JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Eagle Eye: Web Based Reconnaissance Platform

[1]Keshav Kaushik, [1]Akshita Kapoor, [1]Ishita Nanda, [1]Samyukta Sahoo, [1]Ashhar Ali, [2]Gunjan Chhabra

[1]Assistant Professor, [2]Third Year Student, [3]Third Year Student, [4]Third Year Student, [5]Third Year Student
[1] School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
[2] Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun, Uttarakhand, India

Abstract: Web applications have become increasingly popular over the past few years. Additionally, web application vulnerabilities have been increasingly targeted by cyberattacks and thus it becomes necessary to enhance the security of web applications. The process of manually checking all web vulnerabilities is time-consuming and difficult. Several tools, including Shodan, Whois and Censys, which help in the first phase of ethical hacking which is the Reconnaissance phase or the Information Gathering phase, are available. This project focuses on creating a platform which provides all the important reconnaissance tools at a single place which will help the security researchers and penetration testers to gather the required information of the web application. As a result, it will allow them to easily narrow down the specific and valuable information and then analyze known weaknesses to fix them.

IndexTerms **- Reconnaissance, Scraping, Vulnerabilities, Subdomains, Endpoints, Records, Whois, waybackurls**

I. INTRODUCTION

Recent decades have seen rapid growth in the technology industry and a dramatic increase in its use. Several technological advancements have led to unprecedented progress in a variety of fields, including information and communications technology (ICT), artificial intelligence (AI), robotics, machine learning, nanotechnology, space technology, biotechnology, and quantum computing. Data capture, processing, and analysis are integral to numerous areas of research and development. As a result, efficiency and productivity are enhanced across a wide range of sectors. Depending on the technology, it may take the form of websites, Android apps, APIs, etc. Our daily lives are reliant on web applications these days. Due to that, attacks based on web application vulnerabilities and their damage have also increased. So, developing a completely secure web application is imperative, but verifying each vulnerability manually is difficult, time-consuming, expensive, and error prone.

Securing a web application involves various phases, the first one being the reconnaissance or the information gathering phase. Reconnaissance [1] involves discovering and collecting information about the web application and it plays a key role in penetration testing. Many web-based reconnaissance tools are available which security professionals and testers can use. This web-based reconnaissance tools platform aims at bringing some important reconnaissance tools at a single place with a beginner and user-friendly GUI to ease the task of information gathering.

II. LITERATURE REVIEW

The incorporation of technological advancements and evolving techniques by companies leads to more chances of security flaws in new products. In the process of reconnaissance, Linux plays an important part which helps further in the process of pen testing, for which numerous tools are available on different sites [2]. Thus, in this field it can be tedious to gather information as a security researcher, considering the number of tools available online to identify vulnerabilities [3]. To contribute to the same, in this project we have tried to integrate some of the common information gathering tools on a single platform to ease out the task.

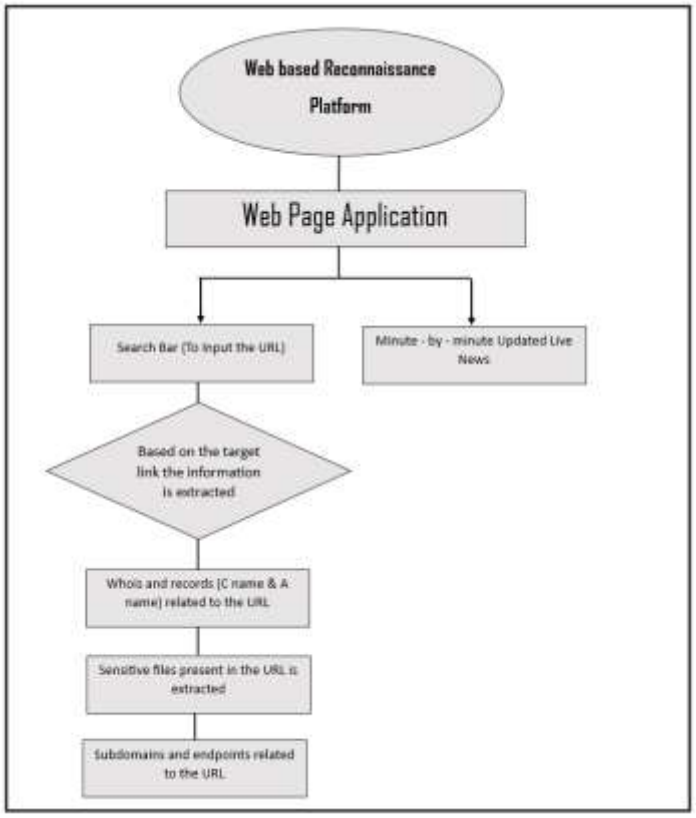

Figure 1. Flow Diagram of Web based Reconnaissance Platform

An important aspect of reconnaissance for a security researcher is not just extracting information about a target, but also gathering information about recent attacks and cybersecurity discoveries. We have added a component to our website containing the latest global security news updates. We have included this through web scraping [4], using requests-HTML which is a Python library used for scraping content from web pages. The news we are presenting here has been scraped from bleepingcomputer.com [5].

Records (A and CNAME records) are types of DNS records that map the domain name to an IP address. A Record or Address record [6] is used to map information that helps the computer to look for an IP address using DNS. CNAME Records or Canonical Name Records [7] are used when one domain name is mapped to another domain. In our project, an inbuilt functionality of Python, dnspython which gives the information about the records that has been displayed on the website [8].

Whois [9] is a query and response protocol, generally used to search databases that contain registered users, such as domain names, IP addresses, and autonomous systems. In our project, an in-built module of Python is used to display information related to the target like registrar, URL, creation date, expiration date, updated date, registrant e-mail, and status.

Sensitive files such as JavaScript and PHP files can provide huge amounts of information. Various attacks can be performed using JavaScript and PHP files on frontend and backend respectively. We have used waybackurls [10] to gather all the past URLs associated with a target which is further filtered using regex to capture JavaScript, PHP, and txt files.

Subdomains and endpoints are both crucial components of reconnaissance. By identifying subdomains and endpoints, security researchers and penetration testers can determine which services are running on the target system, and potentially identify vulnerabilities that can be exploited to gain unauthorized access. We have implemented both subdomains and endpoints finder in this project, using Python requests module.

The tech stack feature of our application tells us the technology used to build the target. It provides information about front-end and back-end technologies used in the development of the target.

Report generation of the data collected through reconnaissance provides a structured and comprehensive summary of the findings. This would help the security researcher or ethical hacker to take actions based on the information that is found to be most crucial in the further stages of penetration testing or ethical hacking.

III. METHODOLOGY

A web-based application is built using HTML, CSS and JavaScript for frontend and Python Flask framework for backend. Numerous python scripts such as subdomain, endpoint enumeration and various other tools are implemented in the backend to produce required results. Python modules such as Requests-HTML, Flask-SocketIO, Requests, whois, Threading, json, re and subprocess are used.

Here the HTML markup language is used for creating and structuring our webpage content, while we have used CSS to style the layout, typography, colors, and other visual aspects of our webpage. The JavaScript programming language is used to create

dynamic and interactive webpages which also helps in communicating with the flask server. Flask is a lightweight web framework for building web applications in Python. It provides a web server, a templating engine, routing, and HTTP handling support for Python web applications.

IV. RESULTS

The web-based platform works in the way that when the webpage gets loaded, it comes with the latest security news, and a search bar.

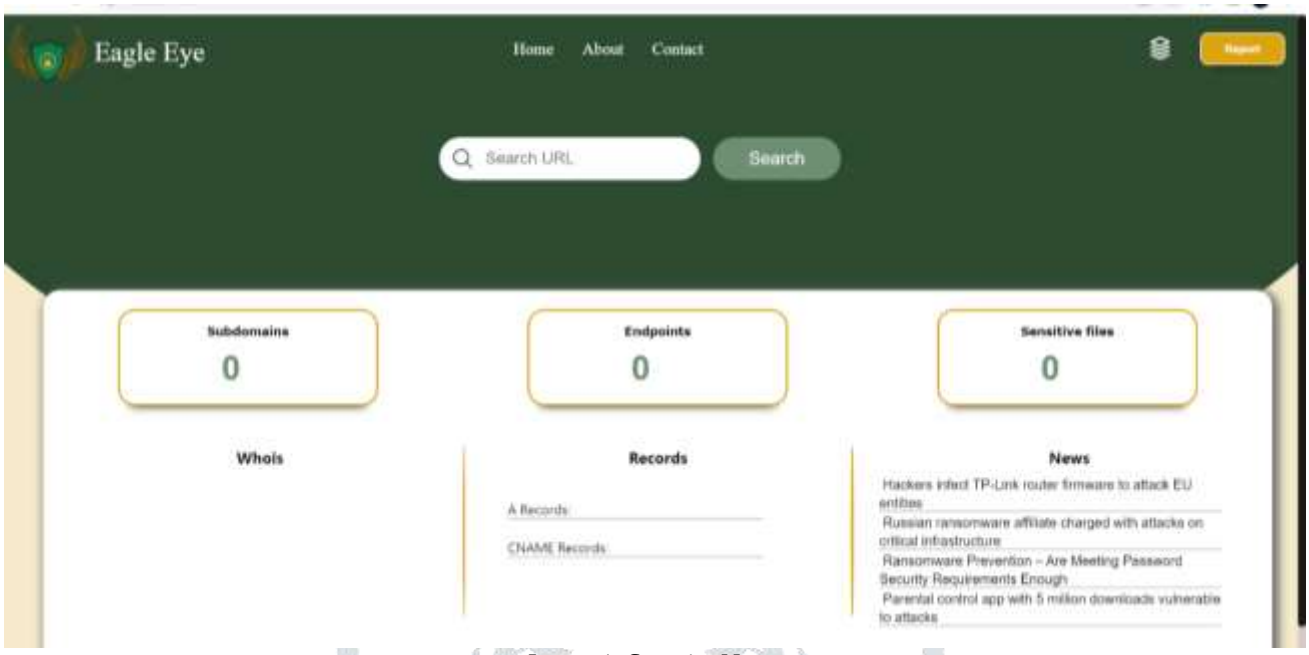

Image 1. Security News

By entering our target URL information in the search box, we can extract information about it. Python flask integrates everything along with Flask-SocketIO to speed up results, and we get information from whois including registrar details, creation and expiration dates, and records related to the domain.

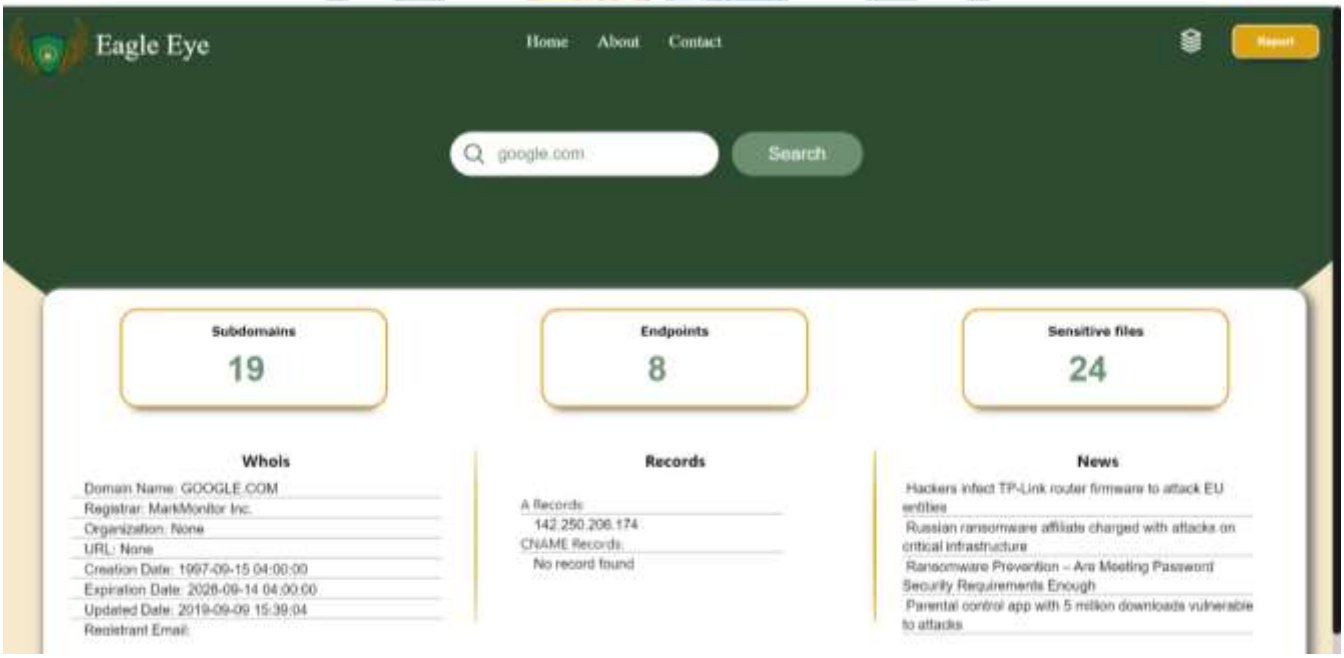

Image 2. Example of a Target URL google.com

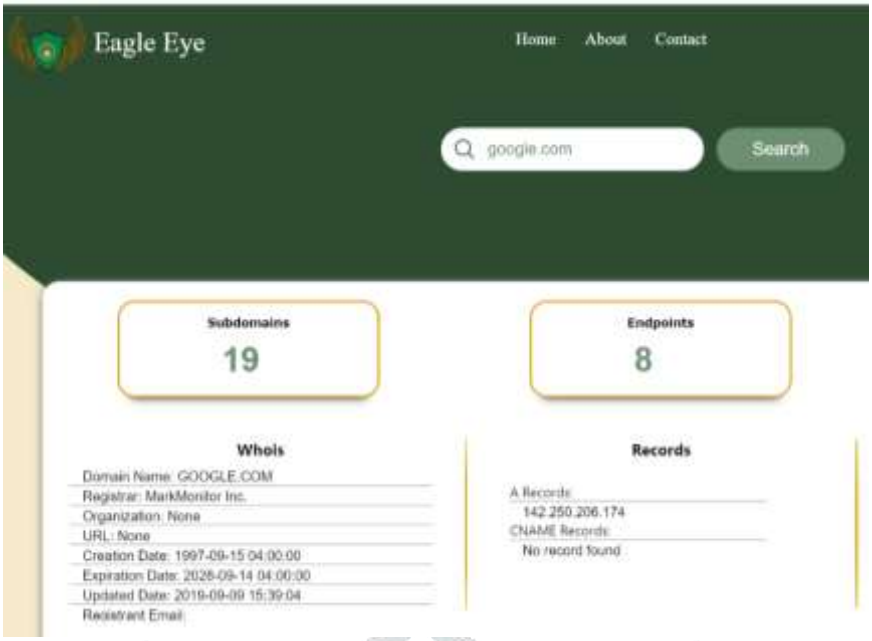

Image 3. Whois and Records information about target URL

Additionally, it displays the subdomains and endpoints associated with the URL, as well as sensitive files extracted from the URL.

Image 4. Subdomains, Endpoints and Sensitive Files related to the target URL

Moreover, it displays the tech stack of the URL Website as well, and finally a report of all these information is generated in pdf format.

Image 5. Tech Stack of the target URL

Image 6. Final Generated Report

V. CONCLUSION

The reconnaissance process or the gathering of information plays an essential role in any cyberattack, regardless of whether you are an attacker or defender. Without it, no further steps can be taken to attack, or defend a system, because as a part of ethical hacking it helps identify vulnerabilities and prepare for potential attacks. To inform decision-making and prioritize remediation efforts, ethical hackers and security researchers can gather information about target environments to gain a deeper understanding of potential risks and vulnerabilities through a platform like ours.

REFERENCES

[1] K. Kaushik, S. A. Yadav, V. Chauhan, and A. Rana, "An Approach for Implementing Comprehensive Reconnaissance for Bug Bounty Hunters," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), pp. 189–193, Dec. 2022, doi: 10.1109/IC3I56241.2022.10072942.

[2] D. Theodorus, M. S. Nabi, and Q. Al-Maatouk, "Web-based Reconnaissance and Vulnerability Scanner: A Review and Proposed Solution," 2021 International Conference on Data Analytics for Business and Industry, ICDABI 2021, pp. 666–670, 2021, doi: 10.1109/ICDABI53623.2021.9655963.

[3] A. S. Laxmi Kowta, K. Bhowmick, J. R. Kaur, and N. Jeyanthi, "Analysis and overview of information gathering tools for pentesting," 2021 International Conference on Computer Communication and Informatics, ICCCI 2021, vol. 2021-January, Jan. 2021, doi: 10.1109/ICCCI50826.2021.9457015.

[4] V. Singrodia, A. Mitra, and S. Paul, "A Review on Web Scrapping and its Applications," 2019 International Conference on Computer Communication and Informatics, ICCCI 2019, Jan. 2019, doi: 10.1109/ICCCI.2019.8821809.

[5] "BleepingComputer | Cybersecurity, Technology News and Support." https://www.bleepingcomputer.com/ (accessed Apr. 17, 2023).

[6] A. Dar, S. Abdullahi, and A. A. Ahanger, "The Silent Art of Reconnaissance: The Other Side of the Hill," Journal of Computer Networks and Communications, 2018.

[7] C.-M. Mathas and C. Vassilakis, "Reconnaissance," Cyber-Security Threats, Actors, and Dynamic Mitigation, pp. 27–80, Apr. 2021, doi: 10.1201/9781003006145-2.

[8] K. Kaushik, S. Aggarwal, S. Mudgal, S. Saravgi, and V. Mathur, "A novel approach to generate a reverse shell: Exploitation and Prevention," International Journal of Intelligent Communication, Computing and Networks Open Access Journal, pp. 2582–7707, doi: 10.51735/ijiccn/001/33.

[9] V. R. Saraswathi, I. S. Ahmed, S. M. Reddy, S. Akshay, V. M. Reddy, and S. M. Reddy, "Automation of Recon Process for Ethical Hackers," 2022 International Conference for Advancement in Technology, ICONAT 2022, 2022, doi: 10.1109/ICONAT53423.2022.9726077.

[10] S. Sharma and K. Shah, "Expedite the Process of Reconnaissance: Eagle's Eye of Security," 2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS), pp. 1–5, Jun. 2022, doi: 10.1109/IC3SIS54991.2022.9885343.