Identifying Subdomains of the Website Using SUBLIST3R and Comparing SUBLIST3R AMASS, KNOCKPY

Arunima Santhosh
Department of Computer Applications
Amal Jyothi College of Engineering
Kanjirappally, India
Arunimasanthosh2021@mca.ajce.in

Rinimol Kurian
Department of Computer Applications
Amal Jyothi College of Engineering
Kanjirappally, India
rinikurian@amaljyothi.a

Abstract : **This research paper discusses a lots of subdomain enumeration tools. There are various number of subdomain enumeration tools in kali. In this paper we discuss about Sublist3r, Amass, and KnockPy. These tools are mainly used to enumerate subdomain of website using OSINT. It helps penetration testers and bug hunters collect and gather subdomains for the domain they are targeting.**

In general, system security has become a crucial element in the digital aspect. Technically, evaluating the system there are aspects, one of which is by conducting a security assessment. Specifically, each system that will be evaluated is the essence of vulnerability search. Reconnaissance technique is mainly used for gathering information about computer systems and the entities they belong to. To make use of any system, a hacker might use various tools and technologies. This information is very useful to a hacker who is trying to crack a whole system. Subdomain enumeration plays a vital role in reconnaissance. Enumeration of subdomains provide an important insight towards the various underlying architecture and enable to find hidden user interfaces and admin panels. The less infrequent and unknown the domain name, the less visitors will visit the site. This enables a blind spot for the easy finding of low hanging vulnerabilities. What i am going to do in this paper is to identify various subdomains of website using Sublist3r ,a subdomain enumeration kali tool. Some of the other most popular tools used for recon on domains are Amass, SubFinder and KnockPy. In this paper also include comparative study and analysis of various functions of these tools on parameters like uniqueness, accuracy, and conclude with work in certain scenarios along with static code analysis to find weak spots within the code infrastructure of each of the tool.

I. INTRODUCTION

Reconnaissance definition means it is a significant instrument used in the beginning of data hacking and penetration testing. The main goal of recon is to retrieve data as quickly as possible without detecting an alarm and without exploiting the system infrastructure. In web infrastructure, subdomain enumeration plays a major role in finding the unknown underlying and unvisited internal domains, it may have a chance to better be vulnerable to 0 day attacks. Numerous tool like Amass, Knockpy, sublist3r are used to perform subdomain enumeration. All these tools uses different attribute for gathering information such as web scraping, API etc.

In the following sections, we discuss how to find out subdomains using these tools sublist3r, ammas, knockpy. In the first section we will discuss about the sublist3r , Sublist3r is a python tool mainly used for enumerate subdomains of website using OSINT. Sublist3r enumerate subdomains using many search engines such as Google, Yahoo, Bing, and Ask. Sublist3r also enumerates subdomains using Netcraft, Virustotal, ThreatCrowd, DNSdumpster and ReverseDNS. In the next section we discuss about Amass, Amass is an asset discovery tool which utilises open source information gathering along with active recon methods to enumerate the subdomain list. It is written in Golang. In the last section we discuss about knockpy, Knockpy is a python tool designed to enumerate subdomain on a target domain through dictionary attack.

II . SUBDOMAIN ENUMERATION

As already mentioned, recon is the maximum commonplace technique use to analyse andmapthe capacity objectives the victim's machine. Duetothefact most systems arelinked toth e outside international thru internet though respecti ve internet infrastructure systems, it endup very critical to examine the internet belong ings. Oneofthedistinguished methods torunreconon

10.5281/zenodo.5091353
ISBN:978-93-5426-386-6@2021 *MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam*

internet assets is bymeansof imposing subdomainen umeration technique.

Subdomain enumeration permits to find and enumerate a listing of subdomains aside from the main area of any organisation. usually, the principle area of any infrastructure is highly secured and could usual adhere to brand new security standards and patches .However it may no longer be a just a blatant bet to signify that all other subdomains may not comply with the strict security protocols set for the primary domain.

Subdomainenumeration approach involves the findi ng of massive listing osofsubdomainsofprimary do main.Subdomains tendtobe greater vulnerableto saf ety and informationdisclosure troubles. considering that theamount of sitevisitors tosubdomainsare less incomparisonto theprimary area,it typically receive s disregarded ofgetting the identical prioritised inter estcompared with the primary area.

Some examples of subdomains of the main domain of google.com are:

images.google.com, mail.google.com

III. ANALYSIS AND DISCUSSION

SUBLIST3R

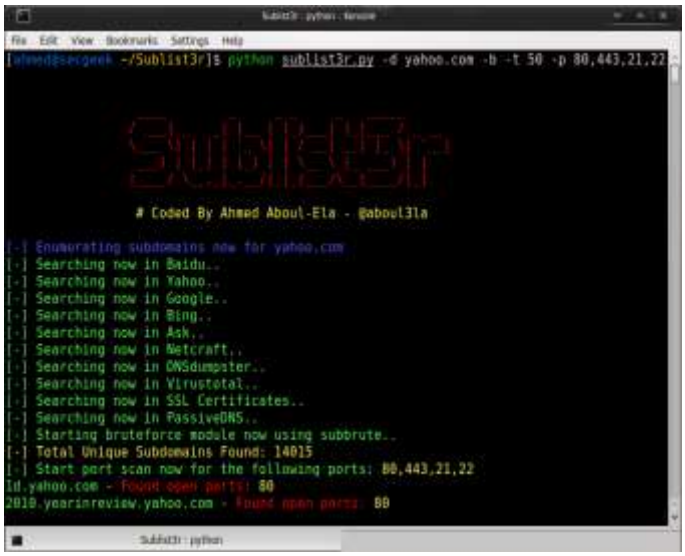
Sublis3r is coded in python and mainly takes use of OSINT strategies to enumerate the subdomain listing thru search engines like Google and yahoo like Google, baidu, bingetc. It has been alsoincluded with another tool subrockto disc over subdomain theusageof brute force froma listin g ofword listing.

Installing Sublist3r

root@kali:~# apt-get install sublist3r

Uninstall Sublist3r

root@kali:~#apt-get remove sublist3



Basic usage :

python sublist3r.py -d yahoo.com

To find specific ports

python sublist3r.py -d yahoo.com -p 88,4433

To enumerate subdomains and use specific engines such Google, Yahoo

Sublist3r –d kali.org –o kali-domains.txt –
v –e google, yahoo

Fig(1). Example of sublist3r

AMASS

10.5281/zenodo.5091353
ISBN:978-93-5426-386-6@2021 *MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam*

Amass is an asset discovery tool that enumerates the subdomain list by combining open source data with active recon methods. It's entirely written in the Golang programming language. The Open Web Application Security Project maintains Amass (OWASP)which is a non-profit organisation dedicated to enhancing software security.

Web scraping (Google, Yahoo, Reddit, Baidu, etc.) is one of the recon techniques used.Dns-zone transfer, brute-force approaches, reverse dns sweeping, and so on.

Basic usage :

amassenum -d yahoo.com

The command line usage of Amass consists of various subcommands with its respective arguments.

Fig(2): ammas tool

'viz' Subcommand

It helps for visualizations and adds a proper structure to the collected information.

ex. amass track –history

'track' Subcommand It shows the differences between enumerations of the same target.

KNOCKPHY

Knockpy is a python script that scans subdomains from a list, lists of words .It also supports DNS zone shift scanning and has the ability toif the wildcard DNS record is allowed, it will be ignored. It also allows you to use Virus Total to conduct searches. Its API keys Gianni Amato is the new maintainer (IT Security Researcher).

Sample Usage Commands

Scan with internal wordlist knockpyyahoo.com

Scan with external wordlist knockpy yahoo.com -w wordlist.txt

fig(3). Knockpy tool

A. ACCURACY

Each each of the five subdomain enumeration tools produces a large number of results and a large number of subdomains, it's critical to check the consistency of the results to see whether the hosts are being resolved. To meet the requirement, we created a Python tool called "Alive" that accepts a list of items of subdomains and decides whether or not the subdomain is currently resolving at the predetermined time.

A python script was created to check whether the resulting domains were resolved.

```
import socket
 def main():
 f=open("read.txt","r")
 count=0
f1=f.readlines()
forsd in f1:
 aas=sd.strip()
 print("CHECKING "+ aas+" --> ",end='')
try:
 w=socket.gethostbyname(aas)
print("Up!",sep='')
count=count+1
 exceptsocket.error:
```

10.5281/zenodo.5091353
ISBN:978-93-5426-386-6@2021 *MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam*

```
print("\033[31m"+ "Down!"+"\033[0m",sep='')

print("\n")

print("\n Total hosts FOUND to be UP ARE : "+str(count))

if __name__=="__main__" :main()
```

Amass is frontrunner, with a number of resolving hosts of around 2161.

B. STATIC CODE ANALYSIS

Static code review is the process of debugging a computer programme without actually running it. Instead, it examines the code structure and looks for flaws. We performed static code analysis on the tools and came to the conclusion that only a few security recommendations about the code quality were necessary. Some of the security risks associated with each of the three tools are mentioned below.

Sublist3r

Insecure hash function

Collision attacks are considered to be vulnerable to SHA1 signature algorithms. Attackers may use this flaw to create a new certificate with the same digital signature as the affected service, enabling them to impersonate it.

Amass

For applications that do not need cryptographic or security-related random data generation, math/rand is much faster. While crypto/rand is suitable for safe and crypto-ready use, it is slower. When you need to be safe with random data, crypto/rand is highly recommended numbers, such as when a web application generates a session ID.

Poor file permissions used when creation file or using chmod

Permissions given to a file/directory that are excessive. When a permission greater than 0600 is given, this alarm is activated.

KnockPy

Use of an insecure method from urllib.

urllib will open URLs that are not only http:// or https:// but also ftp:/ and file:/. It might be possible to open local files on the executing machine using this method, which may pose a security risk if the URL to open can be tampered with by an external user.

IV. CONCLUSION

Reconnaissance is important for gathering both passive and active information about a goal. DNS queries, WHOIS queries, and operating system recognition are some of the different recon methodologies and techniques used in it. Reconnaissance could be achieved both actively and passively. Passive foot printing collects data in an unobtrusive manner that is difficult to detect. Active foot printing makes use of features such as trace route, which pings the target infrastructure directly. The only drawback is that it increases the likelihood of activating IDS programmes. To gather information, all five subdomain enumeration analysis tools use both active and passive recon methods.

A thorough comparison of the various recon tools aids in determining which performs best in terms of precision, time, features, and ease of use, among other factors. We performed a static code analysis, and the results show that sub finder is the most efficient. It has a more efficient and speed-optimized codebase than the others.

Static code review on these recon platforms assisted in identifying the codebase's embedded flaws and proposing requirements for potential development.

V. REFERENCES

1. G. Jaspher Kathrine1 , Ronnie T. Baby2 , V. Ebenzer3, Comparative Analysis of Subdomain Enumeration Tools and Static Code Analysis,2020.

2. A. Kothia, B. Swar and F. Jaafar, "Knowledge Extraction and Integration for Information Gathering in Penetration Testing," 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C), Sofia, Bulgaria, 2019, pp. 330-335.

3. Adiwal, Sanjay &Rajendran, Balaji&Shetty, Pushparaj.(2018). Domain Name System (DNS) Security: Attacks Identification and Protection Methods.

4. Russell, Rebecca & Kim, Louis & Hamilton, Lei &Lazovich, Tomo&Harer, Jacob &Ozdemir, Onur&Ellingwood, Paul &McConley, Marc.

(2018). Automated Vulnerability Detection in Source Code Using Deep Representation Learning. 757-762. 10.1109/ICMLA.2018.00120.

5.Siavvas M., Gelenbe E., Kehagias D., Tzovaras D. (2018) Static Analysis Based Approaches for Secure Software Development. In: Gelenbe E. et al. (eds) Security in Computer and Information Sciences.Euro-CYBERSEC 2018.Communications in Computer and Information Science, vol 821. Springer, Cham

6. Thomassen, P., Benninger, J., &Margraf, M. (2018).Hijacking DNS Subdomains via Subzone Registration: A Case for Signed Zones. OJWT, 5, 6-13.

7. Alka Agrawal, Mamdouh Alenezi, Rajeev Kumar and Raees Ahmad Khan, Securing Web Applications through a Framework of Source Code Analysis, Journal of Computer Science,Volume 15, Issue 12,Pages 1780- 1794

8. P. Harika Reddy Surapaneni Gopi Siva SaiTeja, Cyber Security and Ethical Hacking, International Journal for Research in Applied Science & Engineering Technology (IJRASET),Volume 6 Issue VI, June 2018