Automated Vulnerable Alert System Through Web Scraping from OEM Websites

A PROJECT REPORT

Submitted by

ARUNKUMAR B (411621243007)

GUNASEKARAN D (411621243024)

In partial fulfilment for the award of the degree

Of

BACHELOR OF TECHNOLOGY

IN

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

PRINCE Dr K VASUDEVAN COLLEGE OF ENGINEERING AND TECHNOLOGY,

PONMAR, CHENNAI-600 127

ANNA UNIVERSITY: CHENNAI 600 025

MAY 2025

ANNA UNIVERSITY:: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"Automated Vulnerable Alert System Through Web Scraping from OEM Websites"** is the bonafide of **"ARUNKUMAR B (411621243007)** and **GUNASEKARAN D (411621243024)"**, who carried out the project work under me supervision.

Mrs. M. Vanitha, M.E
HEAD OF DEPARTMENT
Department of AI & Data Science,
Prince Dr. K. Vasudevan College
of Engineering and Technology,
Chennai-600127

Mrs. P. ABIRAMA SUNDARI, M.E.,
SUPERVISOR
Department of AI & Data Science,
Prince Dr. K. Vasudevan college
of Engineering and Technology,
Chennai-6000127

AKOWLEDGEMENT

We wish to express our sincere thanks to our **FOUNDER AND CHAIRMAN, Dr. K. VASUDEVAN, M.A., B.Ed., Ph.D.,** for his endeavour in educating us in his premier institution.

We would like to extend our heartfelt gratitude and sincere thanks to our **VICE CHAIRMAN, Dr. V. VISHNU KARTHIK, M.D.,** for his keen interest in our studies and the facilities offered in this premier institution.

We would like to express our deep gratitude and sincere thanks to our **ADMINISTRATIVE OFFICER, Dr. K. PARTHASARATHY BE.,** for his valuable support.

We wish to express our sincere thanks to our **HONOURABLE PRINCIPAL, Dr. T. SUNDER SELWYN, M.E., Ph.D.,** for permitting to access various resources in the college to complete the project work

We also wish to convey our thanks and regards to our **HOD, Mrs. M.VANITHA M.E.,** Department of Artificial Intelligence and Data Science, for her guidance and support throughout our project.

We wish to express our great deal of gratitude to our Project Guide **Mrs. P. ABIRAMA SUNDARI, M.E.,** Department of Artificial Intelligence and Data Science for the pleasure guidance to finish our project successfully.

We would like to extend our thanks to all teaching and non-teaching staffs of Department of Artificial Intelligence and Data Science for their continuous support.

ABSTRACT

The **"Automated Vulnerability Alert System Through Web Scraping from OEM Websites"** project aims to enhance cybersecurity responsiveness by automating the process of monitoring and notifying users about newly discovered system vulnerabilities. Leveraging web scraping techniques, the system continuously extracts vulnerability data from trusted Original Equipment Manufacturer (OEM) websites such as Microsoft, Cisco, and Adobe. This real-time information is then analysed and forwarded to users via automated email alerts, enabling prompt risk mitigation and patch deployment. The platform reduces dependency on manual tracking, ensuring faster awareness and response to security threats. By integrating automation, data extraction, and notification mechanisms, the project contributes to proactive vulnerability management, promotes digital safety, and supports organizations in maintaining secure IT environments. The system is designed with a focus on reliability, scalability, and user accessibility, making it a valuable tool for institutions aiming to strengthen their cybersecurity framework with minimal human intervention.

INTRODUCTION

DOMAIN INTRODUCTION

1. Cybersecurity:

Cybersecurity is the practice of protecting computer systems, networks, and data from unauthorized access, attacks, damage, or theft. In today's digital age, where businesses, governments, and individuals heavily rely on interconnected systems, ensuring robust cybersecurity has become essential. Cyber threats, including malware, ransomware, phishing, and zero-day vulnerabilities, can lead to serious consequences such as data breaches, financial loss, and reputational damage. As the threat landscape continues to evolve rapidly, proactive security measures, real-time threat detection, and timely mitigation strategies are crucial for protecting digital infrastructure. One of the key components in a strong cybersecurity strategy is timely awareness of known vulnerabilities in software and hardware systems, which allows for preventive actions before exploitation occurs.

2. Automation:

Automation involves the use of technology to perform tasks with minimal human intervention. In cybersecurity, automation is vital for reducing manual workloads, improving response times, and ensuring consistency in monitoring and mitigation activities. From scanning systems for vulnerabilities to generating alerts and reports, automation can streamline the entire process of threat detection and management. Automated systems can work continuously to gather data, detect anomalies, and respond to potential threats—ensuring better security coverage, especially in large-scale or resource-constrained environments.

3. Web Scraping:

Web scraping is a technique used to extract data from websites using automated tools or scripts. It enables the collection of publicly available information from various sources at scale. In the context of this project, web scraping is employed to retrieve real-time updates on security vulnerabilities directly from the official websites of Original Equipment Manufacturers (OEMs) such as Microsoft, Cisco, or Adobe. Since many critical vulnerability disclosures are first published on OEM portals, automating the extraction and monitoring of this information ensures organizations stay updated without relying solely on manual tracking.

4. Alert Systems and Email Notifications:

An alert system is a notification mechanism designed to inform users about significant events or conditions that require immediate attention. When integrated with cybersecurity systems, automated alert systems help security teams respond quickly to newly discovered threats or vulnerabilities. In this project, the alert system is linked with email notifications, enabling instant dissemination of critical vulnerability information to subscribed users. This allows for swift decision-making and patch management, significantly reducing the window of exposure to potential attacks.

PROBLEM DEFINITION

Critical sector organizations face increasing cybersecurity threats due to reliance on IT and OT systems. Timely identification of vulnerabilities is essential to mitigate risks, but existing manual methods for tracking vulnerabilities are slow and inefficient. This project aims to develop an automated system that scrapes vulnerability information from OEM websites and trusted sources, processes it in real-time, and sends immediate alerts with mitigation strategies to stakeholders via email. The goal is to reduce response time, improve awareness, and strengthen cybersecurity posture by automating vulnerability detection and communication, ensuring that organizations stay ahead of potential threats.

PROBLEM DESCRIPTION

In today's digital landscape, critical sector organizations operate complex IT and OT infrastructures that are frequently targeted by cyber threats. Vulnerabilities in software and hardware components—especially those identified by Original Equipment Manufacturers (OEMs)—pose serious risks if not addressed promptly. Currently, many organizations depend on manual monitoring or delayed third-party updates, which increases the window of exposure to cyber-attacks. Furthermore, the dynamic nature of OEM websites and the absence of standardized alerting mechanisms make consistent tracking of vulnerabilities a challenging task.

To solve this, our project proposes an Automated Vulnerable Alert System that uses web scraping techniques to monitor OEM websites and other reputable sources in real time. By leveraging tools like Python's Beautiful Soup, Scrapy, and Selenium, the system extracts vulnerability data, organizes it, and sends instant alerts via email to relevant stakeholders. Each alert includes not just the vulnerability details, but also recommended mitigation actions.

This system minimizes human intervention, speeds up response time, and enables proactive cybersecurity defence. It is especially vital for sectors like energy, healthcare, and finance, where system downtime or breaches can lead to catastrophic consequences. By automating this process, the system enhances both efficiency and resilience.

EXISTING SYSTEM:

In the current setup, most critical sector organizations rely on manual processes or third-party platforms to track cybersecurity vulnerabilities. Security teams often visit OEM websites or subscribe to vendor newsletters and vulnerability databases to stay updated. This method is time-consuming, inconsistent, and prone to delays, as it depends heavily on human intervention and periodic checks. Additionally, vulnerability data is often unstructured and lacks immediate guidance on mitigation. Without real-time alerts or centralized reporting, organizations struggle to respond quickly to emerging threats, which increases their exposure to cyber-attacks and compromises the integrity of IT and OT infrastructure.

DRAWBACKS:

1. Manual Monitoring and Delayed Detection:

Most organizations rely on security personnel to manually visit OEM websites or check vulnerability databases. This process is slow and inconsistent, leading to delays in detecting critical vulnerabilities, which can be exploited before any action is taken.

2. Lack of Real-Time Alerts:

In the absence of an automated alert mechanism, organizations do not receive instant notifications when a new vulnerability is published. This delay reduces

the time available for remediation and increases the window of exposure to threats.

3. Scattered Information Across Multiple Sources:

Vulnerability data is dispersed across various OEM websites, blogs, and forums. Constantly tracking all these sources is impractical and leads to missed or overlooked security announcements.

4. Unstructured Data Format:

Most vulnerability information is available in plain text or embedded in HTML, without a standardized format. This makes it difficult for teams to extract, analyse, and act upon the data quickly, often requiring manual parsing or interpretation.

5. Lack of Mitigation Guidance:

Even when vulnerabilities are found, there is often no direct inclusion of mitigation steps or patches in the alert. Security teams must then spend additional time researching solutions, delaying the remediation process.

6. Heavy Dependence on Human Effort:

The existing system requires continuous manual effort to monitor, extract, validate, and communicate threat information. This not only increases the workload but also raises the risk of human error or oversight in fast-changing threat landscapes.

7. Poor Scalability and Inconsistent Coverage:

As the number of software systems and devices grows, manual methods become increasingly unsustainable. Organizations struggle to keep up with the volume of updates and often miss out on less-publicized vulnerabilities that could be just as dangerous.

PROPOSED SYSTEM AND ADVANTAGES:

1. Real-Time Vulnerability Alerts:

The system continuously monitors OEM websites and sends immediate alerts when new vulnerabilities are published. This ensures timely awareness and faster response, reducing the risk of exploitation.

2. Reduced Human Effort:

By automating the process of data collection and alert distribution, the system significantly minimizes the need for manual monitoring. Security teams can focus more on analysis and response rather than repetitive tasks.

3. Centralized Vulnerability Tracking:

Instead of visiting multiple OEM sites individually, all relevant data is consolidated into a single system. This simplifies monitoring and ensures no critical update is missed.

4. Structured and Cleaned Data Output:

The scraped information is parsed and formatted into structured formats like JSON or CSV, making it easy to store, analyze, and share. This enhances clarity and speeds up decision-making.

5. Integration with Email Notification System:

The system is equipped with an email module to instantly notify stakeholders about high-severity threats. These emails can include detailed descriptions, affected systems, and recommended mitigation steps.

6. Scalability and Flexibility:

The system is modular and can scale to monitor multiple sources simultaneously. It can also be configured to scrape new websites or add support for additional file formats as needed.

7. Mitigation Strategy Included:

Along with vulnerability alerts, the system provides relevant mitigation suggestions or patch links (when available). This helps organizations not just detect, but also act on threats more effectively.

8. Reduces Response Time:

Faster detection and communication mean quicker patch deployment or defensive measures. This minimizes the potential damage or disruption caused by known vulnerabilities.

9. Adaptable to Dynamic Website Structures:

The use of tools like Selenium and Scrapy allows the system to handle JavaScript-rendered or dynamically loaded content, which many modern OEM sites use. This ensures reliability even with complex websites.

10. Supports Cybersecurity Compliance and Auditing:

The system's logs and alert history can be archived for auditing purposes. This helps organizations maintain transparency, demonstrate compliance, and improve cybersecurity governance.

Overall, the proposed system offers a smart and automated way to keep critical organizations safe from cyber threats by instantly alerting them about new vulnerabilities. It removes delays seen in manual tracking and ensures accurate, real-time updates. The system is easy to use, cost-effective, and sends alerts directly to stakeholders through email. It saves time, reduces human effort, and helps organizations take quick action before attackers can exploit weaknesses.