IJCSE International Journal of Computer Sciences and Engineering Open Access

Survey Paper Vol.-7, Special Issue-14, May 2019 E-ISSN: 2347-2693

Information Gathering and FootPrinting Framework for Penetration Testing using Shell Script

[1]Gopichand D, [2]Lakshmikar B, [3]Siva Teja G, [4]Chaitanya Sai G, [5*]Raghavendra Reddy

[1,2,3,4,5]School of Computing and Information Technology, REVA University, Bangalore, India

Corresponding Author: raghavendrareddy@reva.edu.in

DOI: https://doi.org/10.26438/ijcse/v7si14.239244 | Available online at: www.ijcseonline.org

***Abstract*—** Now a days Cyber threats are the costliest threats happening globally. Lets assume a scenario where a user surfs the internet, share files, download files, upload files without any basic security precautions. In this case he/she can get infected with virus and also shares it in form of physical drives or any upload of file, this will also infect other end users. To prevent this stuff in the organisations they conduct a monthly or quarterly security audit which will help them to maintain their systems secure.

***Keywords*—** Information Gathering, Penetration Testing, Automation, Footprinting, Ethical Hacking.

I. PROBLEM STATEMENT

The people who audits the security comes with a moto to perform audit only on particular areas, here they will get less amount of time which will be given them mostly in the midnight to complete their tasks. So there might be a chance of incompletion of the work from the given time and also lacks some tools to Audit Completely. Some organisations do not perform any security audit within the time and adjust their budgets on different operations, and this gap leads to a cyber threat [1].

II. INTRODUCTION

The term Hacking for the most part refers to unapproved interruption into a PC or a system. Hacker is a skilled computer programmer who use computers to gain illegitimate access to data. This hacker may adjust framework or security highlights to achieve an objective that varies from the first reason for the framework. Hacking can likewise allude to non-malignant exercises, more often than not including peculiar or extemporized modifications to gear or procedures [4][7][13].
The generations of hacking stood like [5][7]

A. 1960’s

Curiously, the expression ”hack” did not start from PCs. Or maybe, it started with MIT’s Tech Model Railroad Club route in 1961 when club individuals hacked their cutting edge train sets so as to alter their capacities. They later proceeded onward from toy trains to PCs, utilizing the slippery and costly IBM 704’s at MIT to develop, investigate, make new standards, and attempt to extend the assignments that PCs could achieve.

These MIT substitutes alongside other early programmers were intrigued just with regards to investigating, improving and testing the points of confinement of existing projects. once in a while, these hacks even delivered projects that were extensively exclusive to the prior ones, just like the case with Dennis Ritchie’s and Keith Thompson’s UNIX working framework[5][7].

B. 1970’s

While PC hacking kept on flourishing during the 1970s, the decade also offered path to another kind of programmer: one that toyed with phone frameworks. Named ”phreakers,” telephone programmers, for example, the scandalous John Draper, misused operational qualities in the phone exchanging system, which had as of late gone totally electronic.

Draper legendarily found that a toy whistle found in Cap’n Crunch grain created the accurate tone essential 2600 hertz to demonstrate to long queues that a line was prepared and accessible to highway another call. This enabled him and different phreakers to hoodwink the system and make free long separation calls.

The phreaker subculture offered approach to persuasive programmers like Draper as well as to advanced visionaries, too. Before they went on to establish a standout amongst the best PC organizations on the planet, Steve Wozniak and Steve Jobs were, actually, respectful telephone phreakers[5][7].

C. **1980's**

The 1980s was a watershed decade in the historical back-drop of hacking, as it denoted the acquaintance of turnkey PCs with the overall population. Never again restricted to organizations and renowned colleges, PCs were accessible for everybody to use for their own motivations whatever that might be. Obviously, the vast accessibility of individual PCs prompted a quick increment in programmers.

It was by all account not solitary enormous change to happen in the hacking network. While there were as yet countless intrigued principally in tinkering with working frameworks, another breed rose that was continuously bothered about close to home addition. Rather than utilizing their mechanical ability for improving PCs, they utilized it for crimes, including pirating programming, making infections and breaking into frameworks to take touchy data. It didn't take the law long to react. The rise of digital crooks was quickly met in 1986 with the main enactment identified with hacking, the Federal Computer Fraud and Abuse Act.

In the meantime, this was likewise the time that the possibility of programmers being advanced intellectuals fit for doing both incredible and horrible things entered mainstream culture. Several movies and books were made that advanced the thought, for the most part quite the 1983 flick War Games in which a rural young person finds an crooked access in a military focal PC and about begins World War III [5][7].

D. 1990's

Making the most profit of the huge changes that happened during the 1980s, the 1990s were while hacking truly started to acheive reputation. The term programmer was discolored by a regularly expanding number of digital malpractice executed by ”crackers” (or malicious programmers) and the well known captures that pursued. Kevin Mitnick, Kevin Poulsen, Robert Morris and Vladimir Levin were a portion of the more outstanding wafers to leave the decade, having been captured and sentenced for any semblance of taking appropriateness programming from enormous name companies, tricking radio stations to win extravagance autos, propelling the primary PC worm, and driving the main advanced bank heist. The once affectionate hacking network likewise observed its breakdown in this decade. With an end goal to get serious about PC wrongdoing, the Secret Service propelled sting examinations, directed early morning attacks and captured various program-mers. Attempting to keep away from conviction, individuals in the hacking network started to advise on one another in return for resistance [5][7].

E. 2000's

Moral hackers kept on observing their great name hauled in the soil during the 2000s as assaults propelled by pernicious programmers ruled the features. New and risky kinds of hacks rose that deceived government substances and noticeable organizations. Microsoft, eBay, Yahoo! further more, Amazon were among those brought down in enormous disavowal of-administration assaults, while the Department of Defense and International Space Station had its frameworks broken by a 15- year-old kid [5][7].

F. 2010's

The world now stukked in the computerized age, the hacking network has turned out to be progressively modern, jumbled and complex than at any other time. Lone wolf programmers and little hacking gatherings still exist in each side of the web, either improving programming or propelling ransomware and Wi-Fi assaults relying upon their cap. All things considered, it's ”hacktivist” gatherings, for example, Anonymous, that have become the dominant focal point in this decennium, discharging exceedingly characterized archives, uncovering government mysteries and driving vigilante computerized campaigns for the sake of safeguarding people in general from being hurt, abused, or retained data. In response to both hacktivists and digital crooks, gov-ernment substances and huge partnerships are scrambling to improve security while PC mammoths strive to change their frameworks. Be that as it may, while digital security specialists keep on being selected, frameworks overhauled and innovation developed, programmers great and awful reliably and obviously remain one stage ahead [5][7].

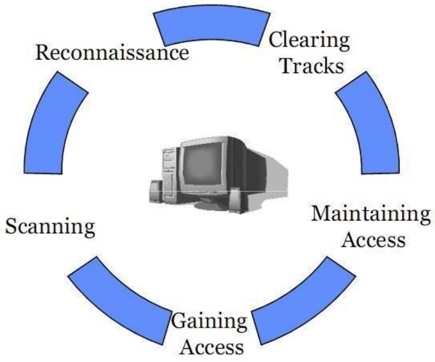


Fig. 1. Phases in hacking

There are predominantly 5 phases in hacking. Not always a hacker has to follow these 5 steps also a pentester in a sequential manner. Its a step by step process and when followed yields a better result.

1. Reconnaissance
2. Scanning
3. Gaining Access
4. Maintaining Access
5. Covering Tracks

In this paper we concentrate on the Reconnaissance, which is formed from the terms Information Gathering, Digital Foot Printing and Enumeration.

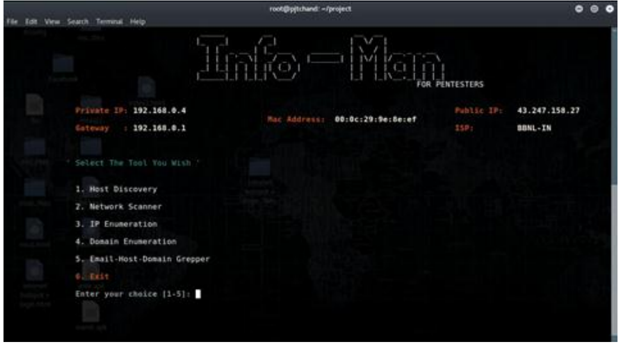
III. OBJECTIVES

a) To Create an environment where a penetration tester can save better time in his work.
b) To reduce the major cyber threats, The vulnerabilities need to be patched according to their updates to minimise the chance of risk.
c) To scan and analyze any network with more speed, efficiency and accuracy compared to use manually.
d) The tool is totally automated where user need t give options to get his work done.
e) This style of Frameworks create the trend for cyber security enthusiasts to work on these type of creations.
f) To create a trust for users that they can perform reconnaissance un-doubtfully using this framework.

The detailed information of the tool as follows:

A. Interface

The interface is created in a manner where user can discover his own Private Public IP addresses, MAC Address, ISP, Gateway.



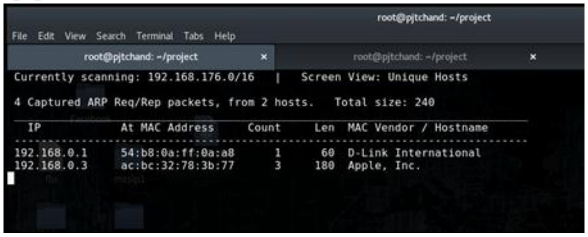
```
root@pjtchand: ~/project
File  Edit  View  Search  Terminal  Help

                     Info-Man
                                  FOR PENTESTERS

Private IP: 192.168.0.4                                        Public IP:   43.247.158.27
Gateway   : 192.168.0.1        Mac Address:  00:0c:29:9e:8e:ef  ISP:         BBNL-IN

* Select The Tool You Wish *

1. Host Discovery
2. Network Scanner
3. IP Enumeration
4. Domain Enumeration
5. Email-Host-Domain Grepper
6. Exit
Enter your choice [1-5]: 
```

Fig. 2. Interface of framework

According to the reconnaissance the tool is divided into five module to spray the fragrance of the Information Gathering, Digital Footprinting and Enumeration.

The modules are as follows:

1. **Host Discovery**
It performs ARP request and pulls out the hosts which are up [2].

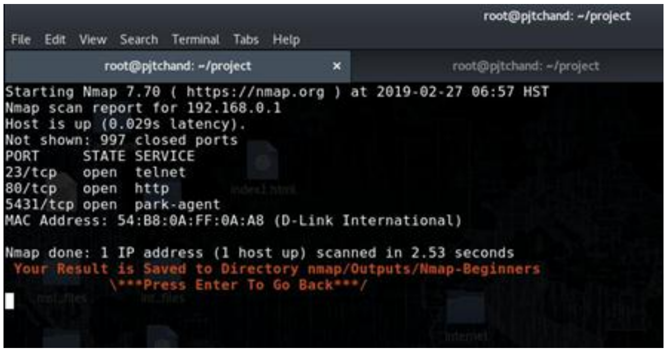


```
root@pjtchand: ~/project
File  Edit  View  Search  Terminal  Tabs  Help
root@pjtchand: ~/project                      root@pjtchand: ~/project
Currently scanning: 192.168.176.0/16   |   Screen View: Unique Hosts

4 Captured ARP Req/Rep packets, from 2 hosts.   Total size: 240
_____________________________________________________________________________
  IP            At MAC Address     Count     Len  MAC Vendor / Hostname
-----------------------------------------------------------------------------
192.168.0.1     54:b8:0a:ff:0a:a8      1      60  D-Link International
192.168.0.3     ac:bc:32:78:3b:77      3     180  Apple, Inc.
```

Fig. 3. Discovering Hosts

2. Network Scanner

It performs operations to discover what works are running on their systems, determining hosts that are available and the services they provide, finding open ports and detecting security risks and threats [2][3][7][8][9].



```
root@pjtchand: ~/project
File  Edit  View  Search  Terminal  Tabs  Help
root@pjtchand: ~/project                      root@pjtchand: ~/project
Starting Nmap 7.70 ( https://nmap.org ) at 2019-02-27 06:57 HST
Nmap scan report for 192.168.0.1
Host is up (0.029s latency).
Not shown: 997 closed ports
PORT     STATE SERVICE
23/tcp   open  telnet
80/tcp   open  http
5431/tcp open  park-agent
MAC Address: 54:B8:0A:FF:0A:A8 (D-Link International)

Nmap done: 1 IP address (1 host up) scanned in 2.53 seconds
 Your Result is Saved to Directory nmap/Outputs/Nmap-Beginners
           \***Press Enter To Go Back***/
```

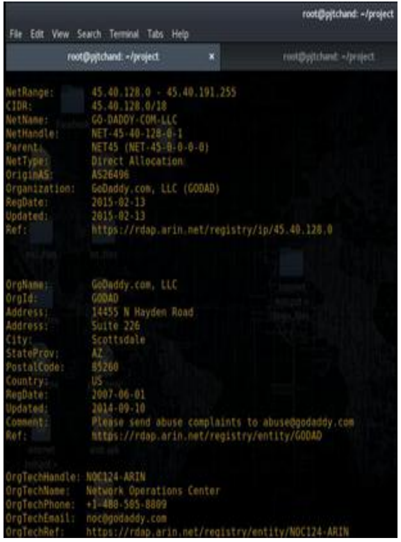
Fig. 4. Scanning Hosts in a network

3. IP Enumeration

This module perform two specific operations:

a). IP/Domain Lookup

We use WHOIS in automated way which is a system that allows users to look up the name and contact information of a registered domain name (website). When someone registers a new domain, the registrar asks for specific contact information, most of which is required by The Internet Corporation for Assigned Names and Numbers (ICANN) [7].

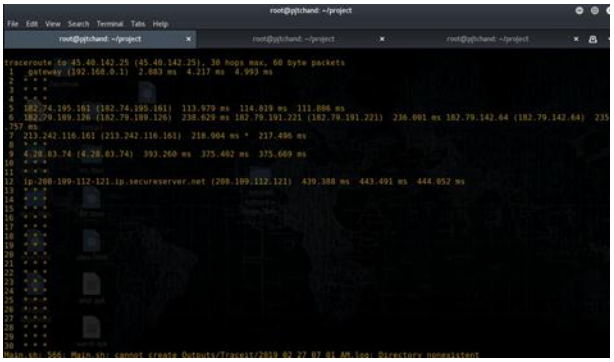


```
root@pjtchand: ~/project
File  Edit  View  Search  Terminal  Tabs  Help
root@pjtchand: ~/project                      root@pjtchand: ~/project

NetRange:       45.40.128.0 - 45.40.191.255
CIDR:           45.40.128.0/18
NetName:        GO-DADDY-COM-LLC
NetHandle:      NET-45-40-128-0-1
Parent:         NET45 (NET-45-0-0-0-0)
NetType:        Direct Allocation
OriginAS:       AS26496
Organization:   GoDaddy.com, LLC (GODAD)
RegDate:        2015-02-13
Updated:        2015-02-13
Ref:            https://rdap.arin.net/registry/ip/45.40.128.0


OrgName:        GoDaddy.com, LLC
OrgId:          GODAD
Address:        14455 N Hayden Road
Address:        Suite 226
City:           Scottsdale
StateProv:      AZ
PostalCode:     85260
Country:        US
RegDate:        2007-06-01
Updated:        2014-09-10
Comment:        Please send abuse complaints to abuse@godaddy.com
Ref:            https://rdap.arin.net/registry/entity/GODAD


OrgTechHandle: NOC124-ARIN
OrgTechName:   Network Operations Center
OrgTechPhone:  +1-480-505-8809
OrgTechEmail:  noc@godaddy.com
OrgTechRef:    https://rdap.arin.net/registry/entity/NOC124-ARIN
```

Fig. 5. Domain/IP Information gathering

b). Traceroute

It is used to trace the IP address from where it is reaching its destination like gateways, firewalls, ISPs, Server farms etc.



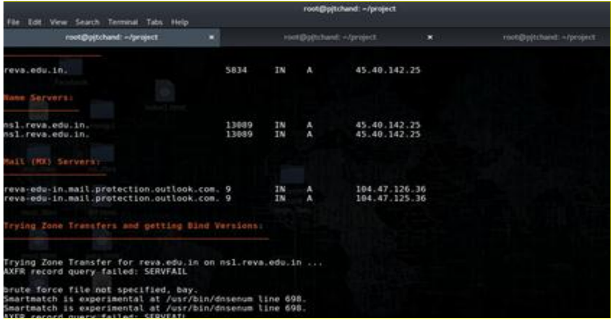
```
traceroute to 45.40.142.25 (45.40.142.25), 30 hops max, 60 byte packets
 1  _gateway (192.168.0.1)  2.883 ms  4.217 ms  4.993 ms
 2  * * *
 3  * * *
 4  * * *
 5  182.74.195.161 (182.74.195.161)  113.979 ms  114.819 ms  111.886 ms
 6  182.79.189.126 (182.79.189.126)  238.629 ms 182.79.191.221 (182.79.191.221)  236.001 ms 182.79.142.64 (182.79.142.64)  235
.757 ms
 7  213.242.116.161 (213.242.116.161)  218.004 ms *  217.496 ms
 8  * * *
 9  4.28.83.74 (4.28.83.74)  393.260 ms  375.402 ms  375.669 ms
10  * * *
11  * * *
12  ip-208-109-112-121.ip.secureserver.net (208.109.112.121)  439.388 ms  443.491 ms  444.052 ms
13  * * *
14  * * *
15  * * *
16  * * *
17  * * *
18  * * *
19  * * *
20  * * *
21  * * *
22  * * *
23  * * *
24  * * *
25  * * *
26  * * *
27  * * *
28  * * *
29  * * *
30  * * *
Main.sh: 566: Main.sh: cannot create Outputs/Traceit/2019_02_27_07_01_AM.log: Directory nonexistent
```

Fig. 6. Tracing IP/Domain

4. Domain Enumeration

This module perform three specific operations:

a). Server enumeration



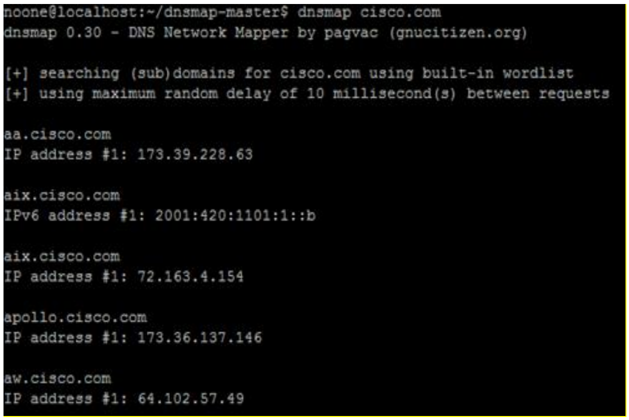
```
reva.edu.in.                          5834     IN    A        45.40.142.25

Name Servers:

ns1.reva.edu.in.                      13089    IN    A        45.40.142.25
ns1.reva.edu.in.                      13089    IN    A        45.40.142.25

Mail (MX) Servers:

reva-edu-in.mail.protection.outlook.com. 9     IN    A        104.47.126.36
reva-edu-in.mail.protection.outlook.com. 9     IN    A        104.47.125.36

Trying Zone Transfers and getting Bind Versions:

Trying Zone Transfer for reva.edu.in on ns1.reva.edu.in ...
AXFR record query failed: SERVFAIL

brute force file not specified, bay.
Smartmatch is experimental at /usr/bin/dnsenum line 698.
Smartmatch is experimental at /usr/bin/dnsenum line 698.
AXFR record query failed: SERVFAIL
```

Fig. 7. Identifying type servers

This uses dnsmap, which is made in a automated way to scan the given domain and find all its servers like Name, Mail servers [7].

b). Sub Domain Enumeration

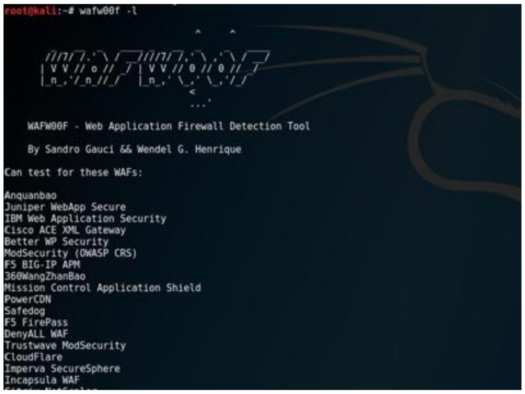
This uses dnsenum, which is made in an automated way to scan the given domain and find all its sub domains linked to the main domain.



```
noone@localhost:~/dnsmap-master$ dnsmap cisco.com
dnsmap 0.30 - DNS Network Mapper by pagvac (gnucitizen.org)

[+] searching (sub)domains for cisco.com using built-in wordlist
[+] using maximum random delay of 10 millisecond(s) between requests

aa.cisco.com
IP address #1: 173.39.228.63

aix.cisco.com
IPv6 address #1: 2001:420:1101:1::b

aix.cisco.com
IP address #1: 72.163.4.154

apollo.cisco.com
IP address #1: 173.36.137.146

aw.cisco.com
IP address #1: 64.102.57.49
```

Fig. 8. Enumerating sub domains

c). WAFW00F

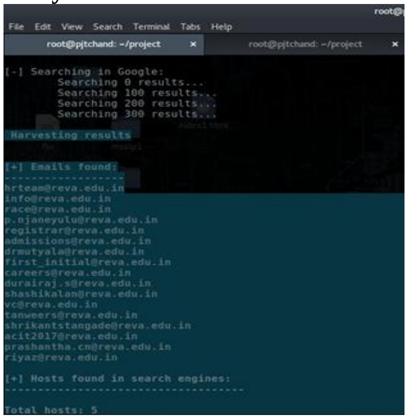


```
root@kali:~# wafw00f -l

    WAFW00F - Web Application Firewall Detection Tool

    By Sandro Gauci && Wendel G. Henrique

Can test for these WAFs:

Anquanbao
Juniper WebApp Secure
IBM Web Application Security
Cisco ACE XML Gateway
Better WP Security
ModSecurity (OWASP CRS)
F5 BIG-IP APM
360WangZhanBao
Mission Control Application Shield
PowerCDN
Safedog
F5 FirePass
DenyALL WAF
Trustwave ModSecurity
CloudFlare
Imperva SecureSphere
Incapsula WAF
```

Fig. 9. Firewall Detection

It is used in finding a Web Application Firewall for a Web Application.

6. Privacy Checker



```
[-] Searching in Google:
        Searching 0 results...
        Searching 100 results...
        Searching 200 results...
        Searching 300 results...

 Harvesting results

[+] Emails found:
------------------
hrteam@reva.edu.in
info@reva.edu.in
race@reva.edu.in
p.njaneyulu@reva.edu.in
registrar@reva.edu.in
admissions@reva.edu.in
drmutyala@reva.edu.in
first_initial@reva.edu.in
careers@reva.edu.in
durairaj.s@reva.edu.in
shashikalan@reva.edu.in
vc@reva.edu.in
tanweers@reva.edu.in
shrikantstangade@reva.edu.in
acit2017@reva.edu.in
prashantha.cn@reva.edu.in
riyaz@reva.edu.in

[+] Hosts found in search engines:
------------------------------------

Total hosts: 5
```

Fig. 10. Privacy lookup from shadon database and pgp key servers

TheHarvester is used here and made it automated for users. The work of this program is to gather emails, subdomains, hosts, employee names, open ports and banners from different public sources like search engines, PGP key servers and SHODAN computer databases [7].

IV. IMPLEMENTATION

This totally made of shell script which is a scripting language.

Here we used linux operating system with combination of number of system packages and GitHub projects, listed as follows:

Operating System: Kali linux/ Debian GNU Linux / Parrat Sec/ BlackArch / BackBox
System Packages Used: Netdiscover, Nmap, Whois, tracer-oute, Dnsenum, Dnsmap, Wafw00f, Theharvester, Zenity.
Github Projects Used: toilet, lolcat.

Internet is mandatory to work on it.

Using the above round of packages we made a beautiful information gathering tool, who can perform duties as follows.

As an pentester, you should utilize similar procedures a programmer uses to look at a system. Infiltration testing as a rule begins with three pretest stages: footprinting, filtering and counting. These pre-test stages are imperative and can have the effect between a fruitful infiltration test that gives a total image of the client's presentation or one that doesn't.

Together, the three pre-test stages are called observation. This procedure tries to assemble however much data about the objective system as could be expected, after these seven stages:

i) Gather initial information
ii) Determine the network range
iii) Identify active machines
iv) Discover open ports and access points
v) Fingerprint the operating system
vi) Uncover services on ports
vii) Map the network

These pre-test stages involve the procedure of revelation, and in spite of the fact that the procedure is regularly executed in a specific order, a great analyzer realizes how to extemporize and head in an alternate course, contingent on the data found.

A. Footprinting

Footprinting is the vital blueprinting of the security profile of an association. It involves collecting data about your client's system to make a one of a kind profile of the association's systems and frameworks. It's a critical route for an attacker to pick up data around an association inactively, that is, without the association's learning [4].

Footprinting employs the initial two stages of surveillance, assembling the underlying target data and deciding the system scope of the objective. Regular devices/assets utilized in the footprinting stage are:

a) Whois
b) SmartWhois
c) NsLookup
d) Sam Spade

We'll investigate these and different apparatuses in the following portion of this arrangement. Footprinting may likewise require manual research, for example, concentrating the organization's Web page for helpful data, for instance [7]:

i) Company contact names, phone numbers and email addresses
ii) Company locations and branches
iii) Other companies with which the target company partners or deals
iv) News, such as mergers or acquisitions
v) Links to other company-related sites

Company privacy policies, which may help identify the types of security mechanisms in place.

B. Scanning

The following four data gathering steps - recognizing dynamic machines, finding open ports and passageways, fingerprinting the working framework, and revealing administrations on ports - are viewed as a major aspect of the examining stage. Your objective here is to find open ports and applications by performing outside or inside system filtering, pinging machines, deciding system ranges and port examining singular frameworks [7].

Some common tools used in the scanning phase are:
i) NMap
ii) Ping
iii) Traceroute
iv) Superscan
v) Netcat
vi) NeoTrace
vii)Visual Route

V. APPLICATIONS

a) Will be useful for penetration testers.
b) It can be used for security audits in organisations & MNCs.
c) It reduces the stress, time and working time of the user.
d) It creates a specific environment like interface, which creates interest for user to work on.
e) This is totally automated where user need to give respective options to get their work done [11][12].

VI. ADVANTAGES & DISADVANTAGES

Advantages:

a) Can use in any linux operating system.
b) It reduces the time of work for the user.
c) User need to enter the options other than huge line of code.
d) Attractive interface.
e) Can swap to other modules by using automated back and home options.
f) Result of the scan will be saved directly into a respective directory without human effort.

Disadvantages:

a) Internet is mandatory.
b) Cannot run on windows Operating system.

CONCLUSION

This is a Tool which is created for the Cyber security Enthusiasts, Penetration Testers. This gives them a environment that they do not feel any kind of work pressure, because this is a tool which is made from combination of Information Gathering and Digital Foot Printing concepts. The Task of the User is to select the IP address and an Option which is appropriate for that scan and major task is to enter a Domain name. So that it reduces the time of user from giving a huge line of commands as input, and also reduces the time of reference.

This tool can be elaborated to next level by adding number of exploits to crack the device that are found as vulnerabilities that are revised from the result of this tool [10][11][12][13].

REFERENCES

[1] Andress, Mandy. ”Network scanners pinpoint problems.” Network World (2002).

[2] O. Arkin, ”ICMP Usage In Scanning”, The SysSecurity Group, June 2001.

[3] R. Farrow, ”System Fingerprinting With Nmap”, Network Magazine, November 2000.

[4] Smith, Yurick, Doss Ethical Hacking IEEE Conference Publication, DOI: 10.1147/sj. 403.0769, pp. 769-780 - 2014.

[5] Behera, Dash Ethical Hacking: A Security Assessment Tool to Uncover Loopholes and Vulnerabilities in Network and to Ensure Protection to the System , International Journal of Innovations & Advancement in Computer Science, Vol 4, pp. 54-61 - 2015.

[6] Digital Defenders Document on Cyber security - 2018.

[7] Hall, Gary, and Erin Watson. Hacking: Computer Hacking, Security Testing, Penetration Testing and Basic Security. CreateSpace Independent Publishing Platform, 2016.

[8] Lin, Huaqing, Zheng Yan, Yu Chen, and Lifang Zhang. ”A survey on network security-related data collection technologies.” IEEE Access 6 (2018): 18345-18365.

[9] Fessi, B. A., S. Benabdallah, M. Hamdi, S. Rekhis, and N. Boudriga. ”Data collection for information security system.” In 2010 Second Inter-national Conference on Engineering System Management and Applica-tions, pp. 1-8. IEEE, 2010.

[10] Guo, Fanglu, Yang Yu, and Tzi-cker Chiueh. ”Automated and safe vulnerability assessment.” In 21st Annual Computer Security Applications Conference (ACSAC’05), pp. 10-pp. IEEE, 2005.

[11] Wotawa, Franz. ”On the Automation of Security Testing.” In 2016 International Conference on Software Security and Assurance (ICSSA), pp. 11-16. IEEE, 2016.

[12] McGraw, Gary. ”Automated code review tools for security.” Computer 41, no. 12 (2008): 108-111.

[13] Urias, Vincent E., William MS Stout, Jean Luc-Watson, Cole Grim, Lorie Liebrock, and Monzy Merza. ”Technologies to enable cyber decep-tion.” In 2017 International Carnahan Conference on Security Technology (ICCST), pp. 1-6. IEEE, 2017.