Information Gathering Tool - Identify Website Pages and Application Information to Easily Apply Open Vulnerability

Dasa Jayshree M[1], Dr. Ravi Sheth[2]
[1]*MTech Student, School of Information Technology & Cyber Security, Raksha Shakti University, Dahegam, Gandhinagar, Gujarat*
[2]*Guide, Assistant Professor, School of Information Technology & Cyber Security, Raksha Shakti University, Dahegam, Gandhinagar, Gujarat*

***Abstract*- Today cyber words are fast developed, approximately a lot of processes are automated by several open sources. While develop anything using open-source to trust their tools which may have many vulnerabilities in the website likewise Programming Language, Language version, CMS name, CMS info, CMS Version, CMS Theme, headers unused pages, unnecessary services, unprotected files, and directory, etc.to identify the first level of find the information in website. Develop a tool to measure websites all links are showing.**

***Index terms*- Software engineering, Requirements Elicitation (RE), System analyst, Information gathering method, System Development.**

I.INTRODUCTION

Data Gathering is the demonstration of get-together various types of data against the focused on casualty or framework. There are different apparatuses, methods, and sites, including open sources, for example, Whois, nslookup that can assist programmers with gathering data.

Data Gathering is the demonstration of social event various types of data against the focused on casualty or framework. It is the initial step or the starting phase of Ethical Hacking, where the infiltration analyzers or programmers (both dark cap and white cap) played out this stage; this is an essential and critical advance to be performed. The more the data accumulated about the objective, the more the likelihood to acquire significant outcomes. Data gathering isn't only a period of security testing; it is a craftsmanship that each infiltration analyzer (pen-analyzer) and programmer should ace for a superior involvement with entrance testing. There are different apparatuses, methods, and sites, including open sources, for example, Whois, nslookup that can assist programmers with gathering data. This progression is important in light of the fact that while performing assaults on any objective, You may require any data, (for example, his pet name, closest companion's name, his age, or telephone number to perform secret phrase speculating assault or different sorts of assaults).

II. REQUIREMENTS METHODS AND TECHNIQUE

Information gathering can be classified into three major categories:

- Foot printing
- Scanning
- Enumeration

What is foot printing?

Foot printing is the strategy to gather however much data as could be expected about the focused on arrange/casualty/framework. It helps programmers in different manners to meddle with an association's framework. This method additionally decides the security stances of the objective. Foot printing can be dynamic just as detached. Detached foot printing/pseudonymous foot printing includes the assortment of information without the proprietor realizing that programmers accumulate his/her information. Interestingly, dynamic impressions are made when individual information gets discharged

deliberately and purposefully or by direct contact of the proprietor.

Tools, Tricks and Technique for information gathering

Tools

- Pentest Tool
- Whois Lookup
- Nmap
- Dmitry
- Th3inspector
- Red Hawk

III. TOOL ANALYSIS

Nmap

Nmap, the Network mapper, is a free, open-source apparatus for defenselessness filtering and arrange revelation. The instrument utilizes crude ip bundles in novel manners that to perceive what hosts are accessible on the system, what administrations (application name and form) those hosts are giving, what usable frameworks (and OS variants) they're running, what sort of parcel channels/firewalls are being used, and many various qualities. It was intended to rapidly filter monstrous systems, anyway works fine against single hosts.

Syntax:

Scan a Single IP: nmap <ip>

Scan a Host: nmap <URL>

Scan a Single Port: nmap –p <port number> <ip>

Scan a Range of a Port: -p <port number> <range> <ip>

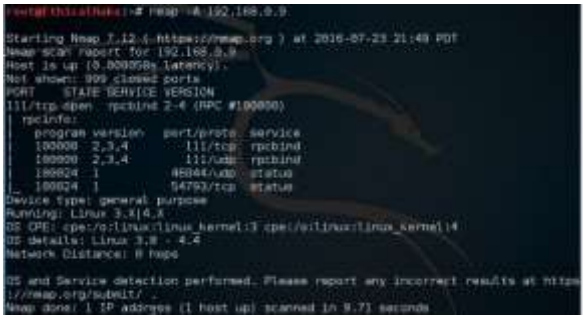

Fig. 1 Nmap

Th3inspector

Th3inspector is an incredible open source data gathering device accessible on GitHub through which you can without much of a stretch discover a lot of data about the objective, for example, server subtleties, whois query, target IP area, telephone number, email address, sub-spaces and so forth. The instrument has numerous choices to count site data.

Syntax:

To Find IP Address and E-mail Server: perl Th3inspector.pl <URL>

To Find Website or IP Address Location: perl Th3inspector.pl -l <URL>

To Get Website Information: perl Th3inspector.pl -i <URL>

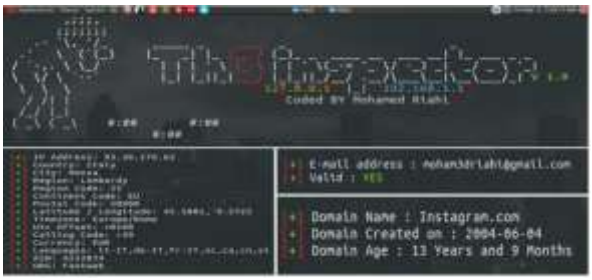

Fig. 2 Th3inspector

Red Hawk

Red Hawk is another open source data gathering device accessible on GitHub. It underpins numerous outputs and highlights like fundamental sweep, web server identification, cms discovery, whois query, invert ip query, snatch flags, dns query, subnet adding machine sub-area scanner, turn around ip query and CMS recognition. Red Hawk additionally bolsters Vulnerability filtering and slithering.

Syntax:

To find the website Information: <URL>

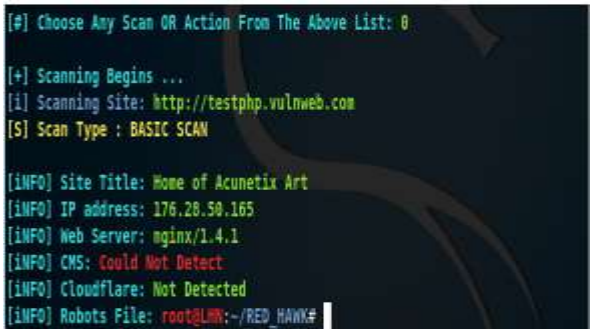

Fig. 3 Red Hawk

Dmitry – Deepmagic Information Gathering Tool

Dmitry (Deepmagic data gathering Tool) is a UNIX/ (GNU) Linux order line program coded absolutely in C with the adaptability to gather as a lot of

information as feasible about a host. The application is viewed as an instrument to help in data gathering when data is required rapidly by evacuating the need to enter numerous orders and the auspicious procedure of looking through data from various sources. Dmitry contains a base usefulness with the adaptability to include new capacities. Fundamental usefulness of Dmitry grants for information to be assembled about an objective host from a straightforward whois search on the objective to uptime reports and tcp port sweeps. The apparatus is open in Kali Linux.

Features

- Performs a whois search.
- Retrieve achievable uptime data, framework and server data.

Plays out a

- Subdomain search on an objective host.
- E-Mail address search on an objective host.
- Tcp Port output on the host target.

Syntax:

To find the website information: Dmitry [options] <host>

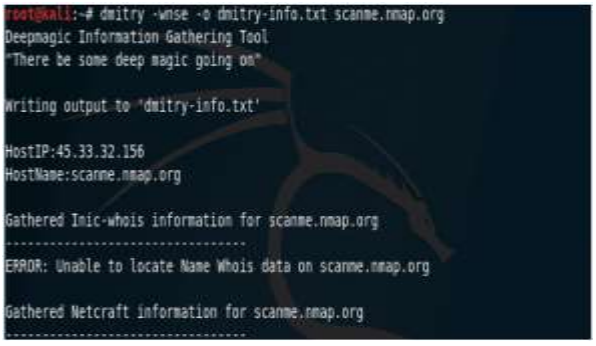

Fig. 4 Dmitry

Techniques

- OS Identification : includes sending unlawful TCP (Transmission Control Protocol) or ICMP (Internet Control Message Protocol) parcels to the casualty's framework to recognize the OS (Operating framework) utilized by the casualty on his server or PC.
- A ping clear is a strategy of setting up a scope of IP delivers that map programmers to live has. Fping, Nmap, Zenmap, SuperScan are a portion of the apparatuses used to ping an enormous number of IP addresses one after another; to create arrangements of hosts for huge subnets.

Tricks

- We can accumulate data from other various sources, for example, interpersonal interaction locales (Facebook, Twitter, LinkedIn, and so on.) are where general clients share their own information and extra data identified with them. Indeed, even web indexes assume a noteworthy job in social affair data.
- Hackers can likewise accumulate data from different monetary administrations about an objective organization, for example, the market estimation of an organization's offers, organization profile, contender subtleties, and so on.
- Hackers can likewise accumulate data from different monetary administrations about an objective organization, for example, the market estimation of an organization's offers, organization profile, contender subtleties, and so on.
- Hackers can likewise gather data from the email header, which incorporates:
- Address from which message was sent.
- Sender's email server.
- Sender's IP address.
- Date and time got by the originator's email server.
- Authentication framework utilized by the sender's letters server.
- Senders complete name.
- Objectives of Foot printing.

Host discovery	Port scanning	Version detection	OS detection	IP-Address	Sub-domain scanner	Show URL subdink
Nmap	✓	✓	✓	✓	✗	✓
Thinspector	✗	✗	✗	✓	✓	✓
Red Hawk	✓	✗	✗	✓	✓	✓
Dmitry	✓	✗	✗	✓	✓	✓
Wireshark	✓	✗	✗	✗	✗	✓
Nikto	✓	✗	✗	✗	✗	✓
Dnstracer	✓	✗	✗	✗	✗	✓
Dnsmap	✗	✗	✗	✓	✓	✓

Table-1 Comparisons of tool

III LITERATURE REVIEW

Abbasi et al. (2015) took a slightly different approach in the categorization of RE techniques/tools. They categorized RE techniques/tools into classic/traditional (interview, survey and questionnaire), cognitive/analytical (card sorting, laddering and repertory grid), modern and group (brain storming, JAD and prototyping), social analysis (ethnography, direct observation and passive observation). Similarly, the comparison of these techniques was based on the type of the elicitation technique (direct or indirect), type of the data (quantitative or qualitative data), communication and understanding of the domain.

Dennis et al. (2012) avoided the categorization of RE techniques/tools. However, the study observed that the most commonly used RE techniques/tools were interviews, JAD sessions, questionnaires, document analysis and observation. Similarly, RE techniques/tools were assessed based on type of information, depth of information, and breadth of information, integration of information, user involvement and cost. Earlier, the authors noted that no one technique is always better than the others, and in practice, most projects benefit from a combination of techniques.

Solution

Many tool comparison after some point noticed like any tool only show language-related information and other tool is only show the how the language used, which language used and target information is provided many tools. After one tool has created some tool all functionality is one tool impose. This tool is all information provided like security ratings, security score, domain score, ssl score, target info., target url, domain name, ip address, ssl expires, CDN, Running on, site links, URLs, internal JavaScript, External JavaScript etc.

IV. PROPOSED WORK

A. Android Code

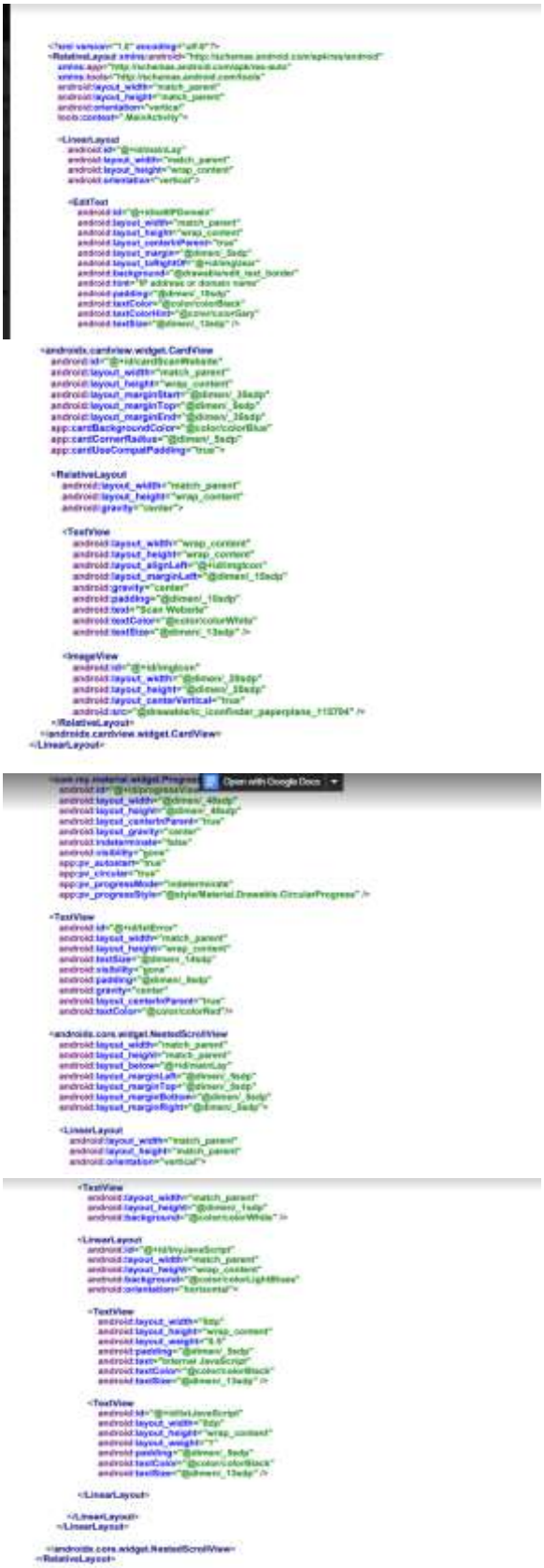

Fig. 5 Android Code

B. Output

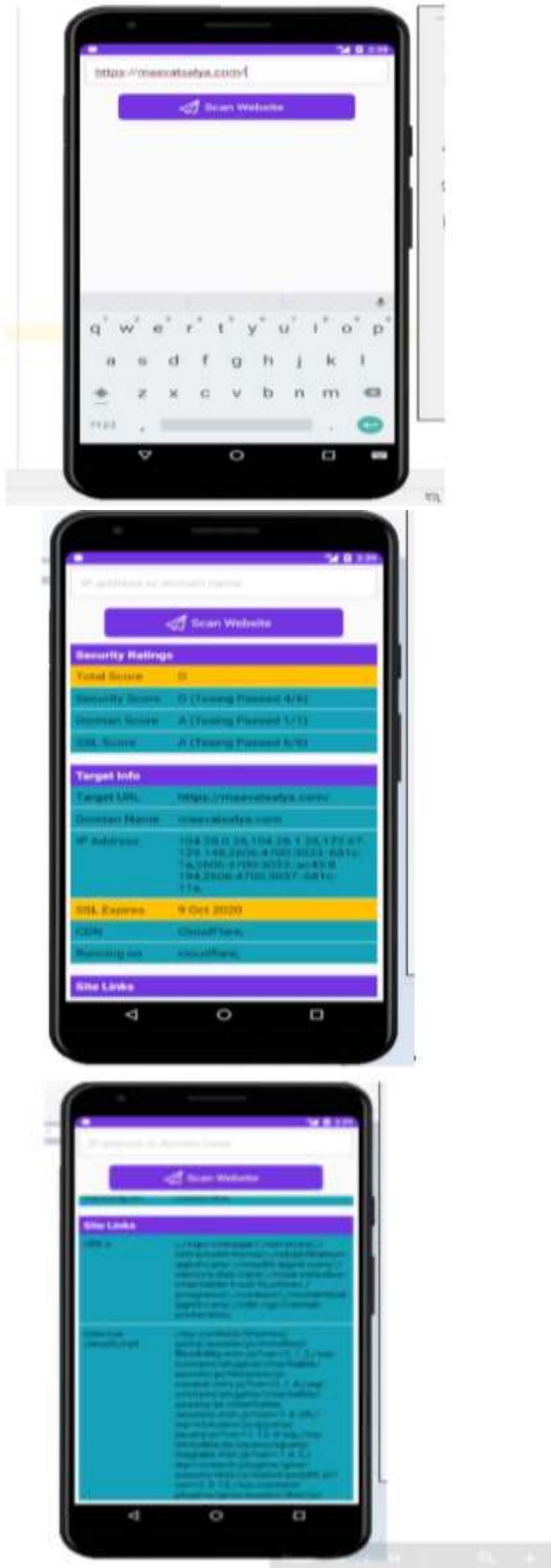

Fig. 6 Output

V. CONCLUSION

Choosing the correct data gathering procedure affects the nature of a product framework. This paper has surveyed the most normally utilized data gathering systems/devices, including host revelation, port filtering, Version identification, OS location, IP-address, sub-area scanner and so on.

The results have shown as that information gathering tool have similar functions but their output is very different. This tool is better than other tools. Many tools are provided many features but not any tool provide in show sub link in website. This tool is one of the features is added all sub link is show in any website.

Future work

- So, I will be developed a tool in show website all URLS sublink.
- So, Ethical hacker is easily referred website all pages and get the information.

VI. ACKNOWLEDGMENT

The author like to thanks IT department of Raksha Shakti University for giving all support and guidance which have enhanced the quality of the paper.

REFERENCES

[1] https://www.researchgate.net/publication/326688869_Information_Gathering_Methods_and_Tools_A_Comparative_Study
[2] https://www.researchgate.net/publication/314235536_Assessment_of_an_IT_network_security_with_information_gathering_tools
[3] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3220324
[4] https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Basic-tools-for-data-collection.pdf
[5] http://www.iosrjournals.org/iosr-jce/papers/Vol15-issue2/I01525965.pdf?id=8451
[6] https://www.w3schools.in/ethical-hacking/information-gathering-techniques/
[7] https://securitytrails.com/blog/information-gathering
[8] https://www.google.com/search?q=wireshark+features&rlz=1C1OWKM_enIN869IN869&oq=wi

reshark+features&aqs=chrome..69i57j0l7.6891j0j7&sourceid=chrome&ie=UTF-8
[9] https://www.google.com/search?q=nikto+features&rlz=1C1OWKM_enIN869IN869&oq=Nikto+fea&aqs=chrome.1.69i57j0.4039j0j7&sourceid=chrome&ie=UTF-8
[10] https://www.google.com/search?q=dnsmap+features&rlz=1C1OWKM_enIN869IN869&oq=dnsmap+feat&aqs=chrome.1.69i57j33.3544j0j7&sourceid=chrome&ie=UTF-8
[11] https://www.cyberpratibha.com/dnsmap-dns-domain-name-brute-forcing-tool/
[12] https://www.google.com/search?q=th3inspector+features&rlz=1C1OWKM_enIN869IN869&sxsrf=ALeKk01kYmTIdSh2h9igVENyimg6kl6NSw:1585308108506&gbv=2&sei=zOF9Xoi-HsPHrQGC0ZrgDQ