*

  LexisNexis provides content-enabled workflow solutions designed specifically for professionals in the legal, risk management, corporate, government, law enforcement, accounting, and academic markets. It maintains an electronic database of legal and public-records related information. It enables customers to access documents and records of legal, news, and business sources.

- **Business Wire**

  Source: *http://www.businesswire.com*

  Business Wire focuses on press release distribution and regulatory disclosure. This company distributes full text news releases, photos, and other multimedia content from various organizations across the globe to journalists, news media, financial markets, investors, information website, databases, and general audiences. The company has its own patented electronic network through which it releases news.

- **FACTIVA**

  Source: *https://www.dowjones.com*

  Factiva is a global news database and licensed content. It is a business information and research tool that gets information from licensed and free sources and provide capabilities like searching, alerting, dissemination and business information management.

Factiva products provide access to more than 32,000 sources such as newspapers, journals, magazines, television and radio transcripts, and photos. It resources are made available to from nearly every country worldwide in 28 languages, including more than 600 continuously updated newswires.

Competitive Intelligence - What Are the Company's Plans?

- **MarketWatch**

  Source: *https://www.marketwatch.com*

  MarketWatch tracks the pulse of markets for engaged investors. The site is an innovator in business news, personal finance information, real-time commentary, and investment tools and data, with journalists generating headlines, stories, videos, and market briefs.

- **The Wall Street Transcript**

  Source: *https://www.twst.com*

  The Wall Street Transcript is a website as well as a paid subscription publication that publishes industry reports. It expresses the views of money managers and equity analysts of different industry sectors. The site also publishes interviews with CEOs of companies.

- **Alexa**

  Source: *https://www.alexa.com*

  Alexa is a great tool to dig deep into the analytics of other companies. It allows users to

  - Discover influencer outreach opportunities by uncovering sites that link to your competitors using Competitor Backlink Checker.
  - Benchmark and track your company's performance relative to your competitors using Competitive Intelligence Tools.

- **Euromonitor**

  Source: *http://www.euromonitor.com*

  Euromonitor provides strategy research for consumer markets. It publishes reports on industries, consumers, and demographics. It provides market research and surveys focused on the organization's needs.

- **Experian**

  Source: *http://www.experian.com*

  Experian helps to gain insights into competitors´ search, affiliate, display, and social marketing strategies and metrics to improve marketing campaign results. It allows the user to:

  - Benchmark the effectiveness of existing customer acquisition strategies
  - Determine what is driving competitors' success
  - Provides historical consumer data to forecast future trends and quickly respond to changing behaviors
  - Measure website ś performance against industry or specific sites

- **SEC Info**

  Source: *http://www.secinfo.com*

  SEC Info offers the U.S. Securities and Exchange Commission ('SEC') EDGAR database service on the web, with many links added to SEC documents. It allows searches by name, industry, and business, SIC code, area code, accession number, file number, CIK, topic, ZIP code and so on.

- **The Search Monitor**

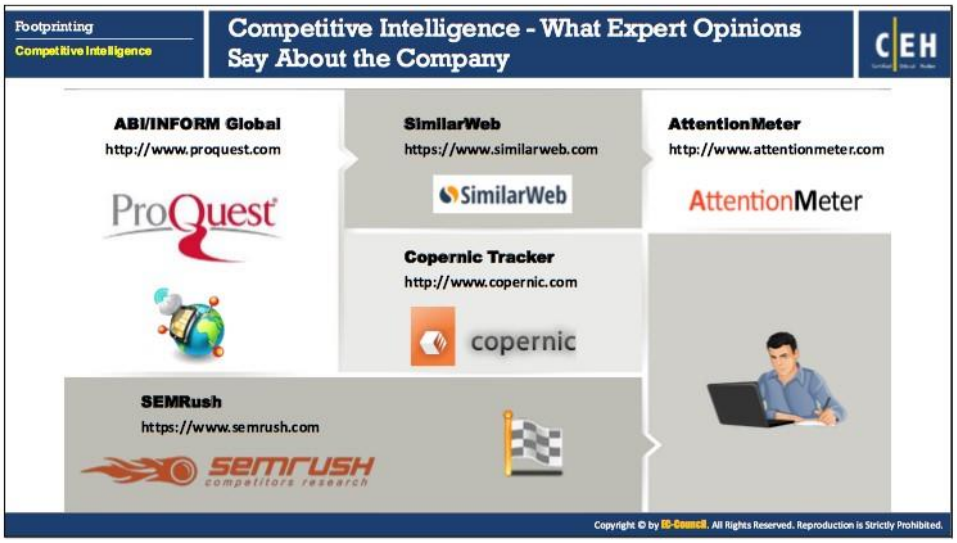
  Source: *https://www.thesearchmonitor.com*

  The Search Monitor provides competitive intelligence to monitor brand and trademark use, affiliate compliance, and competitive advertisers on paid search, organic search, local search, social media, mobile, and shopping engines worldwide. It helps interactive agencies, search marketers, and affiliate marketers to track ad rank, ad copy, keyword reach, click rates and CPCs, monthly ad spend, market share, trademark use, and affiliate activity.

- **USPTO**

  Source: *https://www.uspto.gov*

  The USPTO ('the United States Patent and Trademark Office') provides information related to patent and trademark registration. It provides general information concerning patents and search option for patents and trademark database.



Competitive Intelligence - What Expert Opinions Say About the Company

- **Copernic Tracker**

  Source: *http://www.copernic.com*

  Copernic Tracker is website tracking software. It looks for new content on Web pages and when it detects a change, it notifies the user by email, including a copy of the Web page with the changes highlighted, or by displaying a desktop alert. Copernic Tracker helps to track online forums and social media, auctions, news sites, product updates, new job notices, competitors' Web sites and so on.

- **SEMRush**

  Source: *https://www.semrush.com*

  SEMRush is a competitive keyword research tool. It can provide a list of Google keywords and AdWords for any site, as well as a competitor list in the organic and paid Google search results. It enables an approach for gaining in-depth knowledge about what competitors are advertising and their budget allocation to specific Internet marketing tactics.

- **AttentionMeter**

  Source: *http://www.attentionmeter.com*

  AttentionMeter is a tool used for comparing websites (traffic) by using Alexa, Compete, CrunchBase, and Quantcast. It gives a snapshot of traffic data as well as graphs from Alexa, Compete, CrunchBase, and QuantCast for the specified websites.

- **ABI/INFORM Global**

  Source: *http://www.proquest.com*

  ABI/INFORM Global is a business database. ABI/INFORM Global offers the latest business and financial information for researchers. With ABI/INFORM Global, users can determine business conditions, management techniques, business trends, management practice and theory, corporate strategy and tactics, and the competitive landscape.

- **SimilarWeb**

  Source: *https://www.similarweb.com*

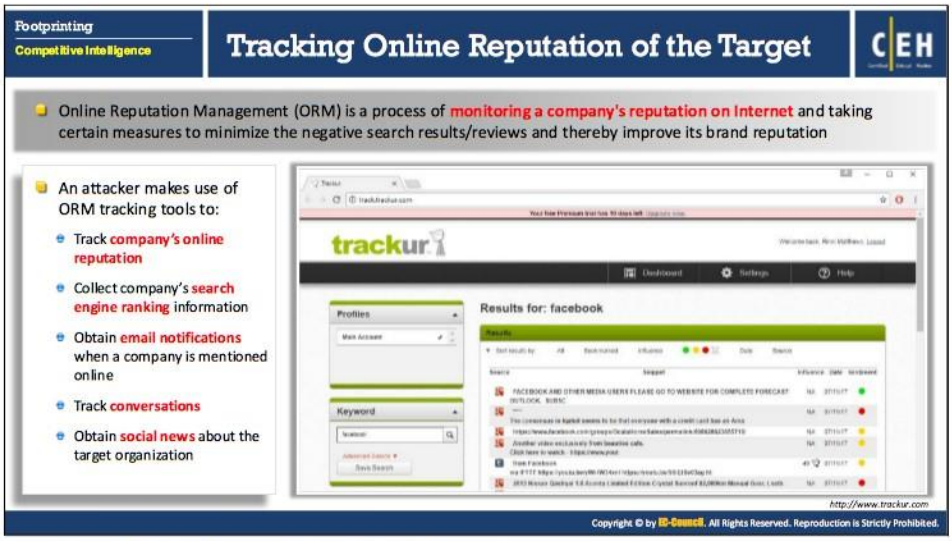
  SimilarWeb aggregates data from multiple sources to estimate traffic, geography, and referral data for company's websites and mobile apps. It also provides a panel through a browser extension, that allows to refine other data sources, by anonymously tracking browser activity across millions of browsers worldwide.

Monitoring Website Traffic of Target Company

Attackers can monitor a target company's website traffic using tools such as Web-Stat, Alexa, and Monitis to collect valuable information. These tools help to collect information about the target's customer base which help attackers to disguise as a customer and launch social engineering attacks on the target.

The information collected includes:

- **Total visitors**: Tools such as Clicky (*https://clicky.com*) find the total number of visitors browsing the target website.
- **Page views**: Tools such as Opentracker (*https://www.opentracker.net*) monitor the total number of pages viewed by the users along with the time stamps and the status of the user on a particular web page (whether the webpage is still active or closed).
- **Bounce rate**: Tools such as Google Analytics (*https://analytics.google.com*) measure the bounce rate of the target company's website.
- **Live visitors map**: Tools such as Web-Stat (*http://www.web-stat.com*) track the geographical location of the users visiting the company's website.
- **Site ranking**: Tools such as Alexa (*https://www.alexa.com*) track a company's rank on the web.
- **Audience geography**: Tools such as Alexa track a company's customers' location on the globe.



Tracking Online Reputation of the Target

Online Reputation Management ('ORM') is a process of monitoring displays when someone searches your company's reputation on the Internet. ORM then takes measures to minimize negative search results or reviews. The process helps to improve brand reputation.

Companies often track the public feedback given to them using ORM tracking tools, and then take measures to improve their credibility and keep their customers' trust. For positive online reputation management, organizations will often try to be more transparent over the Internet. This transparency may help the attacker to collect genuine information about the target organization.

Online Reputation Tracking Tools

Online reputation tracking tools help us to discover what people are saying online about company's brand in real time across the web, social media, and news. They help us in monitoring, measuring, and management of one's reputation online.

An attacker makes use of ORM tracking tools to:

- Track a company's online reputation
- Collect a company's search engine ranking information
- Obtain email notifications when a company is mentioned online
- Track conversations
- Obtain social news about the target organization

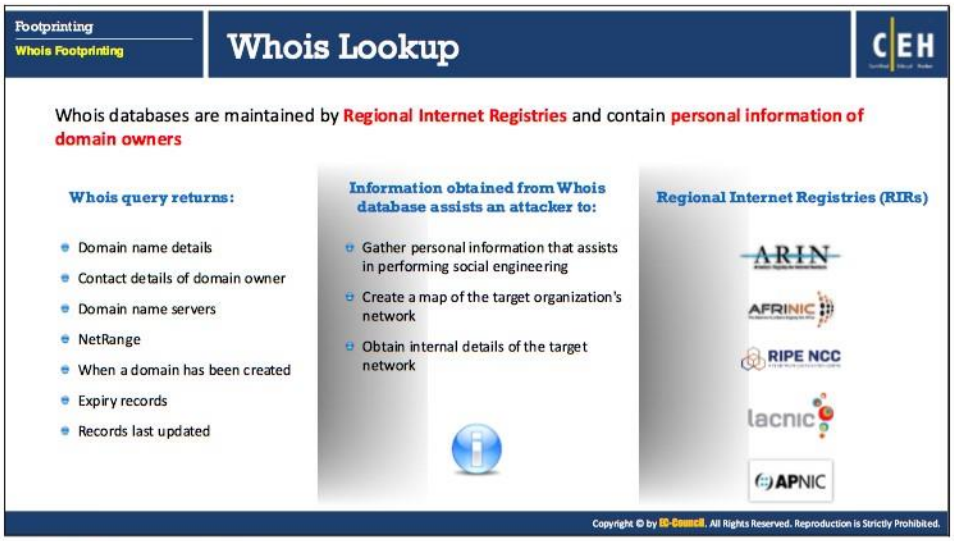
Trackur

Source: *http://www.trackur.com*

Trackur tool provides social media monitoring. It provides full monitoring of all social media and mainstream news, including Twitter, Facebook, Google+ and more. It provides expert social media analytics via executive insights including trends, keyword discovery, automated sentiment analysis and influence scoring.

The following is a list of online reputation tracking tools:

- Brand24 (*https://brand24.com*)
- Social Mention (*http://www.socialmention.com*)
- ReviewTrackers (*https://www.reviewtrackers.com*)
- Rankur (*https://rankur.com*)
- ReputationDefender (*https://www.reputation.com*)
- BrandYourself (*https://brandyourself.com*)
- Google Alerts (*https://www.google.com/alerts*)
- WhosTalkin (*http://www.whostalkin.com*)
- PR Software (*http://www.cision.com*)
- BrandsEye (*https://www.brandseye.com*)
- Talkwalker (*http://www.talkwalker.com*)



Whois Footprinting

Gathering network-related information such as "Whois" information about the target organization is important when planning a hack. In this section, we will discuss Whois footprinting.

Whois footprinting focuses on how to perform a Whois lookup, analyzing the Whois lookup results, and the tools used to gather Whois information.

Whois Lookup

Whois is a query and response protocol used for querying databases that store the registered users or assignees of an Internet resource, such as a domain name, an IP address block, or an autonomous system. This protocol listens to requests on port 43 (TCP). Regional Internet Registries (RIRs) maintain Whois databases and it contains the personal information of domain owners. For each resource, Whois database provides text records with information about the resource itself, and relevant information of assignees, registrants, and administrative information (creation and expiration dates).

Two types of data models exist to store and look up Whois information:

- **Thick Whois** - Stores the complete Whois information from all the registrars for the particular set of data.
- **Thin Whois** - Stores only the name of the Whois server of the registrar of a domain, which in turn holds complete details on the data being looked up.

Whois query returns following information:

- Domain name details
- Contact details of domain owner

- Domain name servers
- NetRange
- When a domain has been created
- Expiry records
- Records last updated

An attacker queries a Whois database server to obtain information about the target domain name, contact details of its owner, expiry date, creation date and so on., and the Whois sever responds to the query with the requested information. Using this information an attacker can create a map of the organization's network, mislead domain owners with social engineering, and then obtain internal details of the network.

Regional Internet Registries (RIRs)

The RIRs include:

- **ARIN (American Registry for Internet Numbers)**

  Source: *https://www.arin.net*

  ARIN provides services related to the technical coordination and management of Internet number resources. ARIN offers its services in form of three areas:

  - Registration - pertains to the technical coordination and management of Internet number resources
  - Organization - pertains to the interaction between ARIN members and stakeholders and ARIN
  - Policy Development - facilitates the development of policy for the technical coordination and management of Internet number resources in the ARIN region

  ARIN also develops technical services to support the evolving needs of the Internet community.

- **AFRINIC (African Network Information Center)**

  Source: *https://www.afrinic.net*

  The acronym, AFRINIC, is the RIR for Africa, responsible for the distribution and management of Internet number resources such as IP addresses and ASN (Autonomous System Numbers) for the African region.

- **APNIC (Asia Pacific Network Information Center)**

  Source: *https://www.apnic.net*

  APNIC is one of five RIRs charged with ensuring the fair distribution and responsible management of IP addresses and related resources required for the stable and reliable operation of the global Internet.

- **RIPE (Réseaux IP Européens Network Coordination Centre)**

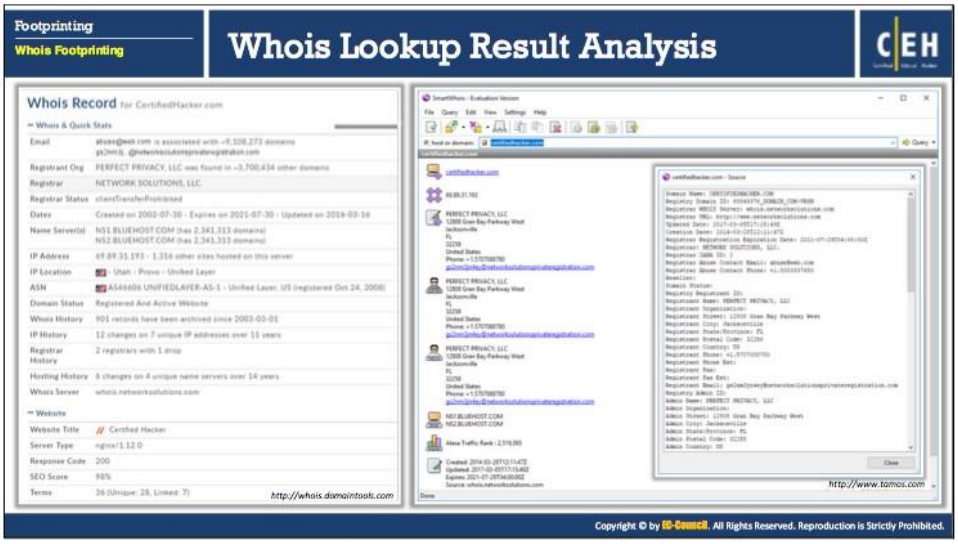
  Source: *https://www.ripe.net*

  RIPE NCC provides Internet resource allocations, registration services, and coordination activities that support the operation of the Internet globally.

- **LACNIC (Latin American and Caribbean Network Information Center)**

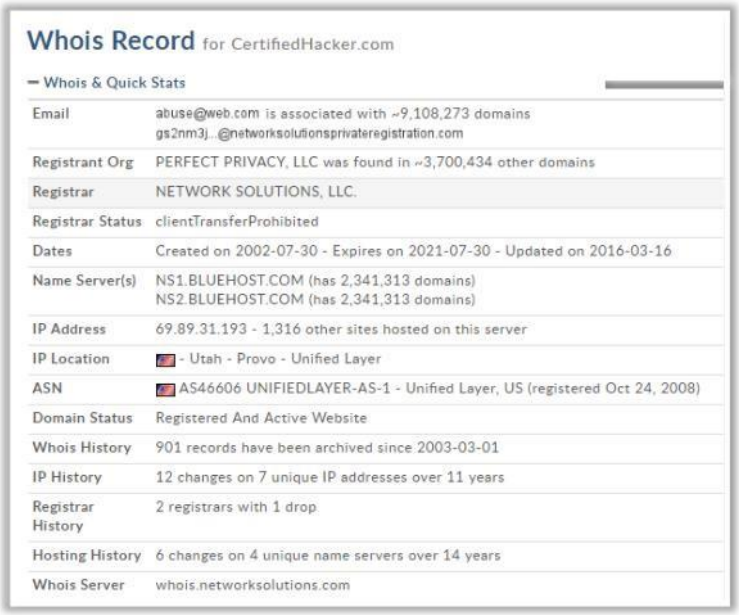
  Source: *http://www.lacnic.net*

  LACNIC is an international non-government organization responsible for assigning and administrating Internet numbering resources (IPv4, IPv6), autonomous system numbers, reverse resolution, and other resources for the Latin America and Caribbean region.



Whois Lookup Result Analysis

Whois services such as *http://whois.domaintools.com* or *http://www.tamos.com* can help to perform Whois lookups. The following figure shows a result analysis of a Whois lookup obtained with the two mentioned Whois services. The services perform Whois lookup by entering the target's domain or IP address. The domaintools.com service provides Whois information such as registrant information, email, administrative contact information, created and expiry date, and a list of domain servers. The SmartWhois available at *http://www.tamos.com* gives information about an IP address, hostname, or domain, including country, state or province, city, phone number, fax number, name of the network provider, administrator, and technical support contact information. It also assists in finding the owner of the domain, the owner's contact information, the owner of the IP address block, registered date of the domain and so on. It supports Internationalized Domain Names (IDNs), which means one can query domain names that use non-English characters. It also supports IPv6 addresses.



Whois Record for CertifiedHacker.com

– Whois & Quick Stats

Email	abuse@web.com is associated with ~9,108,273 domains gs2nm3j...@networksolutionsprivateregistration.com
Registrant Org	PERFECT PRIVACY, LLC was found in ~3,700,434 other domains
Registrar	NETWORK SOLUTIONS, LLC.
Registrar Status	clientTransferProhibited
Dates	Created on 2002-07-30 - Expires on 2021-07-30 - Updated on 2016-03-16
Name Server(s)	NS1.BLUEHOST.COM (has 2,341,313 domains) NS2.BLUEHOST.COM (has 2,341,313 domains)
IP Address	69.89.31.193 - 1,316 other sites hosted on this server
IP Location	- Utah - Provo - Unified Layer
ASN	AS46606 UNIFIEDLAYER-AS-1 - Unified Layer, US (registered Oct 24, 2008)
Domain Status	Registered And Active Website
Whois History	901 records have been archived since 2003-03-01
IP History	12 changes on 7 unique IP addresses over 11 years
Registrar History	2 registrars with 1 drop
Hosting History	6 changes on 4 unique name servers over 14 years
Whois Server	whois.networksolutions.com

FIGURE 2.3: Screenshot of Whois

Whois Lookup Tools

Whois Lookup tools extract information such as IP address, hostname or domain name, registrant information, DNS records including country, city, state, phone and fax numbers, network service providers, administrators and technical support information for any IP address or domain name.

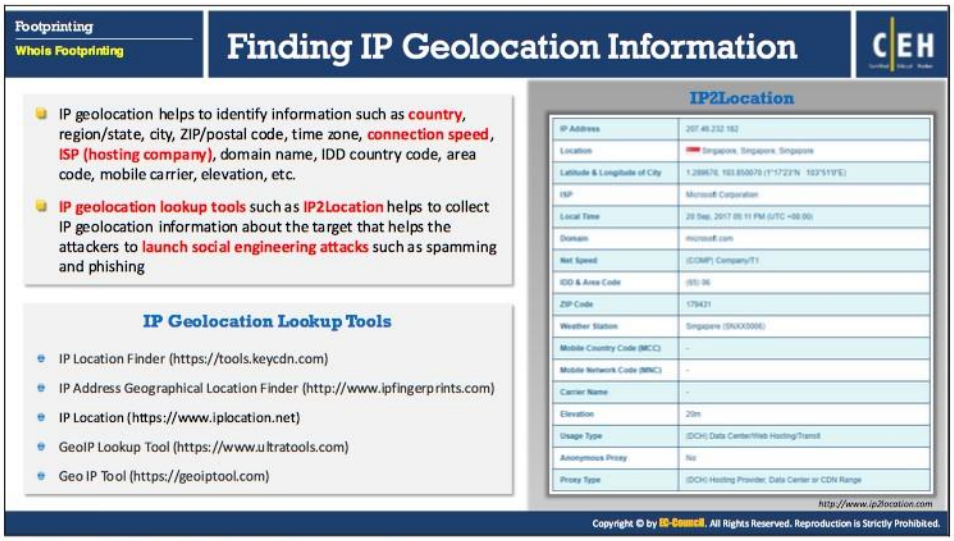
There are numerous tools available to retrieve Whois information, including:

- Batch IP Converter (*http://www.sabsoft.com*)
- WhoIs Analyzer Pro (*http://www.whoisanalyzer.com*)
- Active Whois (*http://www.johnru.com*)
- WhoisThisDomain (*http://www.nirsoft.net*)
- Whois (*https://docs.microsoft.com*)
- WhoIs Lookup Multiple Addresses Software (*https://www.sobolsoft.com*)
- Whois Lookup (*https://pentest-tools.com*)
- ICANN WHOIS (*https://whois.icann.org*)
- IANA WHOIS (*https://www.iana.org*)
- WHOIS Lookup (*https://www.whois.com*)
- HotWhois (*http://www.tialsoft.com*)
- Domain Dossier (*https://centralops.net*)
- BetterWhois (*http://www.betterwhois.com*)
- Whois Online (*http://whois.online-domain-tools.com*)
- Web Wiz (*https://network-tools.webwiz.net/whois-lookup.htm*)
- Network-Tools.com (*http://network-tools.com*)
- WHOis.net (*https://www.whois.net*)

- DNSstuff (*http://www.dnsstuff.com*)
- Whois Lookup (*http://www.webtoolhub.com*)
- UltraTools (*https://www.ultratools.com*)
- AFRINIC WHOIS (*https://www.afrinic.net*)
- APNIC Whois (*http://wq.apnic.net*)
- LACNIC Whois (*https://lacnic.net*)
- Network Solutions Whois (*https://www.networksolutions.com*)

Some of the Whois Lookup tools for cell phones include:

- DNS Tools (*https://www.dnssniffer.com*)
- UltraTools Mobile (*https://www.ultratools.com*)
- Whois® (*https://www.whois.com.au*)
- Deep Whois (*http://happymagenta.com/deepwhois*)
- Whois Lookup Tool (*https://www.znetlive.com*)
- WHOIZ - Domain Name WHOIS Tool (*http://whoiz.link*)
- WHOIS Lookup (*https://play.google.com*)



Finding IP Geolocation Information

IP geolocation help you to identify information such as country, region/state, city, latitude and longitude of 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, elevation and so on.

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 information obtained, attacker can also set up a compromised web server nearby victim's location and if there is detection of the exact location of the victim it can send malicious stuff and infect the victim with a malware designed for that specific area or the attacker can gain an unauthorized access to the target device or may attempt to launch an attack using target device

IP geolocation lookup tools such as IP2Location helps to collect IP geolocation information about the target which help attackers to launch social engineering attacks such as spamming and phishing.

IP Geolocation Lookup Tools

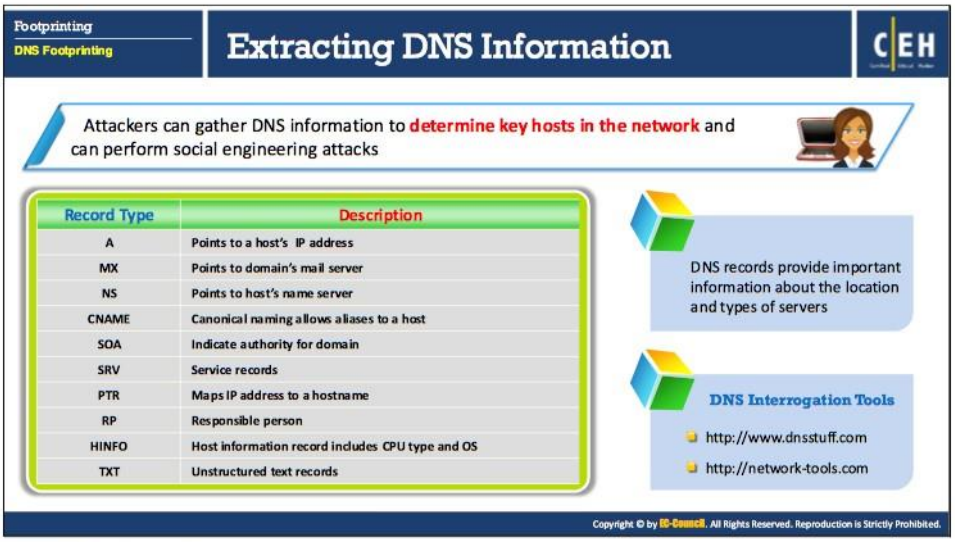
- **IP2Location**

  Source: *http://www.ip2location.com*

  IP2Location is a geo IP solution to help the user identify 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 without invading the Internet user's privacy.

Some of the IP geolocation tools are listed below:

- IP Location Finder (*https://tools.keycdn.com*)
- IP Address Geographical Location Finder (*http://www.ipfingerprints.com*)
- IP Location (*https://www.iplocation.net*)
- GeoIP Lookup Tool (*https://www.ultratools.com*)
- Geo IP Tool (*https://geoiptool.com*)
- IP Address & Geolocation Lookup Tool (*https://www.risk.neustar*)
- GeoIP2 (*https://www.maxmind.com*)
- IP-to-Location Database (*https://www.webhostinghero.com*)
- Check-Host (*https://check-host.net*)
- TextMagic IP Address Location (*https://www.textmagic.com*)



DNS Footprinting

The next phase in footprinting methodology is DNS footprinting. This following section describes how to extract DNS information and the DNS interrogation tools.

Extracting DNS Information

DNS footprinting, namely Domain Name System footprinting, reveals information about DNS zone data. DNS zone data include DNS domain names, computer names, IP addresses, and much more about a particular network. An attacker uses DNS information to determine key hosts in the network, and then performs social engineering attacks to gather even more information.

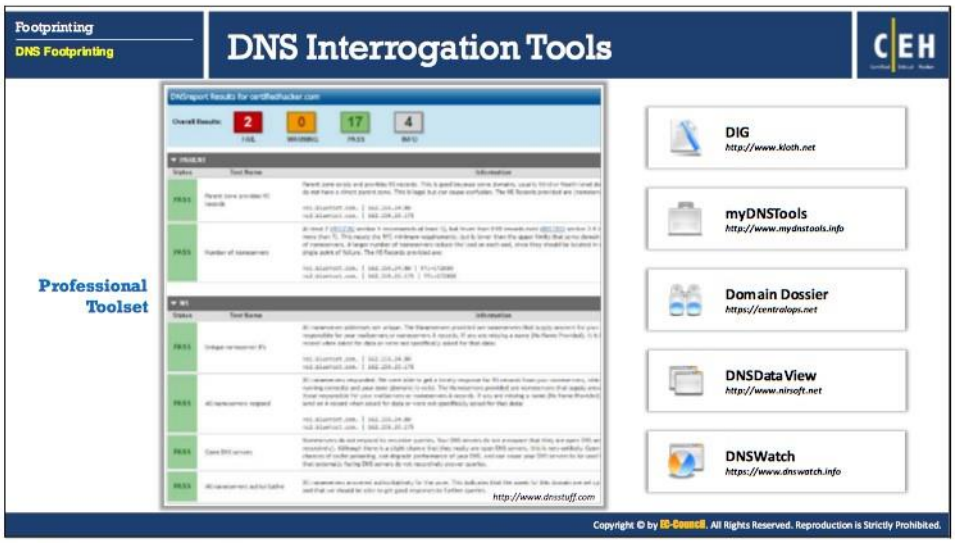
DNS footprinting helps in determining following records about the target DNS:

Record Type	Description
A	Points to a host's IP 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 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

TABLE 2.6: Table showing DNS records and their description

DNS interrogation tools such as *http://www.dnsstuff.com* and DNS Records (*http://network-tools.com*) enable user to perform DNS footprinting. DNSstuff extracts DNS information about IP addresses, mail server extensions, DNS lookups, Whois lookups and so on. It can extract a range of IP addresses utilizing an IP routing lookup. If the target network allows unknown, unauthorized users to transfer DNS zone data, then it is easy for an attacker to obtain the information about DNS with the help of the DNS interrogation tool.

When the attacker queries the DNS server using the DNS interrogation tool, the server then responds with a record structure that contains information about the target DNS. DNS records provide important information about the location and types of servers.



DNS Interrogation Tools

The DNS Lookup tools retrieve the DNS records for a specified domain or hostname. These tools retrieve information such as domains and IP addresses, domain Whois records, DNS records, and network Whois record.

- **Professional Toolset**

  Source: *http://www.dnsstuff.com*

  Professional Toolset assists IT professionals with troubleshooting, managing, and configuring the domain and email.

  Professional Toolset includes Domain/WWW tools, IP tools, Networking tools, and Email tools that assist with:

  - DNS troubleshooting, management and monitoring
  - Network administration and troubleshooting
  - Email troubleshooting and diagnostics
  - Internet/Cybercrime forensics
  - Spam combat
  - Insight into an IP address
  - Internet configuration, connectivity and performance

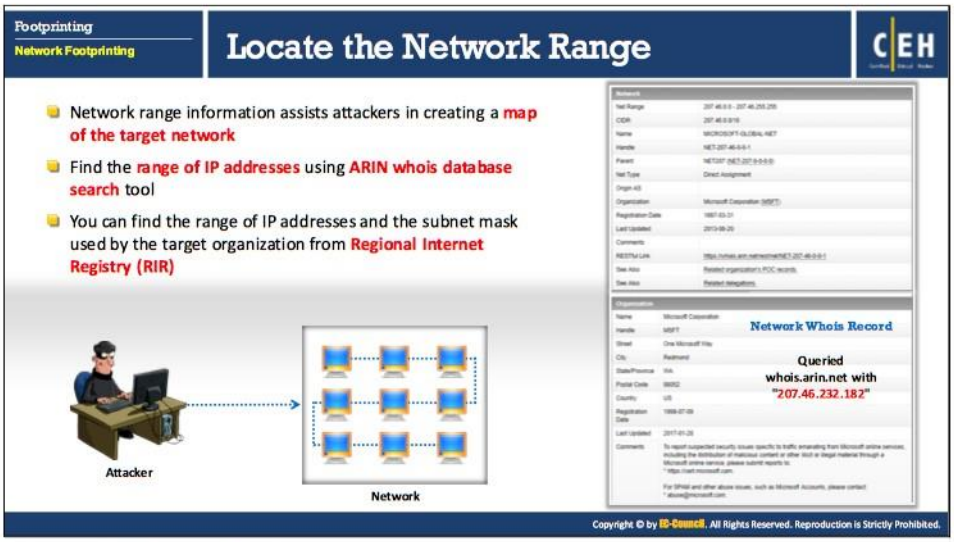
Some of the DNS interrogation tools used to extract DNS information include:

- DIG (*http://www.kloth.net*)
- myDNSTools (*http://www.mydnstools.info*)

- Domain Dossier (*https://centralops.net*)
- DNSDataView (*http://www.nirsoft.net*)
- DNSWatch (*https://www.dnswatch.info*)
- DNS Tools (*http://dnstools.com*)
- DNS Lookup Tool (*https://network-tools.webwiz.net*)
- DomainTools (*http://www.domaintools.com*)
- DNS Query Utility (*http://www.dnsqueries.com*)
- DNS Lookup Tool (*https://www.ultratools.com*)
- DNS Check (*http://dnscheck.pingdom.com*)
- Fierce (*https://github.com*)
- MX Lookup (*https://mxtoolbox.com*)

Some of the DNS interrogation tools for use with smartphones include:

- DNS Lookup Tool (*https://www.theemaillaundry.com*)
- Ping & DNS (*http://www.ulfdittmer.com*)
- IP Tools (*http://www.iptools.su*)
- DNS Lookup (*https://play.google.com*)
- DNS Lookup and Whois (*https://www.networkpanda.com*)
- DNS Tools (*https://www.dnssniffer.com*)



Network Footprinting

The next step after retrieving the DNS information is gathering network-related information. We will now discuss network footprinting, a method of gathering network-related information. This section describes how to locate network range, determine the OS, Traceroute, and the Traceroute tools.

Locate the Network Range

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 they perform in order to perform a network footprinting. The answers to these questions provide information about the internal structure of the target network.

After gathering the information, an attacker can proceed to find the network range of a target system. Detailed information is available from the appropriate regional registry database regarding IP allocation and the nature of the allocation. An attacker can also determine the subnet mask of the domain, and trace the route between the system and the target system. Traceroute tools that are widely used include Path Analyzer Pro and VisualRoute.

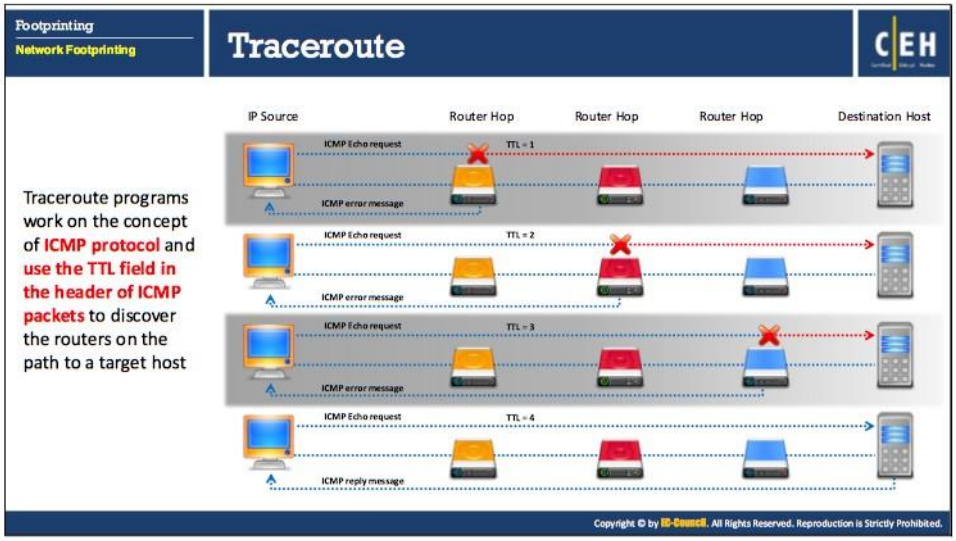
Obtaining private IP addresses can be useful to rogues. The Internet Assigned Numbers Authority ('IANA') has reserved the following three blocks of the IP address space for private Internets: 10.0.0.0–10.255.255.255 (10/8 prefix), 172.16.0.0–172.31.255.255 (172.16/12 prefix), and 192.168.0.0–192.168.255.255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

Using the network range, the attacker can get the information about how the network is structured, which machines in the networks are alive. Using the network range also helps to

identify the network topology, access control device, and OS used in the target network. To find the network range of the target network, one needs to enter the server IP address (that was gathered in Whois footprinting) in the ARIN Whois database search tool. A user can also visit the ARIN website (https://www.arin.net/knowledge/rirs.html) and enter the server IP in the **SEARCH Whois** text box. This gives the network range of the target network. Improperly set up DNS servers offer the attacker a good chance of obtaining a list of internal machines on the server. In addition, sometimes if an attacker traces a route to a machine, it is sometimes possible to obtain the internal IP address of the gateway, which can be useful.

FIGURE 2.4: Screenshot of ARIN Whois Database Search Tool

Attackers typically use more than one tool to obtain network information, as a single tool cannot deliver all of the required information.



Traceroute

Finding the route of the target host on the network is necessary to test against man-in-the-middle attacks and other related attacks. Most operating systems come with a Traceroute utility to perform the task. It traces the path or route through which the target host packets travel in the network.

Traceroute uses the ICMP protocol concept and Time to Live ('TTL') field of IP header to find the path of the target host in the network.

The Traceroute utility can detail the path IP packets travel between two systems. The utility can trace the number of routers the packets travel through, the round trip time (duration in transiting between two routers), and, if the routers have DNS entries, the names of the routers and their network affiliation. It can also trace geographic locations. It works by exploiting a feature of the Internet Protocol called TTL. The TTL field indicates the maximum number of routers a packet may transit. Each router that handles a packet decrements the TTL count field in the ICMP header by one. When the count reaches zero, the router discards the packet and transmits an ICMP error message to the originator of the packet.

The utility records the IP address and DNS name of that router, and sends out another packet with a TTL value of two. This packet makes it through the first router, then times-out at the next router in the path. This second router also sends an error message back to the originating host. Traceroute continues to do this, and records the IP address and name of each router until a packet finally reaches the target host or until it decides that the host is unreachable. In the process, it records the time it took for each packet to travel round trip to each router. Finally, when it reaches the destination, the normal ICMP ping response will be send to the sender. The

utility helps to reveal the IP addresses of the intermediate hops in the route to the target host from the source.

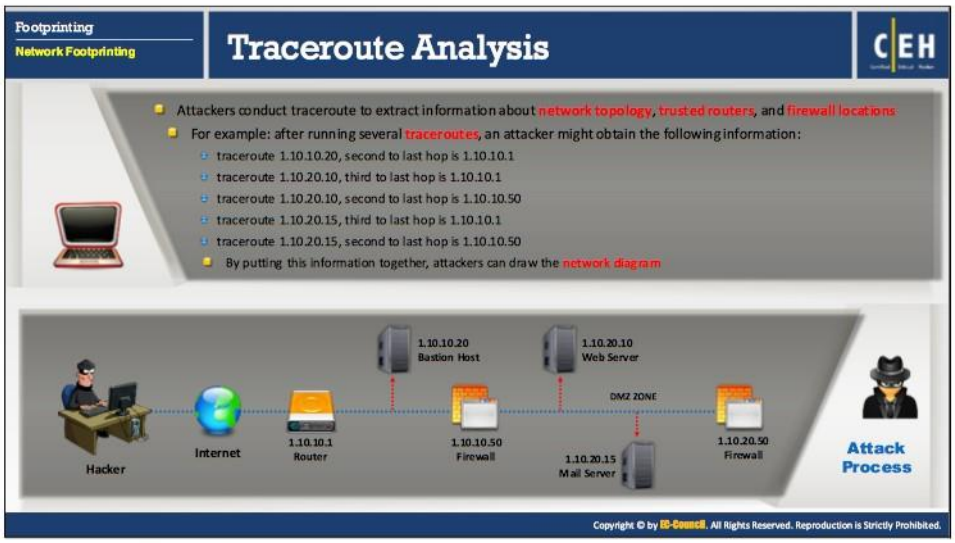
How to use the tracert command?

Go to the command prompt and type the `tracert` command along with the destination IP address or domain name as follows:

```
C:\>tracert 216.239.36.10
```

Tracing route to ns3.google.com [216.239.36.10] over a maximum of 30 hops:

```
 1  1262 ms   186 ms   124 ms  195.229.252.10
 2  2796 ms  3061 ms  3436 ms  195.229.252.130
 3   155 ms   217 ms   155 ms  195.229.252.114
 4  2171 ms  1405 ms  1530 ms  194.170.2.57
 5  2685 ms  1280 ms   655 ms  dxb-emix-ra.ge6303.emix.ae [195.229.31.99]
 6   202 ms   530 ms   999 ms  dxb-emix-rb.so100.emix.ae [195.229.0.230]
 7 609 ms  1124 ms  1748 ms  iar1-so-3-2-0.Thamesside.cw.net [166.63.214.65]
 8 1622 ms  2377 ms  2061 ms  eqixva-google-gige.google.com [206.223.115.21]
 9  2498 ms   968 ms   593 ms  216.239.48.193
10  3546 ms  3686 ms  3030 ms  216.239.48.89
11  1806 ms  1529 ms   812 ms  216.33.98.154
12  1108 ms  1683 ms  2062 ms  ns3.google.com [216.239.36.10]
Trace complete.
```



Traceroute Analysis

We have seen how the Traceroute utility helps to find the IP addresses of intermediate devices such as routers, and firewalls present between a source and its destination. After running several traceroutes, an attacker will be able to find the location of a hop in the target network. Consider the following traceroute results obtained:

- `traceroute 1.10.10.20, second to last hop is 1.10.10.1`
- `traceroute 1.10.20.10, third to last hop is 1.10.10.1`
- `traceroute 1.10.20.10, second to last hop is 1.10.10.50`
- `traceroute 1.10.20.15, third to last hop is 1.10.10.1`
- `traceroute 1.10.20.15, second to last hop is 1.10.10.50`

By analyzing these results, an attacker can draw the network topology diagram of the target network as shown below.

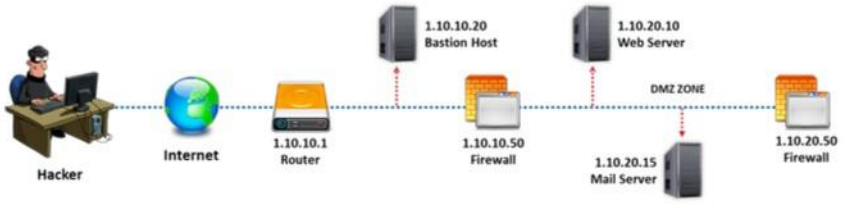


FIGURE 2.5: Traceroute Analysis

Traceroute Tools

Traceroute tools are useful in extracting information about the geographical location of routers, servers and IP devices in a network. Such tools help us to trace, identify, and monitor the network activity on a world map. Some of the features of these tools include:

- Hop-by-hop traceroutes
- Reverse tracing
- Historical analysis
- Packet loss reporting
- Reverse DNS
- Ping plotting
- Port probing
- Detect network problems
- Performance metrics analysis
- Network performance monitoring

Similar to Traceroute, Path Analyzer Pro and VisualRoute are two tools that trace the target host in a network.

- **Path Analyzer Pro**

  Source: *http://www.pathanalyzer.com*

  Path Analyzer Pro delivers network route tracing with performance tests, DNS, Whois, and network resolution to investigate network issues. It shows the route from source to destination graphically. It also provides information such as the hop number, its IP address, hostname, ASN, network name, percentage loss, latency, average latency, and standard deviation for each hop in the path.

  Path Analyzer Pro can:

  - Research IP addresses, e-mail addresses, and network paths
  - Troubleshoot network availability and performance issues

- Determine what ISP, router, or server is responsible for a network problem
- Locate firewalls and other filters that may be impacting connections

- **VisualRoute**

  Source: *http://www.visualroute.com*

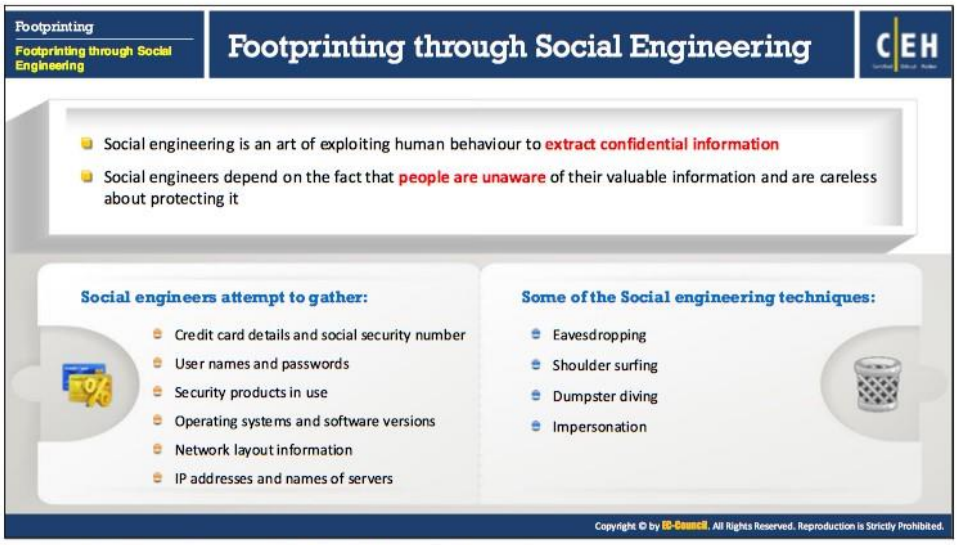
  VisualRoute is a traceroute and network diagnostic tool. It identifies the geographical location of routers, servers, and other IP devices. It provides the tracing information in three forms: as an overall analysis, in a data table, and as a geographical view of the routing. The data table contains information such as hop number, IP address, node name, and geographical location about each hop in the route.

  Features:

  - Hop-by-hop traceroutes
  - Reverse tracing
  - Historical analysis
  - Packet loss reporting
  - Reverse DNS
  - Ping plotting

The following are some of the network trace route tools:

- GEO Spider (*http://www.oreware.com*)
- Trout (*https://www.mcafee.com*)
- Magic NetTrace (*http://www.tialsoft.com*)
- Ping Plotter (*http://www.pingplotter.com*)
- Traceroute Tool (*https://tools.keycdn.com*)
- Network Pinger (*http://www.networkpinger.com*)
- Roadkil's Trace Route (*http://www.roadkil.net*)
- AnalogX HyperTrace (*http://www.analogx.com*)
- Network Systems Traceroute (*https://www.net.princeton.edu*)
- Ping-Probe (*http://ping-probe.com*)



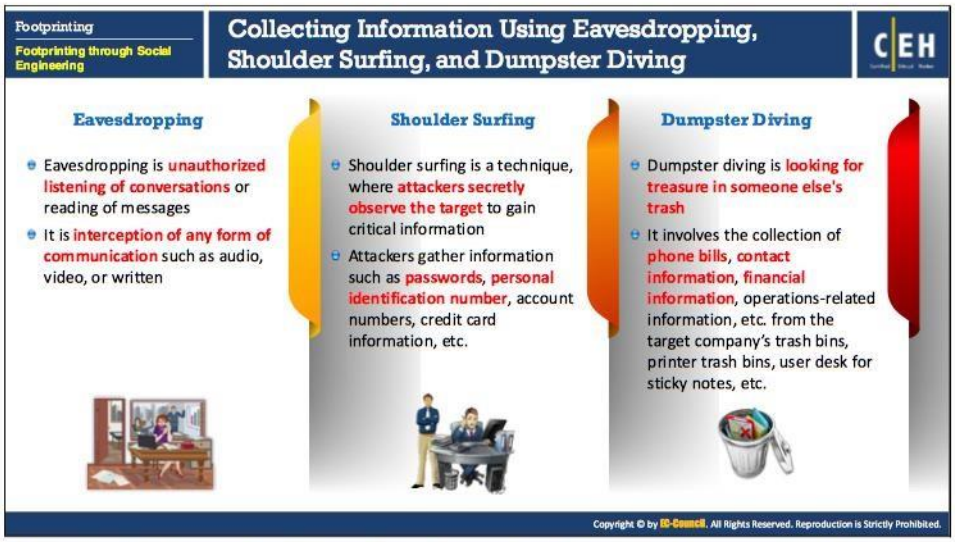
Footprinting through Social Engineering

So far, we have discussed the different techniques for gathering information either with the help of online resources or tools. Now we will discuss footprinting through social engineering, the art of obtaining information from people by manipulating them. This section covers the concept as well as the techniques used to gather information.

Social engineering is a totally non-technical process in which an attacker misleds a person into providing confidential information unknowingly. In other words, the target is unaware of the fact that someone is stealing confidential information. The attacker takes advantage of the helpful nature of people and their willingness to provide confidential information.

To perform social engineering, an attacker first needs to gain the confidence of an authorized user and then mislead that user into revealing confidential information. The goal of social engineering is to obtain required confidential information and then use that information for hacking attempts such as gaining unauthorized access to the system, identity theft, industrial espionage, network intrusion, commits frauds and so on. The information obtained through social engineering may include credit card details, social security numbers, usernames and passwords, other personal information, security products in use, OS and software versions, IP addresses, names of servers, network layout information and so on.

Social engineering can be performed in many ways such as eavesdropping, shoulder surfing, dumpster diving, impersonation, tailgating, third-party authorization, piggybacking, reverse social engineering and so on.



Collecting Information Using Eavesdropping, Shoulder Surfing, and Dumpster Diving

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

- **Eavesdropping**

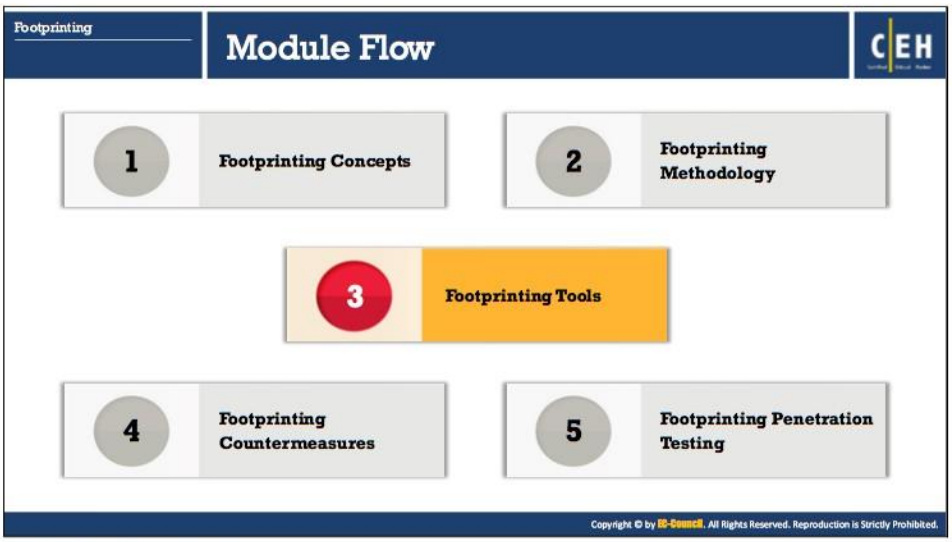
  Eavesdropping is the act of secretly listening to the conversations of people over a phone or video conference without their consent. It also includes reading confidential messages from communication media such as instant messaging or fax transmissions. It is the act of intercepting communication of any form such as audio, video, or written without the consent of the communicating parties. The attacker gains information by tapping the phone conversation, and intercepting audio, video, or written communication.

- **Shoulder Surfing**

  Shoulder surfing is a technique, where attackers secretly observes 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 number, security codes, account numbers, credit card information, and similar data. The 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 looks 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 trash bins, sticky notes at users' desks and so on. The attacker may also gather account information from ATM trash bins. The information can help the attacker to commit attacks.



Footprinting Tools

Attackers are aided in footprinting with the help of various tools. Many organizations offer tools that make information gathering an easy task. This section describes tools intended for obtaining information from various sources.

Footprinting tools are used to collect basic information about the target systems in order to exploit them. Information collected by the footprinting tools contain target's IP location information, routing information, business information, address, phone number and social security number, details about a source of an email and a file, DNS information, domain information and so on.

- **Maltego**

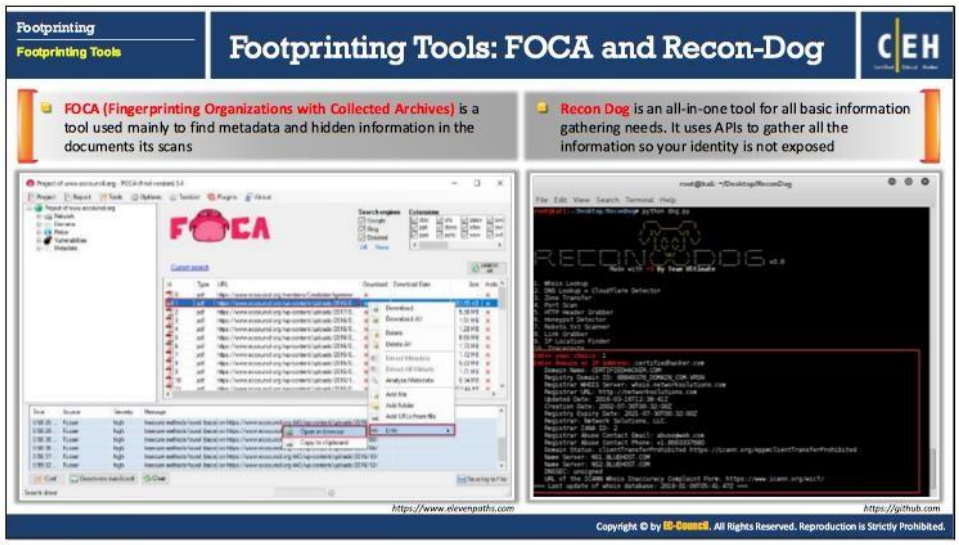
  Source: *https://www.paterva.com*

  Maltego is an open-source intelligence and forensics application. It can be useful during the information gathering phase of all security-related work. Maltego is a platform developed to deliver a clear threat picture to the environment that an organization owns and operates. Maltego demonstrates the complexity and severity of single points of failure as well as trust relationships that exist within the scope of the infrastructure. The unique perspective that Maltego offers to network and resource-based entities is the aggregation of information posted all over the internet. The application can be used to determine the relationships and real-world links between people, social networks, companies, organizations, websites, Internet infrastructure (domains, DNS names, Netblocks, IP addresses), phrases, affiliations, documents, and files.

- **Recon-ng**

  Source: *https://bitbucket.org*

  Recon-ng has a look and feel similar to the Metasploit Framework, reducing the learning curve for leveraging the framework. However, it is quite different. Recon-ng avoids competing with existing frameworks, as it is designed exclusively for web-based open source reconnaissance. It is a Web Reconnaissance framework with independent modules, database interaction, built in convenience functions, interactive help, and command completion, that provides an environment in which open source web-based reconnaissance can be conducted.



- **FOCA**

  Source: *https://www.elevenpaths.com*

  FOCA is capable of scanning and analyzing a wide variety of documents, with the most common being Microsoft Office, Open Office, or PDF files.

  Features:

  - **Web Search** - Searches for hosts and domain names through URLs associated to the main domain. Each link is analyzed to extract from it new host and domain names.
  - **DNS Search** - Checks each domain to ascertain which are the host names configured in NS, MX, and SPF servers to discover a new host and domain names.
  - **IP resolution** - Resolves each host name by comparison to the DNS to obtain the IP address associated with this server name. To perform this task accurately, the tool performs analysis against the organization's internal DNS.
  - **PTR Scanning** - Finds more servers in the same segment of a determined address, IP FOCA executes a PTR logs scan.
  - **Bing IP** -, Launches FOCA which is a search process for new domain names associated with that IP address for each IP address discovered.
  - **Common Names** - Performs dictionary attacks against the DNS.

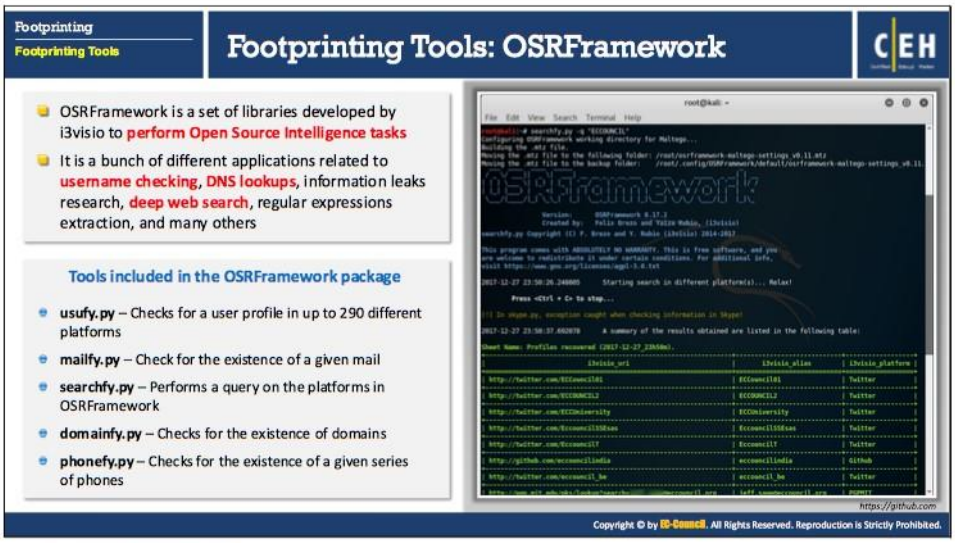
- **Recon-Dog**

  Source: *https://github.com*

  Recon-Dog uses APIs to collect information about the target system.

Features:

- **Whois Lookup** - Searches for information regarding a target domain name.
- **DNS Lookup + Cloudflare Detector** - Checks a target domain using DNS (Domain Name System) lookup in order to find new domain names and hosted connected.
- **Zone Transfer** - Searcher for the vulnerabilities in the DNS zone transfer.
- **Port Scan** - Probes a target system or a server for open ports in order to exploit them.
- **HTTP Header Grabber** - Gathers information about a target system about the type and the version of software it is running.
- **Honeypot Detector** - Detects the presence of honeypot in a target's system. A honeypot contains a data about the system that looks legitimate and is monitored continuously in order to detect any malicious activity which is blocked afterwards.
- **Robot.txt Scanner** - Scans the target system against Robot.txt file that is used to give instruction to web crawlers. Flaws in Robot.txt file can allow an attacker to gain an access to the unauthorized location of a website.
- IP Location Finder, Traceroute, and Link Grabber



- **OSRFramework**

  Source: *https://github.com*

  OSRFramework is a GNU AGPLv3+ set of libraries developed by i3visio to perform Open Source Intelligence tasks. The libraries provide a collection of scripts that can enumerate users, domains, and more across over 200 separate service. They include references to a bunch of different applications related to username checking, DNS lookups, information leaks research, deep web search, regular expressions extraction, and many others. At the same time, by means of ad-hoc Maltego transforms, OSRFramework provides a way of making these queries graphically as well as several interfaces to interact with like OSRFConsole or a Web interface.

  Tools included in the OSRFramework package:

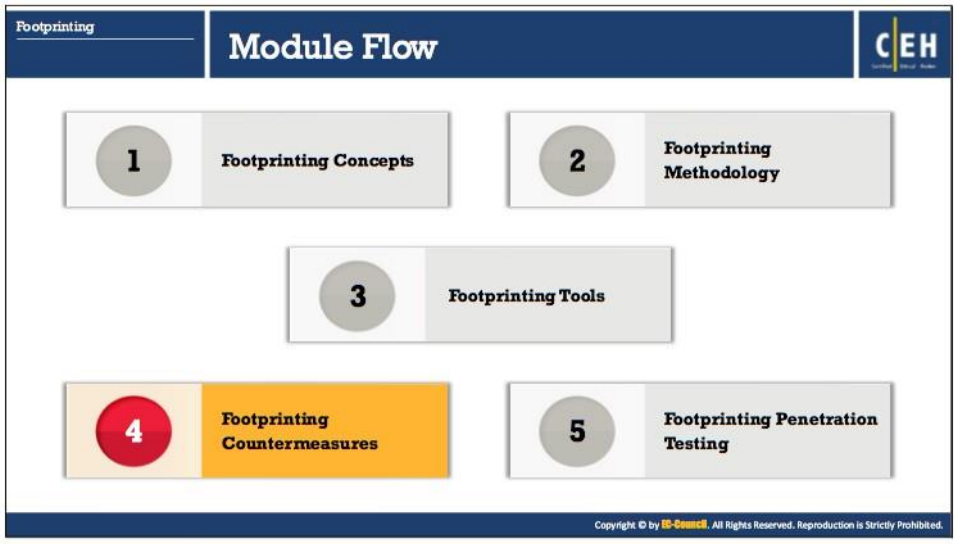
  - **usufy.py** – Checks for a user profile in up to 290 different platforms
  - **mailfy.py** – Check for the existence of a given mail
  - **searchfy.py** – Performs a query on the platforms in OSRFramework
  - **domainfy.py** – Checks for the existence of domains
  - **phonefy.py** – Checks for the existence of a given series of phones
  - **entify.py** – Use regular expressions to extract entities

Additional Footprinting Tools

Some of the other additional footprinting tools that assist in gathering information about the target person or organization include:

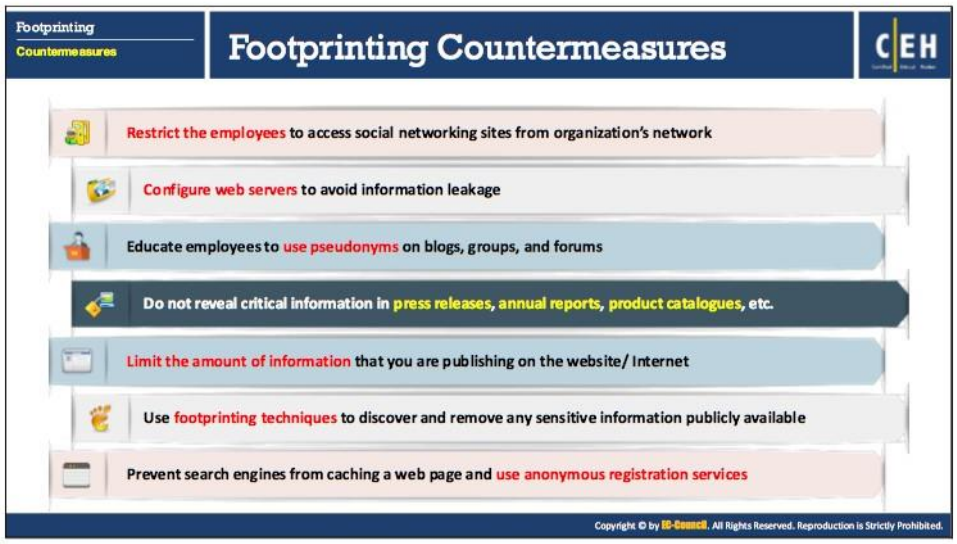
- Prefix Whois (*http://pwhois.org*)
- LHF (Low Hanging Fruit) (*https://github.com*)
- Sni1per (*https://github.com*)
- CloudFail (*https://github.com*)
- Aquatone (*https://github.com*)
- GMapCatcher (*https://github.com*)
- DNS-Digger (*http://www.dnsdigger.com*)
- Reconnoitre (*https://github.com*)
- NSLOOKUP (*http://www.kloth.net*)
- DomainHostingView (*http://www.nirsoft.net*)
- Robtex (*https://www.robtex.com*)
- SearchBug (*https://www.searchbug.com*)
- Zaba Search (*http://www.zabasearch.com*)
- Metasploit (*https://www.metasploit.com*)
- theHarvester (*http://www.edge-security.com*)
- Dig Web Interface (*http://www.digwebinterface.com*)

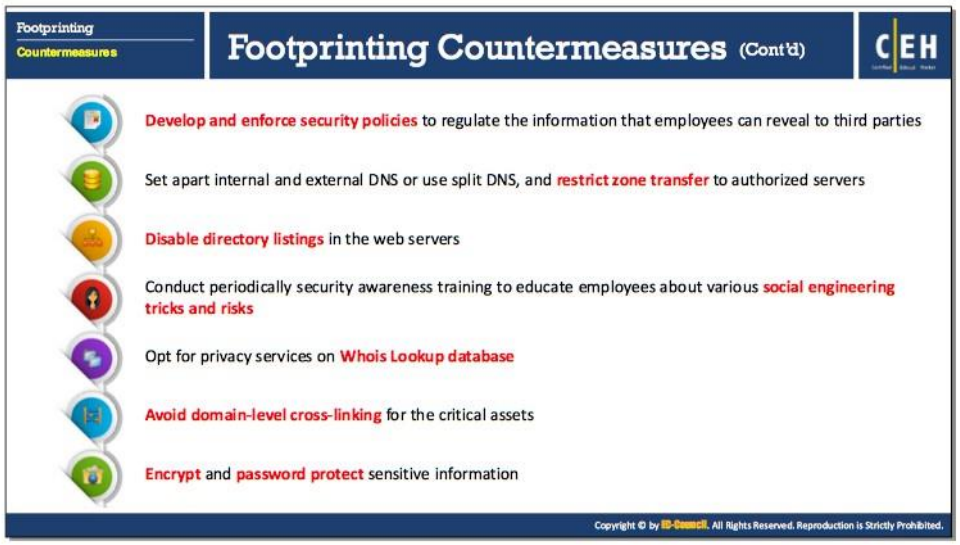
- SearchDiggity (*https://www.bishopfox.com*)
- NetScanTools Pro (*http://www.netscantools.com*)
- Tctrace (*http://www.phenoelit.org*)
- Autonomous System Scanner (ASS) (*http://www.phenoelit.org*)
- Netmask (*http://www.phenoelit.org*)
- TinEye (*https://tineye.com*)
- White Pages (*https://pro.whitepages.com*)
- Ping-Probe (*http://ping-probe.com*)
- SpiderFoot (*http://www.spiderfoot.net*)
- Metagoofil (*https://code.google.com*)
- BiLE Suite (*https://github.com*)



Footprinting Countermeasures

So far, we have discussed the importance of footprinting, various ways to perform the task, and the tools that help to conduct execution. Now we will discuss footprinting countermeasures, the measures or actions taken to prevent or offset information disclosure.



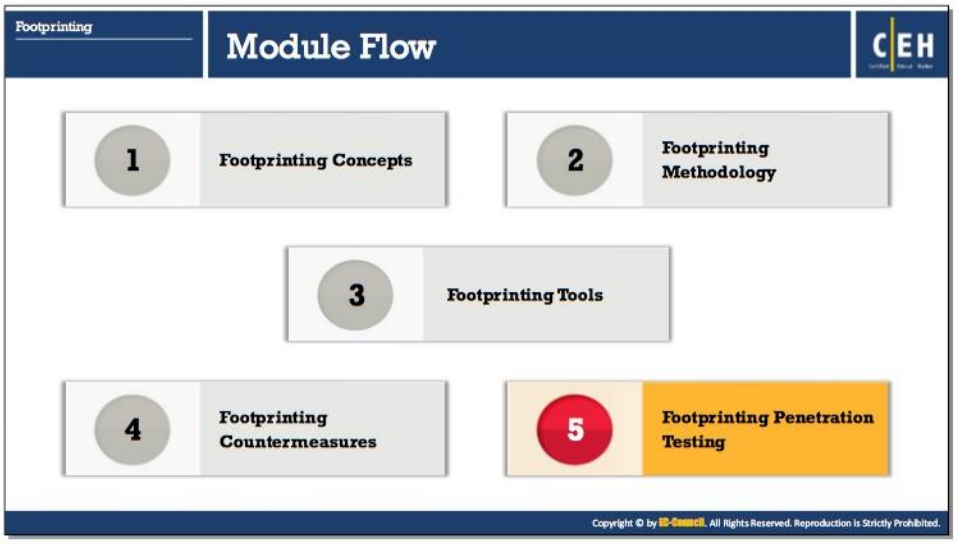


Footprinting Countermeasures

Some of the footprinting countermeasures are as follows:

- Restrict the employees to access social networking sites from 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 catalogues and so on.
- Limit the amount of information that you are publishing on the website/ Internet
- Use footprinting techniques to discover and remove any sensitive information publicly available
- Prevent search engines from caching a web page and use anonymous registration services
- Develop and enforce security policies such as information security policy, password policy and so on. to regulate the information that employees can reveal to third parties
- Set apart internal and external DNS or use split DNS, and restrict zone transfer to authorized servers
- Disable directory listings in the web servers
- Conduct periodically security awareness training to educate employees about various social engineering tricks and risks
- Opt for privacy services on Whois Lookup database
- Avoid domain-level cross-linking for the critical assets
- Encrypt and password protect sensitive information
- Do not enable protocols that are not required
- Always use TCP/IP and IPSec filters for defense in depth
- Configure IIS to avoid information disclosure through banner grabbing
- Hide the IP address and the related information by implementing VPN or keeping server behind a secure proxy
- Request archive.org to delete the history of the website from archive database
- Keep domain name profile a private one



Footprinting Penetration Testing

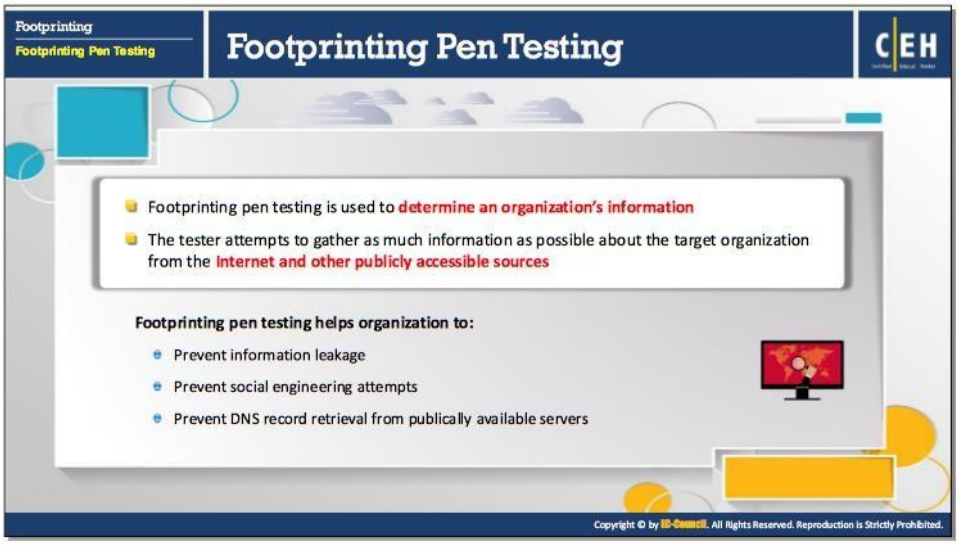
So far, we have discussed the necessary techniques and tools that can be used to footprint a target organization's network. Penetration testing (or pen testing) refers to the process of testing the organization's security posture using similar techniques and tools as that of an attacker, but with the knowledge and approval of the organization. Footprinting is the first step to perform in the pen testing process. Performing footprinting in a systematic manner enables a pen tester to discover potential security liabilities that an attacker may exploit. In the pen testing process, the pen tester acts as a malicious outsider and simulates an attack to find security loopholes.

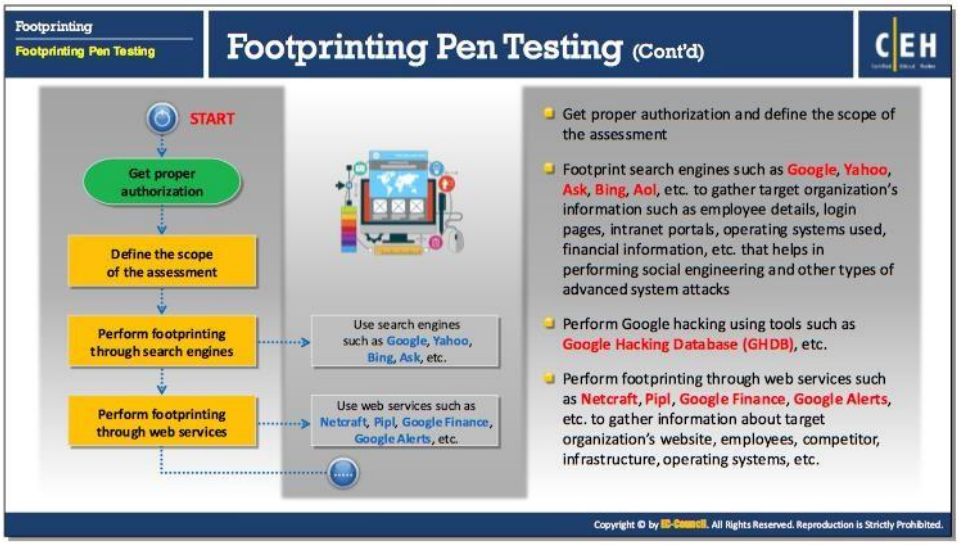
A footprinting pen test helps in determining an organization's information on the Internet such as network architecture, operating systems, applications, and users. The pen tester tries to gather publicly available sensitive information of the target by pretending to be an attacker. The target may be a specific host or a network.

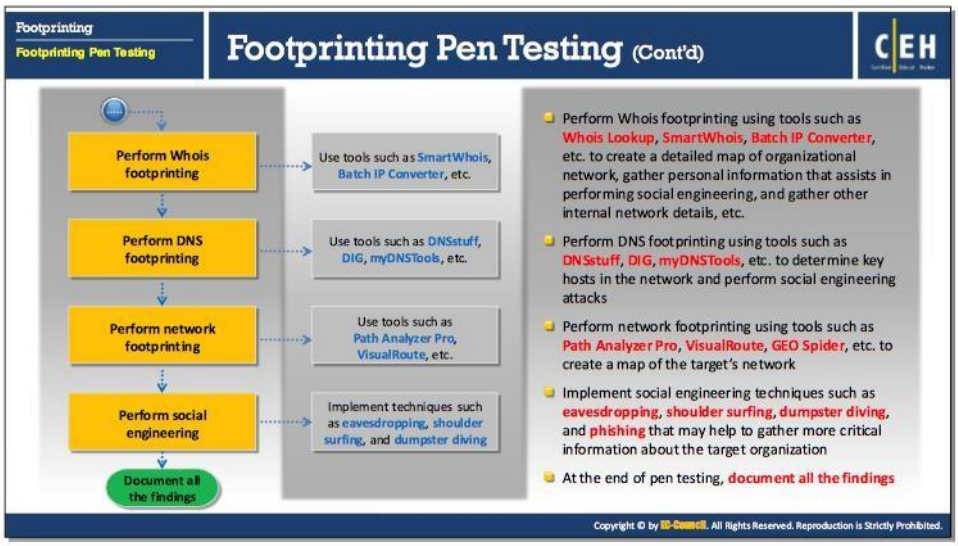
The pen tester can perform the same attacks as an attacker. The pen tester should try all possible ways in which to gather as much information as possible in order to ensure the maximum scope of footprinting pen testing. If the pen tester finds sensitive information on any publicly available information resource, that information should be reported to the organization.

Footprinting pen testing helps organization to:

- Prevent information leakage
- Prevent social engineering attempts
- Prevent DNS record retrieval from publically available servers







Footprinting Pen Testing Steps

Pen testing is a means to examine network security. Steps in the procedure should be followed in order, to ensure maximum scope of testing. The steps involved in footprinting pen testing are:

- **Step 1: Get proper authorization**

  Always perform pen testing with authorization. The first step in a footprinting pen test is to get proper authorization from the organization. This may or may not include the system administrators.

- **Step 2: Define the scope of the assessment**

  Defining the scope of the security assessment is a prerequisite for pen testing. Defining the scope of assessment determines the range of systems in the network to be tested and the resources that can be used to test and so on. It also determines the pen tester's limitations. Once you define the scope, you should plan and gather sensitive information using footprinting techniques.

- **Step 3: Perform footprinting through search engines**

  Use footprint search engines such as Google, Yahoo! Search, Ask, Bing, and Dogpile to gather the target organization's information such as employee details, login pages, intranet portals and so on. that can help in performing social engineering and other types of advanced system attacks.

  Perform Google hacking using tools such as Google Hacking Database (GHDB) and so on. Such use helps to expose security loopholes in the code and configuration of the websites. Google hacking is usually done with the help of advanced Google operators that locate specific strings of text, such as versions of vulnerable web applications.

- **Step 4: Perform footprinting through web services**

  Perform footprinting through web services such as Netcraft, Pipl, Google Finance, and Google Alerts to gather information about target organization's website, employees, competitor, infrastructure, and operating systems.

- **Step 5: Perform footprinting through social networking sites**

  Perform footprinting to gather target organization employee information from personal profiles on social networking sites such as Facebook, MySpace, LinkedIn, Twitter, Pinterest, Google+ and so on. This can assist in performing social engineering. You can also use people search engines to obtain information about a target person.

- **Step 6: Perform website footprinting**

  Perform website footprinting using tools such as Burp Suite, Web Data Extractor, HTTrack Web Site Copier, Metagoofil, and WebSite-Watcher in order to build a detailed map of the website's structure and architecture.

- **Step 7: Perform email footprinting**

  Perform email footprinting using tools such as eMailTrackerPro, Yesware, and ContactMonkey to gather information about the physical location of an individual. Use this to perform social engineering that in turn may help in mapping the target organization's network. Analyzing email headers can help to collect information such as sender's IP address, sender's mail server, sender's address, data and time received by the

originator's email servers, authentication system used by sender's mail server, sender's full name and so on.

- **Step 8: Gather competitive intelligence**

  Gather competitive intelligence using tools such as Hoover's, LexisNexis, or Business Wire. These tools extract competitor information such as its date of establishment, location, progress analysis, higher authorities, product analysis, marketing details and so on.

- **Step 9: Perform Whois footprinting**

  Perform Whois footprinting using tools such as Whois Lookup, SmartWhois, and Batch IP Converter to extract information about particular domains. You can capture information such as IP address, domain owner name, registrant name, and contact details including phone numbers, and email IDs. The information can be used to create a detailed map of organizational network, to gather personal information that assists to perform social engineering, to gather other internal network details and so on.

- **Step 10: Perform DNS footprinting**

  Perform DNS footprinting using tools such as DNSstuff, DIG, and myDNSTools to determine key hosts in the network and to perform social engineering attacks. Resolve the domain name to learn about its IP address, DNS records and so on.

- **Step 11: Perform network footprinting**

  Perform network footprinting using tools such as a Path Analyzer Pro, VisualRoute, and GEO Spider to learn the network range and other information about the target network that helps to draw the network diagram of the target.

- **Step 12: Perform social engineering**

  Implement social engineering techniques such as eavesdropping, shoulder surfing, dumpster diving, impersonation on social networking sites and phishing to gather critical information about the target organization. Through social engineering, you can gather target organization's security products in use, OS and software versions, network layout information, IP addresses and names of servers, and important personnel.

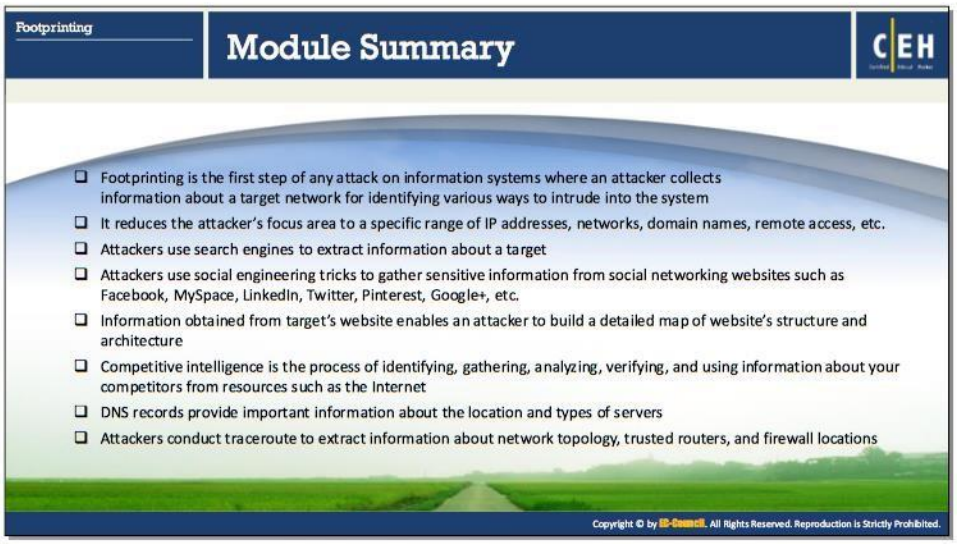
- **Step 13: Document all the findings**

  When finished with the implementation of footprinting techniques, collect and document the information obtained in each stage of testing. You can use this document to study, understand, and analyze the security posture of the target organization. This also enables you to find and fix security loopholes to prevent exploitation.

Footprinting Pen Testing Report Templates

Pen testing helps the organization to enhance its security perimeter. As a pen tester, you should gather sensitive information such as server details, OS and so on. of the target organization by conducting footprinting. Analyze the system and network defenses by breaking into its security with authorization (i.e., ethically) without causing any damage. Find the loopholes and weaknesses in the network or system security and list them along with respective

countermeasures in a pen testing report. Importantly, the pen testing report results from network penetration tests or security audits. It contains all the details such as types of tests performed, the hacking techniques used, and the results of hacking activity. In addition, the report also contains the highlights of security risks and vulnerabilities of an organization. Always keep the report confidential. If this information falls into the hands of attacker, the information in the report could be used to launch attacks.



Module Summary

This module ends with an overview discussion on footprinting methodology. In the next module, we will see how attackers as well as ethical hackers and pen testers perform network scanning to collect information about a target of evaluation before an attack or audit.