# Ethical Hacking: Scanning Report

Jordi Nadeu Ferran

**Abstract**

In this laboratory, we focuses on several cybersecurity and open-source intelligence (OSINT) exercises, particularly those involving scanning techniques.

## 1 Introduction

Here are the main topics covered in this laboratory:

- **Link Analysis**: Explores link analysis for understanding connections between entities, and using the WHOIS database to investigate domain ownership and privacy implications.
- **Maltego for Link Analysis**: Instructions for setting up and using Maltego for visual link analysis, including transforms to discover related domain information and possibly leaked credentials.
- **Password Security**: Exercises on analyzing password strength, checking for compromised accounts, and discussing the ethics around password cracking and leaked databases.
- **Google Dorking**: Awareness exercises on Google Dorking, its use in refining search queries, and how to protect against it.
- **Masscan Vulnerability Scanner**: Steps for setting up and running Masscan for scanning networks, with attention to ethical and legal issues, including exclusion lists to avoid scanning sensitive IP ranges like FBI honeypots.
- **Nessus Vulnerability Scanner**: Instructions for downloading, installing, and using Nessus to perform vulnerability scans.
- **nmap Scanning Techniques**: Covers various nmap scanning methods such as HTTP enumeration, stealthy OS detection, and vulnerability detection using scripts.
- **Discover for OSINT and Vulnerability Scanning**: Instructions for installing and using the Discover framework for both passive and active scanning tools.

# 2 Procedure

## 2.1 Link Analysis Exercise

### 1. Understand link analysis

#### (a) Define link analysis and its purpose in information gathering

Link analysis is a method used to identify relationships and connections between various entities (e.g., individuals, organizations, domains) based on collected data. Its purpose is to reveal hidden connections, patterns, and networks that can provide deeper insights during investigations. It is widely used in cybersecurity, criminal investigations, OSINT (Open-Source Intelligence), and even social network analysis.

#### (b) Discuss examples of first-degree and second-degree connections

- First-degree connections: Direct relationships between two entities. For example, a domain owned by a particular individual.
- Second-degree connections: Connections that require an intermediate entity. For instance, if an A owns a domain and a B is linked to the same organization that owns that domain, then A and B are second-degree connections.

### 2. Explore the WHOIS database

Examining the output of the WHOIS query, there are a lot of information but some relevant are:

- Registrant: Name and organization of the entity that registered the domain.
- Contact Information: Email addresses, phone numbers, and physical addresses, etc...
- Domain Details: Dates when the domain was registered and when it will expire, name servers, and DNS information.

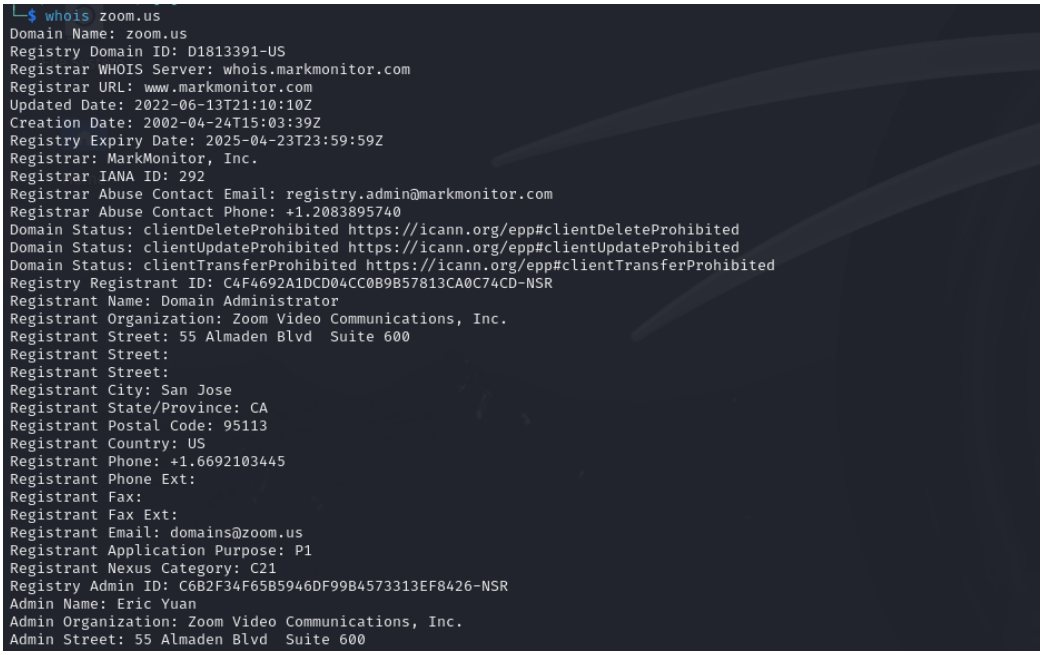


Figure 1: Output command 'whois zoom.us'

**3. Investigate domain privacy**

**(a) Research why some domains have private WHOIS information**

Domain privacy protection (WHOIS privacy) is a service offered by domain registrars to prevent the registrant's personal details (like email and address) from being publicly visible in WHOIS results. It shields personal information to avoid spam, scams, and unwanted solicitations.

**(b) Find out why .us domains are required to publish contact information**

managed The U.S. government manage .us domains, then for U.S. jurisdiction are required by U.S. law to publish accurate contact information as part of efforts to ensure transparency and accountability.

**4. Compare WHOIS results**

**(a) Look up WHOIS information for 3 different domains (.com, .org, .us)**

In the case of .us domain we can reuse the zoom.us like you can see in Figure 1. Then the .com can be google.com and the .org be wikipedia.org as shown in the following figures.


Figure 2: Output command 'whois google.com'

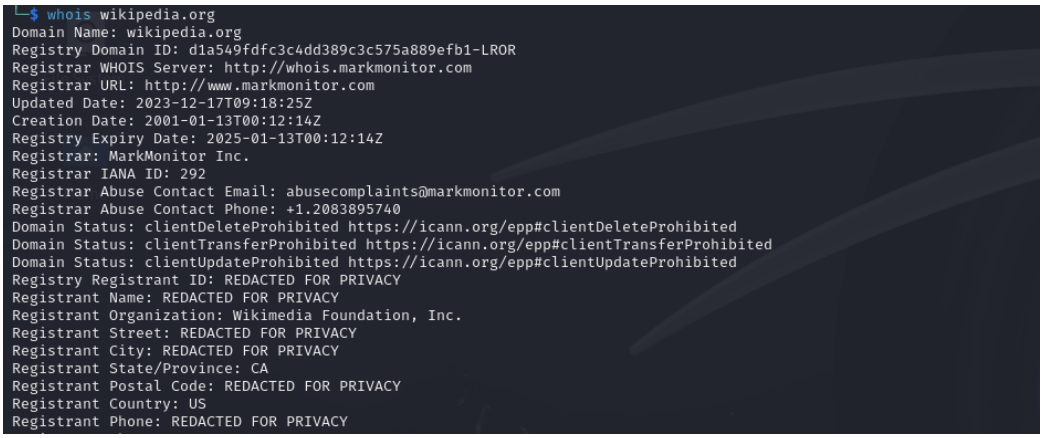


Figure 3: Output command 'whois wikipedia.org'

### (b) Note differences in available information

- zoom.us domain are legally required to publish full contact details, which makes the information more transparent.
- google.com domain are less information than .us domain
- wikipedia.org use domain privacy services, so you might see privacy-protected contact details.

## 5. Discuss ethical considerations

**(a) Consider the implications of publicly available contact information**

Publicly available contact information in WHOIS records can be useful for transparency, accountability, and legal purposes. However, it can also expose individuals or organizations to privacy risks, such as spam, phishing, or even identity theft

**(b) Debate the balance between transparency and privacy**

There's a delicate topic, cause the balance between ensuring transparency for legitimate purposes and respecting the privacy of individuals are too subjective. I suposo depends on the context and the use of the information.

## 2.2 Maltego Link Analysis Exercise

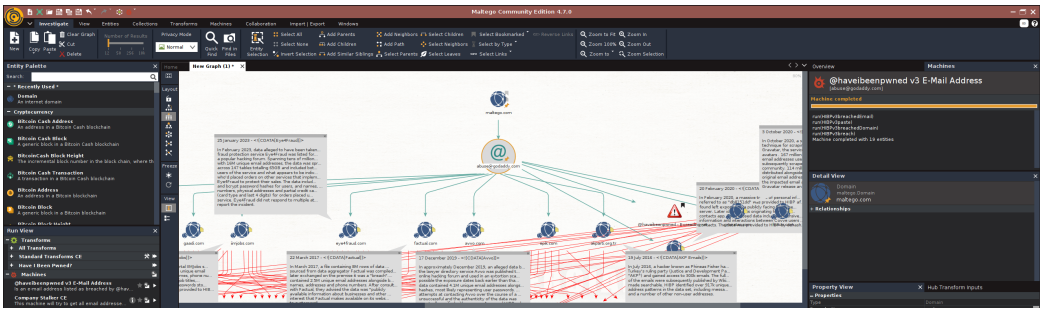


Figure 4: Maltego screenshot after steps done

**(a) What kind of information was easily discoverable?**

The results will often reveal infrastructure details like DNS servers and personal information, like emails.

**(b) How might this information be used by hackers or security researchers?**

- Hackers: Could use the information to craft targeted phishing campaigns, find vulnerabilities and/or gain unauthorized access.
- Security researchers: Could use it to find security flaws, assess exposure, and improve system defenses.

**(c) Consider the ethical implications of using such tools and databases**

Tools like Maltego provide access to vast amounts of information. Using these tools ethically means, have the permission before investigating, respecting privacy laws (like GDPR) and obviously not use the information for illegal activities.

**(d) Discuss the importance of unique passwords and regular password changes**

Using strong, unique passwords for each service and then regularly updating passwords to mitigate risks, especially if data breaches occurs. I think are important things to do if we want to improve the security of the logins.

## 2.3 Password Security Exercise

Large-scale password leaks, like the breach involving over 1.4 billion credentials, showcase the scope of global cybersecurity threats. Databases containing email addresses, usernames, and passwords are often leaked or stolen from compromised systems and circulated on the dark web or other forums. There are potential consequences like identity theft or reputation damage for example.

Understanding whether your accounts have been compromised is critical for taking proactive steps, like changing passwords immediately, enabling two-factor authentication (2FA) or avoid password reuse. Knowing that your data has been breached also encourages better cybersecurity habits and awareness.

A weak password is when use simple words or phrases (e.g., "123456" or "password") that are easy to guess or vulnerable to brute-force attacks. And a strong passwords is when are more complex combinations of upper and lowercase letters, numbers, and special characters (e.g., T8hKf12!).

Some widely used password managers are LastPass, 1Password (I used it in the past) that are privates and Bitwarden which is my current password manager and is open-source. Also ProtonPass are well but is less popular than the previous ones.

Some of the benefits of use a password manager are you can store strong, unique passwords for each service with are encrypted, which can only be unlocked with a master password. And also you can generating strong, random passwords instead of making up your own password.

The main risks are single point of failure, if the master password is compromised, all saved passwords could be exposed. And also another point is the cloud synchronization risks, when the password managers store data in the cloud, introducing potential vulnerabilities.

My personal plan for password security are mainly this following points:

- Use a password manager to store and generate unique, strong passwords.
- Enable MFA on all accounts that support it.
- Check for compromised accounts regularly via services like Have I Been Pwned.
- Change weak or reused passwords immediately.
- Regularly update passwords for critical accounts like email, banking, or social media.

For the ethical implications if you are a security researchers, sometimes use password cracking techniques to test the strength of systems. When done ethically if are with permission, it would helps improve cybersecurity. But acessing leaked password databases without permission or using them for malicious purposes is unethical and illegal.

There's a fine balance between ethical security research, which helps organizations protect against real threats and the privacy credentials expose sensitive personal information, raising concerns over the use of these databases in research. I think organizations should practice responsible disclosure and only engage in research that benefits security without compromising individual privacy.

## 2.4 Google Dorking

Google Dorking involves using advanced search operators in Google to find information that is not easily accessible through typical searches. It leverages specific search queries to uncover sensitive information (e.g., login pages, personal data, vulnerable files) that has been inadvertently made publicly accessible on websites. However, it can also be misused for illegal activities, which raises significant ethical and legal concerns.

While Google Dorking itself is not illegal, using it to exploit vulnerabilities can violate laws like the Computer Fraud and Abuse Act (CFAA), which prohibits unauthorized access to computer systems.

If examine the structure of the following queries, we can say:

- **inurl:"live/cam.html"**: This query targets URLs containing live/cam.html, which may reveal unsecured live camera feeds.

- **"Pop-up" + "Live Image" inurl:index.html**: This looks for pages with "Pop-up" and "Live Image" in the text while restricting the URL to those that contain index.html, possibly identifying pages with live images.

- **filetype:log intext:password after:2019 intext:@gmail.com — @yahoo.com**: This query seeks log files containing the word "password" that were posted after 2019 and include either Gmail or Yahoo email addresses.

The robots.txt file instructs search engines on which pages or directories to ignore during indexing. While it can help prevent accidental exposure of sensitive files to search engines. Tt is not a security mechanism because only tells search engines what to avoid; it does not restrict access. Sensitive data and pages should be protected by proper authentication (e.g., passwords, two-factor authentication).

A list of best practices to protect against Google Dorking:

- Review robots.txt regularly ensure no sensitive data is listed or accessible.

- Conduct regular security audits, identifing and patching vulnerabilities in your websites configurations.

- Monitor publicly accessible files ensure sensitive data (logs, databases) is not exposed.

- Use proper authentication with passwords and MFA.

- Disable directory listings preventing search engines from indexing directories that contain sensitive information.

## 2.5 Internet Scanning Exercise

Example command to scan port 80 on all IP addresses while using the exclusion file:

```
sudo masscan 0.0.0.0/0 -p80 --rate=100000 --excludefile exclude.txt
```

Explain what each part of the command does:

- sudo: Runs the command with superuser privileges, required for network scanning.

- masscan: The tool being used to perform the scan.
- 0.0.0.0/0: Indicates the target range, which includes all IPv4 addresses.
- -p80: Specifies that the scan should target port 80 (HTTP).
- –rate=100000: Sets the packet transmission rate to 100.000 packets per second.
- –excludefile exclude.txt: Specifies the file that contains IP ranges to exclude from the scan.

The exclusion list helps avoid scanning sensitive IP addresses to mitigate potential legal issues and prevent disruption to critical services. Assuming that there are approximately 4.3 billion IPv4 addresses and scanning 100,000 packets per second, the time will be approximately 12 hours.

List potential legal issues with scanning the entire internet:

- **Unauthorized Access**: Scan networks without consent can lead to claims of unauthorized access.
- **CFAA Violations**: Scans can violate the Computer Fraud and Abuse Act if deemed unauthorized, leading to legal penalties.
- **ISP Blocking**: Aggressive scanning trigger defensive measures from ISPs or networks, leading to IP blacklisting.

# 3 Conclusions

Definitely, these exercises illustrate the interconnectivity of various cybersecurity practices, from link analysis to a lot of network scanning techniques and also the ethical considerations of all this. The reflection I come to is the cybersecurity threats continue to evolve and the adoption of responsible practices, ongoing education, and the use of effective tools will be essential for protecting sensitive information and ensuring robust defenses against malicious actors. I think this fact is more important every day and is increasing.

If one day I dedicate myself to cybersecurity or something related to this field, I will strive to maintain ethical standards, stay informed about potential vulnerabilities, and implement proactive security measures to safeguard the integrity of our networks and systems.