International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Website vulnerability scanner using python

Dr. M. SRINIVASAN M.E.Ph.D[1]*, M.DINESH*[2]*, A.MUTHUKUMAR*[3]*, B.SAIGANESH*[4]

[1] **Head of Dept, Dept.of IT, PSV College of Engineering and Technology,Krishanagiri,Anna University, Tamilnadu State,India**
[2] **UGScholar,Dept.of IT,PSV College of Engineering and Technology,Krishanagiri,Anna University, Tamilnadu State,India**
[3] **UGScholar,Dept.of IT,PSV College of Engineering and Technology,Krishanagiri,Anna University, Tamilnadu State,India**
[4] **UGScholar,Dept.of IT,PSV College of Engineering and Technology,Krishanagiri,Anna University, Tamilnadu State,India**

ABSTRACT

Today's software development life cycle should include security elements. Input Validation Attack is one of the most diversified web application attacks. By incorporating security into the software development life cycle, we hope to focus on the detection and prevention of Input Validation attacks such as SQL Injection, Cross Site Scripting File Inclusion (XSS), HTML Injection, LDAP Injection, XPath Injection, Remote File Inclusion, and PHP Code Injection. Take into consideration the hazardous vulnerability that occurred in web apps. When a vulnerability is discovered, the hijacker gains intended access to the authenticated user's web browser and may undertake session hijacking, cookie stealing, or other malicious activities. To prevent such attacks, it is critical to put in place security measures that prevent third-party intrusion. Website vulnerabilities are exploited across the network via web requests using the GET and POST methods. In this project, we are focusing on injection, detection, and prevention information. Shell-Shock, Anonymous Cypher (CVE-2007-1858), Crime (SPDY) (CVE-2012-4929), and Struts-Shock are the vulnerabilities discovered. The lack of automated scanners for detecting vulnerabilities in web applications, which leads to defacement, hijacking, and data theft from servers, causes a security risk for all businesses and government employees.

Keywords: Cross site scripting, remote file inclusion, Local file inclusion, SQL injection, Command injection.

1. Introduction

According to the current state of affairs, cyber hazards pose a significant threat to both small and large organisations, and while large corporations have the resources to prevent and combat these threats, small businesses and start-ups lack the financial and physical resources to do so. So, by removing the high cost component and presenting the bare bones version of complex software used to discover and suggest solutions for loop-holes in any specific company's homepage, we want to bridge the gap and demonstrate the possibilities that affordable software can also deliver. To achieve the goal of attacking websites, hackers can conduct penetration tests on target websites and utilize Web vulnerabilities to elevate privileges on website servers. Here introduce the paper, and put a nomenclature if necessary, in a box with the same font size as the rest of the paper. The paragraphs continue from here and are only separated by headings, subheadings, images and formulae. The section headings are arranged by numbers, bold and 9.5 pt. Here follows further instructions for authors. Based on these security dangers, utilizing vulnerability scanners to detect flaws on websites has some benefit. The Web vulnerability detection analyzer designed in this paper can collect website information in batches to achieve high concurrency between modules, and tasks can be processed simultaneously between crawlers and plug-ins, improving the efficiency of scanning websites, and the vulnerability script of the system has been updated. Scalability is beneficial to system improvement and upgrade.

Among the contributions made in this work are:

- The general architecture of the web vulnerability detection analyzer and the functional needs of the four modules are built based on the process analysis of website vulnerability scanning.
- This study performs vulnerability scanning checks on hundreds of websites and ranks them on three distinct scales. Check the website's entire scan time and the accuracy of its vulnerability results.

2. Framework and Methodology

I. Fingerprint

Server - The usage of a Client/Server architecture does not need the use of a host-based vulnerability scanner. The scanner is only host-based if it runs on the same machine that it is scanning. If it employs the Client/Server paradigm, the client is the host's agent. The server is often a centralised orchestrator in charge of managing the different agents.

a) Web Frameworks

CakePHP - CakePHP is an open-source framework for the rapid construction and maintenance of PHP-based web applications. It is built on the MVC design concept, which makes it easier to create PHP web applications with less code. CakePHP also makes it possible to segregate business logic from the data and presentation layers.

b) ***MVC architecture***

The view is interacted with by the client or user and view the controller's alerts for the relevant occurrence. It sends a database request to the Model, and the Model is updated by the Controller.

CherryPy - Web applications can be constructed faster and more reliably with CherryPy. It's also known as a web application library. Because it is based on object-oriented Python programming, it is noted for its simplicity, resulting in smaller source code in less time. It should be well-versed in Model-View-Controller and Object-Oriented Programming. CherryPy accepts an HTTP request and finds the best Python function or method that fits the URL of the request.

II. Front-end Frameworks

AngularJS - AngularJS is a framework for building dynamic web applications. It allows you to utilise HTML as your template language and extend HTML's syntax to represent your application components in a clear and concise manner. Its data binding and dependency injection capabilities reduce most of the code you're presently writing. And it all takes place within the browser, making it an ideal companion to any server technology.

VueJS - VueJS, like ReactJS, is one of the top JavaScript frameworks. VueJS is used to create the user interface layer; it is also interoperable with other libraries and plugins. VueJS is supported by all major browsers, including Chrome, Firefox, Internet Explorer, and Safari.

III. Content management system

A content management system (CMS) assists businesses in managing digital information. These systems can be used by entire teams to produce, edit, organise, and publish information. It serves as a central repository for content and enables automated processes for collaborative digital content management and creation through the use of built-in (or created) workflows. Individuals are given different rights and responsibilities based on their roles. Authors, for example, can upload and save their work, but editors can revise and publish it. Administrators can accomplish all of these things, as well as grant authorization to others in the organisation to add or revise content.

IV. Content delivery networks

A content delivery network (CDN) is a geographically dispersed network of servers that caches content near end consumers. A CDN enables the rapid delivery of assets required for the loading of Internet content, such as HTML pages, JavaScript files, stylesheets, pictures, and videos.

V. Attacks

Bruteforce - A brute force attack is a well-known breaking technique; according to certain sources, brute force attacks accounted for 5% of confirmed security breaches. A brute force attack involves'speculating' on usernames and passwords in order to gain unauthorised access to a framework. Brute force is a simple attack tactic with a high success rate.

Injection methods :-

HTML injection - Cross-site Scripting (XSS) is closely connected to HTML injection attacks. HTML injection defaces the page by using HTML. As the name suggests, XSS injects JavaScript into the website. Both attacks take advantage of poor validation of user input.

SQL injection - SQL injection (SQLi) is a web security flaw that allows an attacker to tamper with database queries made by an application. It generally enables an attacker to examine data that they would not otherwise be able to retrieve. This could include data belonging to other users or any other data that the programme has access to.

LDAP injection - The Lightweight Directory Access Protocol (LDAP) is a software protocol that allows anyone on a network to find resources such as other people, files, and devices. Intranets benefit from directory services such as LDAP. It can also be used as part of a single sign-on (SSO) system to store usernames and passwords.

XPATH injection - XPath is a query language that assists in finding specific elements, such as attributes, in an XML document by providing relative information. Hackers use XPath injection to exploit programmes that create XPath queries from user input to a browser (navigate) XML document.

Remote file inclusion - Remote file inclusion (RFI) is a type of attack that targets flaws in web applications that dynamically reference external scripts. The perpetrator's purpose is to use an application's referencing feature to upload malware (e.g., backdoor shells) from a remote URL in a separate domain. Local file inclusion (LFI), like RFI, is a vector that involves uploading malicious files to servers via web browsers. In the context of file inclusion attacks, the two vectors are frequently mentioned combined.

PHP code injection - A code injection attack takes advantage of a computer fault created by the processing of incorrect data. The attacker injects code into a susceptible computer programme, causing it to execute incorrectly. Successful code insertion can pose serious consequences.

Access control allow methods - The Access-Control-Allow-Methods response header is of the Cross-Origin Resource Sharing (CORS) type. It specifies which HTTP methods are permitted for accessing resources in response to cross-origin queries.

MutliIndex - Multi-level indexing is quite intriguing since it allows for pretty advanced data analysis and manipulation, particularly when working with larger dimensional data. In essence, it allows you to store and modify data in lower level data structures such as Series and DataFrame with an arbitrary number of dimensions.

Path planning algorithm:

- Dijkstra’s Algorithm
- A-Star(A*)

- D-Star(D*)

Sampling based algorithm:

- Rapidly exploring random tree
- RRT Star(RRT*)
- Informed RRT Star
- Batched informed tress star(BIT*)

Cross site tracing (XST) - Cross-site tracing (XST) is a more sophisticated type of cross-site scripting (XSS), which is a code injection attack in which the attacker executes malicious scripts inserted into a website or application. In XST, attackers can circumvent existing XSS security mechanisms and steal users' cookies. XST scripts make use of HTTP TRACE or TRACK methods that were initially intended for diagnostic purposes. Although server providers deactivate this HTTP request method by default to prevent XST and other similar attacks, some web server administrators continue to use it for debugging.

3. Vulnerabilities

A cybersecurity vulnerability is any flaw in an organization's information systems, internal controls, or system procedures that hackers can exploit. This project includes the vulnerabilities listed below.

Shellshock

Shellshock is the most recent vulnerability that is likely to be as popular, if not more so, than the Heartbleed Vulnerability, and it is already being widely exploited by a worm known as wopbot. The vulnerability was discovered in the Unix Bash shell, which may be found on practically any Unix / Linux based web server, server, and network device. Because Bash does not properly sanitize environment variables before execution, the attacker can submit commands to the server via HTTP requests and have them executed by the web server operating system. Stephane Chazelas found the shellshock vulnerability, which was awarded the CVE designation CVE-2014-6271.

Table 1 - Products affectes by CVE-2007-1858

#	Product Type	Vendor	Product	Version
1.	**Application**	**Apache**	**Tomcat**	**4.1.28**
2.	**Application**	**Apache**	**Tomcat**	**4.1.31**
3.	**Application**	**Apache**	**Tomcat**	**5.0.0**
4.	**Application**	**Apache**	**Tomcat**	**5.0.1**
5.	**Application**	**Apache**	**Tomcat**	**5.0.11**

CVE-2007-1858

The default SSL cypher configuration in Apache Tomcat 4.1.28 through 4.1.31, 5.0.0 through 5.0.30, and 5.5.0 through 5.5.17 employs insecure cyphers, including the anonymous cypher, allowing remote attackers to access sensitive information or have additional, undisclosed consequences.

CVE-2012-4929

The TLS protocol 1.2 and earlier, as used in Mozilla Firefox, Google Chrome, Qt, and other products, can encrypt compressed data without properly obfuscating the length of the unencrypted data, which allows man-in-the-middle attackers to obtain plaintext HTTP headers by observing length differences during a series of guesses in which a string in an HTTP request potentially matches an unknown string in an HTTP header, aka a "CRIME" attack.

Shruts-Shock

Due to erroneously processing an attacker's invalid Content-Type HTTP header, Struts is vulnerable to remote command injection attacks. The Struts vulnerability allows these commands to be executed with Web server capabilities. This is full remote command execution, and it has been actively abused since its initial exposure. The exploited code is found in the Jakarta Multipart parser. If the Content-Type value is invalid, that is, it does not match a legitimate type that was expected, an exception is produced, which is then utilised to display an error message to the user.

4. Maintenance

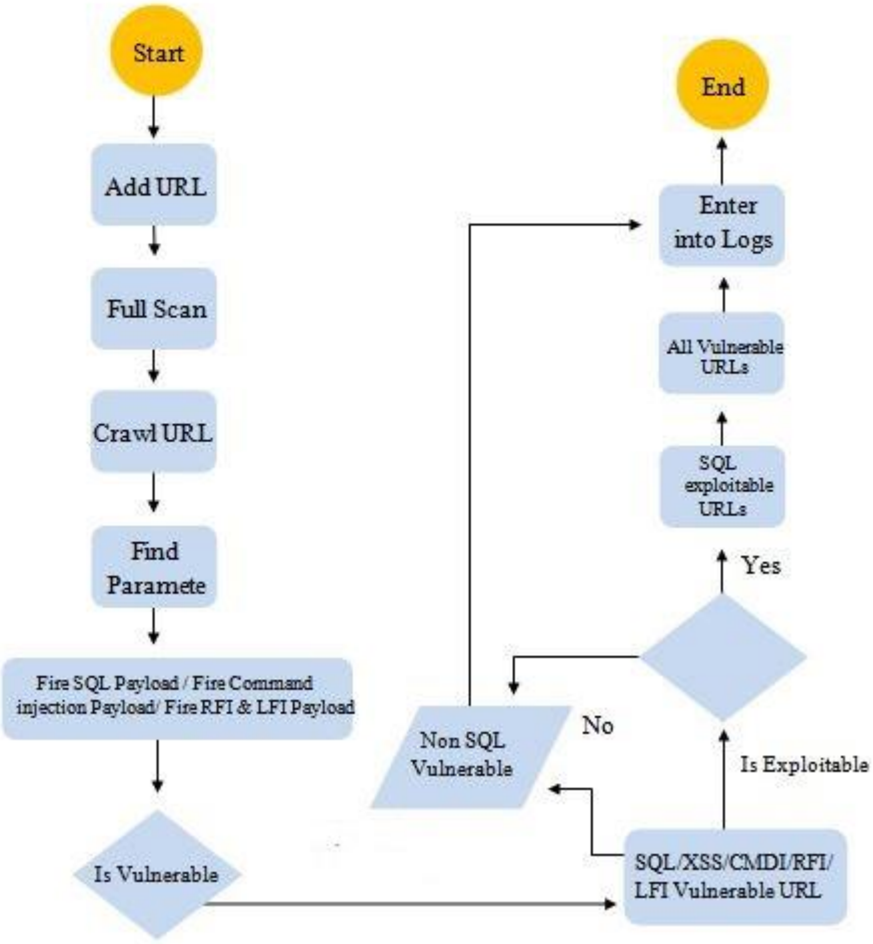
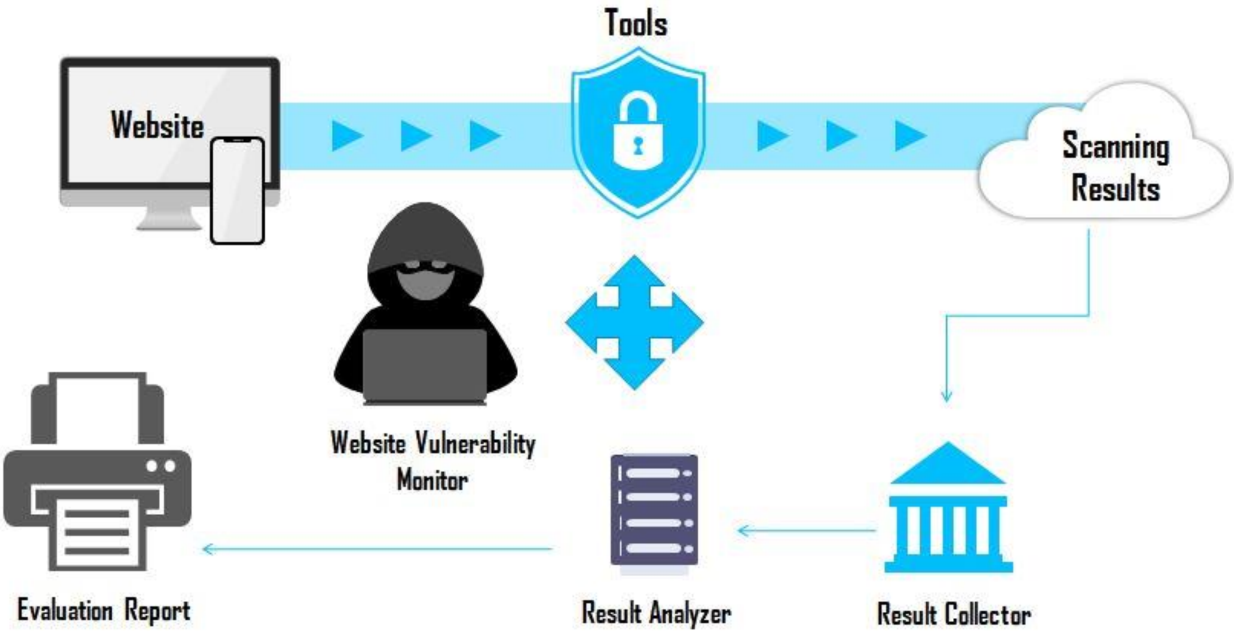


Fig. 1 – Workflow diagram

5. Architecture



6. Conclusion

This study highlighted web application security concepts and basic knowledge that can assist us in preventing web exploits in our system. Web applications are considered to be the most exposed and least protected because standards are not focused on security but rather on supplying the required functionality. Hackers will be a part of this never-ending game as technology advances and introduces new strategies, tools, models, and approaches to raise security levels. The project's goal is to find and exploit vulnerabilities using a preset Python script. Previously, only manual procedures were available, thus the entire process is automated to locate the impacted endpoints to exploit vulnerabilities. Instead of analyzing each endpoint, entire websites are searched to obtain endpoints, which are then exploited using payloads. Python is the programming language utilized to create this too

7. Acknowledgement

I first offer my thanks to all those who are graciously given me the strength and good health during the course of the project at thus pleasing moment of having successfully completed my project, I wish to convey our sincere thanks and gratitude to our beloved chairman,**Dr.P.SELVAM,M.A.,B.Ed.,M.Phil.,Ph.D.,** P.S.V. Group of Institutions. We also acknowledge our deep sense of gratitude to our secretary **Dr.S.VIVEK,M.A.,M.B.A(U.K).,Ph.D.,** P.S.V. Group of Institutions. I would like to express by sincere thanks to my beloved principal **Dr.P.LAWRENCE.,M.E.,Ph.D.,** P.S.V. College of Engineering and Technology, Krishnagiri, for forwarding us to our project and offering adequate duration in completing my project. I also express my sincere thanks to, Head of the Department, **Dr. M. SRINIVASAN M.E.,Ph.D** Department of Information Technology, for providing all the facilities in the successful completion of my project.

REFERENCES

[1] Bai, W., Jun, Y., & Zhao, Y. (2021). Analysis and discussion of XSS vulnerabilities. Electronic World, (20), 89–91. doi:10.19353/j.cnki.dzsj.2021.20.039

[2] Deng. (2020). Analysis of WEB penetration information collection. Electronic Components and Information Technology, (4), 24-25+32. .10.19772/j.cnki.2096-4455.2020.4.009

[3] F.M.Aiysha Farzana, Hameedhul Arshadh. A, Ganesan. J, N. Muthukumaran, 'High Performance VLSI Architecture for Advanced QPSK Modems', Asian Journal of Applied Science and Technology, Vol. 3, No. 1, pp. 45-49, January 2019.

[4] Aruna, Y.Bibisha Mol, G.Delcy, N. Muthukumaran, 'Arduino Powered Obstacles Avoidance for Visually Impaired Person', Asian Journal of Applied Science and Technology, Vol. 2, No. 2, pp. 101-106, April 2018.