International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET218317

Prevention to Sensitive Information Disclosure via OSINT

Shweta Sondarva[1], Dr. Priyanka Sharma[2], Prof. Dharti Dholariya[3]

[1]Student, School of Information Technology and Cyber Security, Rashriya Raksha University, Lavad, Gandhinagar, Gujarat,

[2]School of Information Technology, Artificial Intelligence, and Cyber Security, Rashtriya Raksha University, Gandhinagar, Gujarat, India

[3]School of Information Technology, Artificial Intelligence, and Cyber Security, Rashtriya Raksha University, Gandhinagar, Gujarat, India

Article Info
Volume 8, Issue 3
Page Number: 109-114

Publication Issue :
May-June-2021

Article History
Accepted : 10 May 2021
Published: 16 May 2021

ABSTRACT

This paper describes OSINT Tools and Approaches to find out sensitive information of any organization's Web Application or network. The paper contains the steps for gathering information and how to secure the web application, organization or network. There are many automated and paid tools available for vulnerability finding and penetration testing. In this paper we are performing recon with the help of OSINT to gather information and give the solution, before an attacker uses this vulnerability and exploits it. Nowadays lot many vulnerabilities are on the web application. I already learned the many cases in the security programs, where a Sensitive data leakage was happening on many reputed websites. So I will start to find out a web-application in which such types of information's are disclosed, the Problem was that if we find out such information leaking like credentials, Token, API key we can easily get authorization to admins/users account. I found a lot many well-known websites where we can easily use this sensitive data. To perform such kind of attack you just need to perform reconnaissance with the help of various open source tools available on internet.

Keywords: OSINT, Open source Intelligence, Reconnaissance, Recon Report

I. INTRODUCTION

People now use the internet to pass valuable information or files, for entertainment purpose, purchasing various products, make connection with other people, also using social networking websites to communicate with global word without any geographical barriers.

This dissertation intends to understand how internet is facing Data exposure discovered by attacker. As mentioned by cyber security ventures, by 2030, 90 percent of the world population will make internet usage, small and older aged will be found online, this means more than 8 billion Internet surfer. As the world sustains to be a digital globe, computerized social orders will deliver colossal sum of

computerized information created from individuals and trade intelligent in the internet.

Misusing this information within the right course will open various openings for public and trade organizations to extend benefits and work more proficiently within the unused data age.

Open Source Insights (OSINT) alludes to all data that can be found freely – generally by means of the web – without breaching any copyright or security laws. Beneath this definition, a wide cluster of sources can be considered a portion of OSINT. For occurrence, data posted freely on social media websites, posts on talk gatherings and bunch chats, unprotected websites registries and any piece of data that can be found by looking online. Be beyond any doubt that most OSINT assets cannot be found utilizing normal look motors such as Google or Yahoo!, as numerous assets are buried profound within the profound and dark net and such assets constitute more than 96% of the net substance.

II. OBJECTIVES

A. Defining OSINT

OSINT is open source intelligence that is publicly available information. Which is gathered from publicly available data, or open sources. Techniques of OSINT is helpful to security researcher, hackers or any state/national government performing a security operations. The Internet has reconceive how people passes information with each other and transformed how corporations do business. Nowadays, the major of world communicate in what, known as Cyberspace. While talking about OSINT one should understand that it is not limited to search data that is found through various search engines. Whereas 90% of data is to be dug from what we call is deep web.

B. Phases of OSINT

- RECONNAISSANCE PHASE

In reconnaissance phase gather all the information about the target. This is very important stage as it can bring you lots of important and confidential data. So this phase plays a key role into sensitive data collection.

- PLANNING PHASE

In this phase the collected information is processed to check how the information could be abused by looking at the attack scenario and then it is pass to investigator for further process

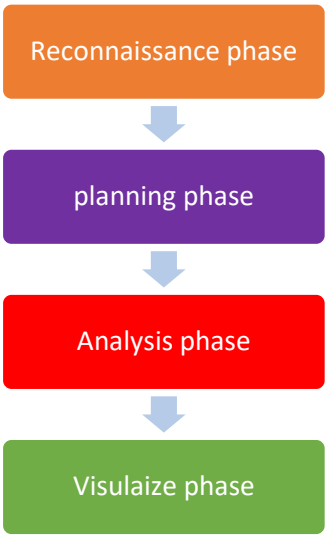

Fig 1 : Phases of OSINT

- Analysing phase

In analysis phase examiner thoroughly understand the requirement and design of the new system. Here the processed information is passed to investigator who have a deep understanding of various data and examiner inspect all possible scenario where the attack can be performed.

- Finalised phase

This is the last step where documentation is done on the basis of analysis made by investigator. Here whole report is prepared describing final system, improving their weakness.

C. Sensitive Information Disclosure

Sensitive information disclosure is leakage of confidential data which should be kept hidden and this data can be used to perform unauthorized access to someone's intellectual property.

Sensitive information could be anything like bank details, financial information, personal information, password, Account information, university administrative computer data, and many more. Sensitive data should never be stored in such way that unauthorized person can get easy access to it.

Here we are talking in context of web application. So most of the time information disclosure occurs when website /application fails to protect their sensitive data by allowing an unauthorized access.

This kind of security incidents cannot be misused as a rule, however, still considered beneath web application attacks as they permit cybercriminals to pick up access to data, result in gaining unauthorized access. The obtained data may vital for the organization or might contain area of other delicate files.

The web application which are vulnerable leads to reveal many restricted information which makes attackers work easier, and this kind of web application attack is known as sensitive information disclosure via OSINT.

D. Types of Information Disclosure attack

- Banner grabbing

Banner grabbing usually leaks the information of the version of the system, services running on the system, also known as service fingerprinting. This could be exploitable if found a matching CVE.

- Source code disclosure:

Source code disclosure occurs when backend code disclose information in public. There could be many hardcoded secrets like username, email-ids, password, API keys, secret keys and token.

- File path disclosure:

Many a time's error message also displays path to some sensitive file which leads door open to an attacker. Path disclosure can also happen by fuzzing URL or improper handling of user request.

- Directory listing

Directory listing occurs when there is no index file in any directory. It list out directories and files available on the server due to that directory listing can be dangerous.

III. RELATED WORK

Here is the study to some research work that is conducted to prevent information disclosure and their methodology.

In the paper **Open source intelligence: An intelligence sustenance** author discusses about how OSINT is widely used for information gathering source in an intelligent manner with the help of available intelligence tools. Author also mentioned how Intelligence tools plays a vital role in the exploitation of information. Tools analyses in such a critical manner that it fulfils the purpose of intended user. Tools can be less costly, time preventing and can be used for financial purpose as well as social welfare. It helps confirms the security and benefits of the OSINT at every phase.

In the paper **Privacy Preserving Detection of Sensitive Data Exposure** author discuss the leakage of information. He proposed a model for fingerprint called fuzzy fingerprint. Their model is based on bloom filter library in python. They displayed a data-leak detection arrangement which can be outsourced and conveyed in a semi-honest discovery environment. During data-leak detection operations, they used a fuzzy fingerprint technique that improves data privacy. The data owner prepares and cross fuzzy fingerprints before issuing them to the DLD provider. The DLD provider creates fingerprints from network traffic and looks for possible leaks. To keep the DLD provider from learning everything there is to know about the sensitive data, exact knowledge about the sensitive data, the collection of potential leaks is composed of real leaks. Data owner then post-processes the potential leaks sent back by the DLD provider and decides whether there is any real data leak.

As discussed in **ShoVAT: Shodan-based vulnerability assessment tool for Internet-facing services** it is possible to use Shodan to detect sensitive data exposure and prevent from data breach. Here ShoVAT takes output of Shodan queries and perform depth analysis of that specific given data. It runs algorithm which automatically reconstruct common enumeration names and extract vulnerabilities from national Vulnerability database.

Shodan is found to be one of the very well-known search engines available nowadays, developed to crawl the Internet and to index discovered services. This paper expands the features exposed by Shodan with advanced vulnerability assessment capabilities implanted into a novel tool called Shodan-based vulnerability assessment tool (ShoVAT).

In the **paper the not yet exploited goldmine of OSINT** author describes today's status and importance of OSINT. He shows how robust and self-managed OSINT is. He shows the technique on how one can change the profile of culprit from cyber incident with the help of OSINT. The article also represents some OSINT tools for advanced investigation.

Also going through the paper **Open Source Intelligence: Performing Data Mining and Link Analysis to Track Terrorist Activities**, author's main objective was to collect various information and conduct link analysis. He shows us how OSINT data were useful for terrorist organization. His analysis mainly focused on data and links he found were mapped with terrorist and most of this data was from news reporter and their agencies

IV. PROPOSED SYSTEM

During developing phase in many website data are disclosed and developers are not much aware of it. Not all developers have proper practice of security management. For that reason many paid security scanning methods are available in the market. But this tools are very much costly and provide scans for limited number of application. It cannot be afforded by every organizations. So to overcome such issues we will discuss about some easy methodology and open source tools for preventing information disclosure with the help of OSINT.

In this project with the help of OSINT techniques vulnerabilities can be discovered and data leak assessment can be performed. Their main job is recording what information someone could publicly find on or about company assets without resorting to hacking. An organization can identify leaked sensitive data such as breached account credentials, open ports, insecure network services, and outdated software and operating system versions already in use.

Finding Information with following techniques:

1. Github Recon

Github recon is not just limited to password. Here one can find information's like credit cards, secret access keys and token and many other credentials. Github recon is important as most of the organizations data is shared with github. Here by simply searching the name of the website we may find types of documents and repositories placed into github. There are many simple dorks shown below to find sensitive data.

"Website.com" aws key

Instead of 'aws key' one may use API_key, password, secret_key and many more.

2. Subdomain finder

Some of the times main website is secure due that attacker search for the backdoor in subdomains to attack the website. For that many subdomain finder tools available for free like Dnsdumpster, virustotal, sublist3r, amass and many more. With the help of this we can identify before launching the website if any backdoor left for an attacker.

3. Shodan favicon hash

Shodan is a search engine like Google but it help searches for device within all ports over the internet. Favicon is an icon image that is collected by Shodan all across the internet. Favicon hash is the IP address owned by the company. It is very much helpful since different tools uses different favicons and services. Shodan calculate hash value for all favicon it discover and let us search through them. It would make easy to find specific devices and services this way. Hash of the website can be easily calculated using Python code that can be found on Github.

5. Directory brute force and 403 bypass techniques

Directory brute forcing is the best techniques to find the hidden directories and pages of the website. There are many fuzzer and tools available for free scan like Dirsearch, DirBuster, WFuzz, DirB and many more. It brute force with the help of wordlist and gives many interesting result. It also shows forbidden, not found, or moved webpages. Here one can use 403 bypass techniques on the forbidden page to enter the webpage without permission. 403 bypass techniques work by manipulating HTTP response header or content type.

6. Version detection with known CVE's

The common vulnerabilities and exposure are publicly known vulnerabilities. This are the vulnerabilities which are already found by other security researcher. One may check for versions, technology, product related vulnerabilities, and then patch that vulnerability by upgrading it.

7. Google dorks

One of the very useful and easy techniques is finding sensitive data on internet using Google dorks. Here we can use Google search engine to find data that is accidently exposed to internet. By combining few dorks one can make their own search list for example:
site: http://sony.com confidential | top secret | classified | undisclosed |

8. Past Archives

Past archives is the techniques to retrieve deleted page or old content that could be helpful in investigation. This archives could help find out important information's like links for other sites, other employs details, contact information's, some useful documents and many more.

9. Third party service providers

Sometimes websites is secure but vulnerable third party vendor open the door of attack to attackers. Many organization relies on third party services like cloud service, telecommunication services, DNS providers etc. This could be vulnerable and the data could be at risk. So it can be prevented using third

party risk assessment tools and automated scanner to find the vulnerable technologies.

10. OSINT framework

This is very useful framework contains a vast number of open source tools. This framework was mainly created for IT security but due to its popularity it is used to find information for other industries also. Most of the information here are for free but some may require low cost fee also.

V. RESULT

This Paper mainly focuses on how with the help of low cost one can easily identify lots of sensitive and confidential information over the internet. With the help of certain OSINT tools and framework we may collect tons of important information which may cause an organization a huge loss. Preventing measures with low cost by performing Assessment with the help of OSINT, is the main aim of this thesis.

VI. CONCLUSION

Preventing Sensitive information disclosure is method through which data breaching can be prevented from attackers. Various open source tools are been used to find the loopholes where attacker can gain access to the data.

Reconnaissance is the first method where targets information is collected. Here one can get idea of what technologies are used, on which server the website is running, which port is open and many more. Of course Information gathering is a long process and it may take time but Attacker always finds confidential data and the loopholes with this process

VII. REFERENCES

[1] Annie ahuja, "Open source intelligence: An intelligence sustenance," in International Journal of Recent Trends in Engineering & Research, volume 04, Issue 04; Apr- 2018 [ISSN: 2455-1457]

[2] Béla Genge and Calin Enachescu,"shoVAT: shodan-based vulnerability assessment tool for internet-facing services" in Willey online library. (May 2015), DOI: 10.1002/sec.1262

[3] Xiaokui Shu and Danfeng Yao "Privacy preserving detection of sensitive data exposure" in IEEE transactions on information forensics and security, volume 10, May 2015

[4] Jacob Hedges, "closing gap between data and open source intelligence" in ieworldconference, 2 May 2019.

[5] J. pastor-galindo, "The not yet exploited goldmine of OSINT" in IEEE Access, volume 8, Jan 2020.

[6] Maurice Dawson "Open Source Intelligence: Performing Data Mining and Link Analysis to Track Terrorist Activities" in Information Technology - New Generations (pp.159-163), July 2017, DOI:10.1007/978-3-319-54978-1_22

[7] Ashleigh Powell, "Social Media data in digital forensic investigation" in book digital forensic education (pp.281-303), Jan 2020 DOI:10.1007/978-3-030-23547-5_14

[8] Aishwarya baby "A literature survey on data leak detection and prevention method" in International Journal of Advanced Research in Computer Science, Volume 8, No. 5, May-June 2017

Cite this article as :

Shweta Sondarva, Dr. Priyanka Sharma, Prof. Dharti Dholariya, "Prevention to Sensitive Information Disclosure via OSINT", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 8 Issue 3, pp. 109-114, May-June 2021. Available at doi : https://doi.org/10.32628/IJSRSET218317

Journal URL : https://ijsrset.com/IJSRSET218317