Analysing Phishing Techniques That Bypass Spam Detection

Cybersecurity, Email Security

CDAC, Noida

CYBER GYAN VIRTUAL INTERNSHIP PROGRAM

Submitted By:

Rahul Kumar Sinha

Project Trainee, (July-August) 2025 BIT SINDRI, DHANBAD
(Information Technology)

BONAFIDE CERTIFICATE

This is to certify that this project report entitled "**Analysing Phishing Techniques That Bypass Spam Detection**" submitted to CDAC Noida, is a Bonafede record of work done by "**Rahul Kumar Sinha**" under the supervision of Project Supervisor "**Mr. Varun Mishra**" from 1StJuly,2025 to 6thAug, 2025.

Declaration by Author

This is to declare that this report has been written by me/us. No part of the report is plagiarized from other sources. All information included from other sources have been duly acknowledged. I/We aver that if any part of the report is found to be plagiarized, I/we are shall take full responsibility for it.

Name of Author: RAHUL KUMAR SINHA

TABLE OF CONTENTS

ACKNOWLEDGEMENT

It is with a profound sense of gratitude that I acknowledge the contributions of several individuals who have made the successful completion of this project possible.

Firstly, I extend my sincere gratitude to **Mr. Varun Mishra**, my project supervisor, for their invaluable guidance, constant encouragement, and insightful feedback throughout the course of this project. Their expertise was instrumental in shaping and completing this project work.

I am also grateful to the faculty members of **Information Technology** at **Birsa Institute of Technology, Sindri, Dhanbad** for providing necessary facilities and fostering a supportive academic environment. Their teachings and advice have been instrumental in shaping my understanding of the subject matter.

I would like to acknowledge the contributions of the developers and researchers behind the success of the projects completed by **CDAC,Noida** which have been pivotal in the implementation and success of the Projects.

Lastly, I wish to extend my deepest appreciation to my family for their unwavering love, support, and patience, without which this endeavour would not have been possible.

Thank you.

Analysing Phishing Techniques That Bypass Spam Detection

PROBLEM STATEMENT

- This project aims to thoroughly investigate the ways in which phishing attacks evade spam detection filters, a critical challenge in maintaining email security. Despite the advancements in spam detection technologies, attackers continually adapt their techniques to bypass these defences, making it essential to understand their methods and limitations.

- The project will begin with an analysis of common techniques employed by phishers. Phishers often craft highly convincing email content that mimics legitimate organizations, utilizing psychological manipulation to reduce suspicion. Additionally, they use technical methods like embedding malicious code within attachments or using image-based messages to evade text-based filters.

- Next, the project will explore the limitations of current spam detection methods. Most filters rely on keyword analysis, reputation scoring, and pattern recognition, which can be circumvented by adapting techniques such as dynamic content generation and URL obfuscation. Machine learning models, although powerful, can be limited by their training data, leading to false negatives or positives. Further, many filters struggle to keep pace with the ever-evolving tactics of cybercriminals, highlighting the need for adaptive and multilayered security strategies.

- Finally, the project seeks to identify potential improvements to enhance email security. Overall, this investigation aims to contribute to developing more robust, intelligent, and adaptive spam filters that can better identify and block sophisticated phishing attacks, thereby safeguarding users and organizations from potential cyber threats.

Learning Objectives

- **Enhanced Understanding:**

  This project focuses on acquiring in-depth knowledge of the various techniques used by cybercriminals to conduct phishing attacks. By dissecting how phishing messages are crafted attachments, and image-based content—the study aims to understand the underlying mechanics of these tactics. This comprehensive analysis will reveal how these techniques are designed to deceive recipients and evade detection by spam filters. Understanding the intricacies of phishing methods enables security professionals develop more effective countermeasures, and stay ahead of cybercriminals who continually refine their approaches.

- **Improved Detection:**

  A critical aspect of the project involves identifying the vulnerabilities and weaknesses in existing spam detection systems. This entails scrutinizing current technologies—such as keyword filtering, reputation scoring, machine learning models, and heuristic analysis—and understanding how cybercriminals exploit these systems to bypass security. By developing and testing these improvements, the goal is to create more resilient and intelligent email security systems capable of accurately detecting and blocking even the most sophisticated phishing campaigns.

- **Awareness for Users & Organizations:**

  "**Building awareness is essential for preventing successful phishing attacks**". This part of the project aims to develop practical, user-friendly strategies to educate individuals and organizations about the nature of phishing threats. This includes designing targeted training programs, awareness campaigns, and clear guidelines on recognizing suspicious emails. By fostering a culture of vigilance and proactive security measures, organizations can significantly reduce their vulnerability to phishing attacks and enhance their overall cybersecurity posture.

APPROACH WITH IMPLEMENTATION

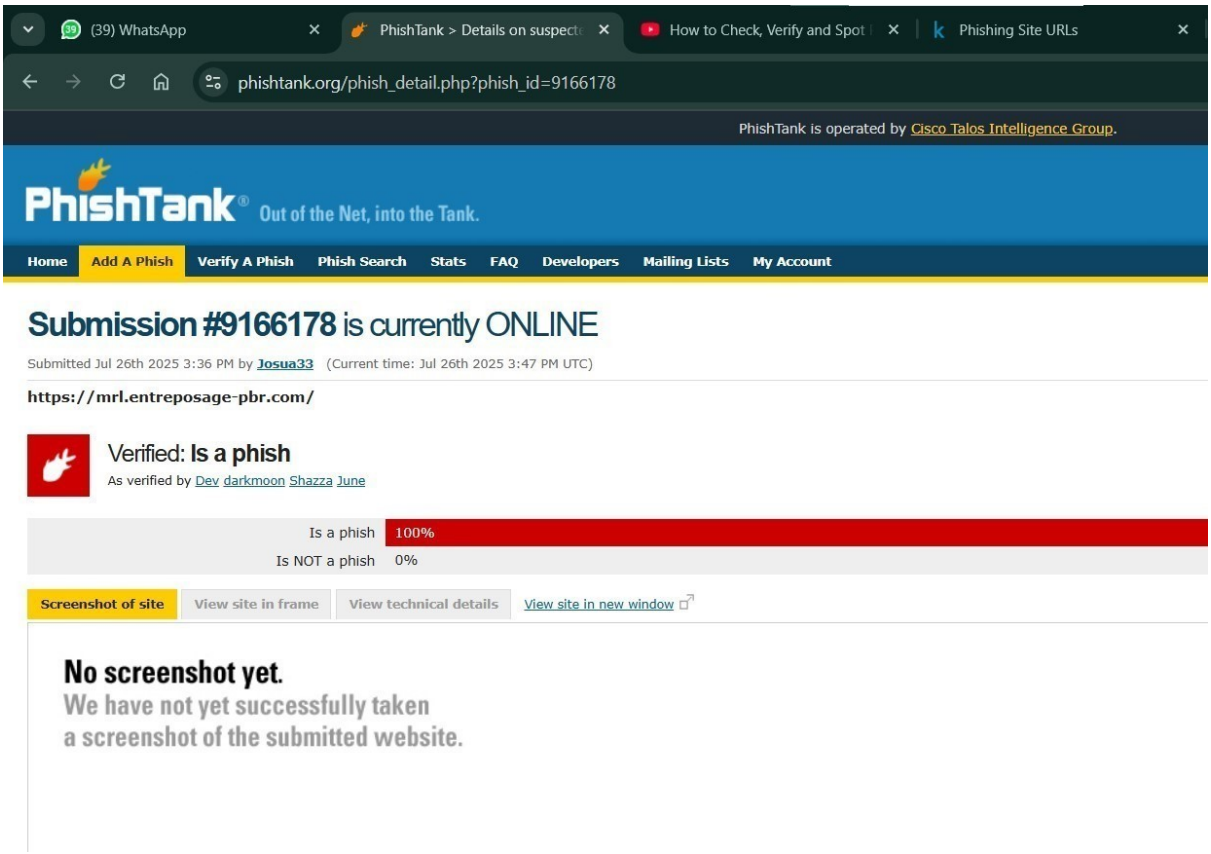
Firstly, lets analyse some of the methods by which we can analyse and stay safe from the phishing Techniques: -

1. Phishing Tank.org

- PhishingTank.org is an online platform designed to educate users, cybersecurity professionals, and organizations about phishing threats through interactive simulations and real-world examples. It provides a comprehensive database of phishing examples, tools for analysing phishing techniques, and resources to help users recognize and prevent phishing attacks.

- The platform aims to raise awareness about the evolving tactics used by cybercriminals and strengthen defences against such threats**.**

- **Examples;**



- Fig: - **The above figure shows the usage of this website**.

- # Advantages Of The Phishing Tank:

1. **Educational Resource:** PhishingTank.org offers detailed information on various phishing techniques, including email spoofing, malicious links, fake websites, and social engineering tactics.

2. **Real-World Examples:** The platform maintains a database of actual phishing emails and websites, illustrating common tactics, patterns, and security flaws. Analysing these examples helps users recognize similar threats in their environments.

3. **Tools and Resources:** It offers various tools for analysing URLs, email headers, and phishing techniques, enabling cybersecurity professionals to better understand attack vectors and develop targeted defences.

4. **Community and Updates:** PhishingTank.org often collaborates with cybersecurity communities and provides updated threat intelligence, ensuring users stay informed about the latest phishing scams and trends.

2. Open Web Application Security Project (OWASP):

- OWASP is a globally recognized nonprofit organization dedicated to improving the security of software applications. It provides a wide range of free and open-source resources, including tools, documentation, guidelines, and best practices, to help developers, security professionals, and organizations identify and mitigate security vulnerabilities in web applications.

- OWASP's mission is to make software security visible so that developers and organizations can make informed decisions about application security.

OWASP ZAP (Zed Attack Proxy):

- OWASP ZAP is an open-source security testing tool designed to help find vulnerabilities in web applications during development and testing phases.

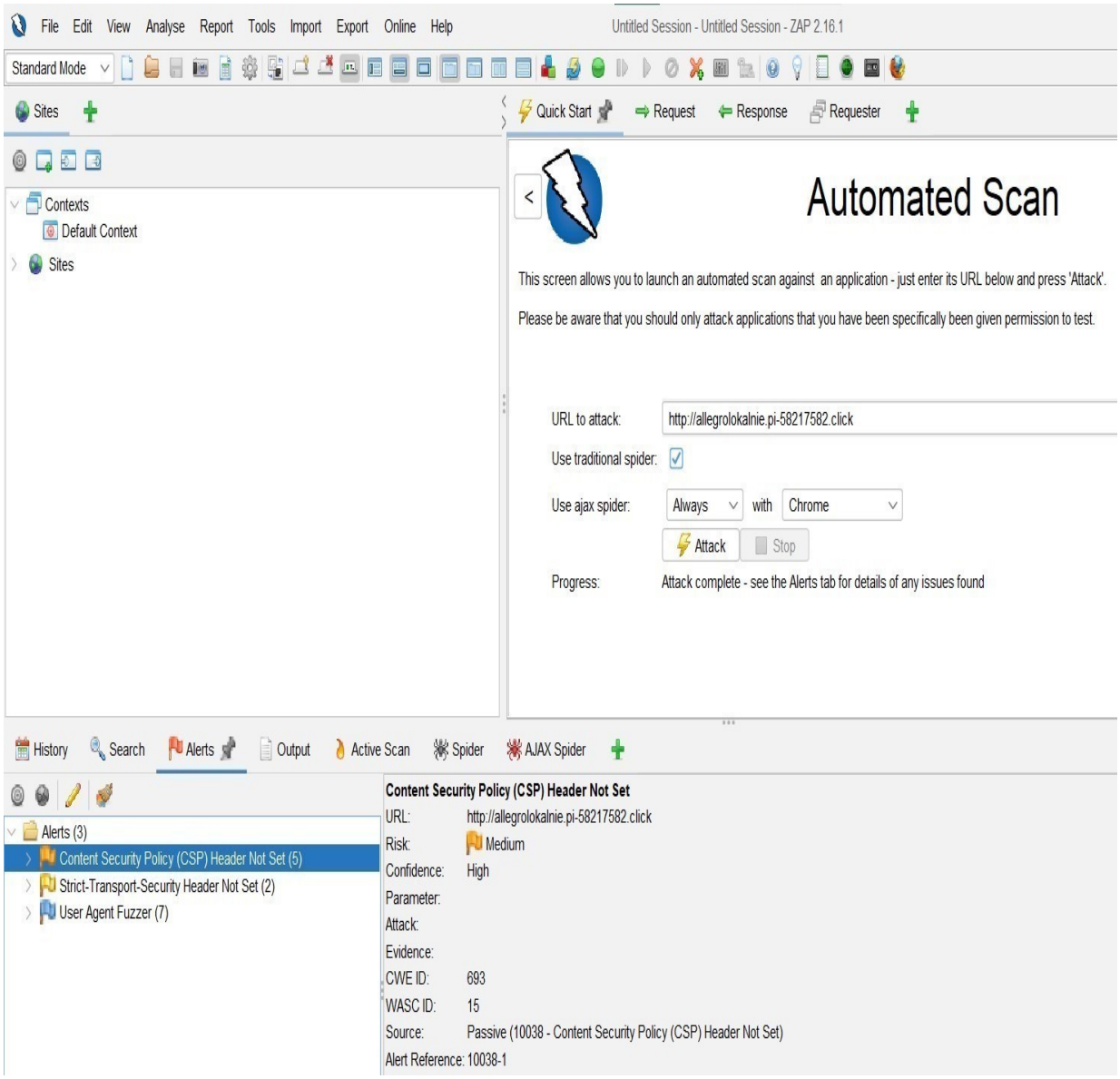
- It is widely used by security professionals and developers for automated and manual security assessments.

- ZAP offers features such as intercepting proxy, automated scanners, various attack tools, and a user-friendly interface that makes it accessible for both experts and beginners.

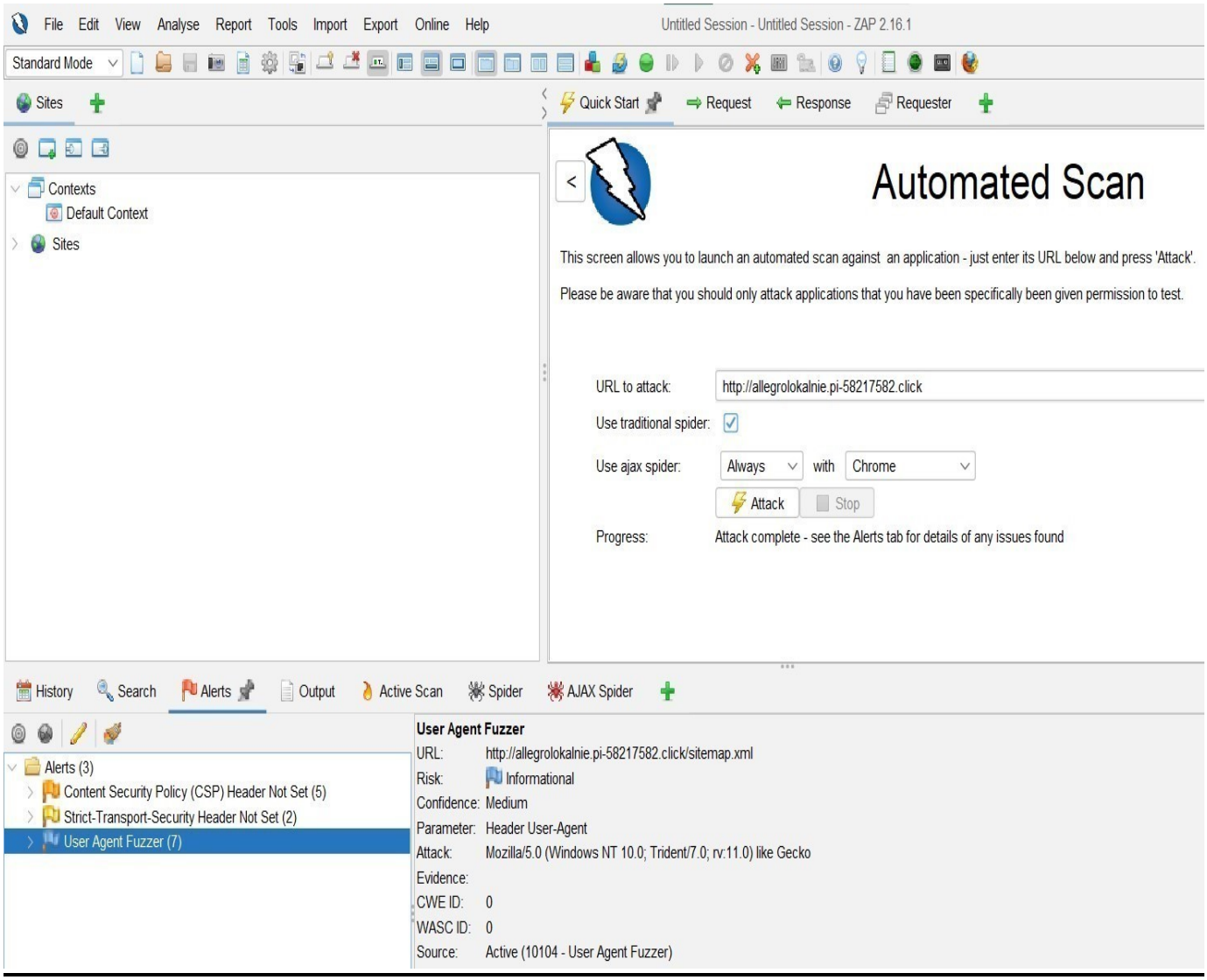
- Its primary purpose is to identify security flaws like SQL injection, Cross-Site Scripting (XSS), and other common web vulnerabilities, enabling organizations to strengthen their applications against potential threats.

- Hence, it becomes one of the most important open-source tools for web applications security testing.

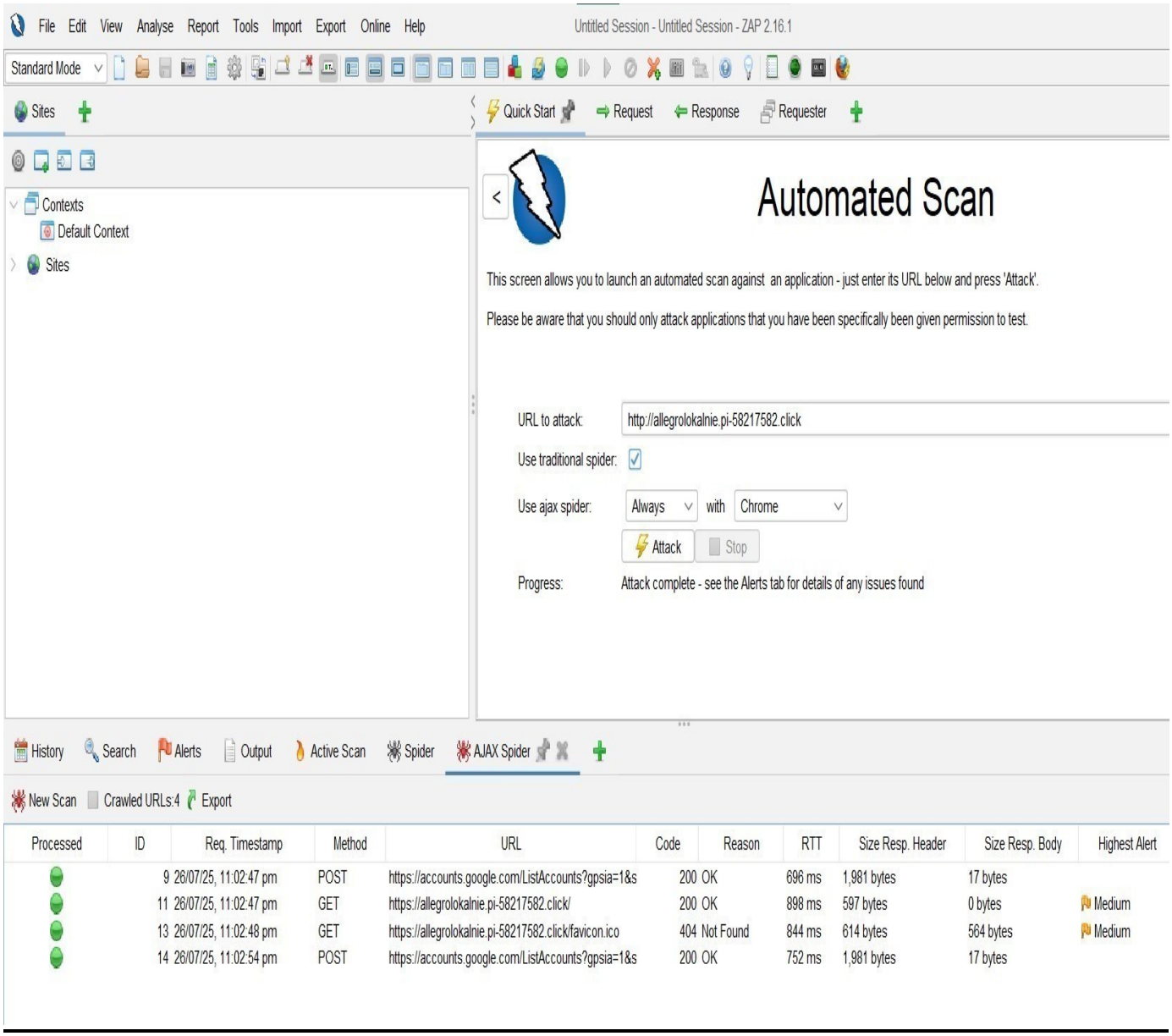
- **Examples of usage of OWASP(ZAP): -**



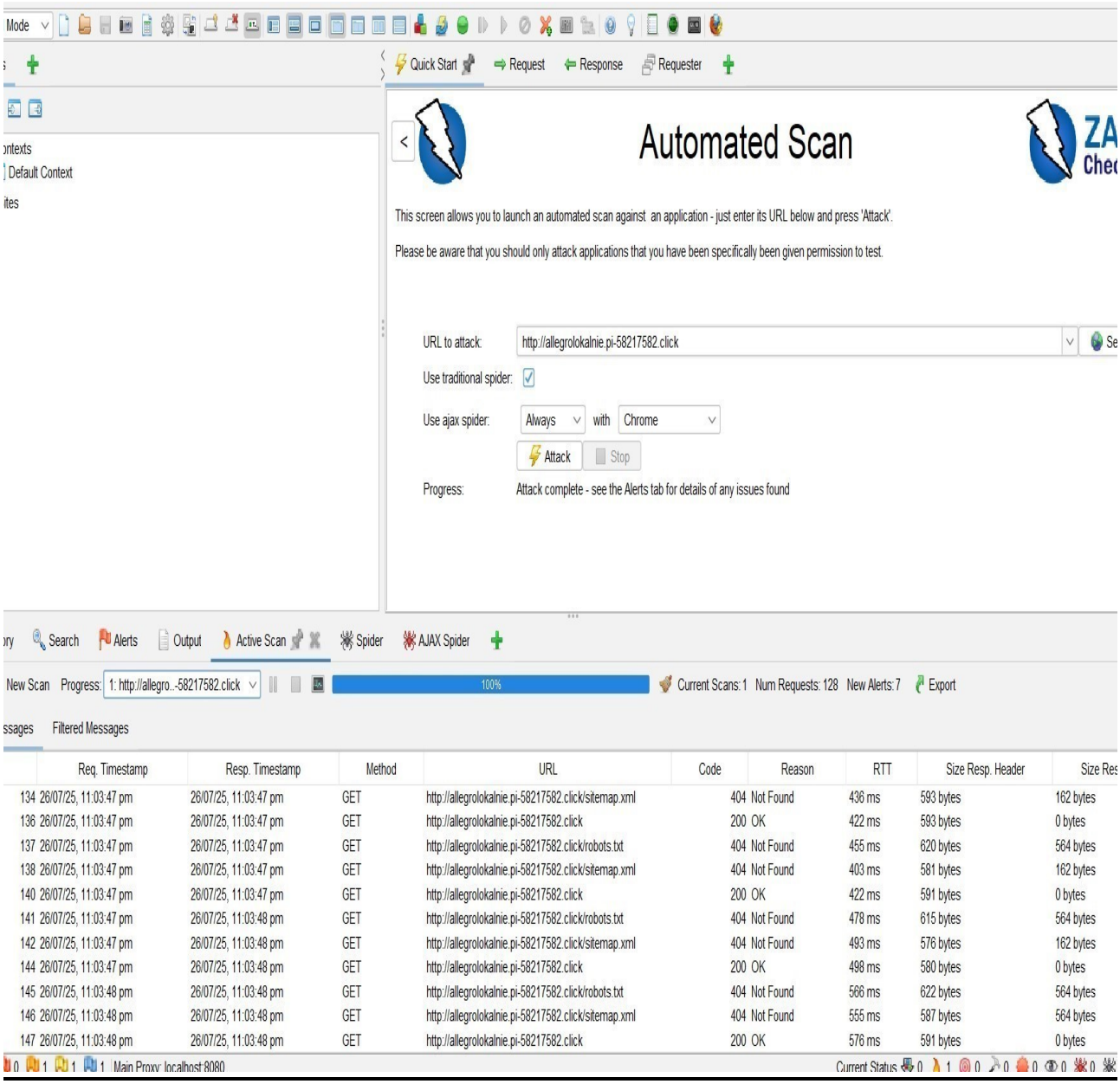
- Fig:- **Getting the Content Security Policy(CSP) using the OWASP ZAP.**



- Fig:- **OWASP ZAP giving the User Agent Fuzzer of the malicious URL.**



- Fig:- **OWASP ZAP giving the AJAX Spider of the Malicious URL.**



- Fig:- **OWASP ZAP giving the active scan of the malicious URL.**

3. Anti-Phishing Working Group(APWG):

- The Anti-Phishing Working Group (APWG) is a global industry- and government-driven organization dedicated to combating phishing, cybercrime, and email-based fraud. Established in 2003.

- The APWG serves as a collaborative forum where cybersecurity experts, law enforcement agencies, financial institutions, technology companies, and researchers come together to share information, intelligence, and best practices.

- The primary goal of the APWG is to coordinate efforts to reduce the prevalence of phishing attacks and mitigate their impact on individuals and organizations.

- They achieve this by collecting and analyzing phishing attack data, issuing alerts and warnings, and developing effective countermeasures.

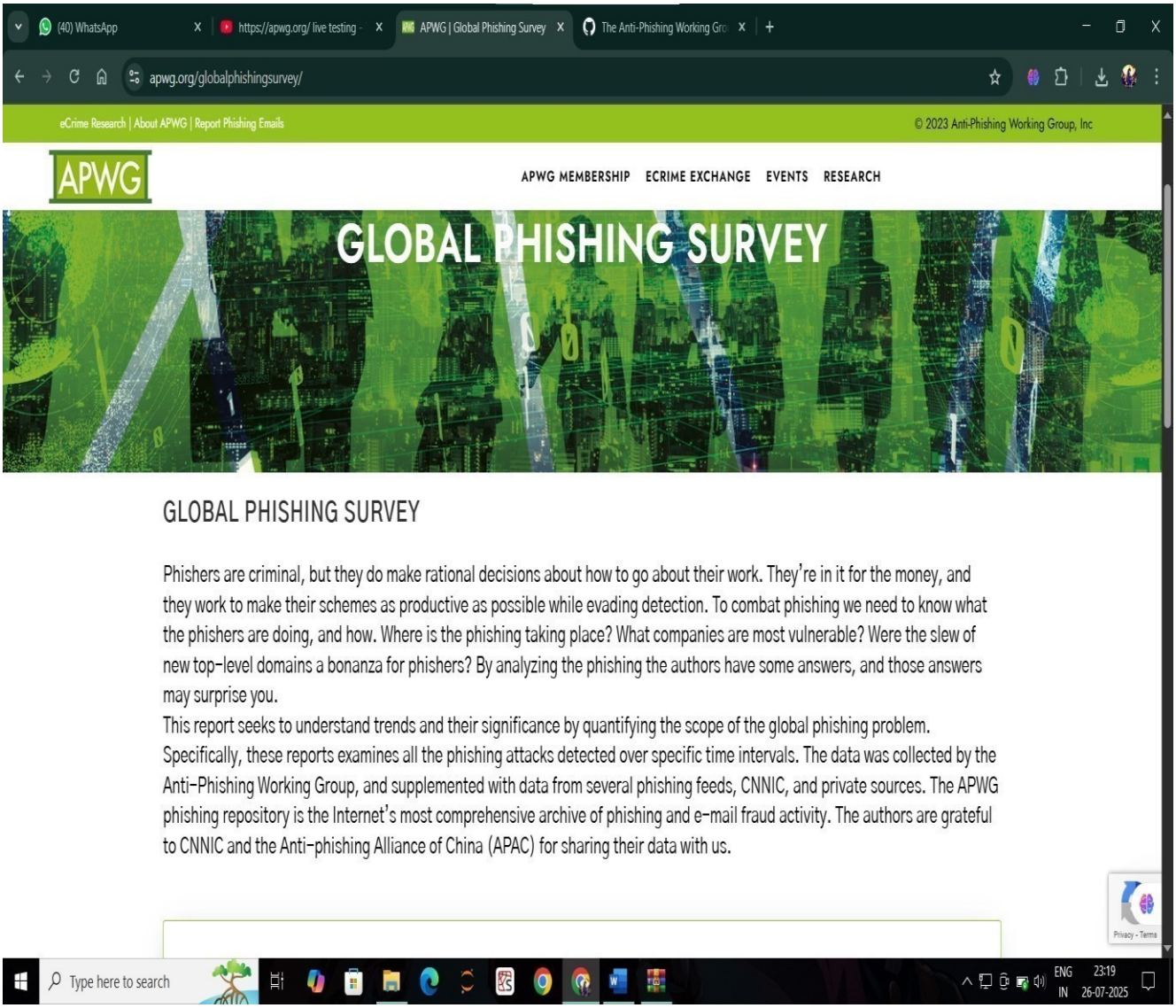
- The organization also conducts research and provides educational resources to raise awareness about phishing tactics and defenses.

- One of the notable contributions of the APWG is the publication of reports, such as the quarterly Phishing Activity Trends report, which offers insights into the evolving landscape of phishing threats.

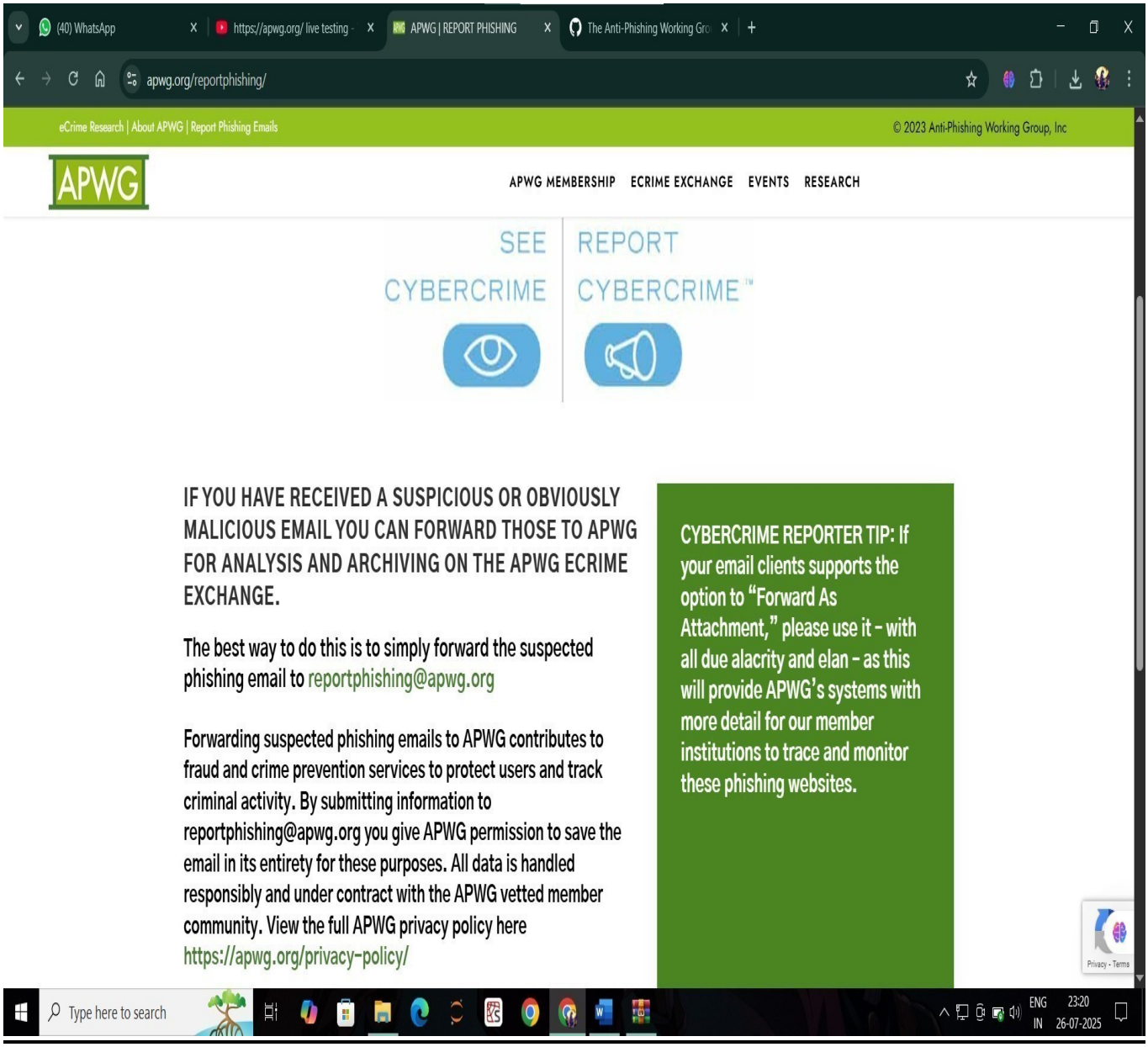
- The APWG maintains various working groups that focus on specific issues such as email security, malware, and cybercrime law enforcement collaboration.

- Overall, the APWG plays a crucial role in strengthening global responses to phishing and cyber threats by fostering collaboration among stakeholders and promoting proactive security measures.

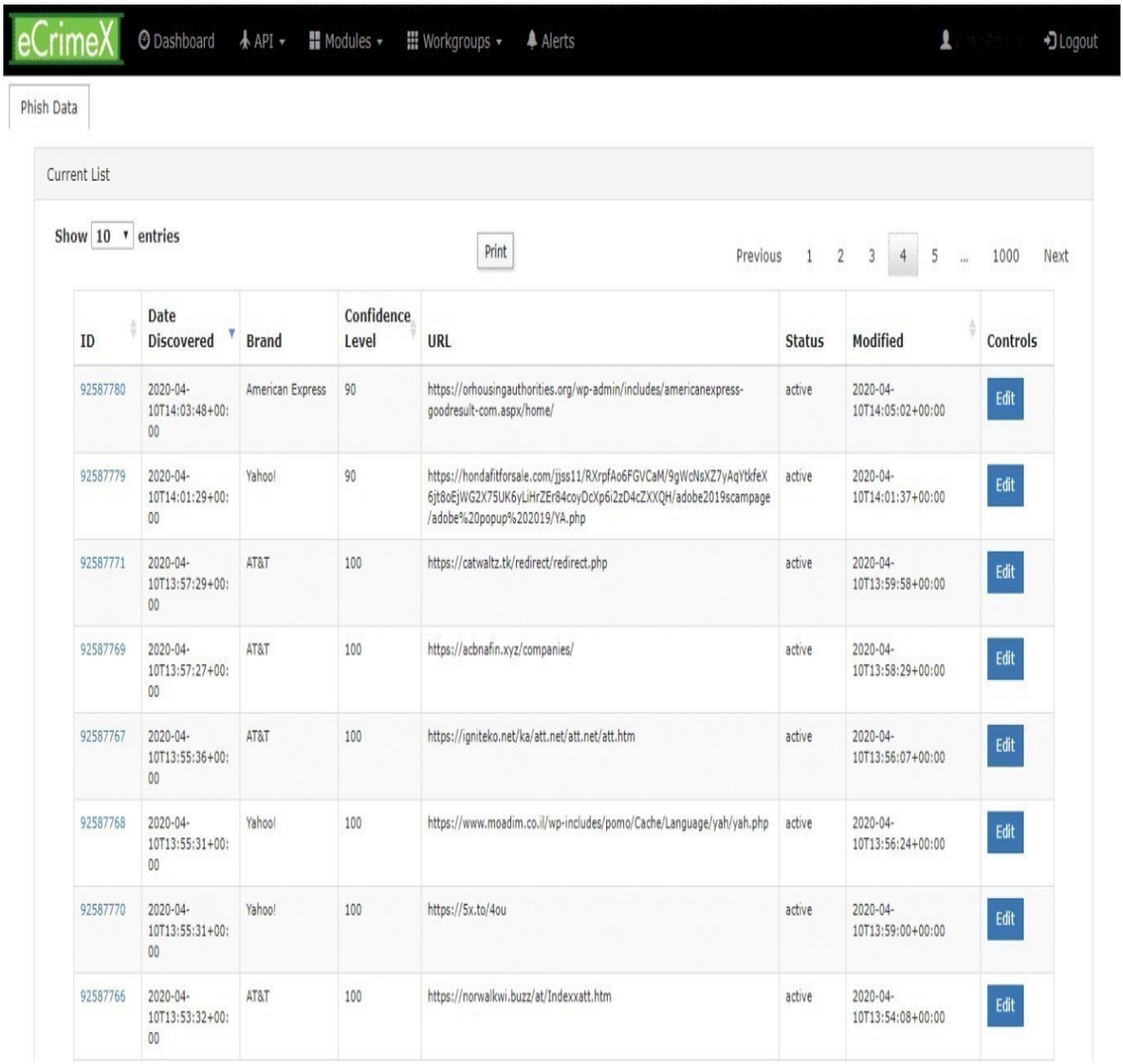
- **Some of the works of the Anti-Phishing Working Group (APWG) are as the follows: -**



- **Fig:- The APWG conducting the Global Phishing Survey.**



- **<u>Fig:-</u>The APWG Combating against the phishing and cybercrime by giving the victims space for putting their problems and taking effective measures against the crime.**



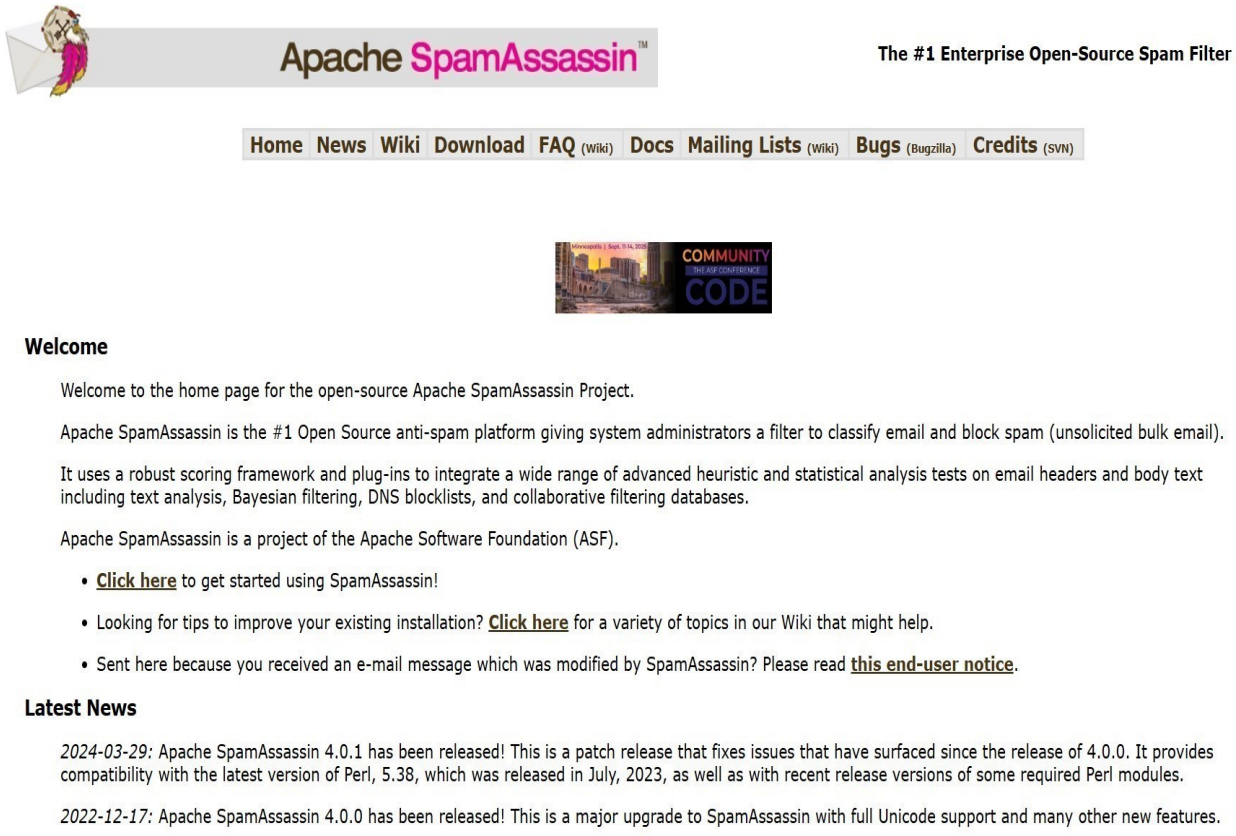
- Fig:-_**The APWG providing the phish data to the public so that they could be aware about them and stay conscious about the crimes**.

4. SPAM ASSASSIN: -

- **Spam Assassin**, an open-source spam filtering platform developed by the Apache Software Foundation. Spam Assassin uses a **combination of rulebased scoring, machine learning, and Bayesian filtering techniques to identify and block unwanted email spam**.

- It covers topics such as setting up spam filters, tuning performance, integrating with mail servers, and understanding the scoring system used to evaluate incoming messages.

- Overall, this resource is essential for administrators and developers aiming to deploy or manage Spam Assassin effectively in their email security infrastructure.



- Fig :- **Apache Spam Assassin major website outlook.**

Usage of Spam Assassin: -

- **Spam Assassin** is widely used in email security systems to detect and filter out spam messages effectively.

Its primary usage includes:

1. **Spam Detection and Filtering**: Spam Assassin scans incoming emails based on a set of predefined rules, scoring each message to determine the likelihood of it being spam. Emails that exceed a certain score are marked or rejected, reducing the amount of unwanted email reaching users' inboxes.

2. **Integration with Mail Servers**: It can be integrated with various mail transfer agents (MTAs) such as Postfix, send mail, or Exim, acting as a plugin or filter to automatically evaluate and filter email traffic in real-time.

3. **Custom Rule Creation**: Administrators can create custom rules tailored to their organization's specific needs, enhancing the accuracy of spam detection based on unique patterns or threats.

4. **Training and Adaptation**: Spam Assassin uses **Bayesian filtering which can be trained periodically with users' emails to improve detection of new or evolving spam tactics**.

5. **Email Security Enhancement**: By filtering out spam and malicious emails (such as phishing attempts or malware attachments), Spam Assassin enhances overall email security and prevents users from falling victim to scams.

6. **Automated Maintenance and Updates**: Its open-source nature allows continuous updates and community-driven improvements, ensuring the filter remains effective against emerging spam techniques.

HOW PHISING ATTACKS BYPASS SPAM FILTERS

- Phishing attacks successfully bypass spam filters and network controls by meticulously testing and exploiting organizational defences, emphasizing the importance of proactive validation and testing by security teams.

Methods by which Cyber Attackers By pass the safety measures by which we can save ourselves from phishing attacks are: -

1. Firstly, Bypassing the Firewall / Proxy:

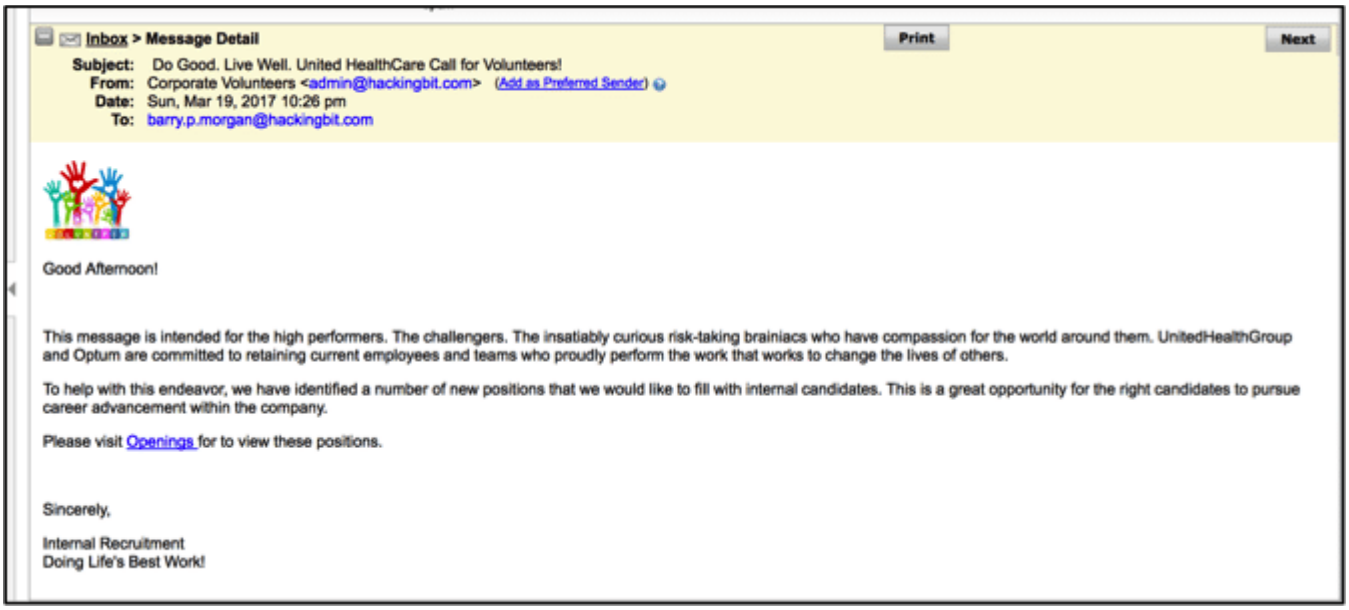
- Web proxy and firewall controls are often similar in their rules, so if an attacker can **create an attack that bypasses spam filters, then they are confident that the network is allowing inbound SMTP, so now they just need to test the firewall and proxy**.

- A common method used for such testing is to **host HTML images on the malicious web server, and embed those images into the phishing message.**

- This allows the attacker **to view the GET request of the embedded image, which confirms that the firewall and proxy are allowing communication back to the malicious web server**.

2. Implementing the Attack: -

- Now the attacker is confident that his messages are bypassing spam filters and the payload is bypassing firewalls and proxies. The image below represents a typical message that has utilized these pre-attack methods.



- Notice the image at the top of the message. Not only does the image add to the authenticity of the message, but it is also providing the phisher a confirmation that the message has been rendered in a user's inbox. With the appropriate attack settings validated, the malicious actor can now develop and deploy a slew of attacks against the organization.

- Below is the attacker's view of his web server logs.

```
03/19 22:27:42 visit from: 71.65.221.46
        Request: GET /volunteer.jpg
        page Serves /home/ubuntu/cobaltstrike/uploads/colourful-volunteer-vector.jpg
        Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_3) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/56.0.2924.87 Safari/537.36
```

- Knowing the level of preparation that hackers are taking, its paramount that security teams perform the same level of preparation and testing.

3. Email Filtering Evasion:

- Criminals use obfuscation techniques, image-based emails, or encoding to bypass spam filters.

4. SSL Encryption:

- Some phishing sites obtain SSL certificates to appear secure and trustworthy.

5.Domain Spoofing:

- Registering similar domain names using homoglyphs (e.g., using Cyrillic characters) to deceive users.

6.Social Engineering:

- Exploiting human psychology over reliance on technological defences.

7.Use of Malware:

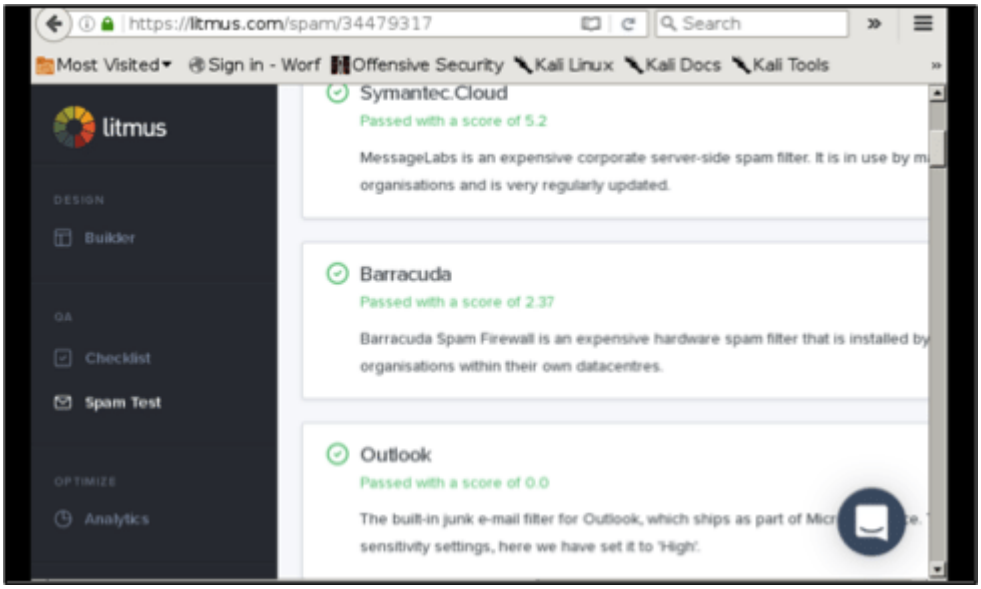
- Employing malware to bypass email filters or security scans.

Protecting from the Phishing Attacks using spam filters

Spam Filters:

- Spam filters are the first line of defence that protects organizations from phishing attacks. These filters can be configured to alert on anything from keywords to untrusted sending domains or IP addresses, depending on your spam-filtering solution.

- The image below shows an example of how one application can be used to test message delivery across multiple major email providers.



- To help counter this attack method, you want to make sure that your organization's spam filters are configured to be as strict as operationally possible. Whenever feasible you should whitelist trusted domains and sandbox all others for human review. At a minimum, your enterprise email solution should be configured to black hole any untrusted messages which will deny the attacker any bounces that may hint to the security stack that you have in place.

- The best way to combat targeted phishing attacks would be to whitelist all authorized domains; of course, this may not be practical for organizations that depend on external customers. At a minimum,

security teams should not allow the download images within messages, which will help prevent the attacker from knowing if his malicious message was delivered.

- Key strategies include configuring strict spam filters, whitelisting trusted domains, and denying requests to newly registered domains, while understanding attacker tactics such as embedding images to confirm message delivery. To mitigate these threats, organizations must enhance technical controls and adopt a security mindset that anticipates attacker methods, thus strengthening defenses against targeted phishing campaigns.

CONCLUSION & RECOMMENDATIONS:

For Ensuring Robust Spam Filter Configuration the followings techniques are adopted.

1. Make Spam Filters as Strict as Operationally Feasible.

To minimize the risk of phishing emails slipping through, your organization should fine-tune spam filters to operate at the highest practical level of strictness. This includes:

- Implementing advanced filtering rules: Use heuristics, anomaly detection, and machine learning-based filters to identify suspicious patterns and behaviors typical of phishing emails.
- Enforcing strict sender verification: Require SPF, DKIM, and DMARC validation for incoming messages to authenticate legitimate sender domains and detect spoofed addresses.
- Disabling automatic delivery for suspicious emails: Configure filters to quarantine or hold suspected messages for manual review, rather than delivering them directly to inboxes.

2.Whitelist Trusted Domains

To reduce false positives and ensure critical communications aren't blocked, maintain a curated list of trusted domains. This helps:

- Streamline legitimate business communications: Trusted domains are automatically delivered, reducing delays.
- Reduce attack surface: By explicitly trusting certain domains, malicious actors are less likely to bypass filters pretending to be trusted sources.

5. Sandbox Untrusted Messages

All messages that don't come from whitelisted domains should be subjected to sandboxing:

- Isolate and analyse suspicious emails: Use sandbox environments to safely render emails, embedded links, and attachments to identify malicious activity before delivery.
- Require human review: emails flagged as suspicious are routed for manual inspection, reducing the likelihood of malicious content reaching users.

6. Blackholing Untrusted Messages

Configure your email security solution to "black hole" or outright reject (reject or drop) messages from unknown or untrusted sources:

- Prevents bounce-back signals: Attackers often use bounce messages to verify email validity, thus denying such replies can help obscure your email infrastructure.
- Reduces information leakage: By not accepting or providing feedback on untrusted emails, you prevent attackers from confirming whether your system accepts certain addresses or domains.

Additional Best Practices

- **Constantly update filter rules and signatures**: Cybercriminal tactics evolve rapidly. Regular updates ensure your filters recognize new attack vectors.

- **Implement multi-layered defences**: Combine spam filtering with URL filtering, malware scanners, and behavioural analysis for comprehensive protection.

- **User education**: Train users to recognize and report phishing emails, reinforcing technical defences.

- **Incident response plans**: Prepare to quickly respond to detected attacks and inform stakeholders.

- **Be cautious with emails and messages:** Don't click on links or open attachments from unknown or suspicious sources.

- **Verify the sender:** Check the email address or sender's details carefully—phishers often use fake addresses that look similar to legitimate ones.

- **Use multi-factor authentication:** Add extra security layers to your accounts.

- **Look for signs of phishing:** Poor spelling, urgent language, or mismatched URLs can be red flags.

- **Keep software updated:** Regular updates help protect against security vulnerabilities.

- **Use security tools:** Employ spam filters and anti-phishing browser extensions.

Summary

- Phishing remains a persistent threat, employing sophisticated tools and techniques to evade detection. While technology provides various protective measures — such as email filters, website authentication (like SSL/TLS), and multi-factor authentication — criminals continually adapt to bypass these defenses through domain spoofing, obfuscation, and social engineering.

- Therefore, an effective defense requires a combination of technological solutions, user education, and vigilant monitoring.

- Continued research and proactive security practices are essential to mitigate the evolving landscape of phishing attacks.

- By configuring your email security infrastructure to be as tight as operationally feasible—whitelisting trusted domains, sandboxing all others, and blackholing untrusted messages—you significantly reduce the chances of malicious emails reaching end-users

  .

- This approach, combined with ongoing update practices and user training, creates a resilient defence against evolving phishing threats.

LIST OF REFERENCES:

- https://phishtank.org/
- www.Google.com
- https://owasp.org/
- [https://apwg.org/]
- https://spamassassin.apache.org/doc.html]
- https://archive.ics.uci.edu/ml/datasets/Phishing+Websites
- https://www.infosecinstitute.com/resources/phishing/please-volunteer