Module III

Information Gathering Techniques

Information Gathering Techniques

“The more information you have about the target, the more is the chance of successful exploitation.”

Information gathering is the **first phase of hacking**.

We gather as much information as possible regarding the **target’s online presence**, which in turn r**eveal useful informati**on about the target itself.

The required information will depend on whether we are doing a **network pentest or a web application pentes**t.

In network pentest, our **main goal** would be to **gather information on the network.**

information gathering techniques can be **classified into two main categories:**

1. **Active information gathering**
2. **Passive information gathering**

Active Information Gathering

we would directly engage with the target

example,

gathering information about what ports are open on a particular target,

what services they are running?

what operating system they are using?

The techniques involving active information gathering would be very noisy at the other end. As they are easily detected by **IDS, IPS, and firewalls and generate a log of their presence**, and hence are not recommended sometimes

Passive Information Gathering

we do not directly engage with the target

Instead, we use search engines, social media, and other websites to gather information about the target.

it does not generate any log of presence on the target system

Example

would be to **use LinkedIn, Facebook, and other social networks to gather information about the employees and their interests**

Sources of Information Gathering

There are many sources of information; the most important ones are as follows:

Social media website

Search engines

Forums

Press releases

People search

Job sites

So let's discuss some of these sources in detail along with some tools of the trade

Information Gathering Tools

Assessment is about gathering the necessary information about who you are as a literacy agency, your current strengths, weakness, opportunities and threats. Remember, at this point, you are simply gathering information on the current internal and external environment facing your literacy agency; you are not making any judgments or drawing any conclusions about what the information means for the future.

There are many different methods of information gathering that people have used to good advantage and here are a few:

1. Questionnaires, surveys and checklists
   - Used when you want to collect a lot of information from people in a non-threatening way.
2. Personal interviews
   - Used when you want to fully understand a person's opinions or point of view or to get additional information to a questionnaire.
3. Documentation review
   - Used when you want to gather information on current practices without interrupting the program by examining program monitoring reports, program statistics, learner progress reports, annual reports, performance appraisals, board evaluations, written policies and procedures, memos, minutes, financial records, etc.
4. Observation
   - Used to watch the program in operation to gather information about what actually happens day-to-day.
5. Focus group
   - Used to explore a topic in depth with key stakeholders to learn what the common understanding is on various issues.
6. Case Studies
   - Used to depict experiences, processes or practices with a view to developing understanding through examination and cross comparisons.

Copying Websites Locally

There are many tools that can be used to copy websites locally

however, one of the **most comprehensive tool is httrack**.

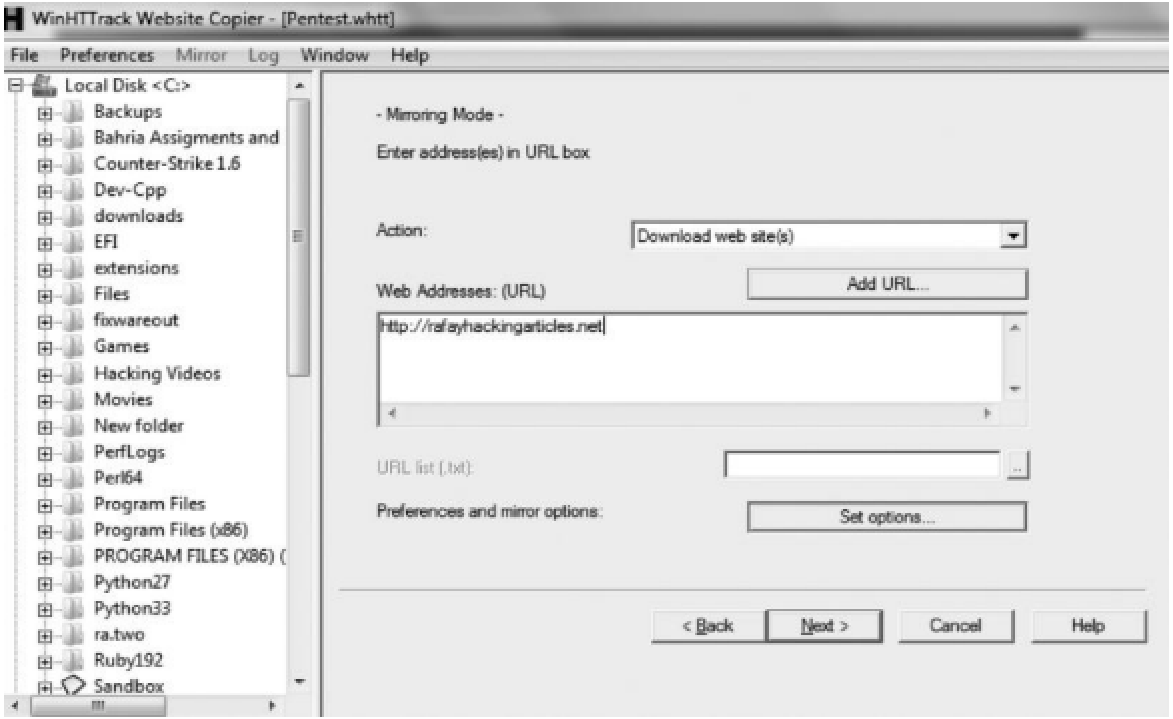
It can be used to investigate the website further.

example,

let’s suppose that the **file permissions of a configuration file are not set properly**.

The **configuration might reveal some important information**

example, **username and password, about the target**



- Mirroring Mode -

Enter address(es) in URL box

Action: Download web site(s)

Web Addresses: (URL)

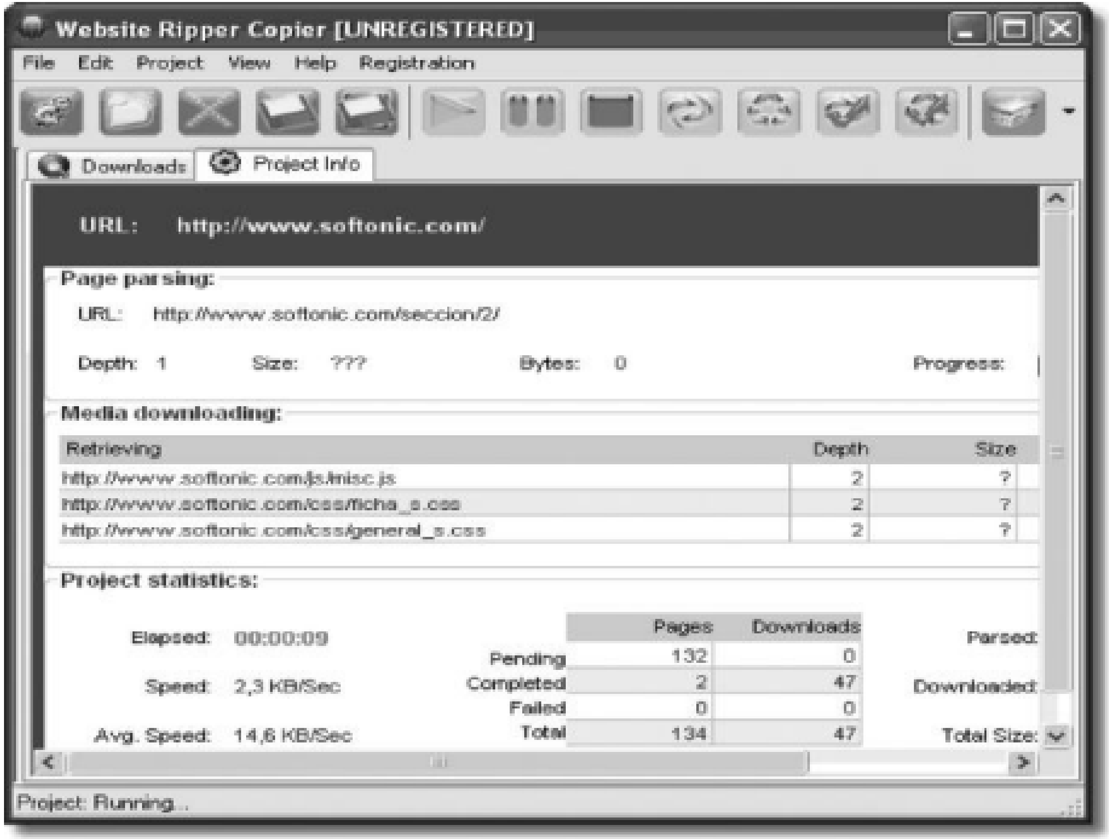
http://rafayhackingarticles.net

URL list (.txt):

Preferences and mirror options:

If you are on Linux, you can use Wget command to copy a webpage locally. Wget http:// www.rafayhackingarticles.net

Another great tool is Website **Ripper Copier,** which has a few additional functions than httrack.



URL: http://www.softonic.com/

Page parsing:

URL: http://www.softonic.com/seccion/2/

Depth: 1 Size: ??? Bytes: 0 Progress:

Media downloading:

Retrieving	Depth	Size
http://www.softonic.com/js/misc.js	2	?
http://www.softonic.com/css/ficha_s.css	2	?
http://www.softonic.com/css/general_s.css	2	?

Project statistics:

Elapsed: 00:00:09

Speed: 2,3 KB/Sec

Avg. Speed: 14,6 KB/Sec

	Pages	Downloads
Pending	132	0
Completed	2	47
Failed	0	0
Total	134	47

Parsed:

Downloaded:

Total Size:

Project: Running...

1. **Information Gathering with Whois** - our goal in the information gathering and enumeration phase is to gather as much information as possible about the target.

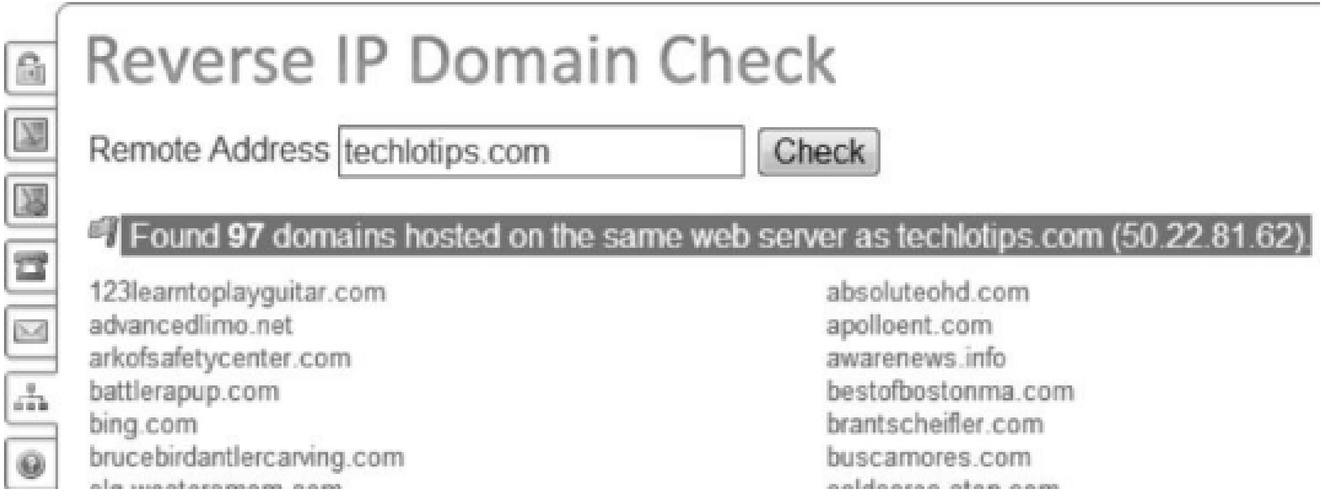
Whois holds a huge database that contains information regarding almost every website that is on the web, most common information are **“who owns the website”** and **“the e-mail of the owner,”** which can be used to perform social engineering attacks.

2.**Finding Other Websites Hosted on the Same Server**

which will show you exactly how an attacker can use a single website in order to compromise every website on the same server

Yougetsignal.com

Yougetsignal.com allows you to perform a reverse IP lookup on a webserver to detect all other websites present on the same server. All you need to do is enter the domain.



Reverse IP Domain Check

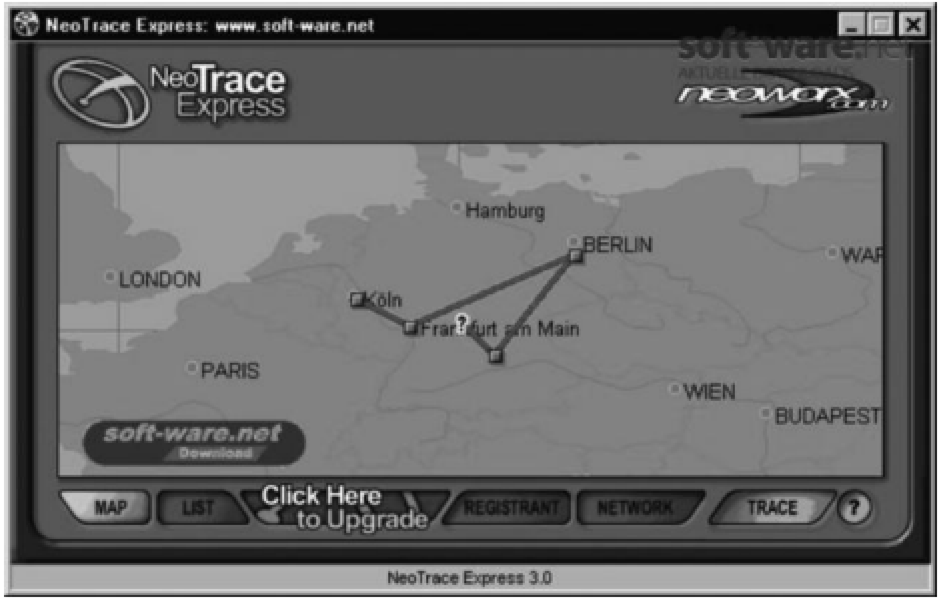
Remote Address: techlotips.com

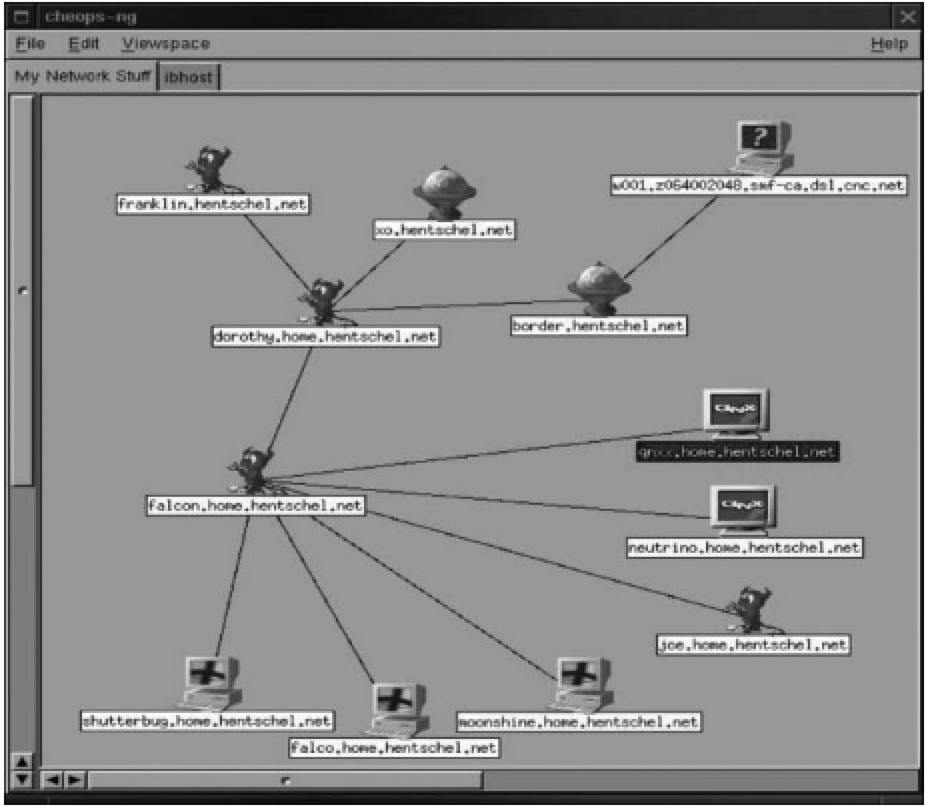
Found **97** domains hosted on the same web server as techlotips.com (50.22.81.62).

```
123learntoplayguitar.com          absoluteohd.com
advancedlimo.net                  apolloent.com
arkofsafetycenter.com             awarenews.info
battlerapup.com                   bestofbostonma.com
bing.com                          brantscheifler.com
brucebirdantlercarving.com        buscamores.com
```

NeoTrace:

NeoTrace is a very fine GUI-based tool for mapping out a network.




cheops-ng
File
Edit
Viewspace
Help
My Network Stuff
ibhost
franklin.hentschel.net
xo.hentschel.net
w001.z064002048.swf-ca.dsl.cnc.net
dorothy.home.hentschel.net
border.hentschel.net
gnxx.home.hentschel.net
falcon.home.hentschel.net
neutrino.home.hentschel.net
joe.home.hentschel.net
shutterbug.home.hentschel.net
falco.home.hentschel.net
moonshine.home.hentschel.net

Enumerating and Fingerprinting the Webservers

For successful target enumeration,

it's necessary for us to figure out what webserver is running at the back end.

We will look at both active and passive information gathering methods.

As a reminder, in **active information gathering, we directly interact with the target;**

in **passive information gathering, we do not interact with the target**, but use the information available on the web in order to obtain details about the target.

Netcraft

Netcraft contains a huge online database with useful information on websites and **can be used for passive reconnaissance against the target**. It is als**o capable of fingerprinting the webserv**ers.

What's that site running? | Netcraft

Google Hacking

Google searches can be more than a treasure for a pentester, if he uses them effectively.

With Google searches, an attacker may be able to gather some very interesting information, including passwords, on the target.

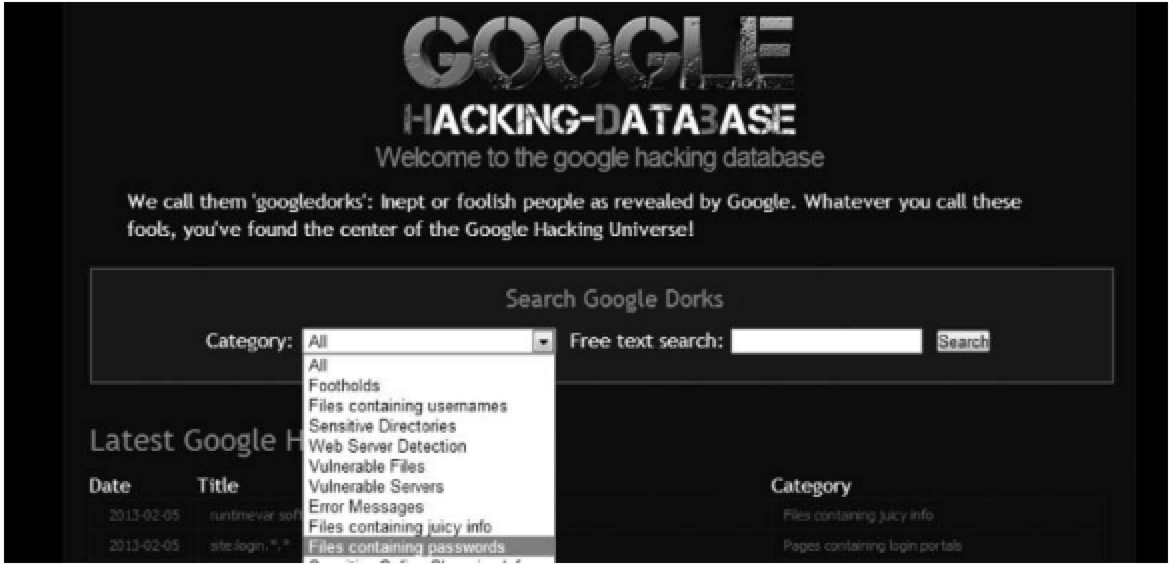
Google has developed a few search parameters in order to improve targeted search.

However, they are abused by hackers to search for sensitive information via Google.

Google Hacking Database

Google hacking database is set up by the offensive security guys, the ones behind the famous BackTrack distro.

Google hacking database has a list of many Google dorks that could be used to find usernames, passwords, e-mail list, password hashes, and other important information.



We call them 'googledorks': Inept or foolish people as revealed by Google. Whatever you call these fools, you've found the center of the Google Hacking Universe!

Category: All

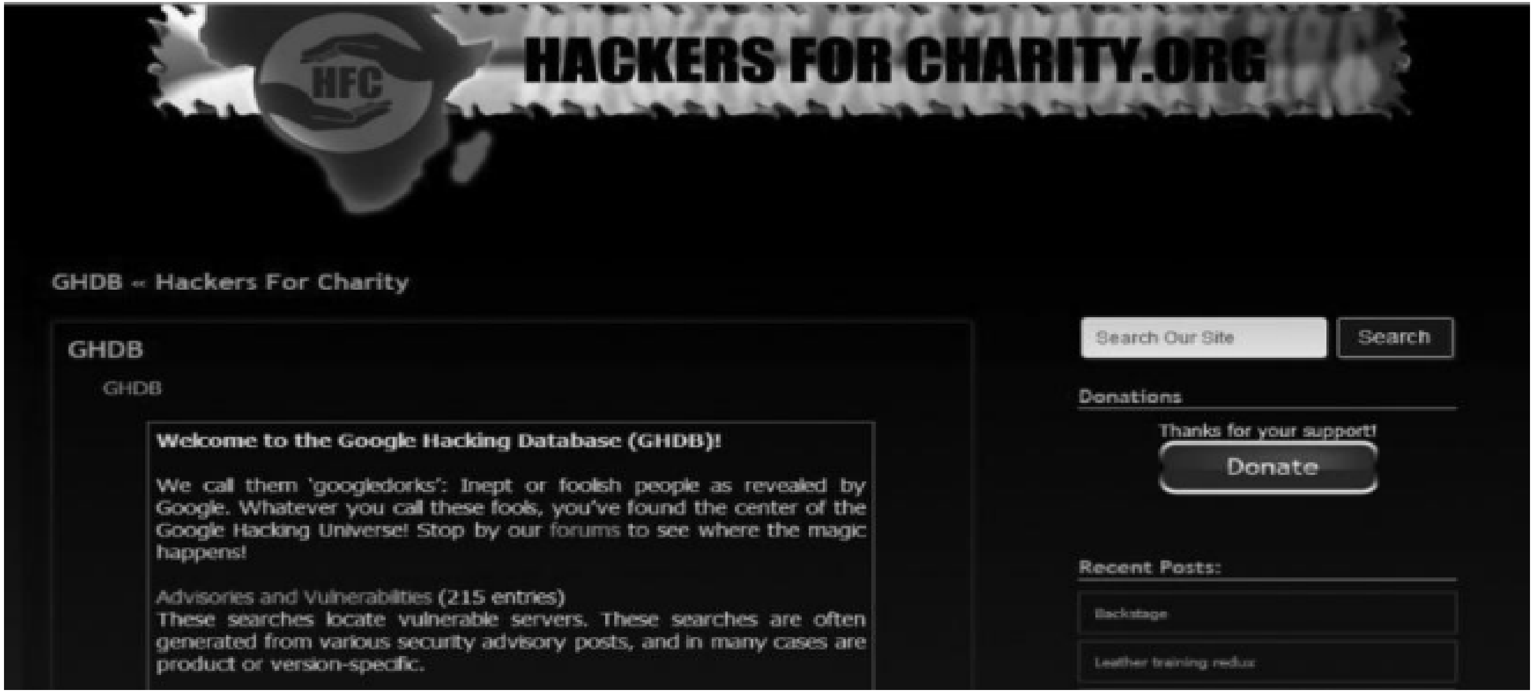
Free text search:

Latest Google H

Date	Title	Category
2013-02-05	runtmevar soft	Files containing juicy info
2013-02-05	site:login.*.*	Pages containing login portals

Hackersforcharity.org/ghdb

Another database that contains a collection of some interesting Google dorks.



GHDB « Hackers For Charity

GHDB

GHDB

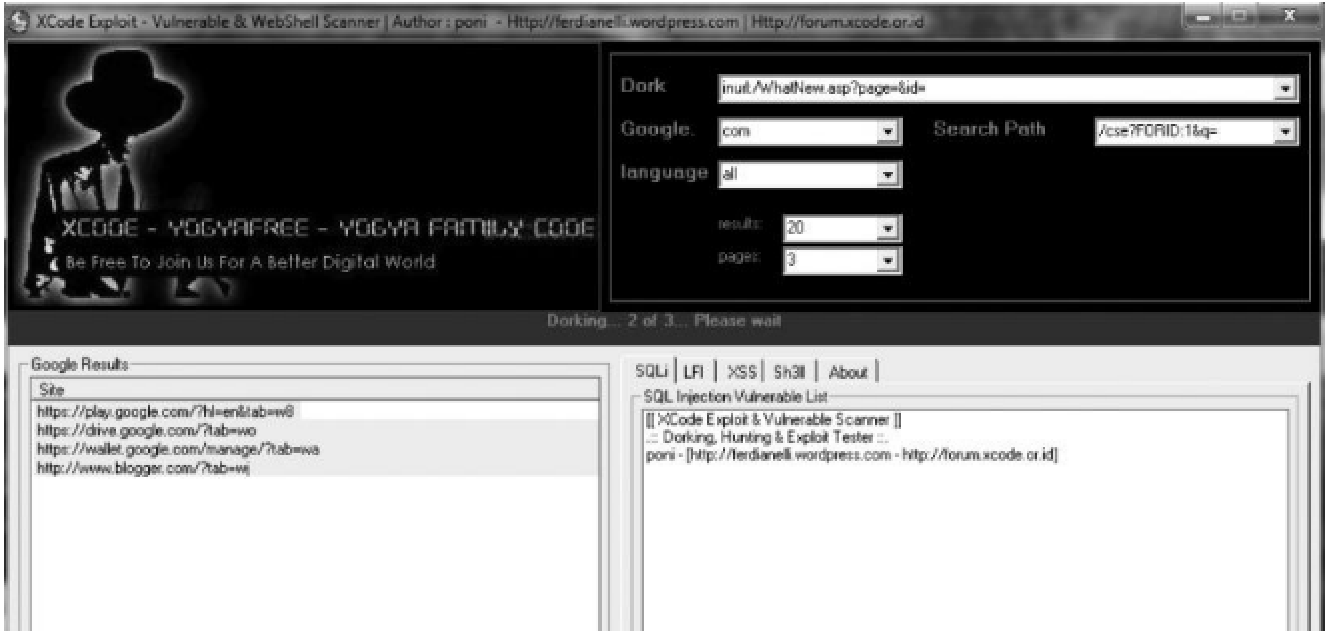
Welcome to the Google Hacking Database (GHDB)!

We call them 'googledorks': Inept or foolish people as revealed by Google. Whatever you call these fools, you've found the center of the Google Hacking Universe! Stop by our forums to see where the magic happens!

Advisories and Vulnerabilities (215 entries)
These searches locate vulnerable servers. These searches are often generated from various security advisory posts, and in many cases are product or version-specific.

Xcode Exploit Scanner

Xcode exploit scanner is an automated tool that uses some common Google dorks to scan for vulnerabilities such as SQLI and XSS.



XCode Exploit - Vulnerable & WebShell Scanner | Author : poni - Http://ferdianelli.wordpress.com | Http://forum.xcode.or.id

XCODE - YOGYAFREE - YOGYA FAMILY CODE

Be Free To Join Us For A Better Digital World

Dork: inurl:/WhatNew.asp?page=&id=

Google.: com

Search Path: /cse?FORID:1&q=

language: all

results: 20

pages: 3

Dorking... 2 of 3... Please wait

Google Results

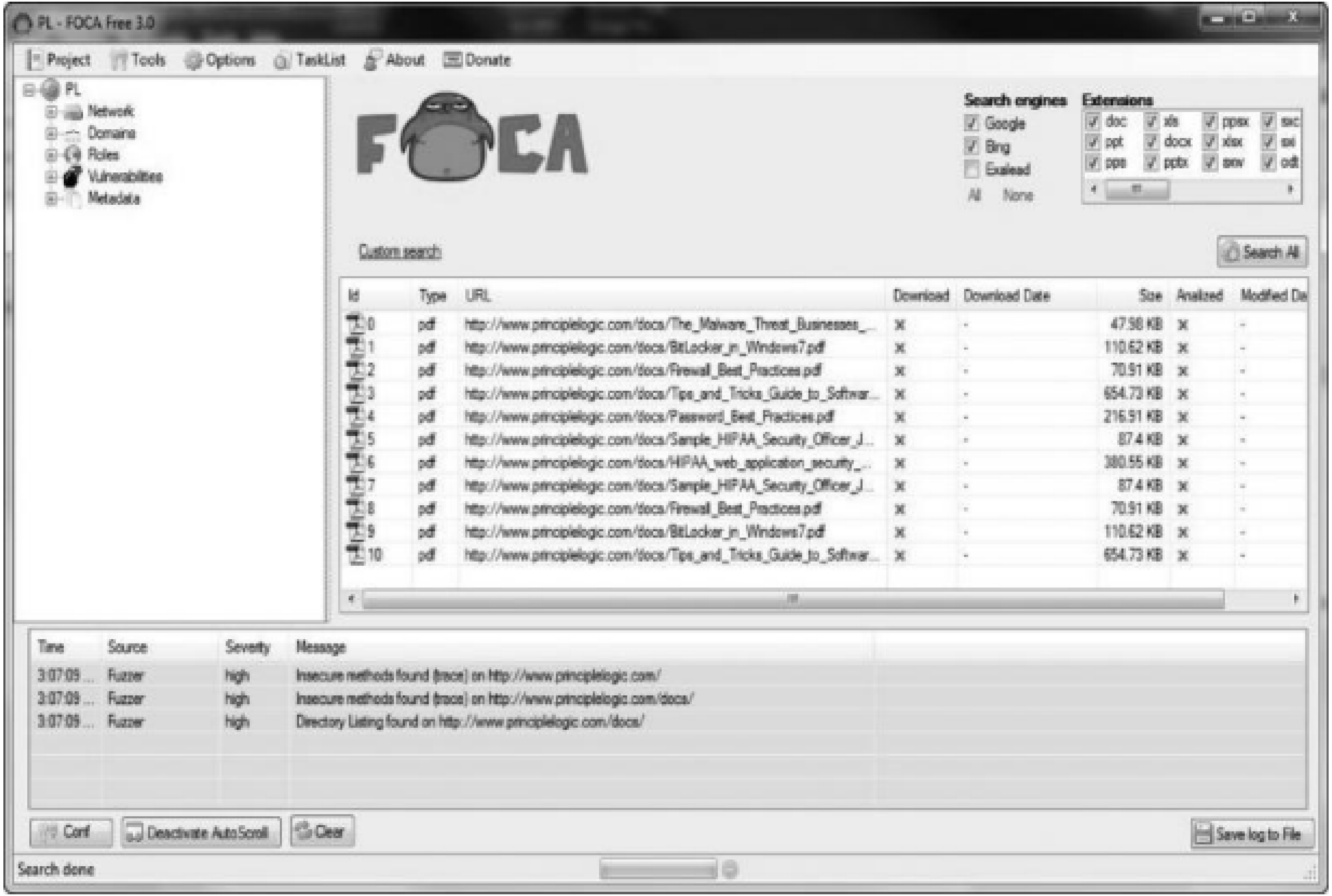
Site
https://play.google.com/?hl=en&tab=w8
https://drive.google.com/?tab=wo
https://wallet.google.com/manage/?tab=wa
http://www.blogger.com/?tab=wj

SQL Injection Vulnerable List

```
[[ XCode Exploit & Vulnerable Scanner ]]
.:: Dorking, Hunting & Exploit Tester ::.
poni - [http://ferdianelli.wordpress.com - http://forum.xcode.or.id]
```

Foca

Foca is a very effective tool that is **capable of analyzing files without downloading** them. It can search a wide variety of extensions from all the three big search engines (Google, Yahoo, and Bing). It's also capable of finding some vulnerabilities such as directory listing and DNS cache snooping.



Search engines
☑ Google
☑ Bing
□ Exalead

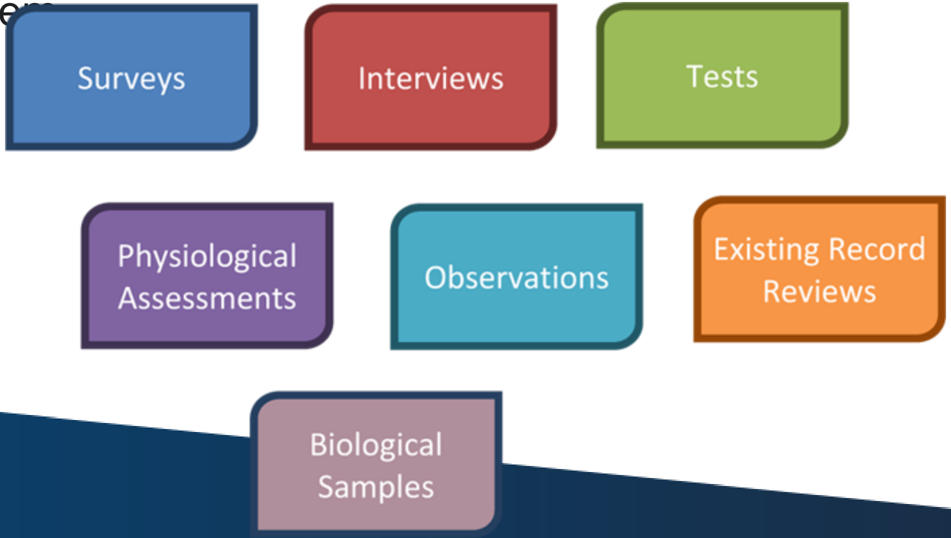
Extensions
☑ doc ☑ xls ☑ ppsx ☑ sxc
☑ ppt ☑ docx ☑ xlsx ☑ sxi
☑ pps ☑ pptx ☑ sxw ☑ odt

Id	Type	URL	Download	Download Date	Size	Analized	Modified Da
0	pdf	http://www.principlelogic.com/docs/The_Malware_Threat_Businesses_...	✕	-	47.58 KB	✕	-
1	pdf	http://www.principlelogic.com/docs/BitLocker_in_Windows7.pdf	✕	-	110.62 KB	✕	-
2	pdf	http://www.principlelogic.com/docs/Firewall_Best_Practices.pdf	✕	-	70.91 KB	✕	-
3	pdf	http://www.principlelogic.com/docs/Tips_and_Tricks_Guide_to_Softwar...	✕	-	654.73 KB	✕	-
4	pdf	http://www.principlelogic.com/docs/Password_Best_Practices.pdf	✕	-	216.91 KB	✕	-
5	pdf	http://www.principlelogic.com/docs/Sample_HIPAA_Security_Officer_J...	✕	-	87.4 KB	✕	-
6	pdf	http://www.principlelogic.com/docs/HIPAA_web_application_security_...	✕	-	380.55 KB	✕	-
7	pdf	http://www.principlelogic.com/docs/Sample_HIPAA_Security_Officer_J...	✕	-	87.4 KB	✕	-
8	pdf	http://www.principlelogic.com/docs/Firewall_Best_Practices.pdf	✕	-	70.91 KB	✕	-
9	pdf	http://www.principlelogic.com/docs/BitLocker_in_Windows7.pdf	✕	-	110.62 KB	✕	-
10	pdf	http://www.principlelogic.com/docs/Tips_and_Tricks_Guide_to_Softwar...	✕	-	654.73 KB	✕	-

Time	Source	Severity	Message
3:07:09 ...	Fuzzer	high	Insecure methods found (trace) on http://www.principlelogic.com/
3:07:09 ...	Fuzzer	high	Insecure methods found (trace) on http://www.principlelogic.com/docs/
3:07:09 ...	Fuzzer	high	Directory Listing found on http://www.principlelogic.com/docs/

Methods of Gathering Information

1. Conduct interviews.
2. Identify and study statistics.
3. Send questionnaires out to employees, customers, or other people concerned with the problem.
4. Conduct technical experiments.
5. Observe the procedures or processes in question first hand.
6. Create focus groups to discuss the proble



Surveys

Interviews

Tests

Physiological Assessments

Observations

Existing Record Reviews

Biological Samples

Harvesting E-Mail Lists

Gathering information about e-mails of employees of an organization can give us a very broad attack vector against the target.

This method can be classified under passive reconnaissance since we are not engaging with the target in any way, but would be using search engines to gather a list of e-mails.

These e-mail lists and usernames could be used later for social engineering attacks and other brute force attacks.

Gathering Wordlist from a Target Website

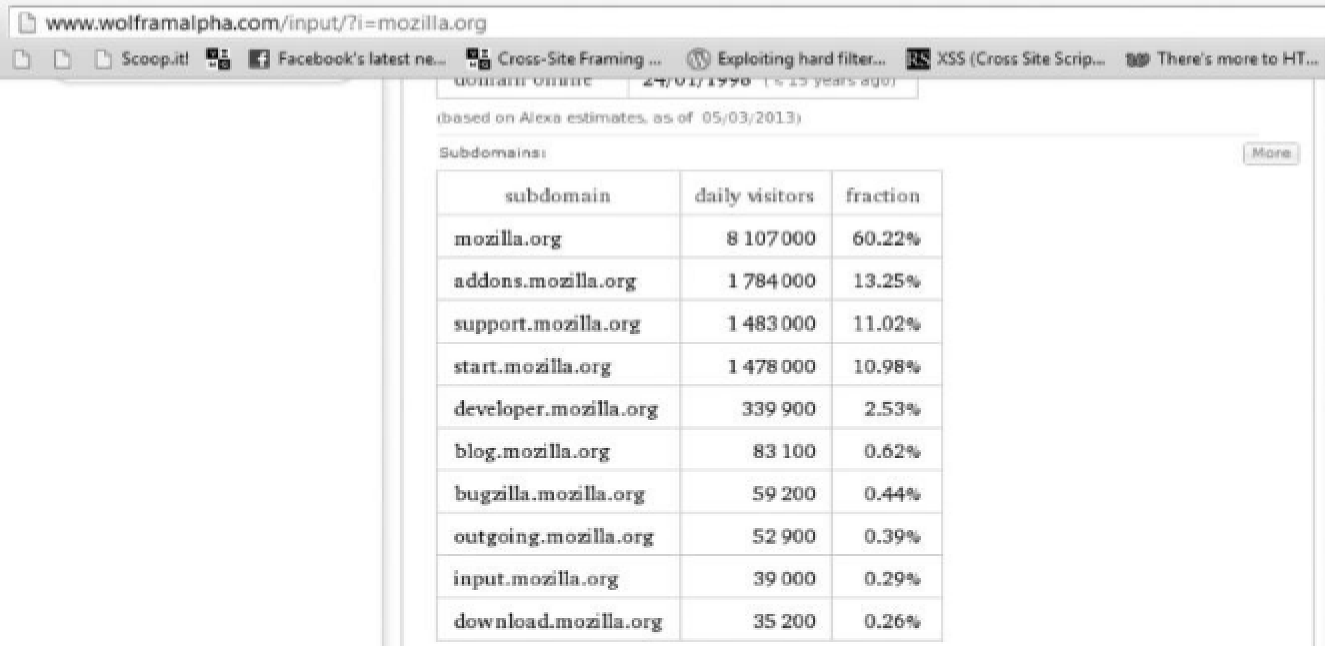
Scanning for Subdomains

TheHarvester - TheHarvester can also be used for this task, which uses Google to search for subdomains

Example

Knock.py - is a tool that has capabilities similar to fierce for determining subdomains

Wolframaplha - The following **website** also gives a **decent amount of subdomains**. It returns the most important subdomains that get the most traffic. If you want to save time, you can try wolframaplha.



(based on Alexa estimates, as of 05/03/2013)

Subdomains:

subdomain	daily visitors	fraction
mozilla.org	8 107 000	60.22%
addons.mozilla.org	1 784 000	13.25%
support.mozilla.org	1 483 000	11.02%
start.mozilla.org	1 478 000	10.98%
developer.mozilla.org	339 900	2.53%
blog.mozilla.org	83 100	0.62%
bugzilla.mozilla.org	59 200	0.44%
outgoing.mozilla.org	52 900	0.39%
input.mozilla.org	39 000	0.29%
download.mozilla.org	35 200	0.26%

Scanning for SSL Version

SSL stands for secure socket layer. It is used for encrypting communication. Since **an attacker on the local network could easily sniff the traffic**, most highly sensitive communications such as “login pages” use https (Port 443).

SNMP

SNMP stands for **Simple Network Mapping Protocol**

it is **widely used for the purpose of management and remote configurations of the devices**.

SNMP runs on UDP port 161. It has three versions: SNMP V1, SNMP V2, and SNMP V3

Problem with SNMP

1 was **developed in 1980.**

The **problem with this protocol was that there was no authentication system of any kind**, so anyone could access the SNMP server and gain access to the details present on it, as at that time, they did not consider securing it. Later, they developed SNMP and added some security features.

However, SNMP V2 was not backward compatible, the reason it was not widely adopted. Therefore, SNMP V3 was developed to become backward compatible with SNMP V1 and also to reduce the complexity of implementation. In an SNMP protocol, there are two types of community strings: a public community string and a private community string.

Sniffing SNMP Passwords

Most of the times, the SNMP passwords would be unencrypted if the devices are on SNMP V1.

An attacker can simply set up a sniffer to intercept the traffic on the network.

SMTP Enumeration

SMTP stands for Simple Mail Transfer Protocol. Sometimes, this could be a very useful source of information.

Knowing the valid usernames that exist would aid us immensely when brute-forcing them.

Module IV

Target Enumeration and Port Scanning Techniques

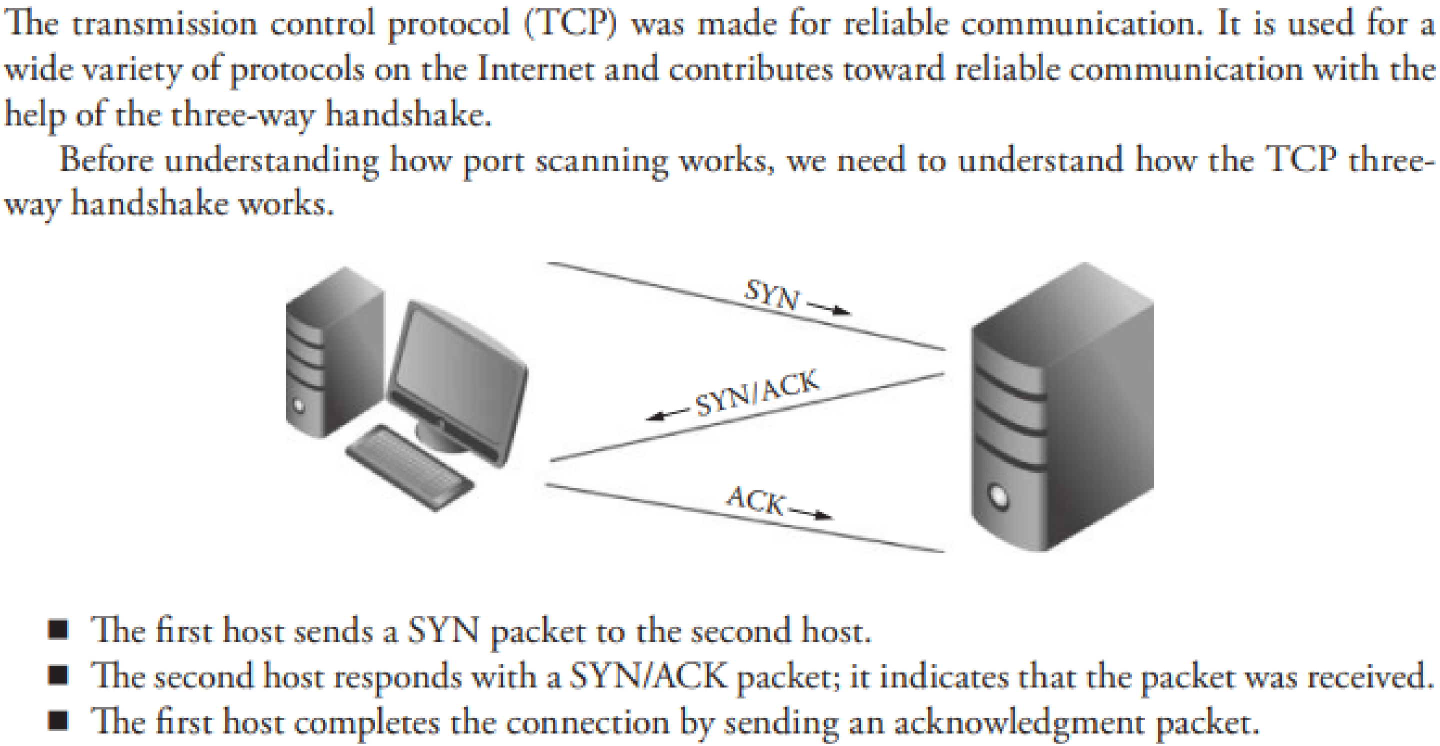
main goal of this chapter is to learn the following:

- Host discovery
- Scanning for open ports
- Service and version detection
- OS detection
- Bypassing firewalls

OWZAP 2.12.0

The **Open Web Application Security Project** (OWASP) is a nonprofit foundation dedicated to improving software security.

Understanding the TCP Three-Way Handshake



The transmission control protocol (TCP) was made for reliable communication. It is used for a wide variety of protocols on the Internet and contributes toward reliable communication with the help of the three-way handshake.

Before understanding how port scanning works, we need to understand how the TCP three-way handshake works.

SYN →
← SYN/ACK
ACK →

- The first host sends a SYN packet to the second host.
- The second host responds with a SYN/ACK packet; it indicates that the packet was received.
- The first host completes the connection by sending an acknowledgment packet.

TCP Flags

SYN—Initiates a connection.

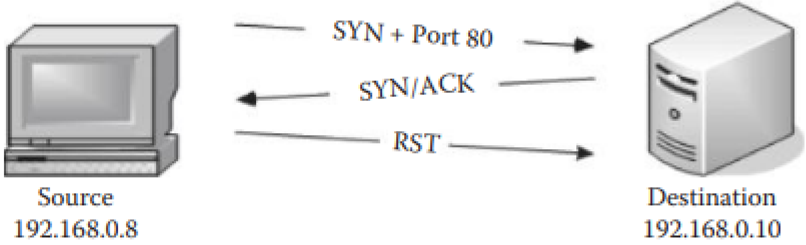
ACK—Acknowledges that the packet was received.

RST—Resets the connections between two hosts.

FIN—Finishes the connection

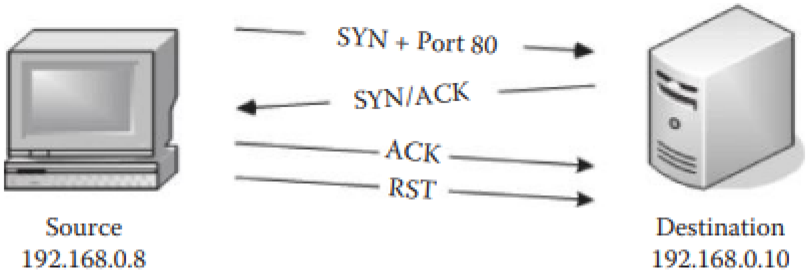
TCP SYN Scan

The **TCP SYN scan is the default scan that runs against the target machine.** It is the fastest scan. You can tweak it to make it even faster by using the –n option, which would tell the nmap to skip the DNS resolution

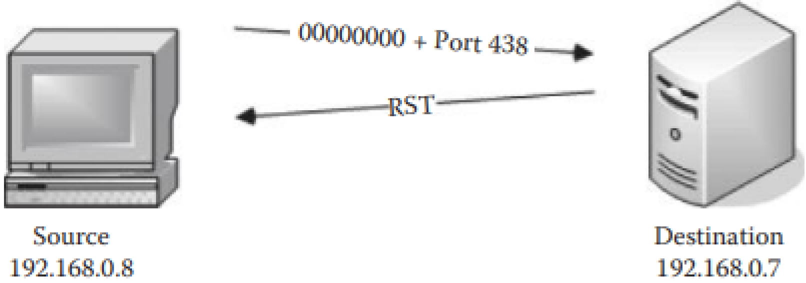


TCP Connect Scan

The TCP connect scan is similar to the SYN scan, with a slight difference in that it completes the three-way handshake. The TCP connect scan becomes the default scan if the SYN scan is not supported by the machine. A common reason for that could be that the machine is not privileged to create its own RAW packet

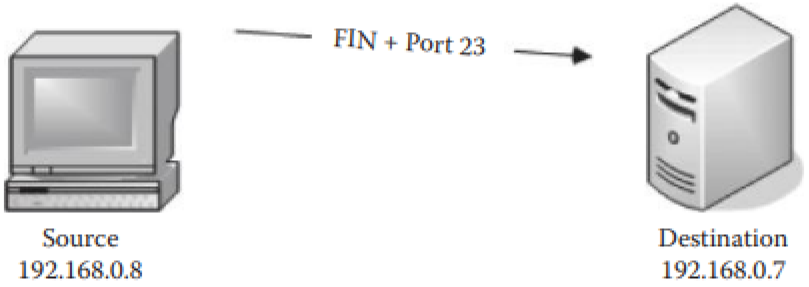


NULL Scan



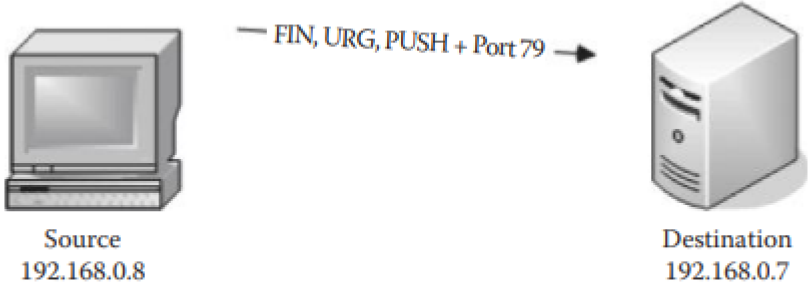
A null scan is accomplished by sending no flags/bits inside the TCP header. If no response comes, it means that the port is *open*; if a *RST* packet is received, it means that the port is *closed* or *filtered*.

FIN Scan



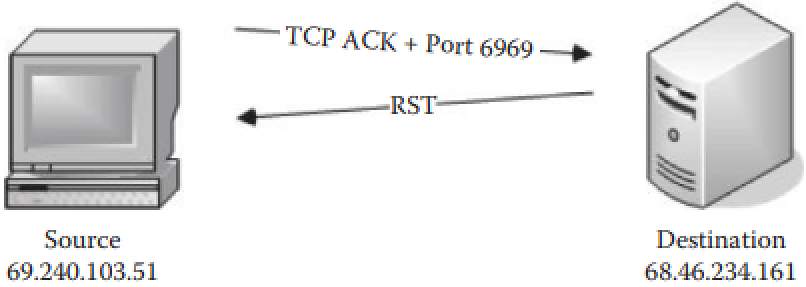
A FIN flag is used to close a currently open session. In a FIN scan the sender sends a FIN flag to the target machine: if no response comes from the target machine, it means that the port is *open*; if the target machine responds with a *RST*, it means that the port is *closed*.

XMAS Scan



The XMAS scan sends a combination of FIN, URG, and PUSH flags to the destination. It lightens the packet just like a Christmas tree and that is why it is called an XMAS scan. It works just like the FIN and null scans. If there is *no* response, the port is *open*; if the target machine responds with a *RST* packet, the port is *closed*.

TCP ACK Scan



The TCP ACK scan is not used for port scanning purposes. It is commonly used to determine the firewall and ACL rules (access list) and whether the firewall is able to keep track of the connections that are being made.