IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 2, February 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202054 |

A Survey on Email Phishing Detection Techniques

Aaron Horta[1], Dr. A. Rengarajan[2]

Student of MCA, Dept. of CS & IT, Jain (Deemed-to-be-University), Bengaluru, India [1]

Professor, Dept. of CS & IT, Jain (Deemed-to-be-University), Bengaluru, India[2]

ABSTRACT: Detecting email phishing, a persistent cyber threat, is essential for safeguarding against potential damages. This survey comprehensively examines a spectrum of techniques employed for phishing email detection, ranging from conventional methods to advanced approaches like machine learning and behavioral analysis. Through an evaluation of each technique's merits and limitations, including considerations such as detection accuracy and scalability, this study aims to provide a nuanced understanding of their effectiveness in mitigating phishing risks. Furthermore, the survey delves into the influence of human factors, such as user awareness training, on enhancing overall detection efficacy. By synthesizing insights from diverse research efforts, this survey offers valuable insights into the current landscape of email phishing detection, informing future research directions and guiding the implementation of robust cybersecurity measures.

KEYWORDS: Email Phishing Detection, Cybersecurity, Machine Learning Algorithms, Behavioral Analysis, Systematic Literature Review

I. INTRODUCTION

Social engineering involves manipulating individuals through human interaction to bypass security measures. It relies on psychological tactics to deceive people into giving out private information or performing actions that could compromise the security. The attack exploits human trust and willingness to assist others, making victims susceptible to manipulation. Social engineering, as opposed to technical assaults on systems, focuses on manipulating individuals with access to information, coaxing them into revealing sensitive data or engaging in malicious activities. Traditional security measures are often ineffective against such attacks

Phishing involves cybercriminals deceiving individuals into sharing sensitive data through methods like fraudulent emails or websites. Email phishing specifically targets victims through email, masquerading as legitimate entities like businesses or banks. These deceptive emails often appear authentic, making it challenging for recipients to identify the fraud. It involves the concept of Social Engineering where the primary objective of email phishing is to coerce individuals into clicking on malicious links or divulging personal information. Despite advancements in technology and user awareness, email phishing remains a significant cybersecurity concern, demanding effective detection and prevention measures. The problem statement revolves around the efficacy of existing email phishing detection techniques in countering increasingly sophisticated attacks. Despite the availability of various detection methods, such as rule-based filtering, machine learning algorithms, and Behavioral analysis, the evolving nature of phishing tactics presents ongoing challenges. This paper aims to explore the landscape of email phishing detection techniques, examining their strengths, weaknesses, and potential areas for enhancement. By addressing this issue, the study seeks to contribute to the advancement of email security measures and strengthen defense against phishing threats.

II. PRIMARY INFORMATION AND RELEVANT WORKS

Social engineering attacks are often conducted by attackers personally, with impersonation being a common tactic, such as feigning distress or urgency. Such assaults may also be carried out via computers or automated means, such as crafting counterfeit websites that bear striking resemblance to authentic ones. Tempting victims with complimentary downloads or substantial price reductions, and urging them to furnish personal information using their official credentials, is a commonly employed tactic.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202054 |

A. Methods of Social Engineering Attacks

Attackers employ various methods in social engineering endeavours. Initially, they aim to gather sensitive data such as bank account details or personal information like names and addresses. If initial attempts fail, subsequent attacks are launched using previously obtained data. The commonly used methods involve phishing, pretexting, baiting, tailgating, spear phishing and email attachments.

Pretexting involves crafting elaborate lies to gather information about individuals or organizations, often under false pretences. Phishing, a widely recognized term, involves deceiving users into divulging personal information through fraudulent emails or websites. Spear phishing targets specific individuals within organizations, often using messages from trusted sources to gain credibility.

Vishing, or Voice Phishing, involves phone calls prompting users to call a specified number, often posing as legitimate organizations like banks. Emails with enticing subjects or attachments may contain malicious software, downloaded unknowingly by the user.

Emails with intriguing subjects often prompt employees to open them, sometimes containing attachments with appealing names. Upon opening these attachments, malicious software may be downloaded onto the employee's device without their awareness. Additionally, some emails may falsely warn employees of a virus attack and offer a solution in the form of an attachment, further compromising their system's security. Phishing attacks typically mimic legitimate organizations, utilizing logos and content styles to appear authentic and deceive users. These methods exploit human curiosity and trust, making users vulnerable to identity theft and other malicious activities.

B. Web Phishing attacks

A successful phishing can be considered as a combination of phishing emails and a phishing website as represented in the Figure 1. In the pursuit of victimizing individuals online, perpetrators employ spoofed emails to create an illusion of authenticity in their communication. These emails mimic addresses associated with credible entities like credit card agencies, banks, or even governmental bodies. The spoofing technique is employed to such precision that the source addresses of these emails closely resemble those originating from legitimate sources.

For Example, In the context of email spoofing, an attacker might attempt to mimic the email address of a bank manager, such as bankmanager@jkl.co.in, to deceive users into believing the legitimacy of the email and complying with requests outlined by the person who is attacking rather knows as a phisher. Typically, these requests involve actions like clicking on web links and divulging personal or transactional data either through website forms or by replying to the email. This deception is facilitated through the use of open SMTP (Simple Mail Transfer Protocol) servers, allowing the phishers to send a set of spoofed emails to their targets. Recognizing that many users are cautious about revealing personal information via email replies, attackers employ another deceptive tactic: crafting phishing websites that closely resemble the appearance and functionality of legitimate websites targeted for impersonation.

Upon clicking the website link given in the spoofed emails, the user is redirected to the phishing website set up by attacker. Due to its close resemblance to the authentic site, unsuspecting users, particularly those lacking experience, may fail to recognize the fraudulent nature of the website and proceed to enter requested data, which consequences in a triumphant phishing attack. In addition to email-based tactics, attackers may also disseminate the links which are malicious by promoting the links as advertisements on legitimate sites.

Moreover, in certain instances, compromised DNS (Domain Name System) servers can redirect users to abnormal or phishing websites, further enhancing the effectiveness of the attack. Phishers employ various methods to conduct web phishing attacks:

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

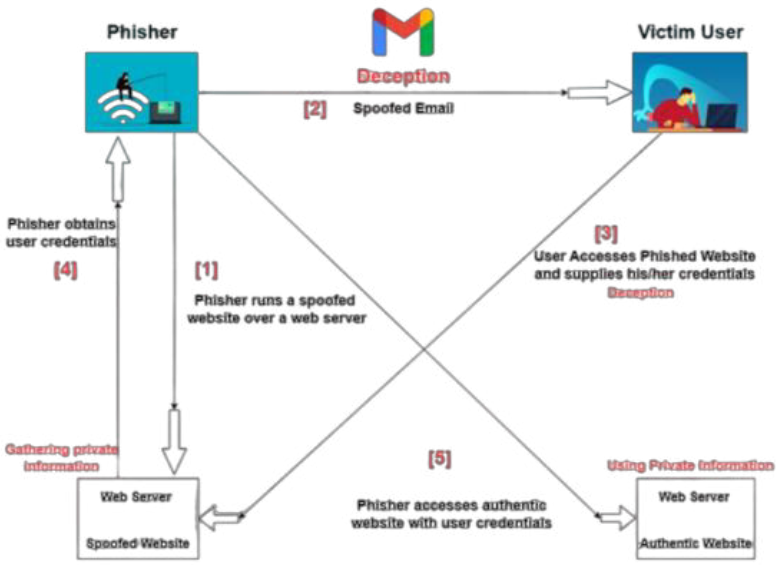
| DOI: 10.15680/IJIRCCE.2024.1202054 |

1. Creation of Informative Websites: Phishers develop websites offering valuable content and seemingly legitimate links to popular platforms like Facebook, Gmail, and Twitter. However, these links redirect users to malicious websites instead of the legitimate ones. Phishers often employ link manipulation techniques to maintain user trust in these deceptive links.
2. Covert Phishing via Compromised Sites: In covert phishing attacks, attackers exploit compromised legitimate websites to present users with fake login prompts, soliciting their personal information unknowingly.
3. Compromised DNS Redirection: Phishers may hijack DNS systems to redirect users' requests for benign URLs to their malicious counterparts, leading unsuspecting users to land on phishing sites.

A Practical Scenario of phishing

In a phishing scenario depicted in Figure 2, spoofed email communication is utilized. The typical procedure undertaken by a phisher during a phishing endeavour involves several steps:

1. Setting up a spoofed rendition of the desired website on a server, then dispatching spoofed emails to target recipients.
2. These emails usually convey an urgent message, compelling immediate action, such as logging into the recipient's bank account and furnishing details, under the threat of account termination.
3. The fraudulent link embedded in the email redirects the recipient to a server hosting a visually replicated login page of the targeted website.
4. Unaware, the recipient submits their credentials on the counterfeit website, which the phisher then captures.
5. The phisher exploits this pilfered data for deceitful purposes, like conducting unauthorized transactions using the victim's banking credentials or making online purchases with stolen credit card information.



Phishing Tactics

Spear phishing is a kind of selected form of phishing targeted at explicit individuals or companies. Figure 3 shows the types of phishing tactics that an attacker uses on victims. When high-profile individuals are targeted, it's referred to as Whaling. Another variation, Tab nabbing, capitalizes on user tab switches to load a phishing page when they return, assuming it's an authentic page left open. Deception is central to phishing's success, with techniques like email and website spoofing and exploiting browser vulnerabilities. Phishers employ various tricks, including manipulating links to lead to malicious URLs while appearing authentic, evading detection filters using images instead of text, hiding browser address bars with scripting languages, using pop-up windows to collect credentials, and exploiting browser vulnerabilities like Tab nabbing.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202054 |

III. RESEARCH METHODOLOGY

Phishing URL Detection

A URL (Uniform Resource Locator) functions as a string identifying a web resource, frequently susceptible to fraudulent alterations through the incorporation of new pages where they should not exist. locating those lurking URLs proves challenging as they evade normal navigation and typically escape search engine indexing. Prior research in this domain primarily concentrated on detecting phishing URLs and those associated with spam-advertised websites, extending to include hidden defacement URLs for attack detection. Detection techniques generally fall into three types: Approaches categorized as URL-centric, host-centric, and content-centric.URL-centric approaches, broadly speaking, offer faster and more flexible and scalable detection capabilities as they classify URLs based on their inherent structure. Hidden URLs refer to pages hosted within a website unbeknownst to the administrator, while fraudulent URLs denote defacements or phishing attacks.

Lexical Analysis of URLs: One prevalent technique for recognizing deceitful URLs hinges on their linguistic attributes, enabling user notifications before page access. Security risks commonly arise from illicit modifications to established pages or the insertion of new pages where they shouldn't be. Identifying these irregularities proves challenging, as administrators may overlook them, and users rarely access such URLs. Linguistic traits are extracted from the URL, excluding the domain section, and fed into a Support Vector Machine for categorization. The efficacy of this method is assessed across concealed fraudulent URL types, encompassing phishing pages and webpage defacements.

Email Data Features

Phishing emails typically lure users to divulge private information on external websites. To detect such emails, we examine their structure and embedded links. Our approach, comprising 27 basic features, expands upon previous work by researchers. by including additional attributes directly derivable from the email content. These features are categorized as follows:

- **Body Structure Attributes (4):** These relate to the organization of the email body, encompassing the total count of body segments and diverse body segment types as per the MIME standard.
- **Link Characteristics (8):** These cover attributes of links embedded within the email, including the overall count of links, differentiation between internal and external links, presence of IP addresses in links, deceptive link indicators, links concealed behind images, and traits of link anchor text.
- **Web Technology Elements (4):** These signify the utilization of web technologies within the email, such as HTML, scripting languages (especially JavaScript), and form elements.
- **Spam Filter Indicators (2):** These are generated by an offline SpamAssassin tool, comprising a spam rating and a Boolean value denoting whether the email is flagged as spam.
- **Terminology List Attributes (9):** These attributes leverage a positive terminology list linked with phishing, detecting the occurrence of terms like "account," "update," "confirm," etc., within the email content.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202054 |

Phishing Detection Classification Schemes

A comprehensive classification of phishing detection methods is presented in Figure 6, based on the technique used for identifying phishing websites. These methods are broadly categorized as follows:

1. ***Search Engine Based (SEB):***
   These techniques involve extracting website features and searching for them using search engines. The assumption is that normal websites rank higher in search results compared to phishing websites. Search engine-based techniques involve utilizing search engines to detect phishing by extracting webpage text, images, or URLs as search strings. These techniques vary in their features extracted, search engine usage, and decision-making algorithms.

   - One approach focused on lightweight phishing detection using minimal features like page title and domain name, implemented in an anti-phishing Chrome extension.
   - An alternative method involved gathering and cross-referencing domains linked to the suspicious webpage to identify phishing attempts.
   - A method entailed querying the webpage URL across renowned search engines such as Google, Bing, and Yahoo, leveraging the quantity and rankings of search outcomes for classification purposes.
   - One strategy involved capturing screenshots of webpages and conducting a Google image search on the website logo. Subsequently, the returned keywords were cross-referenced with text search results.
   - A technique used specific Google search queries to check if returned results indicated the same domain as the visited webpage, determining phishing based on search rankings.

   Research gaps:
   - Examining prolonged benign domains that unexpectedly transition into phishing activities, as they might retain visibility in top search results.
   - Mitigating false positives for transient benign domains that evade detection in top search results, potentially via enhanced filtering mechanisms.
   -

2. ***Heuristics and Machine Learning Based (HMLB):***
   These methods extract features from websites and utilize heuristics or machine learning algorithms for anomaly detection. Heuristics and machine learning-based techniques involve extracting features from webpages, URLs, or network data and applying machine learning or classification algorithms to create detection models. These methods vary in the features used, algorithms applied, and optimization techniques employed.

   - One method suggested employing Levenshtein Distance alongside Support Vector Machines (SVM) for classification purposes.
   - Another method utilized Adaline network training for neural network-based phishing detection.
   - Self-structuring neural networks and backpropagation training were suggested for improving classification accuracy.
   - Six heuristics, including primary domain and page rank, were assigned weights to optimize phishing detection thresholds.
   - Rule-based data mining techniques were applied to identify important phishing detection features.
   - Proposing a blend of heuristic and Naive Bayes classifiers for classification, supplemented by extra filters to minimize false positives.
   - Important features were identified and utilized with the TF-IDF approach and SVM classifier for phishing detection.
   - Utilizing logistic regression, lexical and host-based features were extracted and classified.

   Research gaps:
   - Identifying lightweight machine learning techniques that demand fewer resources without compromising accuracy.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202054 |

- Constructing scalable and dynamic solutions capable of adapting to shifts in phishing tactics, potentially through self-learning mechanisms to recognize and integrate novel features.

3. ***DNS Based:***
DNS (Domain Name System) is utilized to authenticate the IP address of a website, establishing its legitimacy. If the IP address does not correspond to an authentic website, it is flagged as phishing. DNS-based methods employ DNS data to validate the legitimacy of domain names and their associated IP addresses for detecting phishing. These approaches differ in their utilization of DNS to gather pertinent information for identifying phishing attempts.
   - One proposed system features a client-side component that extracts the unique signature of the current webpage, sending it as a DNS query to a distant server for comparison against known phishing webpage signatures. A client-side policy enforcer then takes action based on the received response.
   - Another method entails sending the domain name of the visited URL to both default and third-party DNS servers. If the default IP address aligns with those provided by the third-party DNS server, the site is deemed authentic. Otherwise, a webpage content similarity assessment is conducted to identify phishing attempts.
   - Recursive DNS query logs are utilized to identify visited hosts and detect suspicious phishing hosts.
   - A scheme involves an information server storing bank-related data and allowed DNS server IP addresses for bank login. Packets are monitored to detect card number entries on unauthorized websites, signaling phishing attempts.

   Research gaps and future needs include:
   - Developing methods to reduce communication burden on DNS servers, possibly through caching and smart storage techniques, to minimize network overhead and delays.
   - Exploring strategies to optimize network communication costs associated with DNS queries, enhancing efficiency and scalability of DNS-based phishing detection methods.

4. ***Proactive Phishing URL Detection Based (PPUDB):***
This method detects potential phishing URLs by generating diverse combinations from existing legitimate URLs and scrutinizing their engagement in phishing activities. Proactive phishing URL detection-based techniques employ methods to generate or identify potential phishing URLs pre-emptively. These schemes vary in their approach and technique for generating probable phishing URLs and mining the web for such URLs.
   - One approach focuses on identifying newly registered malicious domain names drawing from the insight that legitimate domain names often contain meaningful English words, this method employs a second-order Markov model to pinpoint valuable features. Subsequently, random forest classification is utilized for detection purposes.
   - An alternative method entails categorizing phishing emails, then conducting reverse lookups of web links within the emails using WhoIs queries to collect details about the server hosting the phishing links. Following this, proactive warning notifications are dispatched to the server administrator for necessary action.

IV. EVALUATION METRICS FOR PHISHING EMAIL DETECTION

Assessing the effectiveness of phishing detection models is crucial for selecting the most appropriate model for different scenarios. Many studies in the literature have investigated four primary evaluation metrics: accuracy, precision, recall, and F-measure. These metrics gauge the model's ability to identify phishing emails effectively.

1. ***True Positive (TP):*** This metric indicates the percentage of phishing emails accurately identified by the model. TP is calculated as the ratio of correctly classified phishing emails (NP) to the total number of phishing emails in the dataset (P), expressed as TP = NP/P.

2. ***True Negative (TN):*** TN represents the percentage of legitimate emails correctly recognized as such by the model. It is calculated as the ratio of correctly classified legitimate emails (NL) to the total number of legitimate emails (L), given by TN = NL/L.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202054 |

3. ***False Positive (FP):*** FP reflects the percentage of legitimate emails incorrectly classified as phishing by the model. It is calculated as the ratio of incorrectly classified legitimate emails (Nf) to the total number of legitimate emails (L), expressed as FP = Nf/L.

4. ***False Negative (FN):*** FN indicates the percentage of phishing emails inaccurately labeled as legitimate by the model. It is calculated as the ratio of incorrectly classified phishing emails (Npl) to the total number of phishing emails (P), given by FN = Npl/P.

These metrics are then used to compute the following evaluation measures:

- ***Accuracy:*** This parameter embodies the overall ratio of accurately classified emails throughout the dataset. It's computed using the formula: Accuracy = (TP + TN) / (TP + FP + FN + TN).
- ***Precision:*** Precision gauges the precision of the classifier in identifying BEC phishing emails. The calculation is: Precision = TP / (TP + FP).
- ***Recall:*** Evaluating the comprehensiveness of the classifier's outcomes in detecting phishing emails, Recall is computed as: Recall = TP / (TP + FN).
- ***F-measure (or F1 score):*** Representing the balanced assessment of the model's efficacy, the F1 score is the harmonic mean of Precision and Recall. It's determined as: F1 Score = 2 * TP / (2 * TP + FP + FN).

V. CONCLUSION AND FUTURE WORK

In conclusion, the study on email phishing detection techniques underscores the constant threat of cyber-attacks, particularly through social engineering. Exploring various attacker methods like pretexting, phishing, spear phishing, and vishing highlights the urgent need for robust prevention measures. The widespread nature of email phishing stresses the importance of proactive cybersecurity efforts. Vigilance and proactive strategies are crucial as attackers become more sophisticated in their deception techniques. The research underscores the importance of ongoing user education and technological solutions like advanced email filtering systems. Combining these approaches enhances resilience against phishing attacks, protecting against potential financial and reputational harm.

REFERENCES

1. Xiang, G., & Hong, J. I. (2009, April). A hybrid phish detection approach by identity discovery and keywords retrieval. In *Proceedings of the 18th international conference on World wide web* (pp. 571-580).
2. Miyamoto, D., Hazeyama, H., & Kadobayashi, Y. (2007, August). A proposal of the AdaBoost-based detection of phishing sites. In *Proceedings of the joint workshop on information security*..
3. Sarika, S., & Paul, V. (2014, March). An anti-phishing framework to defend Tabnabbing attack. In *International conference on security and authentication* (pp. 132-135).
4. Ferolin RJ,Kang C-U. Phishing attack detection, classification and proactive prevention using fuzzy logic and data mining algorithm. onlinepresent.org, 12, 2012, 2012.
5. Wu, L., Du, X., & Wu, J. (2015). Effective defense schemes for phishing attacks on mobile computing platforms. *IEEE Transactions on Vehicular Technology*, *65*(8), 6678-6691.
6. Bin, S., Qiaoyan, W., & Xiaoying, L. (2010, April). A DNS based anti-phishing approach. In *2010 Second International Conference on Networks Security, Wireless Communications and Trusted Computing* (Vol. 2, pp. 262-265). IEEE.
7. Zhang, Y., Hong, J. I., & Cranor, L. F. (2007, May). Cantina: a content-based approach to detecting phishing web sites. In *Proceedings of the 16th international conference on World Wide Web* (pp. 639-648).
8. Varshney, G., Misra, M., & Atrey, P. K. (2016). A survey and classification of web phishing detection schemes. *Security and Communication Networks*, *9*(18), 6266-6284.
9. Şentürk, Ş., Yerli, E., & Soğukpınar, İ. (2017, October). Email phishing detection and prevention by using data mining techniques. In *2017 International Conference on Computer Science and Engineering (UBMK)* (pp. 707-712). IEEE.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details