



University of Babylon College of
Information Technology
Department of Information Security

Ethical Hacking

Lecture 3: Footprinting and Reconnaissance I

Asst.Lect. Rasha Hussein



What is Footprinting ?

- **Footprinting** is the first step of any attack on information system in which an attacker collects information about a target network to identify various ways to intrude into the system.

Types of Footprinting

- **Passive Footprinting**

Gathering information about the target without direct interaction

- **Active Footprinting**

Gathering information about the target with direct interaction

Why Footprinting?

Footprinting helps to:

1. Understand Security Posture: Provides an overview of the organization's security, such as firewall presence and security configurations.
2. Reduce Attack Surface: focuses on specific systems or targets, minimizing the scope of potential attacks.
3. Identify Vulnerabilities: Creates a database of vulnerabilities, threats, and system loopholes.
4. Draw Network Map: Develop a network map showing topology, routers, servers, and other critical details.

Information Obtained in Footprinting

- **Organization Information**

Employee details , telephone numbers , location , background of the organization , web technologies , etc.

- **Network Information**

Domain and sub-domains , network blocks, IP addresses of the reachable systems, Whois record , DNS , etc.

- **System Information**

OS and location of web servers, users and passwords , etc.

Comparison Between Information Gathering, Footprinting, Reconnaissance, and Enumeration

Term	Description
Information gathering	A general process of collecting data about a target system or network for various purposes. Information gathering can be performed using different methods and tools, such as web search engines, web services, social media platforms, email headers, network scanners, and more.
Footprinting	A specific type of information gathering that focuses on identifying and understanding the security risks present in an organization. It involves gathering information about the target's domain name, IP address range, network infrastructure, security policies, web server type, operating system type, and more.
Reconnaissance	A specific type of information gathering that is part of the ethical hacking methodology. It involves finding potential attack vectors or vulnerabilities on the target system or network. Reconnaissance can include footprinting as well as other techniques such as port scanning and vulnerability assessment.
Enumeration	A further step in information gathering that involves extracting more detailed information from the target system or network. It involves identifying the hosts, ports, services, users, groups, shares, files, directories, and other resources that are available on the target system or network. Enumeration can be done using various tools and protocols, such as SNMP, NetBIOS, LDAP, NTP, SMTP, and more.

Footprinting Threats

1. Social Engineering

Without using any intrusion methods, hackers directly and indirectly collect information through persuasion and various other means. Here, crucial information is gathered by the hackers through employees without their consent.

2. System and Network Attacks

Through footprinting, attackers can gather information related to the target organization's system configuration, operating system running on the machine, and so on. Using this information, attackers can find the vulnerabilities present in the target system and then can exploit those vulnerabilities. Thus, attackers can take control over a target system. Similarly, attackers can also take control over the entire network.

3. Information Leakage

If sensitive organizational information falls into the hands of attackers, then they can build an attack plan based on the information, or use it for monetary benefits.

4. Privacy Loss

With the help of footprinting, hackers are able to access the systems and networks of the company and even escalate the privileges up to admin levels. Whatever privacy was maintained by the company is completely lost.

5. Corporate Espionage

Corporate espionage is one of the major threats to companies as competitors can spy and attempt to steal sensitive data through footprinting. Due to this type of espionage, competitors are able to launch similar products in the market, affecting the market position of a company.

6. Business Loss

Footprinting has a major effect on businesses such as online businesses and other ecommerce websites, banking and financial related businesses, etc. Billions of dollars are lost every year due to malicious attacks by hackers.

Footprinting Methodology

1. Footprinting through search engines
2. Footprinting through web services
3. Website footprinting
4. Email footprinting
5. Footprinting through social networking sites
6. Whois footprinting
7. DNS footprinting
8. Network footprinting
9. Footprinting through social engineering

1- Footprinting through search engines

- Attackers use search engines to **extract information about a target**, such as employed technology platforms, employee details, login pages, and intranet portals, which help the attacker to perform social engineering and other types of advanced system attacks

- Major search engines:

Google

Bing

YAHOO!

Ask

Aol.

Baidu 百度

DuckDuckGo

- Attackers can use **advanced search operators** available with these search engines and create complex queries to find, filter, and sort specific information about the target

- Search engines are also used to find other sources of **publicly accessible information resources**, e.g., you can type “top job portals” to find major job portals that provide critical information about the target organization

Google search Command (Operators)

Filter	Description	Example
allintext	Searches for occurrences of all the keywords given.	allintext:"keyword"
intext	Searches for the occurrences of keywords all at once or one at a time.	intext:"keyword"
inurl	Searches for a URL matching one of the keywords.	inurl:"keyword"
allinurl	Searches for a URL matching all the keywords in the query.	allinurl:"keyword"
intitle	Searches for occurrences of keywords in title all or one.	intitle:"keyword"
allintitle	Searches for occurrences of keywords all at a time.	allintitle:"keyword"
site	Specifically searches that particular site and lists all the results for that site.	site:"www.google.com"
filetype	Searches for a particular filetype mentioned in the query.	filetype:"pdf"
link	Searches for external links to pages.	link:"keyword"
numrange	Used to locate specific numbers in your searches.	numrange:321-325
before/after	Used to search within a particular date range.	filetype:pdf & (before:2000-01-01 after:2001-01-01)
allinanchor (and also inanchor)	This shows sites which have the keyterms in links pointing to them, in order of the most links.	inanchor:rat
allinpostauthor (and also inpostauthor)	Exclusive to blog search, this one picks out blog posts that are written by specific individuals.	allinpostauthor:"keyword"
related	List web pages that are “similar” to a specified web page.	related:www.google.com
cache	Shows the version of the web page that Google has in its cache.	cache:www.google.com

Google search Command (Operators) Example

"CEO" "email" "@" "Name" "Phone" filetype:csv OR filetype:xls OR filety



All

Images

Maps

News

Videos

More

Settings

Tools

About 14,700 results (0.61 seconds)

[XLS] [fortune 1000](#)

[assets.time.com/cm/fortune-data.../2016_FORTUNE_1000_w_Contacts_Sample.xls](#) ▼

... CORPORATE WEBSITE, CEO NAMERETURN TO MAIN DATA, CEO TITLE, Email, Office Phone, Office Ext, Direct Dial, CFO NAME, CFO TITLE, Email, Office ...

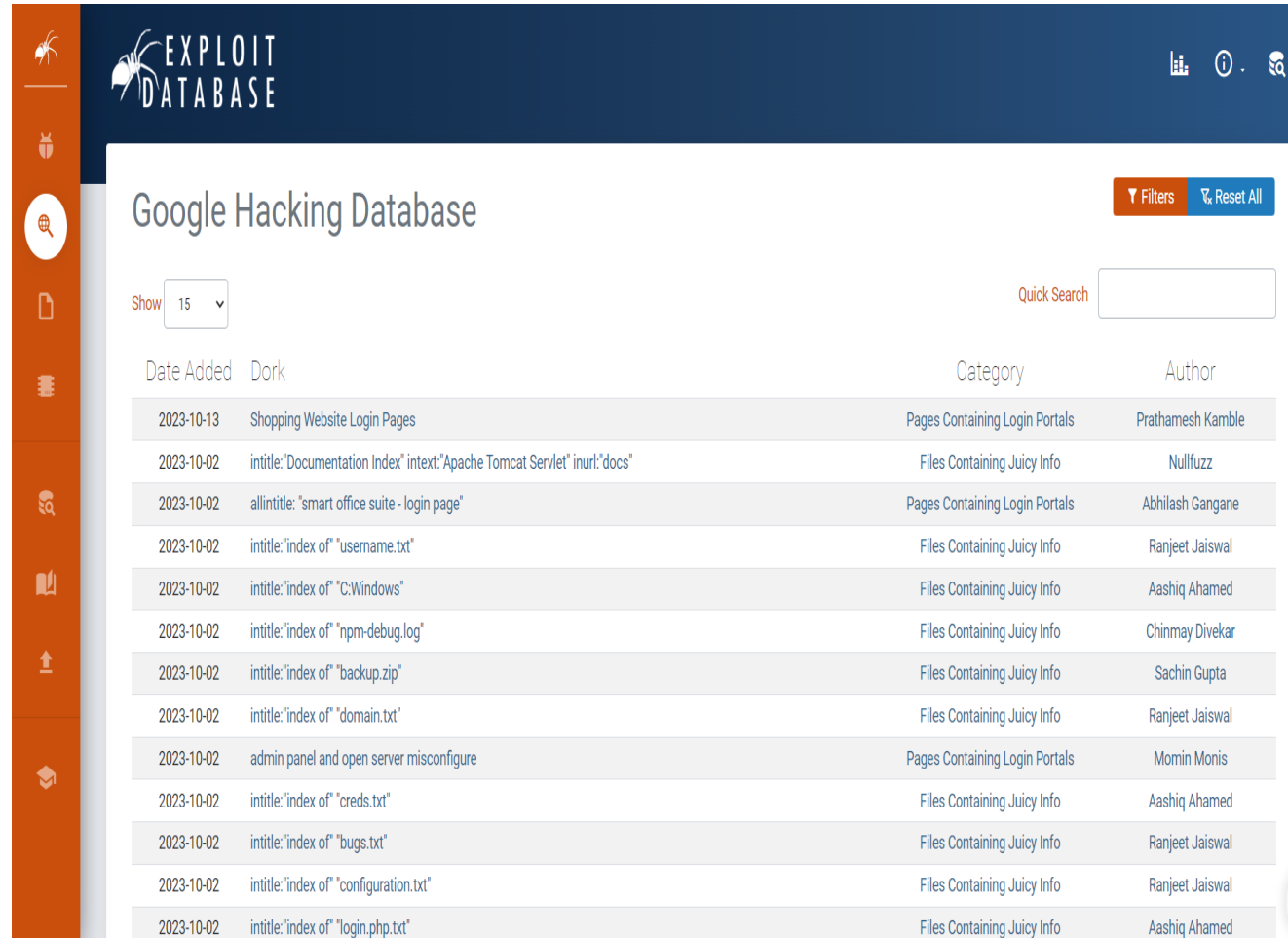
[XLS] [Fortune 1000 Companies List and Contact Info - Boolean Strings](#)

[booleanstrings.com/wp-content/uploads/2014/01/fortune1000-2012.xls](#) ▼

6, Company, Phone, Email Format, Email Format 2, General Email, CEO Name, CEO Email, Website, Address, City, State, Zipcode. 7, Chevron, 925-842-1000 ...

1- Footprinting through search engines

- Google Hacking Database (GHDB) : is very good for learn Google Dorks and how it's done in real world scenario , <https://www.exploit-db.com/google-hacking-database>
- The Google Hacking Database (GHDB) is an authoritative source for querying the everwidening reach of the google search engine
- Attackers use Google dorks in Google advanced search operators to extract sensitive information about their target,such as vulnerable servers, error messages, sensitive files, login pages,and websites
- **Metagoofil** - Command line interface that uses **Google hacks** to find information in meta tags (domain, filetype, etc; Is a google dorks for terminal).



The screenshot displays the Exploit Database website interface. At the top, the logo "EXPLOIT DATABASE" is visible. Below it, the page title "Google Hacking Database" is shown. A search bar with a "Quick Search" button and a "Filters" button is present. A "Show 15" dropdown menu is also visible. The main content is a table with the following columns: "Date Added", "Dork", "Category", and "Author".

Date Added	Dork	Category	Author
2023-10-13	Shopping Website Login Pages	Pages Containing Login Portals	Prathamesh Kamble
2023-10-02	intitle:"Documentation Index" intext:"Apache Tomcat Servlet" inurl:"docs"	Files Containing Juicy Info	Nullfuzz
2023-10-02	allintitle: "smart office suite - login page"	Pages Containing Login Portals	Abhilash Gangane
2023-10-02	intitle:"index of" "username.txt"	Files Containing Juicy Info	Ranjeet Jaiswal
2023-10-02	intitle:"index of" "C:Windows"	Files Containing Juicy Info	Aashiq Ahamed
2023-10-02	intitle:"index of" "npm-debug.log"	Files Containing Juicy Info	Chinmay Divekar
2023-10-02	intitle:"index of" "backup.zip"	Files Containing Juicy Info	Sachin Gupta
2023-10-02	intitle:"index of" "domain.txt"	Files Containing Juicy Info	Ranjeet Jaiswal
2023-10-02	admin panel and open server misconfigure	Pages Containing Login Portals	Momin Monis
2023-10-02	intitle:"index of" "creds.txt"	Files Containing Juicy Info	Aashiq Ahamed
2023-10-02	intitle:"index of" "bugs.txt"	Files Containing Juicy Info	Ranjeet Jaiswal
2023-10-02	intitle:"index of" "configuration.txt"	Files Containing Juicy Info	Ranjeet Jaiswal
2023-10-02	intitle:"index of" "login.php.txt"	Files Containing Juicy Info	Aashiq Ahamed

Gathering Information from IOT Search Engines

As shown in the screenshot, attackers can use Shodan to find all the IoT devices of the target organization that are having open ports and services.

The screenshot displays the Shodan search engine interface. The search query is 'SCADA', and the total number of results is 843. The interface includes a navigation bar with 'SHODAN' and a search input field containing 'SCADA'. Below the search bar, there are tabs for 'Exploits' and 'Maps'. The main content area is divided into several sections:

- TOTAL RESULTS:** 843
- TOP COUNTRIES:** A world map showing the distribution of results by country. The top countries listed are Belgium (186), Italy (95), Spain (93), United States (72), and Bulgaria (46).
- TOP SERVICES:** A list of services found on the devices, including HTTP (478), FTP (70), S061 (39), Modbus (33), and HTTPS (11).
- TOP ORGANIZATIONS:** A list of organizations, including Euphony Benelux n.v. (185), Telecom Italia Mobile (70), Vodafone Spain (53), Verizon Wireless (33), and Com4 AS (21).
- TOP OPERATING SYSTEMS:** A list of operating systems, including Linux 2.5.x (4), Windows 7 or 8 (2), and Unix (2).

The search results are displayed in a grid format. Each result includes an IP address, the organization name, the date it was added, and the service details. For example, the first result is for IP 2.195.234.126, identified as Telecom Italia Mobile, added on 2019-06-14 03:46:36 GMT, located in Italy. The service details for this IP are: HTTP/1.1 307 Temporary Redirect, Server: CirCarLife Scada v4.2.4, Connection: keep-alive, Date: Fri, 14 Jun 2019 3:46:36 GMT, Content-Length: 0, and Location: html/index.html.

```
HTTP/1.1 307 Temporary Redirect
Server: CirCarLife Scada v4.2.4
Connection: keep-alive
Date: Fri, 14 Jun 2019 3:46:36 GMT
Content-Length: 0
Location: html/index.html
```

The second result is for IP 109.111.118.113, identified as Andorra Telecom, added on 2019-06-14 01:49:29 GMT, located in Andorra. The service details for this IP are: HTTP/1.1 307 Temporary Redirect, Server: CirCarLife Scada v4.2.4, Connection: keep-alive, Date: Fri, 14 Jun 2019 1:49:29 GMT, Content-Length: 0, and Location: html/setup.html.

```
HTTP/1.1 307 Temporary Redirect
Server: CirCarLife Scada v4.2.4
Connection: keep-alive
Date: Fri, 14 Jun 2019 1:49:29 GMT
Content-Length: 0
Location: html/setup.html
```

The third result is for IP 149.13.154.5, identified as Z.G. Electronica, added on 2019-06-14 04:42:51 GMT, located in Italy. The service details for this IP are: HTTP/1.1 200 OK, Expires: Thu, 27 Jan 2009 16:00:00 GMT, Set-Cookie: ID=50524088C1C2; path=/, and a meta tag: <meta http-equiv="content-type" content="text/html; charset=windows-1252">.

```
HTTP/1.1 200 OK
Expires: Thu, 27 Jan 2009 16:00:00 GMT
Set-Cookie: ID=50524088C1C2; path=/

<html>
<head>
<meta http-equiv="content-type" content="text/html; charset=windows-1252">
<meta name="robots" content="noindex">
<title>IS-Lite access verification</title>
<link re...
```

2- 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 carry out other types of advanced system attacks.
- **Tools**
- Netcraft : For getting information about target organization domains and subdomains
- Sublist3r
- Pentest-Tools
- theHarvester
- Deep and Dark web By Tor Browser
- Censys
- Google Earth
- Tineye : <https://tineye.com/>

Sublist3r

- Source: <https://github.com>
- Sublist3r is a Python script designed to enumerate the subdomains of websites using OSINT. It enables you to enumerate subdomains across multiple sources at once. Further, it helps penetration testers and bug hunters in collecting and gathering 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.
- **Syntax:** `sublist3r [-d DOMAIN] [-b BRUTEFORCE] [-p PORTS] [-v VERBOSE][-t THREADS] [-e ENGINES] [-o OUTPUT]`

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

Table 2.4: Sublist3r options with description

Examples

```
Parrot Terminal
File Edit View Search Terminal Help

[~]-[root@parrot]-[~]
#sublist3r -d google.com

SUBLIST3R
# Coded By Ahmed Aboul-Ela - @aboul3la

[-] Enumerating subdomains now for google.com
[-] Searching now in Baidu..
[-] Searching now in Yahoo..
[-] Searching now in Google..
[-] Searching now in Bing..
[-] Searching now in Ask..
[-] Searching now in Netcraft..
[-] Searching now in DNSdumpster..
[-] Searching now in Virustotal..
[-] Searching now in ThreatCrowd..
[-] Searching now in SSL Certificates..
[-] Searching now in PassiveDNS..
[-] Total Unique Subdomains Found: 853
www.google.com
alt.aspmx.1.google.com
client.1.google.com
clients.1.google.com
gmail-smtp-mas.1.google.com
misc-anycast.1.google.com
31.google.com
360suite.google.com
Cliens-2.google.com
www1.google.com
aboutme.google.com
```

```
Parrot Terminal
File Edit View Search Terminal Help

[~]-[root@parrot]-[~]
#sublist3r -d google.com -p 80 -e Bing

SUBLIST3R
# Coded By Ahmed Aboul-Ela - @aboul3la

[-] Enumerating subdomains now for google.com
[-] Searching now in Bing..
[-] Total Unique Subdomains Found: 57
[-] Start port scan now for the following ports: 80
accounts.google.com - Found open ports: 80
admin.google.com - Found open ports: 80
aboutme.google.com - Found open ports: 80
console.actions.google.com - Found open ports: 80
adssettings.google.com - Found open ports: 80
careers.google.com - Found open ports: 80
calendar.google.com - Found open ports: 80
business.google.com - Found open ports: 80
chrome.google.com - Found open ports: 80
code.google.com - Found open ports: 80
classroom.google.com - Found open ports: 80
ssh.cloud.google.com - Found open ports: 80
console.cloud.google.com - Found open ports: 80
cse.google.com - Found open ports: 80
contacts.google.com - Found open ports: 80
crowdsourcing.google.com - Found open ports: 80
appmaker.google.com - Found open ports: 80
artsandculture.google.com - Found open ports: 80
analytics.google.com - Found open ports: 80
datastudio.google.com - Found open ports: 80
adwords.google.com - Found open ports: 80
```

Harvesting Email Lists

- Gathering email addresses related to the target organization acts as an **important attack vector during the later phases of hacking**
- Attackers use automated tools such as **theHarvester** and **Email Spider** to collect publicly available email addresses of the target organization that helps them perform social engineering and brute-force attacks

```
Parrot Terminal
File Edit View Search Terminal Help
[~]-[attacker@parrot ~]-
theHarvester -d microsoft.com -l 200 -b baidu
table results already exists

.....
theHarvester
.....

theHarvester 3.1.0
Coded by Christian Martorella
Edge-Security Research
cmartorella@edge-security.com
.....

[*] Target: microsoft.com
[*] Searching Baidu.
```

```
Parrot Terminal
File Edit View Search Terminal Help
[+] Emails found:
-----
msdneg@microsoft.com
tonas@contoso.onmicrosoft.com
user@contoso.onmicrosoft.com
Rone.Li@microsoft.com
v-lanz@microsoft.com
support@microsoft.com
delist@messaging.microsoft.com
honeypage@microsoft.com
postmaster@ul.onmicrosoft.com
TheWebInterfaceShouldBeRadicallyRefactoredJohnR.DouceurJonHowellBryanParnoJohndo
howellparno@microsoft.com
quarantine@messaging.microsoft.com
hicwhql@microsoft.com
ctcwhql@microsoft.com
age3support@microsoft.com
...STOR.WW.00.EN.MSF.RMD.TS.T15.SPT.00.EM@css.one.microsoft.com
pexdata@microsoft.com
brohrer@microsoft.com
rightlicense@microsoft.com
```

```
Parrot Terminal
File Edit View Search Terminal Help
[*]-[attacker@parrot]-[-]
$ theHarvester -d microsoft.com -l 200 -b baidu
table results already exists

*****
*
* [HARVESTING]
*
* theHarvester 3.1.0
* Coded by Christian Martorella
* Edge-Security Research
* cmartorella@edge-security.com
*
*****

[*] Target: microsoft.com
[*] Searching Baidu.
```

- The harvester -d microsoft.com -l 200 -b baidu
theHarvester tool, which is a tool that can help you collect information about a target domain from different public sources. The command has three parameters: **-d microsoft.com**: This specifies the target domain name as microsoft.com. The tool will search for information related to this domain, such as email addresses, subdomains, hosts, and names. **-l 200**: This limits the number of results to 200. The tool will stop searching after finding 200 results or exhausting the source. **-b baidu**: This specifies the source as baidu. The tool will use the Baidu search engine to find information about the target domain.
The command will output a summary of the results, such as the number and list of email addresses, subdomains, hosts, and names found for the target domain.

Deep and Dark Web Footprinting

Deep web

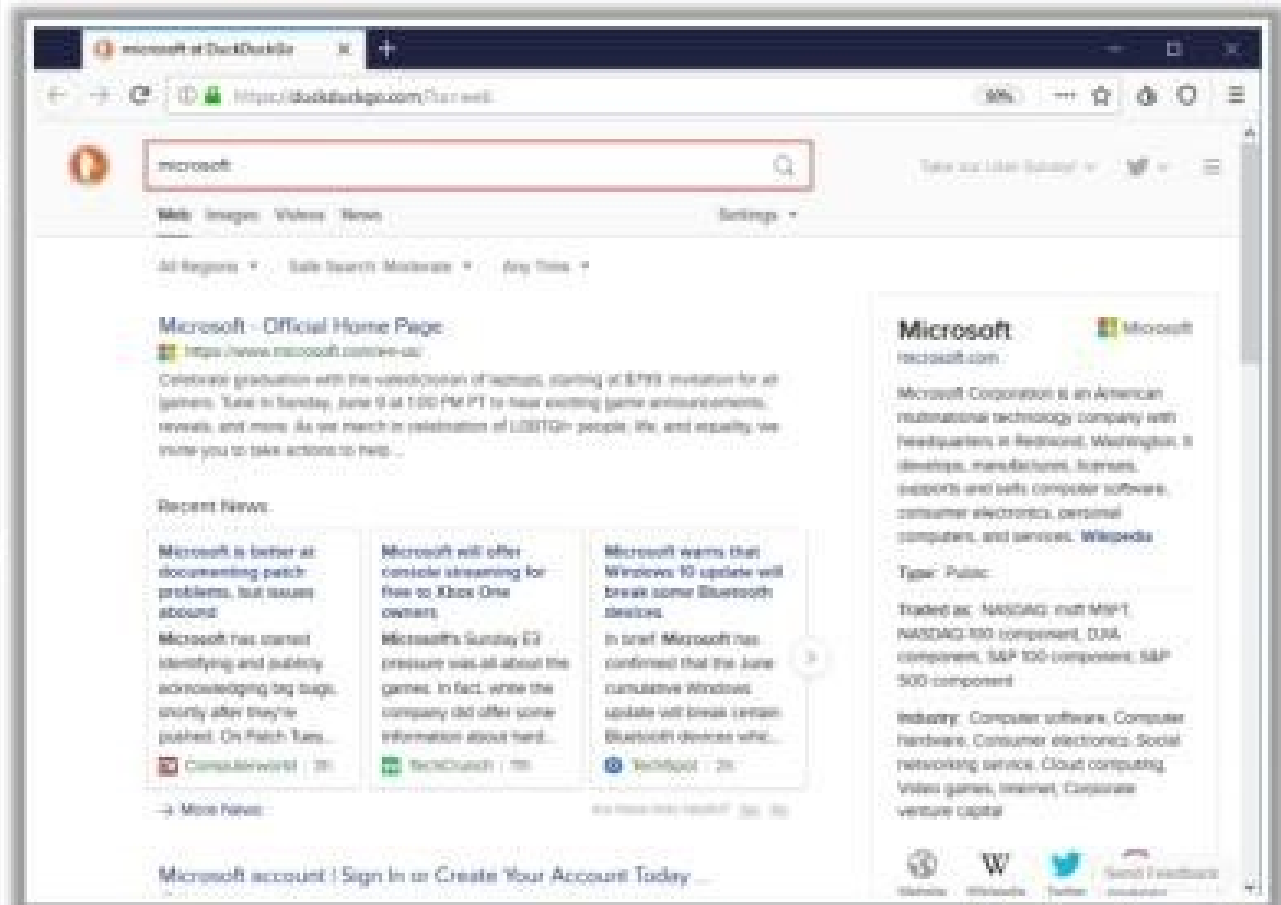
- It consists of web pages and contents that are **hidden and unindexed** and cannot be located using traditional web browsers and search engines
- It can be accessed by **search engines** like Tor Browser and The WWW Virtual Library

Dark web or Darknet

- It is the subset of the deep web that enables anyone to **navigate anonymously** without being traced
 - It can be accessed by **browsers**, such as TOR Browser, Freenet, GNUnet, I2P, and Retroshare
-
- Attackers use deep and dark web searching tools, such as **Tor Browser** and **ExoneraTor**, to **gather confidential information about the target**, including credit card details, passport information, identification card details, medical records, social media accounts, Social Security Numbers (SSNs), etc.

TOR Browser

It is used to access the deep and dark web where it acts as a **default VPN** for the user and bounces the network IP address through several servers before interacting with the web



<https://www.torproject.org>

Tor Browser

- Source: <https://www.torproject.org/>
- Tor Browser is used to access the deep and dark web, where it acts as a default VPN for the user and bounces the network IP address through several servers before interacting with the web. Attackers use this browser to access hidden content, unindexed websites, and encrypted databases present in the deep web. As shown in the screenshot, by using Tor Browser, attackers can obtain more detailed and hidden information about the target organization.

3- Website Footprinting

- Web mirroring | Website Cloning - allows for discrete testing offline
 - HTTrack - you can use the CLI version or Web Interface version
 - Wget - Linux command
 - `wget -mk -w 10 http://hackthissite.org/`
 - Black Widow
 - WebRipper
 - Teleport Pro
 - Backstreet Browser
- Archive.org / [Wayback machine](https://archive.org/web/) : Provides cached websites from various dates which possibly have sensitive information that has been now removed.

The screenshot shows the top navigation bar of the Internet Archive website, including links for WEB, BOOKS, VIDEO, AUDIO, SOFTWARE, and IMAGES. Below this is a dark navigation menu with links for ABOUT, BLOG, PROJECTS, HELP, DONATE, CONTACT, JOBS, VOLUNTEER, and PEOPLE. The main content area features the Internet Archive logo and the text "Explore more than 846 billion web pages saved over time". A search bar with the text "http://" and a "BROWSE HISTORY" button is visible. Below the search bar is a "DONATE" button. A row of ten small thumbnail images of various websites is displayed. At the bottom, there are three sections: "Tools" with links for "Wayback Machine Availability API", "WordPress Broken Link Checker", and "404 Handler for Webmasters"; "Subscription Service" with a description of Archive-It; and "Save Page Now" with a search bar, a "SAVE PAGE" button, and a description of the service.

Website Footprinting

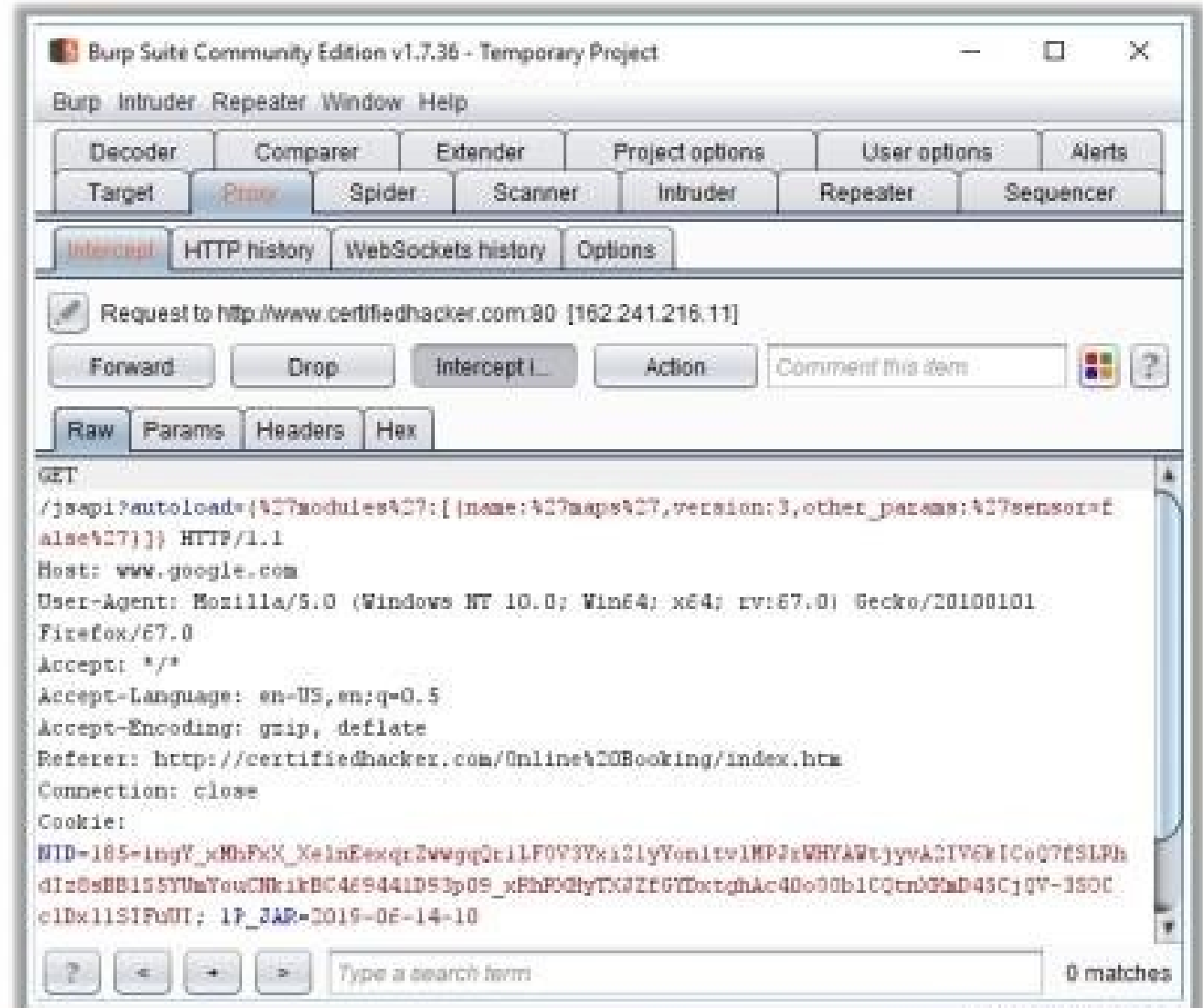
- Website footprinting refers to the **monitoring and analysis of the target organization's website** for information

Browsing the target website may provide the following information:

- Software used and its version
- Operating system used and its scripting platform
- Sub-directories and parameters
- Filename, path, database field name, or query
- Technologies used
- Contact and CMS details

Attackers use **Burp Suite**, **Zaproxy**, **Wappalyzer**, **Website Informer**, etc. to view headers that provide the following information:

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



Website Footprinting (Cont'd)

Examining the HTML source code may provide

- Comments present in the source code
- Contact details of the web developer or admin
- File system structure and script type

Examining cookies may provide

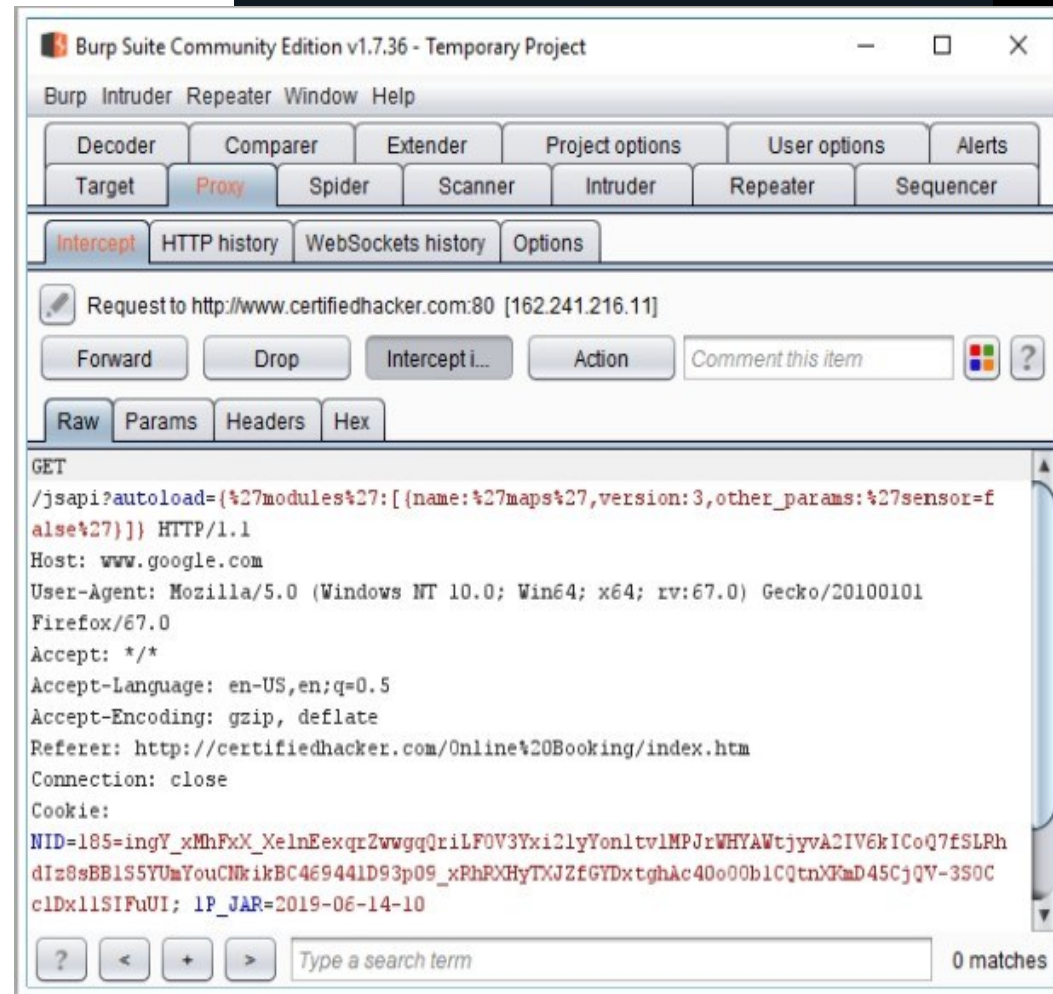
- Software in use and its behavior
- Scripting platforms used

```
1 <!DOCTYPE html>
2 <html lang="en-us" dir="ltr">
3 <head data-bbox="0 0 1000 1000">
4 <meta charset="UTF-8" />
5
6 <meta http-equiv="x-ua-compatible" content="ie=edge" />
7 <meta name="viewport" content="width=device-width, initial-scale=1" />
8 <title>Microsoft - Official Home Page</title>
9
10 <meta name="twitter:url" content="https://www.microsoft.com/en-us" />
11 <meta property="og:url" content="https://www.microsoft.com/en-us" />
12 <meta name="twitter:title" content="Microsoft - Official Home Page" />
13 <meta property="og:title" content="Microsoft - Official Home Page" />
14 <meta name="twitter:description" content="At Microsoft our mission and values are to help people and businesses throughout the world realize their full potential." />
15 <meta property="og:description" content="At Microsoft our mission and values are to help people and businesses throughout the world realize their full potential." />
16 <meta name="twitter:card" content="summary" />
17 <meta property="og:type" content="website" />
18 <meta name="description" content="At Microsoft our mission and values are to help people and businesses throughout the world realize their full potential." />
19
20 <link rel="SHORTCUT ICON" href="https://s.a-microsoft.com/favicon.ico?v1" type="image/x-icon" />
```

Domain	Storage
answers.microsoft.com	Database Storage, Local storage
azure.microsoft.com	Database Storage, Local storage
blogs.microsoft.com	Database Storage, Local storage
bn2.ftl.microsoft.com	Local storage
c1.microsoft.com	2 cookies
cy2.ftl.microsoft.com	Local storage
docs.microsoft.com	1 cookie, Local storage

Burp Suit

- 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 to finding and exploiting security vulnerabilities.
- Burp Proxy allows attackers to intercept all requests and responses between the browser and the target web application and obtain information such as web server used, its version, and web-application-related vulnerabilities.



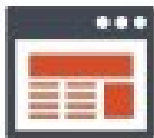
Website footprinting can be performed by examining HTML source code and cookies.

- 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 those that the CMS system creates. The comments may provide clues as to what is running in the background. They may even provide contact details of the web developer or administrator. Observe all the links and image tags 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 cookie name, value, and domain size can also be extracted.

Mirroring Entire Website

HTTrack Web Site Copier

- Mirroring an entire website onto a local system enables an attacker to browse website offline; it also assists in finding **directory structure** and other valuable information from the mirrored copy without sending multiple requests to web server
- Web mirroring tools, such as HTTrack Web Site Copier, and NCollector Studio, allow you to **download a website to a local directory**, recursively building all directories, HTML, images, flash, videos, and other files from the server to your computer



The screenshot shows the HTTrack Web Site Copier application window. The title bar reads "Site mirroring in progress (2/30 (-+25), 720381 bytes) - (bestwhit)". The interface is divided into several sections:

- File Explorer:** On the left, a tree view shows the local file system. A callout box points to the "test" directory under "Local Disk (C:)", with the text "Location of Mirrored Website in C: drive".
- Information:** A box on the right displays statistics: "In progress: Transferring data...", "Bytes saved: 712,294B", "Time: 28s", "Transfer rate: 24,39KB/s (23,61KB/s)", "Active connections: 4", "Links scanned: 2/30 (-+25)", "Files written: 25", "Files updated: 0", and "Errors: 0".
- Actions:** A table lists the files being mirrored, each with a progress bar and a "SKIP" button. The files are:

scanning	www.certfiedhacker.com/js	██████████	SKIP
receive	www.certfiedhacker.com/ages/hibidshere/slide-1.png	██████████	SKIP
receive	www.certfiedhacker.com/ages/hibidshere/slide-2.png	██████████	SKIP
receive	www.certfiedhacker.com/ages/hibidshere/slide-3.png	██████████	SKIP
			SKIP
			SKIP
			SKIP
			SKIP
			SKIP
			SKIP
			SKIP
			SKIP
			SKIP

At the bottom right of the window, the text "Mirroring target website" is written in red. The bottom of the window contains buttons for "Back", "Next", "Cancel", and "Help", along with a "NUM" indicator.

4- Email Footprinting

- Email footprinting is the process of gathering information from email headers, such as sender's name, email address, IP address, date and time, subject line, etc. It can also involve analyzing the email content and attachments for clues or malicious code. Some examples of tools that can perform email footprinting are:
- Email Tracker: A tool that can track when an email is opened by the recipient and provide information such as location, device type, browser type, etc.
- MailSniper: A tool that can perform various attacks on email servers and accounts, such as password spraying, enumeration, harvesting, etc.
- The Harvester: A tool that can extract email addresses from various sources such as search engines, social media platforms, websites, etc.

4- Email Footprinting

Tracking Email Communications

Email tracking is a method that helps you to monitor as well as to track the emails of a particular user.. A lot of email tracking tools are readily available in the market, using which you can collect information such as IP addresses, mail servers, and service provider from which the mail was sent. Attackers can use this information to build the hacking strategy.

By using email tracking tools you can gather the following information about the victim:

- Recipient's system IP address
- Geolocation
- Read duration
- Proxy detection
- Links
- Operating system
- Forward email
- Device type
- Path travel

Collecting Information from Email Headers

- Email headers are hidden metadata of an email message that show the sender, recipient, date, subject, and the route of the email. They can help you to troubleshoot, verify, and identify emails. To view them, you need to access the raw message source in your email client. To understand them, you need to know the meaning of the different fields and values they contain, such as From, To, Subject, Date, Received, Message-ID, Content-Type, MIME-Version, DKIM-Signature, SPF, and DMARC. You can use online tools such as [Message header analyzer] to help you parse and interpret them more easily.
- **Some of the common fields are:**
- **Message-ID:** A unique identifier for the email message.
- **Content-Type:** The type and format of the content of the email, such as text/plain or text/html.
- **MIME-Version:** The version of Multipurpose Internet Mail Extensions (MIME) used by the email, which is a standard for encoding and transferring different types of data via email.
- **DKIM-Signature:** A digital signature that verifies that the email was sent by an authorized sender and that the content was not altered in transit.
- **SPF:** A mechanism that verifies that the sender's domain name matches the IP address of the sender's server.
- **DMARC:** A mechanism that specifies how the receiver should handle emails that fail DKIM or SPF validation.

4:06 PM (1 hour ago)



Reply



Reply

Forward

Filter messages like this

Print

Add Help Communities to Contacts list

Delete this message

Block "Help Communities"

Report spam

Report phishing

Show original

Translate message

Download message

Mark unread from here

Delivered-To: wwwbirdie@gmail.com

Received: by 2002:a8a:a99:0:0:0:0:0 with SMTP

Sun, 9 Jun 2019 21:09:48 -0700 (PDT)

Return-Path: <riniimathews@gmail.com>

Received: from mail-sor-f41.google.com (mail-sor-f41.google.com. [209.85.220.41])

by mx.google.com with SMTPS id v17sor28

for <wwwbirdie@gmail.com>

(Google Transport Security);

Sun, 09 Jun 2019 21:09:48 -0700 (PDT)

Received-SPF: pass (google.com: domain of riniimathews@gmail.com designates 209.85.220.41 as

permitted sender) client-ip=209.85.220.41;

Authentication-Results: mx.google.com;

dkim=pass header.i=@gmail.com header.s=20161025 header.b=s6SMnvzN;

spf=pass (google.com: domain of riniimathews@gmail.com designates 209.85.220.41 as

permitted sender) smtp.mailfrom=riniimathews@gmail.com;

dmarc=pass (p=NONE sp=QUARANTINE dis=NONE) header.from=gmail.com

DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;

d=gmail.com; s=20161025;

h=mime-version:from:date:message-id:subject:to;

bh=nheQC6dgq1LhKwkOyk8x4gYw0VwtRRaK2KrErWhvfCg=;

b=s6SMnvzNwWAeedUZf5r7LGPdGSiUyxSKDxvLIBGHvEcF/pIIqx8KkNR2JGfOMPVXAL

e7630+SPbK+M54CPx9hkvdbyhbcVgUZFuEvp3J/fPvIliT7Blf8jGXwqvwxwQhTH4+/g

XeIE0g6h98SYL4lvePj8I9hw1xvjym8QYRoCgEqWE8JVRfqmNcOxNBa6yoxuOVIJRT0A

aFdUZS3KJMwbG8gBU6hS+bHrr3no370YJgLlh/YwkLTx76h7BgDYBzHcyg+ZPA+HvKSK

3BwvrqeaGvGeZWh6xaS6LNmhf7CIuuxa/skSls1pfsKIEJv1qeCAV0Cqi34JC292HRn2

YCxw==

MIME-Version: 1.0

From: [riniimathews](mailto:riniimathews@gmail.com) <riniimathews@gmail.com>

Date: Mon, 10 Jun 2019 09:39:37 +0530

Message-ID: <CA++=zy1VzQ1gFmUDByZzqE90SbjwFYK7jes@com>

Subject: Check Out Daily News Feed

To: wwwbirdie@gmail.com

The address from which the message was sent

Date and time received by the originator's email servers

Sender's IP address

Sender's mail server

Authentication system used by sender's mail server

Sender's full name

Date and time of message sent

File Help

My Inbox My Trace Reports Trace Headers Trace Address Email Accounts Settings Export Rules Trial Edition

Home Subject: THYBNKCRD CRE... x

The trace is complete, the information found is displayed on the right.

New Trace View Report



#	Hop IP	Hop Name	Location
Start	10.10.10.2		
End	208.78.224.20	server.economist.co.in	West Chester, Pennsylvania, USA

Email Summary

From: [redacted]
To: [redacted]@gmail.com
Date: Mon, [redacted] 09:18:09 +0000
Subject: THYBNKCRD CREDIT CARD (XX2917) WILL BE DEI
Location: West Chester, Pennsylvania, USA

Misdirected: No
Abuse Address: abuse@privatesystems.net
Abuse Reporting: To automatically generate an email ab
From IP: 208.78.224.20

System information:

- The system is running a mail server (ESMTP Exim- This means that this system can be used to send e
- The system is running a web server on port 80 (cl This means that this system serves web pages.
- The system is running a secure web server (Apach (link here to view it). This means that this system

- Network Whois
- Domain Whois
- Email Header

For 24 hours only you can get up to 20% off eMailTrackerPro! [Click Here](#)

Activate Windows

Infoga

- Source: <https://github.com>
- Infoga is a tool used for gathering email account information (IP, hostname, country, etc.) from different public sources (search engines, pgp key servers, and Shodan), and it checks if an email was leaked using the haveibeenpwned.com API. For example, the command
- `python infoga.py --domain microsoft.com --source all --breach -v 2 --report ../microsoft.txt`
- will retrieve all the publicly available email addresses related to the domain microsoft.com along with email account information.
- `python infoga.py --info m4110k@protonmail.com --breach -v 3 --report ../m4110k.txt`
- The above command will retrieve email account information for a specified email address.

```
Parrot Terminal
File Edit View Search Terminal Help
~[root@parrot]-[~/infoga]
~#python infoga.py

==| Infoga - Email OSINT
==| Momo (m4ll0k) Outaadi
==| https://github.com/m4ll0k

Usage: infoga.py [OPTIONS]

  -d --domain          Target URL/Name
  -s --source          Source data, default "all":

                        all      Use all search engine
                        google   Use google search engine
                        bing     Use bing search engine
                        yahoo    Use yahoo search engine
                        ask      Use ask search engine
                        baidu    Use baidu search engine
                        dogpile  Use dogpile search engine
                        exalead  Use exalead search engine
                        pgp      Use pgp search engine

  -b --breach          Check if email breached
  -i --info            Get email informations
  -r --report          Simple file text report
  -v --verbose         Verbosity level (1,2 or 3)
  -H --help            Show this help and exit
```

5- Footprinting through Social Engineering on Social Networking Sites

Attackers use **social engineering tricks** to gather sensitive information from social networking websites

Attackers create a **fake profile** and then use the false identity to lure employees into revealing their sensitive information

Attackers collect information about the employees' **interests** and tricks them into revealing more information

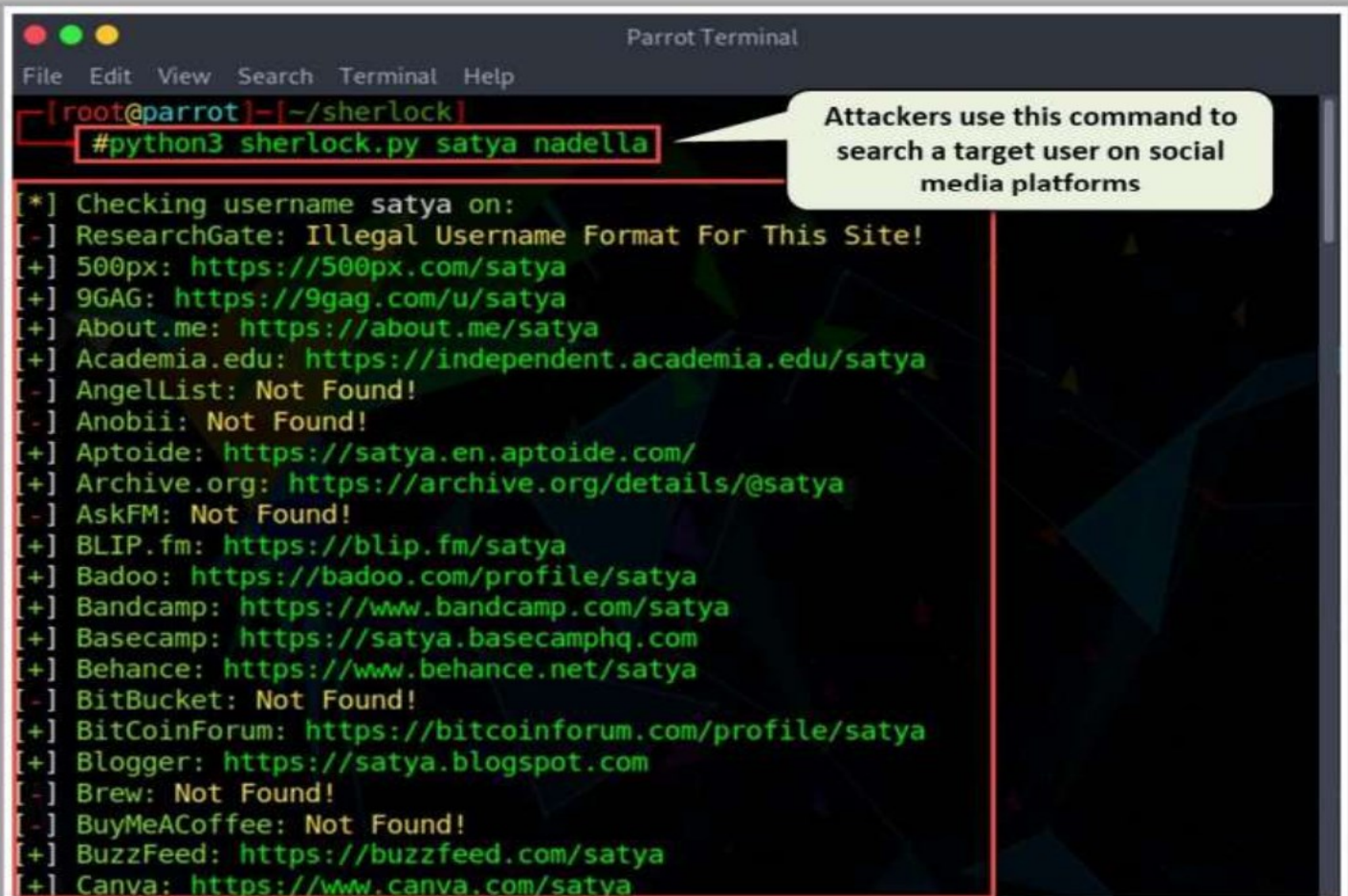
What Users Do	What Attacker Gets	What Organizations Do	What Attacker Gets
Maintain profile	Contact info, location, etc.	User surveys	Business strategies
Connect to friends, chat	Friends list, friends' info, etc.	Promote products	Product profile
Share photos and videos	Identity of family members, interests, etc.	User support	Social engineering
Play games, join groups	Interests	Recruitment	Platform/technology
Create events	Activities	Background check to hire employees	Type of business

Numerous online tools enable attackers to gather valuable information from social media sites, including popular content, account tracking, and email acquisition. This data is exploited for phishing and social engineering attacks. Tools like **BuzzSumo**, **Google Trends**, **Hashatit**, and **Ubersuggest** are commonly used for this purpose.

Attackers use various tools such as Sherlock, Social Searcher, and UserRecon to footprint social networking sites such as Twitter, Instagram, Facebook, and Pinterest to gather sensitive information about the target such as DOB, educational qualification, employment status, name of the relatives, and information about the organization that they are working for, including the b

- **Sherlock**

Source: <https://github.com> As shown in the screenshot, attackers use Sherlock to search a vast number of social networking sites for a target username. This tool helps the attacker to locate the target user on various social networking sites along with the complete URL.



```
Parrot Terminal
File Edit View Search Terminal Help
[ root@parrot ] - [ ~/sherlock ]
#python3 sherlock.py satya nadella

[*] Checking username satya on:
[-] ResearchGate: Illegal Username Format For This Site!
[+] 500px: https://500px.com/satya
[+] 9GAG: https://9gag.com/u/satya
[+] About.me: https://about.me/satya
[+] Academia.edu: https://independent.academia.edu/satya
[-] AngelList: Not Found!
[-] Anobii: Not Found!
[+] Aptoide: https://satya.en.aptoide.com/
[+] Archive.org: https://archive.org/details/@satya
[-] AskFM: Not Found!
[+] BLIP.fm: https://blip.fm/satya
[+] Badoo: https://badoo.com/profile/satya
[+] Bandcamp: https://www.bandcamp.com/satya
[+] Basecamp: https://satya.basecamphq.com
[+] Behance: https://www.behance.net/satya
[-] BitBucket: Not Found!
[+] BitCoinForum: https://bitcoinforum.com/profile/satya
[+] Blogger: https://satya.blogspot.com
[-] Brew: Not Found!
[-] BuyMeACoffee: Not Found!
[+] BuzzFeed: https://buzzfeed.com/satya
[+] Canva: https://www.canva.com/satya
```

Attackers use this command to search a target user on social media platforms

Thank
You

