Acceptance clause: by completing this project the student accepts that he/she is responsible for what he/she will do and that he/she will not use these tools outside the virtual one created for that purpose. The student also accepts the legal repercussions that some of these actions may be subject to if these actions are carried out on external sites.

Project 2: Information Gathering (footprinting)[1]

Level: 1

Difficulty: easy

Objective: Install and analyze the indicated tools obtaining information about worker1 (DO NOT USE outside the virtual network).

Task: The student will have to collect information about the tools, install them from the Debian repository and use it on worker1 to analyze its functionality and extract information from the worker1 VM.

Information Gathering is the act of gathering different kinds of information against the targeted victim or system. It is the first step or the beginning stage of Ethical Hacking, where the penetration testers or hackers (both black hat or white hat) performed this stage; this is a necessary and crucial step to be performed. The more the information gathered about the target, the more the probability to obtain relevant results. Information gathering is not just a phase of security testing; it is an art that every penetration-tester (pen-tester) and hacker should master for a better experience in penetration testing.

There are various tools, techniques, and websites, including public sources such as **whois, nslookup, dig** that can help hackers gather information. This step is necessary because you may need any information (such as his pet name, best friend's name, age, or phone number to perform password guessing attack or other kinds of attacks) while performing attacks on any target.

Information gathering can be classified into three major categories:

1. **Footprinting**
2. Scanning
3. Enumeration

Footprinting is the technique to collect as much information as possible about the targeted network/victim/system. It helps hackers in various ways to intrude on an organization's system. This technique also determines the security postures of the target. Footprinting can be active as well as passive. Passive footprinting/pseudonymous footprinting involves collecting data without the owner knowing that hackers gather their

1 Some parts are based on: https://www.w3schools.in/ethical-hacking

data. In contrast, active footprints are created when personal data gets released consciously and intentionally or by the owner's direct contact.
This type of footprinting is the safest, holding all legal limitations, and hackers can do it without fear because it is legal and, hence, coined the term Open-source. Examples of this type include: finding someone's email address and phone numbers, scanning IP through automated tools, searching for age, DOB, house address, etc. Most companies provide information about themselves on their official website without knowing that hackers can take advantage of it.
Using this footprinting category, hacktivists can retrieve information such as user name, information within a group, shared data among individuals, network services, etc.

After gathering the information from the different areas using various techniques, the hacker usually queries the DNS using pre-existing tools. Many freeware tools are available online to perform DNS interrogation.

Useful tools:

1. **whois** is a renowned Internet record listing tool to identify who owns a domain or who registers for that particular domain along with their contact details. The Internet Corporation for Assigned Names and Numbers (ICANN) regulates domain registration and ownership details. Whois records have proven extraordinarily beneficial and have developed into an essential resource for maintaining the integrity of domain name registration and website ownership process.
2. **harvester** is also an information-gathering tool that helps you extract a particular target's email address and subdomains. harvester is coded using a simple python script that searches information from giant search engines like Google, Yahoo, Bing, and much more (https://gitlab.com/kalilinux/packages/theharvester).
3. **metagoofil** is another information gathering or footprinting tool used for extracting information or data publicly available on the internet belonging to the company (https://github.com/laramies/metagoofil).
4. **netifera** (https://github.com/netifera/netifera) is a potent tool that gives a complete platform to gather information regarding the targeted website you want to attack . This software will give information such as IP address, the Programming language used for website development, the number of websites hosted, and DNS.
5. **OS Identification**: involves sending **illegal** TCP (Transmission Control Protocol) or ICMP (Internet Control Message Protocol) packets to the victim's system to identify the OS (Operating system) used by the victim on his server or computer. A ping sweep establishes a range of IP addresses that map hackers to live hosts. **fping**, **nmap**, **nmapsi4**, **hping3**, and **doscan/masscan/pnscan/nbtscan** are some of the tools used to ping a large number of IP addresses at a time; to generate lists of hosts for large subnets.
6. We can gather information from other sources, such as social networking sites (Facebook, Twitter, LinkedIn, etc.), where general users share their personal data and additional related information. Even search engines play a significant role in

gathering information.

7. **Kali distribution (https://www.kali.org/)**: Kali Linux is an open-source, Debian-based Linux distribution geared towards various information security tasks, such as Penetration Testing, Security Research, Computer Forensics and Reverse Engineering. This distribution has the indicated tools installed.

Type of information to get:

- Email (header); Address from which message was sent, Sender's email server, Sender's IP address, Date and time received by the originator's email server, The sender's mail server uses the authentication system, Sender's full name.
- Collect Network Information: such as Domain name, Internal domain names, IP addresses of the reachable systems, rogue websites/private websites within the domain, Access Control Mechanisms, protocols used, existing VPNs, analog and digital telephone numbers, authentication mechanisms, and system enumeration.
- Collect System Information: such as users and group names, system banners, routing tables, and the routing protocols it is using, SNMP information, system architecture, operating system used, remote system type, username, and passwords.
- Collect Organizations' Information: such as Employee details, organization's website, company directory, local details, address and phone numbers, comments in HTML Source code within an organization's website, security policies implemented, web server links relevant to the organization, news articles and press release.

Classify the type of information which is needed to be kept public. Don't put unnecessary information into any profile, social networking account, or website. Don't keep a personal contact number in any company or organization's phone book to prevent war-dialing. Keep internal DNS and external DNS separate. Restrict and disable zone transfer to authorized servers.

Countermeasures:

- For personal information: education and recommendations for not publish any type of information that may serve these purposes.
- For information systems: Firewalls (**iptables**).

References: Ethical Hacking, Cibersecurity, TutorialsPoint, MygreatLearning, Coursera, Portal-Offensive.

criminal charges be brought against any individuals misusing the information in this document/tools to break the law.

The examples, tools, tests or any other activity shown or suggested in this document should NEVER be carried out outside the private network created for this purpose, since criminal and/or punishable responsibility may be incurred.