World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:18, No:12, 2024

Enhancing Email Security: A Multi-Layered Defense Strategy Approach and an AI-Powered Model for Identifying and Mitigating Phishing Attacks

Anastasios Papathanasiou, George Liontos, Athanasios Katsouras, Vasiliki Liagkou, Euripides Glavas

Abstract**—**Email remains a crucial communication tool due to its efficiency, accessibility and cost-effectiveness, enabling rapid information exchange across global networks. However, the global adoption of email has also made it a prime target for cyber threats, including phishing, malware and Business Email Compromise (BEC) attacks, which exploit its integral role in personal and professional realms in order to perform fraud and data breaches. To combat these threats, this research advocates for a multi-layered defense strategy incorporating advanced technological tools such as anti-spam and anti-malware software, machine learning algorithms and authentication protocols. Moreover, we developed an artificial intelligence model specifically designed to analyze email headers and assess their security status. This AI-driven model examines various components of email headers, such as "From" addresses, 'Received' paths and the integrity of SPF (Sender Policy Framework), DKIM (Domain Keys Identified Mail) and DMARC (Domain-based Message Authentication, Reporting and Conformance) records. Upon analysis, it generates comprehensive reports that indicate whether an email is likely to be malicious or benign. This capability empowers users to identify potentially dangerous emails promptly, enhancing their ability to avoid phishing attacks, malware infections and other cyber threats.

Keywords**—**Email security, artificial intelligence, header analysis, threat detection, phishing, Sender Policy Framework, Domain Keys Identified Mail, Domain-based Message Authentication, Reporting and Conformance, AI, Artificial Intelligence.

I. Introduction

EMAIL has become an indispensable part of our daily lives, serving as a digital conduit for communication that spans the globe. Its inception dates back to the 1960s and 1970s, when it was first developed as a means for researchers to communicate over long distances. With the widespread adoption of the internet in the 1990s, email quickly became a unique tool for both business and personal communication, offering a level of immediacy and connectivity that was previously unattainable.

The advantages of email are numerous. It provides an efficient way to send and receive messages and documents, allowing for rapid decision-making and timely responses. Furthermore, the convenience of email is unparalleled; it can be accessed from almost anywhere, enabling people to stay connected without the need for real-time interaction. Additionally, email is cost-effective, eliminating the expenses associated with traditional mailing methods. It also serves as a reliable form of documentation, creating a record of exchanges that can be easily stored, searched and referenced. However, with these benefits come potential risks. Email communication, while simplifying many aspects of interaction, should not replace face-to-face conversations in certain contexts, such as when delivering sensitive information or discussing confidential matters.

Moreover, the realm of email is not immune to malicious activities. Cybercriminals often exploit email to commit fraud, engage in phishing scams and distribute malware. Email security is a critical concern, as it involves safeguarding accounts and communications from unauthorized access and threats.

The very attributes that make email an asset—its ubiquity and accessibility—also render it a target for various exploits. Cybercriminals, armed with a sophisticated arsenal, relentlessly seek to exploit various vulnerabilities. Malware, cloaked in the form of attachments, can unleash havoc on unsuspecting systems, compromising sensitive data and corrupting systems.

Phishing attacks, with their deceptive techniques, lure individuals into divulging confidential information, masquerading as trustworthy entities. BEC assaults aims at the heart of corporate communication, impersonating high-level executives to orchestrate fraudulent financial transactions. These digital deceptions are considered calculated strikes that can dismantle the foundations of trust and security upon which email is built.

In this era of digital dominance individuals and organizations must fortify their defenses with security protocols, dedicated software and continuous education in order to be able to protect their communication against the aforementioned malicious cyber activities.

II. Email Flow

An email flow is the flow path that an email follows and the transitions that the email passes when sent from the sender until it is delivered to the recipient. The email crosses multiple transitions between the sender and the recipient before it is delivered [1], [2]. More analytically:

Anastasios Papathanasiou* is with Cyber Crime Division, Hellenic Police, 173 Alexandras Avenue, 11522 Athens, Greece and with Department of Informatics and Telecommunications, University of Ioannina, Kostaki Artas, 47150 Arta, Greece (corresponding author, e-mail: author@ anastasios.papathanasiou@gmail.com).

George Liontos and Athanasios Katsouras are with Department of Materials Science and Engineering, University of Ioannina, 45110 Ioannina, Greece

Vasiliki Liagkou and Euripides Glavas are with Department of Informatics and Telecommunications, University of Ioannina, Kostaki Artas, 47150 Arta, Greece.

- Mail User Agent (MUA): The MUA represents the client-side interface used by the sender to compose and send emails. Examples include desktop email clients like Outlook or web-based interfaces like Gmail or Yahoo Mail. Analyzing the MUA involves considering user experience, interface design, and compatibility across different platforms.
- Mail Submission Agent (MSA): Once the email is composed and sent from the MUA, it is transmitted to the MSA. The MSA is a server responsible for receiving emails from MUAs and processing them for further transmission. Analyzing the MSA involves examining protocols for accepting emails, security measures such as authentication, and error handling mechanisms.
- Mail Transfer Agent (MTA): The MTA, also known as the SMTP relay server, plays a crucial role in routing emails across the internet. It receives emails from the MSA and forwards them to the recipient's mail exchange server. Analyzing the MTA involves assessing network infrastructure, routing algorithms, and reliability measures to ensure efficient email delivery.
- Mail Exchange (MX): The MX server represents the MX server responsible for receiving emails destined for a specific domain. It is identified by MX records in the Domain Name System (DNS) of the recipient domain. Analyzing the MX server involves understanding DNS configuration, server administration, and redundancy strategies to ensure continuous email service.
- Mail Delivery Agent (MDA): The MDA is responsible for delivering emails to the recipient's mailbox after successful authentication. It retrieves emails from the recipient's MX server and stores them for access by the recipient. Analyzing the MDA involves examining authentication protocols, storage systems, and access control mechanisms to protect the confidentiality and integrity of email communications.

In analyzing the email flow, considerations such as security, reliability, scalability and performance optimization are crucial. By understanding each component's function and interactions, organizations and individuals can enhance their email infrastructure to meet the evolving demands of modern communication. Additionally, ongoing monitoring and optimization are critical to ensure the continuous flow of email communication and mitigate potential issues such as delays, delivery failures or security breaches [3].

III. Most Prevalent Email Threats

Widespread adoption of the email use has also made it a promising field for various cybersecurity threats. Among the most prevalent threats are phishing, spear phishing, spam, malware and BEC attacks, each exploiting the unique nature of email for malicious purposes. More analytically:

- Phishing is arguably the most well-known email threat, characterized by broad, non-targeted attacks. Cybercriminals send deceptive emails that mimic the appearance and language of legitimate institutions such as banks, government agencies or popular online platforms. The goal is to trick users into revealing sensitive information like passwords, credit card numbers or social security numbers. These emails often contain links that lead to fake websites or direct recipients to download attachments infected with malware.
- Spear phishing is a more insidious variant of phishing, where attackers personalize their attacks. By using information that is specific to the recipient -gathered from social media activities, professional profiles or public records- the attacker crafts an email that is highly convincing and tailored to that individual. The personalized approach significantly increases the likelihood of the recipient taking the bait, as the email might reference a recent transaction, an ongoing project, or a mutual contact making the fraudulent request appear legitimate [4].
- BEC is a more direct financial threat. In this scam, attackers focus on employees with access to company finances and attempt to trick them into making wire transfers to fraudulent accounts. BEC scams are sophisticated and often involve identity theft. Attackers might impersonate senior executives, suppliers or vendors and request urgent wire transfers often in a confidential or time-sensitive way.
- Spam is another prevalent threat, involving the mass distribution of unsolicited messages. While often considered merely annoying, spam can also be a vehicle for spreading malware and scams. Spam filters have evolved to block many of these messages but spammers continually refine their strategies to bypass these defenses by using sophisticated techniques, such as modifying email content or misusing legitimate email servers.
- Malware distribution via email remains a critical threat. This method involves embedding malicious software in an email attachment or link. When the attachment is opened or the link clicked the malware is installed on the user's device. This can lead to a range of problems, from the annoying exhibition of ads to severe issues like identity theft, data loss or encryption of files for ransom and financial gain [5].

These threats are becoming more and more dangerous due to the increasing sophistication of attackers who constantly develop new strategies to exploit human errors and system vulnerabilities. The human factor often proves to be the weakest link in cybersecurity and thus cybersecurity education is critical for the abovementioned reason. Individuals and organizations must be trained to recognize the signs of malicious emails, such as slight irregularities in domain names, unusual sender addresses, grammatical errors and suspicious attachments.

Organizations are also strongly encouraged and advised to implement several technical measures to combat these threats. Advanced spam filters, anti-malware programs and regular system updates are essential. Additionally, deploying advanced threat protection solutions that analyze email content for malicious intent and authenticate email senders can significantly mitigate the risks associated with email threats.

Overall, as cybercriminals refine their techniques, the need for vigilance and advanced security measures becomes ever

more critical.

IV. Defense Strategies Against Email Threats

While preventive measures alone offer a good foundation, they are not sufficient to counter email attacks and their continuously advancing tactics. Various technical measures must be considered in order to effectively protect email communication and more specifically:

- Anti-spam and anti-malware tools, which provide essential protection from spam and phishing, common components of BEC strategies. These tools are capable of detecting malicious software that might be embedded in emails [6].
- Time-of-click protection rewrites URLs within email messages to provide security when a link is clicked, utilizing various reputation services to assess potential risks.
- Executive tracking features, which leverage synchronization with Active Directory to automatically recognize users' real names in the email header and address fields.
- Domain similarity checks, which compare the sender's domain against valid domain names to spot domains that are only slightly altered - often by just one or two characters.
- Directory harvest attack (DHA) prevention, which rejects emails sent to invalid or fabricated email addresses.
- Multi-factor authentication (MFA), requiring multiple forms of verification to access an email account or other resources, enhancing security beyond traditional password methods.
- Email protocol updates, ensuring that older and potentially insecure protocols like POP, IMAP, and SMTP are replaced with more secure options to prevent them from bypassing MFA protections.
- DKIM which is a cryptographic technique to verify email integrity and authenticity, involving digital signatures verified via public keys published in DNS.
- SPF which authenticates an email's originating IP address against a list of approved senders to prevent unauthorized email transmission.
- DMARC which integrates DKIM and SPF to thwart domain spoofing, although it primarily processes only header information and is less effective against direct impersonation [7].
- Encryption, which secures email content by requiring cryptographic keys for both sender and receiver, safeguarding against data breaches.
- Verification tools, such as invoice verifiers that can scan QR codes on invoices to confirm their authenticity providing an additional layer of security for email attachments.
- Machine learning (ML) technologies, part of the broader field of artificial intelligence, using algorithms to analyze email patterns, building models that classify email behavior and identify anomalies indicative of BEC attempts.
- Artificial intelligence (AI) which transforms email security by employing sophisticated algorithms that can detect and neutralize threats more efficiently than traditional methods. By analyzing extensive datasets, AI identifies complex behavioral patterns and inconsistencies indicative of email fraud or attacks such as those seen in BEC scams. These systems evolve by learning from new and emerging threats, ensuring that they remain effective against cybercriminals' evolving strategies. Additionally, AI enhances the accuracy of spam filters minimizing the occurrence of false positives and ensuring that legitimate communications are not disrupted [8].
- Anti-phishing solutions like email identification tokens. For example, Binance uses an anti-phishing identification code that is defined by the user in the Binance platform and it is send along with all costumer emails while Papathanasiou et al. [9] propose a QR code authentication scheme based on an encrypted QR code that contains the MAC address of the sender as an authentication token.

These technical defenses are critical in enhancing an organization's resilience against sophisticated email-based attacks, complementing preventive measures to create a more robust security. Nevertheless, technical methods should be also be accompanied by non-technical measures -in the case of enterprises- like:

- Continuous Employee Training: Employees need ongoing training to effectively identify, report and react to various red incidents. Regular training sessions should focus on social engineering tactics and schemes, particularly for departments like finance that are more susceptible to these threats. Training should include caution against suspicious hyperlinks, email attachments, misspellings in names, and last-minute instructions regarding wire transfers or changes in account information [10].
- Formation of Tactical Teams: For companies, the formation of both blue and red teams is highly recommended. Blue teams are tasked with assessing and safeguarding the security environment, whereas red teams play the adversary role, aiming to breach these defenses to pinpoint weaknesses.
- Policy Definition: Companies should establish internal rules, policies and guidelines that enforce stringent measures for sharing information and conducting financial transactions. Policies should prohibit initiating fund transfers via email and require multiple approvals or verbal confirmations for financial transactions. Phone communications should include identity verification procedures to prevent data leaks.
- Fraud Risk Evaluation: Conducting evaluations of fraud risk can help identify and deal with potential vulnerabilities that cybercriminals might exploit. Utilizing frameworks like the fraud risk management guide by the Committee of Sponsoring Organizations of the Treadway Commission (COSO), in partnership with the Association of Certified Fraud Examiners (ACFE) and referencing standards such as ISO 22380:2018 can be particularly effective [11].
- Regular Real-World Checkups: Given that employees are thoroughly educated about social engineering and that

appropriate policies are in place, regular real-world assessments should be conducted. These checkups should be performed by professionals such as penetration testers and social engineers to ensure defenses remain robust against actual threats.

V. Literature Review for Header Analysis

A range of studies have explored the use of email header analysis to detect potential threats. Odunibosi [12] and Banday [13] both highlight the importance of email header information in detecting email spoofing and for forensic investigation. Beaman [14] further demonstrated the effectiveness of machine learning algorithms, such as Support Vector Machine and Random Forest, in classifying email headers to detect spam and phishing emails. Charalambou [15] provides a comprehensive review of existing email forensic tools for header analysis while Garba et al. [16] propose an email content-based insider threat detection model using anomaly detection algorithms. These studies collectively underscore the potential of email header analysis in identifying and mitigating various email threats.

A range of studies have also explored the use of AI models for email header analysis. Salcedo-Campos et al. [17] and Beaman et al. [18] both demonstrated the effectiveness of using email header information for spam detection, with Beaman specifically highlighting the success of supervised learning algorithms. Karim et al. [19] and Khamis et al. [20] further emphasized the importance of email header features in spam detection, with Khamis specifically using Support Vector Machine (SVM) for classification. Yang et al. [21] and Al-Jarrah et al. [22] both found that email categorization based on header information is comparable or superior to that based on the entire message. These studies collectively highlight the potential of AI models for email header analysis in various applications, particularly in spam detection and email management.

VI. Trained AI Model for Spoofing Detection

Artificial intelligence (AI) technology is revolutionizing the field of email security by providing advanced tools that enhance the detection and prevention of email threats. AI systems leverage complex algorithms to learn from vast amounts of data, enabling them to identify subtle patterns and anomalies that may elude traditional security measures. For example, AI can analyze the writing style of communications to detect inconsistencies that might indicate a compromised email account or a sophisticated impersonation attempt in a BEC scheme. Furthermore, AI-driven systems can continuously adapt to new threats by learning from ongoing attacks, thus staying ahead of cybercriminals who constantly refine their tactics. AI also improves the efficiency of filtering systems, reducing false positives and ensuring legitimate emails are not mistakenly blocked, which is crucial for maintaining smooth business communications. Overall, the integration of AI into email security frameworks significantly bolsters an organization's defenses by providing dynamic, intelligent and responsive solutions to the evolving landscape of email threats [23].

With the use of AI and specifically ChatGPT 4 and inspired by our prior analysis on how to enhance email security, we trained an algorithm to analyze the header of an email. This model can be found at this link: https://chat.openai.com/g/g-cwuFvYUEm-email-header-analysis.

Examining an email header can help identify potential signs of email spoofing and general red flags, although it may not always provide conclusive evidence. Here's how the trained AI model can look for signs of spoofing in an email header:

- Check the "From" Address: Sometimes, the display name may look legitimate, but the actual email address could be from a suspicious domain. This is a common spoofing tactic.
- Analyze the "Received" Paths: This part of the header shows the journey of the email across different servers. If the email claims to be from a well-known domain but the received paths show transitions through unrelated, suspicious, or frequently changing domains, it might be spoofed.
- Review SPF, DKIM and DMARC Results: These technologies are designed to validate the sender:
- SPF checks if the IP address sending the email is authorized by the domain's DNS records.
- DKIM checks for a digital signature that should match a key published in the domain's DNS records.
- DMARC leverages SPF and DKIM to enhance email security, specifying how receivers should handle emails that do not pass these checks.
- If these checks fail (spf = fail, dkim = fail), it' is a strong indication of spoofing, especially if the email claims to be from a reputable domain that usually has strict SPF and DKIM policies.
- Look for Mismatching Information: If the "Return-Path", "Reply-To", and "From" addresses are all different, it could indicate a spoofed email, especially if none of them align with the expected domain.
- Examine IP Address Locations: If the IP addresses in the header (especially in the first "Received from" line closest to the actual text of the email) are from countries or networks not typically associated with the sender, this might be a sign of spoofing.

Training an AI model to analyze email headers offers a strong indication against email spoofing which is a prevalent tactic among cybercriminals. By verifying "From" addresses, analyzing the path of email transmission and conducting thorough checks on sender authentication protocols like SPF, DKIM and DMARC, this AI-driven approach effectively identifies and flags potential threats. This not only enhances the security of email communications but also ensures that organizations and individuals can maintain trust and integrity in their digital interactions. Adopting such advanced technologies is essential in safeguarding against sophisticated cyber threats and protecting sensitive information from being compromised.

VII. Header Analysis

We provided the AI model with a variety of benign and

malicious headers, aiming to enhance its capacity to discern and identify potential threats effectively. Through this training process, our goal was to equip the model with the ability to distinguish between harmless and harmful elements within headers, thereby enhancing its threat detection capabilities across diverse scenarios.

In the following paragraphs we present a scenario with a benign header for analysis in the trained AI model with instructions to identify any possible red flags and provide the user with a report. The report of the header analysis by the AI model was the following:

Email Header Analysis Summary

Here's a detailed analysis of the provided email header by the AI model:

1. **Received Headers**
- **IP Information and Route:**
- The email has been routed through Google's infrastructure, as indicated by the IP addresses (`209.85.220.41`) and the domain (`mail-sor-f41.google.com`). This is consistent with emails sent from Gmail accounts, suggesting the mail servers are part of the legitimate Google email network.
- **Received Timestamps:**
- The timestamps in the received headers are consistent and show that the email was processed through Google servers without unusual delays, typically indicative of a direct path from the sender to the recipient within the expected timeframe for this type of service.
2. **SPF, DKIM, and DMARC Records:**
- **SPF:**
- `Pass` status for the sender domain (`emailsender@gmail.com`) with the IP `209.85.220.41` designated as a permitted sender. This matches the sending IP noted in the received paths.
- **DKIM:**
- `Pass` status with a signature that matches the `@gmail.com` domain. This is an authentication method to help prevent email spoofing.
- **DMARC:**
- `Pass` status with a policy of none for the domain `gmail.com`, which is standard for Gmail and does not enforce strict actions against failed DMARC checks but allows for monitoring.
3. **From Address Verification:**
- **Email Sender:**
- `Email SENDER <emailsender@gmail.com>`
- This is consistent throughout the header analysis, showing no discrepancies in the displayed name and the email address, which helps verify the sender's identity.
4. **X-Headers and Google-Specific Headers:**
- These headers are specific to Google's email handling and security practices. They include encryption indicators and internal routing information that supports the legitimacy of the email's journey through Google's infrastructure.
5. **ARC (Authenticated Received Chain):**
- The ARC protocol is used to maintain email authentication results and verify the identity of the intermediaries handling the emails. The `ARC-Seal` indicates that there has been no modification since it was first sent and the ARC authentication aligns with passing results, which is a positive indicator of a secure email chain.

Conclusion (Written by the AI Model)

- The analysis indicates that the email has a **high likelihood of being legitimate**, with all major authentication checks passing and the email being routed through expected, secure channels as per Google's infrastructure.
- **No immediate red flags or threats** have been detected in this email header, suggesting that it is safe based on the details provided. However, always remain cautious with email links and attachments, even if the email itself appears to be legitimate.

In the following paragraphs we present a second scenario with a malicious header for analysis in the trained ai model with instructions to again identify any possible red flags and provide the user with a report. The report of the header analysis by the ai model was the following:

- "The email header you've provided contains several details that can be analyzed to assess the legitimacy and security of the message. Here's a detailed analysis based on the received paths, SPF, DKIM, and other headers:
- Received Paths Analysis
1. **Chain of Custody**: The email has traversed from `cloudflare.net` through several Microsoft Outlook servers before reaching the final destination. This indicates the use of legitimate email services.
2. **IP Addresses**: The path shows both IPv4 and IPv6 addresses belonging to `outlook.com` and `cloudflare.net`. The final sender's IP is 89.47.165.109 which should be scrutinized.
- SPF
- **Result**: The SPF result is `pass` for the IP 89.47.165.109 with the helo identifier of `app.webberit.us` and mail from `luichang@webberit.us`. A passing SPF result indicates that the sending server is authorized to send emails on behalf of the domain `webberit.us`.
- DKIM
- **Signature**: The DKIM signature uses SHA-1 with relaxed headers, which is a valid configuration but less secure than SHA-256. The presence of a valid DKIM signature suggests that the message has not been altered in transit.
- DMARC
- **Missing Information**: The header does not explicitly mention a DMARC verification result. Typically, the absence of DMARC verification or a policy enforcement could allow potential spoofing issues to go undetected.
- Other Key Headers
- **From and Return Path**: Both fields show the email address `luichang@webberit.us`, which is consistent, reducing the likelihood of address header forgery.
- **X-MS-Exchange-Organization-AuthAs**: Marked as `Anonymous`, which often occurs in legitimate emails but

can also indicate a message routed through different servers or gateways without authentication.

- **Message Consistency**: The `Message-ID` domain and the `From` address domain match, suggesting consistency in the origin domain.

- Red Flags and Concerns

1. **IP Address Origin**: The originating IP address (89.47.165.109) is not well-known and should be verified for any malicious associations. Without more specific geolocation or reputation data, it is a point of scrutiny.
2. **Multiple MIME Versions**: The presence of two `MIME-Version` headers is unusual and might indicate header tampering or misconfiguration.
3. **Subject Matter and Urgency**: The subject about "Overdue Invoices" could be a tactic often used in phishing to create urgency and prompt quick action from the recipient.

Conclusion (by the AI Model)

The email passes basic SPF and DKIM checks, indicating it was likely sent from an authorized server and was not altered in transit. However, the lack of explicit DMARC verification, the unusual presence of multiple `MIME-Version` headers, and the use of an IP address that is not well-known warrant a cautious approach. It is advisable to verify the identity of the sender and the context of the email before taking any action, especially given the subject matter which could be a potential phishing indicator."

In both scenarios, the AI model demonstrated its proficiency by identifying signs indicative of malicious or benign content within the headers. Its discerning capabilities allowed it to accurately differentiate between benign and malicious headers, effectively flagging potential threats and safeguarding against harmful activities. This success underscores the model's effectiveness and reliability in threat detection, showcasing its capacity to contribute significantly to cybersecurity efforts.

VIII. Conclusion

Email remains an indispensable aspect of daily digital interactions offering unmatched efficiency and reach. However, its widespread use also renders it a prime vector for cyberattacks, including phishing, spear phishing, malware distribution, spam and BEC attacks. These threats not only compromise personal and organizational security but also pose significant risks to the integrity of information and financial assets. Understanding the mechanics of email -from the MUA to the MDA- is crucial in identifying where vulnerabilities may occur and how they can be addressed.

In our research, we highlighted the critical role of technological defenses in protecting against email threats. Tools such as anti-spam and anti-malware software, machine learning algorithms and authentication protocols (SPF, DKIM and DMARC) form the backbone of a well-crafted defense strategy. These technologies help in filtering out malicious content, authenticating the sources of emails and ensuring that messages are not altered in transit. Moreover, the evolution of AI has significantly enhanced the capability of these tools, enabling them to adapt to new threats dynamically and provide real-time security updates.

In addition to technological solutions, the human element of cybersecurity cannot be overlooked. Education and continuous training are paramount in cultivating an environment of awareness and vigilance. Users equipped with the knowledge of potential email threats and the skills to recognize suspicious elements are the first line of defense against cyber intrusions. As cybercriminals continually refine their tactics, the ongoing education of all email users is essential for maintaining effective defenses.

Furthermore, our development of an AI model to analyze email headers marks a significant advancement in this field in the concept that regular users without technical knowledge can us the model in order to characterize the email as benign or malicious. This model serves as a critical tool in the identification of email-based threats, providing users with detailed reports on the legitimacy of their emails. By examining key aspects of email headers, the model assesses the security status of incoming messages and makes an estimation about the type of the email (benign or malicious). This capability is crucial in preventing the exploitation of email systems and safeguarding sensitive information. More specifically, the AI model analyzes several critical components of email headers to determine the security status of incoming messages. It analyzes the "From" address to verify sender authenticity, evaluates the "Received" paths to trace the email's journey across servers, and assesses the implementation of SPF, DKIM and DMARC records to ensure sender integrity and message security. This comprehensive examination helps in identifying red flags that might indicate phishing or spoofing attempts. Additionally, the model reviews IP address origins and analyzes any mismatches in the email's routing information, providing a thorough inspection that is vital for detecting and flagging potentially malicious emails.

In conclusion, while email will continue to be a fundamental component of global communications, the security challenges it presents are ever-evolving. Our comprehensive approach, which combines advanced technological tools with non-technical measures, offers a secure framework for mitigating these risks. By integrating sophisticated AI technologies with a strong emphasis on user training and awareness, we can enhance the resilience of email systems against a wide range of cyber threats. This dual strategy not only protects individual users but also reinforces the security of broader digital infrastructures, ensuring that email remains a safe, reliable and effective tool for communication.

References

[1] M. Lui, M. Gray, A. Chan, J. Long, "Message Flow: Routing and Filtering", in: *Pro Spring Integration*. Apress, Berkeley, CA. 2011, pp. 245-290, https://doi.org/10.1007/978-1-4302-3346-6_8

[2] T. Karagiannis, M. Vojnovic, *Email information flow in large-scale enterprises*. Microsoft Research. 2008.

[3] Y. Mostafa, *Effective Threat Investigation for SOC Analysts: The ultimate guide to examining various threats and attacker techniques using security logs*, Packt Publishing Ltd, 2023, pp. 15–64.

[4] M.A. Butavicius, K. Parsons, M.R. Pattinson, A. McCormac, *Breaching the Human Firewall: Social engineering in Phishing and Spear-Phishing*

Emails. ACIS, 2016

[5] A. Ayodele, J. Henrydoss, W. Schrier, T.E. Boult, "Study of Malware Threats Faced by the Typical Email User", in: D.C. Wyld, M. Wozniak, N. Chaki, N. Meghanathan, D. Nagamalai, (eds) *Advances in Network Security and Applications. CNSA 2011*. Communications in Computer and Information Science, vol 196. Springer, Berlin, Heidelberg. 2011. https://doi.org/10.1007/978-3-642-22540-6_50

[6] P. Bruce, D. Greg, "The effectiveness of anti-malware tools", *Computer Fraud & Security*, Volume 2009, Issue 3, 2009, https://doi.org/10.1016/S1361-3723(09)70033-8.

[7] J. Nightingale, "Email Authentication Mechanisms: DMARC, SPF and DKIM, Technical Note (NIST TN)", *National Institute of Standards and Technology*, Gaithersburg, MD, 2017. online, https://doi.org/10.6028/NIST.TN.1945 (Accessed April 24, 2024)

[8] S. Mca, "AI and ML techniques to Analyze Communication Emails and Text patterns To Secure from Attacks". *SSRN Electronic Journal*. 8. 2020, pp 2084-2087.

[9] A. Papathanasiou, G. Liontos, G. Paparis, V. Liagkou, E. Glavas, "BEC Defender: QR Code-Based Methodology for Prevention of Business Email Compromise (BEC) Attacks". *Sensors*, 24(5), 1676, 2024.

[10] D. Negussie, "Importance of cybersecurity awareness training for employees in business", *VIDYA - A Journal of Gujarat University*, 2, 104-107, 2023. https://doi.org/10.47413/vidya.v2i2.206

[11] N. Jaffar, H. Haron, T. Iskandar, A. Salleh, "Fraud risk assessment and detection of fraud: The moderating effect of personality", *International Journal of Business and Management*, 6, 2011. https://doi.org/10.5539/ijbm.v6n7p40

[12] O. Odunibosi, "Classification of email headers using Random Forest algorithm to detect email spoofing", Doctoral dissertation, Dublin, National College of Ireland, 2019.

[13] M. T. Banday, "Technology Corner: Analysing e-mail headers for forensic investigation", *Journal of Digital Forensics, Security and Law*, 6(2), 5. 2011.

[14] C. Beaman, H. Isah, "Anomaly detection in emails using machine learning and header information". *arXiv* preprint arXiv:2203.10408. 2022.

[15] E. Charalambou, R. Bratskas, G. Karkas, A. Anastasiades, "Email forensic tools: A roadmap to email header analysis through a cybercrime use case". *Journal of Polish Safety and Reliability Association*, 7(1). 2016.

[16] N. Garba, S. Rakshit, C. D. Mang, N. R. Vajjhala, "An email content-based insider threat detection model using anomaly detection algorithms" in Proceedings of the *International Conference on Innovative Computing & Communication* (ICICC). 2021.

[17] F. Salcedo-Campos, J. Díaz-Verdejo, P. García-Teodoro, "Segmental parameterization and statistical modelling of e-mail headers for spam detection". *Information Sciences*, 195, 2012, pp 45-61.

[18] C. Beaman, H. Isah, "Anomaly detection in emails using machine learning and header information". *arXiv* preprint arXiv:2203.10408. 2022.

[19] A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti, M. Alazab, "A comprehensive survey for intelligent spam email detection". *Ieee Access*, 7, 2019. 168261-168295.

[20] S. A. Khamis, C. F. Foozy, M. Aziz, N. Rahim, "Header based email spam detection framework using Support Vector Machine (SVM) Technique", in Recent Advances on Soft Computing and Data Mining: Proceedings of the *Fourth International Conference on Soft Computing and Data Mining (SCDM 2020)*, Melaka, Malaysia, January 22– 23, 2020

[21] J. Yang, S. Y. Park, "Email categorization using fast machine learning algorithms", in *International Conference on Discovery Science* Berlin, Heidelberg: Springer Berlin Heidelberg. 2002, pp. 316-323.

[22] O. Al-Jarrah, I. Khater, B. Al-Duwairi, "Identifying potentially useful email header features for email spam filtering", in the *sixth international conference on digital society* (ICDS), vol. 302012, p. 140.

[23] A. Hemalatha, "AI-driven marketing: Leveraging artificial intelligence for enhanced customer engagement". 2023. https://doi.org/10.47715/JPC.B.978-93-91303-61-7