Shodan Search Engine

Author: Paula A. Nevárez Román
Advisor: Dr. Jeffrey Duffany
Electrical & Computer Engineering and Computer Science

Abstract

Search engines have become a crucial component of the Internet. Shodan is a unique type of search engine. Unlike conventional search engines that index the World Wide Web, Shodan offers users the ability to explore all publicly accessible devices on the Internet. Shodan explores the realm of IoT devices, servers, webcams, and more, providing valuable insights into the digital landscape. This paper delves into Shodan's functionalities, including IP scanning and other services, highlighting its applications in network security, market research, and cyber risk assessment. However, alongside its benefits, Shodan also poses significant risks due to its accessibility to anyone with internet access and basic search query knowledge. This paper emphasizes the responsible and ethical use of Shodan to prevent security vulnerabilities and privacy breaches, providing a comprehensive analysis of its features, implications, and ethical considerations in today's interconnected digital landscape.

Introduction

Search engines are essential tools used daily for personal and work-related tasks, enabling users to search and retrieve data from the World Wide Web. They operate in three phases: crawling, indexing, and ranking, to analyze, organize, and present relevant search results. [1] Early search engines like Archie, and Yahoo! paved the way for modern ones like Google. [2] Unlike common search engines, Shodan focuses on indexing information from internet-connected devices rather than just web pages, offering a broader view of the Internet but requiring search query and programming knowledge. [3]

Background

Shodan, created by computer scientist John Matherly in 2009, originated from his idea to track internet-connected devices. Named after a character from System Shock, it collects data from devices via their publicly-available metadata, known as banners, using ports. Shodan requires users to understand its search query syntax, involving banners and search filters, to effectively retrieve information about devices and their vulnerabilities. The platform supports detailed searches with no limit on filters applied, enhancing its utility for security purposes. [4]

Shodan's features vary based on account type, from free to academic and paid memberships, offering different levels of query credits, scan credits, and IP monitoring capabilities. Users can leverage numerous tools, such as Shodan Monitor, Trends, and Private Firehose, for comprehensive network analysis and real-time data access. Additionally, Shodan provides detailed IP lookups, batch IP lookups, bulk data access, and an Internet scanning API, making it a versatile resource for understanding and securing internet-connected devices. With advanced features like InternetDB and full firehose access, Shodan offers extensive capabilities for cybersecurity professionals and researchers. [5]

Problem

Shodan offers valuable services in areas like network security, market research, cyber risk assessment, IoT tracking, and ransomware impact measurement. However, its public accessibility poses significant security risks to companies and individuals, as it can be misused for malicious purposes despite legal restrictions in some US states.

Methodology

The methodology of this paper is structured into three phases. The first phase involves an extensive literature review and background research, examining academic journals, industry reports, technical documentation, and credible internet sources to establish a comprehensive understanding of Shodan, including its historical development, technical aspects, and practical applications. The second phase employs a hands-on approach to apply the acquired knowledge by creating a Shodan account and actively using the platform to conduct searches, analyze results, and explore its various functionalities, thus gaining deeper insights into its user interface, search syntax, and CLI. The final phase entails a critical analysis of the ethical and legal frameworks associated with Shodan's use, including privacy laws, data protection regulations, cybersecurity ethical guidelines, and industry standards. This phase emphasizes the importance of ethical conduct, data privacy protection, and legal compliance when using tools with significant privacy and security implications.

Results and Discussion

In this project, the search feature of Shodan was tested. I wanted to search if any of the systems in my household were compromised which is why the search was based on the cameras exposed in the Internet that were located in Puerto Rico and the organization was Liberty. I started by making a comparison between the results obtained of two simple searches that will look for banners that have “camera” and “webcam” in their data property. Both of this searches include information like the amount of results obtained, the amount on each of the top countries, ports, organizations, products, and operating systems. Also, they provide access to reports, results, historical trends, images and maps. All the results are not shown because of the amount of results obtained. After the initial searches, I opted to focus on the camera search because it provided more results to be evaluated. Figure 1 shows the search results for “camera”.

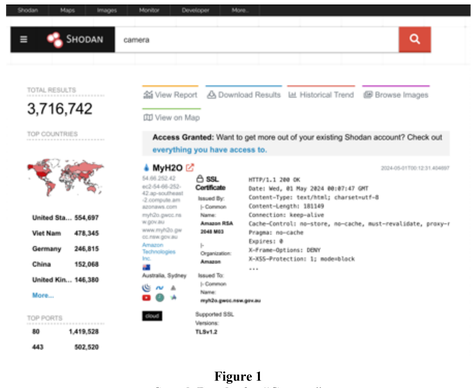

Figure 1
Search Results for “Camera”

When opening the browse images tab and you click on an image images of the camera feed are shown. If you click on an image you can find all it’s data. In Figure 2 we can see part of the results including a regular view and the raw data which is no other than the general information written in code. This includes hostnames, domains, country, city, organization, ISP, and ASN. It also provides all the open ports and detailed information on them.

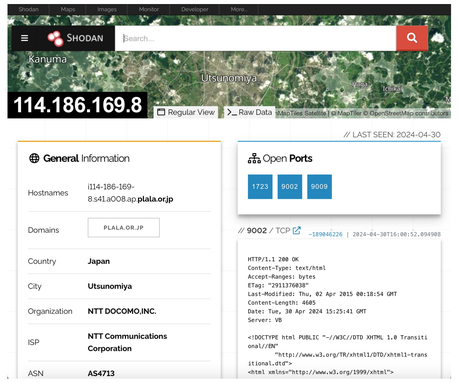

Figure 2
Image Data for “Camera” Search

Figure 3 contains the historical trends of the search of cameras exposed in the Internet that were located in Puerto Rico and the organization was Liberty are shown. This section shows a time line going back to 2017. It includes an analysis of the amount of results obtained during that period of time. This analysis shows how much the internet has grown and how security has increased. The world map section provides a similar analysis but instead of using graphics it is focused on how the numbers have increased or decreased along time with a live graphic.

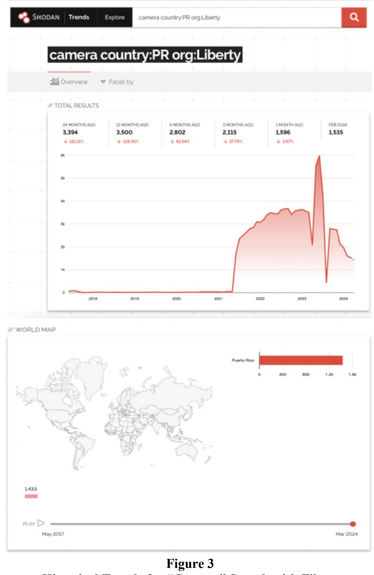

Figure 3
Historical Trends for “Camera” Search with Filters

I tested Shodan's IP address search feature, which provides information like country, city, organization, and open ports. However, the results provided in figure 4 didn't match the previous searches results seen in figure 2 due to discrepancies in open ports. Additionally, attempting to search for my private IP address was unsuccessful due to privacy restrictions.

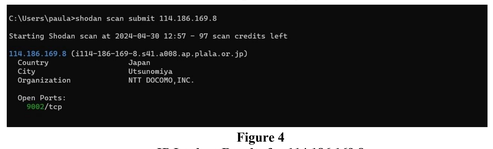

Figure 4
IP Lookup Results for 114.186.169.8

Conclusions

In conclusion, Shodan provides their users a glimpse into the interconnected world of Internet of Things (IoT) devices by offering unique insights and capabilities. Its distinctive features, such as the ability to search for specific devices, ports, vulnerabilities, and IP addresses makes it an extremely powerful tool for cybersecurity professionals, researchers, and individuals alike.Shodan contributes to a better understanding of the complex nature and potential risks associated with our interconnected devices by allowing users access to information that may not be easily accessed through traditional search engines.

However, in addition to its amazing features, Shodan also presents major concerns surrounding security, privacy, and ethical standards. Shodan's IoT device indexing approach allows their users to uncover sensitive information about individuals, companies, and infrastructure. This presents a serious risk, especially when it is being misused or exploited by malicious actors for unauthorized access, data breaches, or surveillance purposes. For this reason, the responsible use of Shodan and similar tools requires strict adherence to ethical guidelines, data protection measures, and cybersecurity best practices.

The constantly evolving world of IoT and interconnected devices emphasize the importance of creating a balance between innovation and security. While Shodan offers extremely valuable insights into the state of our interconnected world, its potential risks cannot be ignored. Therefore, as we continue to leverage advanced technologies like Shodan, it is imperative to prioritize cybersecurity awareness, education, and proactive measures to safeguard individuals, companies, and critical infrastructure from potential vulnerabilities and threats in the digital age.

Future Work

While this paper on Shodan provides valuable insights into its functionality and basic features, several advanced features remain to be evaluated in future work. Future research could focus on Batch IP lookups, Bulk Data, InternetDB, Full firehose, Internet scanning API, and Hostname scans. Analyzing these features could involve assessing their effectiveness in cybersecurity, evaluating bulk data quality across industries, and exploring historical IoT data trends. Additionally, it could study real-time data streams for emerging threats and assess the integration of scanning APIs and hostname scans with existing tools.

References

[1] (2024, March 28). *An Introduction to Search Engines and How They Work.* WebAlive. [Online]. Available: https://www.webalive.com.au/an-introduction-to-search- engines/
[2] Rose, C. (2023, July 16). *The Complete History of Search Engines. SEO Mechanic*. [Online]. Available: https://www.seomechanic.com/complete-history-search-
[3] *What is Shodan*? Shodan Help Center. [Online]. Available: https://help.shodan.io/the-basics/what-is-shodan
[4] *Search Query Fundamentals*. Shodan Help Center. [Online]. Available: https://help.shodan.io/the- basics/search-query-fundamentals
[5] *Choose Your Plan*. Shodan Account. [Online]. Available: https://account.shodan.io/billing