INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 2, February 2026 | ISSN: 2456-4184 | IJNRD.ORG

PHISHING ATTACKS: A SURVEY OF DETECTION TECHNIQUES AND CHALLENGES

[1]Varshini Rangu, [2]Padmavathi Marri, [3]Siri Pasunuti

[1]Student,
Department of Computer Science and Business Systems,
VNR Vignana Jyothi Institute of Engineering &Technology, Hyderabad, Telangana, India

[2]Student,
Department of Computer Science and Business Systems,
VNR Vignana Jyothi Institute of Engineering &Technology, Hyderabad, Telangana, India

[3]Student,
Department of Computer Science and Business Systems,
VNR Vignana Jyothi Institute of Engineering &Technology, Hyderabad, Telangana, India

Abstract : Phishing has been a very prominent, widespread, and harmful phenomenon amongst all kinds of cybercrimes. On the one hand, the cybercriminals keep exploiting different technical loopholes. On the other hand, they psychologically manipulate the users, and both such situations lead to the final revealing of sensitive information by the victims to the attackers, including login credentials, financial data, and other personal information. Attacks have become more sophisticated mainly due to the presence of such an infrastructure as Phishing as a Service (PaaS) and social engineering. Thus, the traditional defense mechanisms against phishing attacks are not simple and have failed to produce positive results with attackers. In this article, the authors provide a complete survey of phishing attacks and identify their vectors, typologies, and detection mechanisms by combining a large-scale real-world dataset and recent academic research. Moreover, the authors also discuss the limitations of the defense mechanisms equipped with the present state of the art and the important gaps identified, which include zero-second detection, Explainable Artificial Intelligence, as well as Multimodal architectures. The intent of this paper is to facilitate researchers and practitioners in understanding the dynamically changing phishing threats and in developing efficient detection systems.

Keywords **- Phishing Attacks, Cybersecurity, Machine Learning, Deep Learning, Threat Intelligence.**

1. INTRODUCTION

Phishing is a type of social engineering attack where the victim is tricked into sharing sensitive information or releasing malware by clicking on malicious links, which look like legitimate links or services. This attack has been extensively spread through digital communication tools such as email, social media, SMS, and VoIP, turning it into a technical as well as a human problem in terms of security. The first phishing attacks were very basic and easy to spot, but now they have become more complicated. Attackers register domains that are very similar to the real one, send personalized messages with a certain pattern, and mostly generate URLs dynamically. Besides, they also use subscription-based phishing platforms where they launch large-scale attacks with automated tools. Thus, conventional forms of defense like signature-based and blacklist-based approaches against these threats are not effective anymore. To tackle the current phishing threats efficiently, contextualized intelligence information along with machine learning models should be combined. The major contributions of this work are a well-structured taxonomy of phishing attacks, a thorough review of phishing detection methods, an analysis of new threats, and the identification of unresolved research problems and possible future directions.

2. LITERATURE REVIEW

The detection of phishing attacks, also called phishing filtering, has become a hot topic lately because the number of phishing attacks has increased significantly. At first, the experiments for detecting phishing attacks were more concerned with the implementation of blacklist and whitelist methods. Even though such methods involved low computational complexities, their inability to detect zero-day or short-lived phishing attacks made them limited in effectiveness. In order to overcome such identified weaknesses, methods which involve the use of heuristics for detecting phishing attacks were developed by considering characteristic information in URLs, webpage contents, and appearances. However, such methods were related to high false positives. With the advent of data-centric security methods, machine learning approaches have gained more and more attention as tools to increase the accuracy of phishing detection processes. Various experiments with classifiers like Logistic Regression, Support Vector Machines, Decision Trees, and Random Forests have shown the machine learning models to be more accurate and adaptable; nevertheless, ensemble models turned out to be more robust than the others. Despite these enhancements, the focus of

recent research has shifted towards devising deep learning models such as CNN, LSTM, and hybrid models that enable the auto-feature extraction of the given URLs' contents, which in turn results in higher accuracy in the detection of phishing attacks. Nevertheless, these models also bring about new issues concerning interpretation, processing costs, and their suitability for use in real-time. Latest publications underline the importance of real-time intelligence for detecting phishing attacks that target compromised domains hosting phishing sites. Even though there have been significant advances in the field of information security, challenges related to one-time phishing URLs, adversary in the middle, and zero-second explanation mechanisms still remain and call for further research.

3. EVOLUTION OF PHISHING ATTACKS

The first forms of attacks in the phishing domain are based on traditional email scams that incorporate poorly written content, including easily identifiable malicious links. The attacks focused on a broad scope of the population, incorporating a low technical level, making the detection of phishing attacks easy. As a result, phishing gradually transitioned into a sophisticated form of cybercriminal practice, complete with phishing kits, auto-distributors of email, and thriving underground economies. Attackers began effectively incorporating sophisticated social engineering tactics, domain spoofing, and legitimate-looking decoy websites, making phishing attacks more effective. Similar to this, the introduction of 'Phishing as a Service' could be regarded as another factor that completely changed the threat landscape, as it helps attackers who are not so technically skilled to launch large-scale or sophisticated phishing campaigns with the aid of the already existing infrastructure. Currently, the growth in phishing has led to the employment of advanced evasion techniques such as use of HTTPS in the attack, legitimate domains, use of cloud infrastructure in the attack, and use of AI in the attack. Essentially, these are the new phishing techniques that failed with the old defense techniques. These developments make adaptive detection techniques necessary and pose a serious threat to conventional phishing defenses.

PHISHING THREAT MODEL AND ATTACK LIFECYCLE

A typical phishing attack follows a structured lifecycle:

1. Reconnaissance – Gathering target-specific information
2. Preparation – Creating Spoofed Domains, Emails, or web services
3. Delivery – Distributing phishing content via email, SMS, or social media
4. Exploitation – Victim interaction and credential submission
5. Data Exfiltration – Harvesting and storage of stolen information
6. Monetization – Financial fraud, identity theft, or resale

Understanding this lifecycle is essential in order to design appropriate mitigation, mechanisms and to develop the detection mechanisms at various stages of the attack.

Fig. 1. "Phishing attack lifecycle showing reconnaissance, delivery, exploitation, and monetization stages"[18]

4. TAXONOMY OF PHISHING ATTACKS

Phishing attacks can primarily be categorized based on their attack channels, level of targeting, and deception techniques. Understanding and classifying phishing attacks is the first step toward the development of efficient detection systems.

4.1. Email Phishing

Phishing attacks can primarily be categorized based on their attack channels, level of targeting, and deception techniques. Understanding and classifying phishing attacks is the first step toward the development of efficient detection systems.

4.2. Spear Phishing

Phishing attacks that are performed using targeted emails are referred to as spear phishing. The attacker prepares such emails based on publicly available information or data from previous breaches. Therefore, spear phishing emails are highly personalized and less likely to be detected by automated security systems, resulting in a better conversion rate.

4.3. Whaling

Phishing attacks that are performed using targeted emails are referred to as spear phishing. The attacker prepares such emails based on publicly available information or data from previous breaches. Therefore, spear phishing emails are highly personalized and less likely to be detected by automated security systems, resulting in a better conversion rate.

4.4. Smishing (SMS Phishing)

Smishing is a variety of phishing that attackers use leveraging SMS to deliver the malicious content or lure the targets into divulging the sensitive information. The attackers use fake text messages to trick the users into clicking a link to a malware site or providing sensitive information. Because mobile devices are nearly universal, and SMS messages do not show much of a URL preview, smishing attacks are very effective.

4.5. Vishing (Voice Phishing)

Vishing mainly exploits voice calls to extract confidential information from the victims. Attackers acting as customer support, bank officials, or government agents, for instance. By making it urgent or fearful, vishing attacks exploit real-time human interaction, which makes detection challenging.

4.6. Clone Phishing

Clone phishing means duplicating an email or web page that has previously been legitimate. It involves copying the logos, layouts, and content of a legitimate message. The cloned version contains malicious links that replace the legitimate ones. Because this kind of message is so like your regular correspondence, you are very likely to trust it and even respond to it.

4.7. Pop-Up Phishing

Pop-up phishing is the use of fake pop-up windows that randomly appear while browsing. These types of pop-ups usually portray that the computer is infected, or the user has an account problem or software update that requires signing in or downloading an executable file.

4.8. Pharming

Pharming causes redirection of users from genuine websites to fake ones by compromising the DNS settings or local host files. Pharming is different from phishing in that it does not need the user to click on a malicious link; hence, it is much more dangerous and challenging to detect.

4.9. Social Media (Angler) Phishing

Angler phishing is a kind of social media phishing where hackers pretend to be customer support profiles and other trusted figures. The hackers use either direct messages or comments to lure the victims into clicking on harmful links or disclosing their personal data.

4.10. Adversary-in-the-Middle (AITM) Phishing

Adversary-in-the-Middle (AITM) phishing works by the attacker interposing in the communication between the real users and service providers, usually to steal credentials and session IDs, thus bypassing MFA security systems. These types of phishing pose the greatest threat to all currently implemented security solutions. The different types of phishing attacks are a good reason for the diverse nature of the attack. The attackers also can combine several techniques to carry out the attack. The attack then becomes complicated, and the attacker may decide to use a variety of methods. Here, the notion of phishing taxonomy can be very helpful.

Table 1: Phishing Attack Types and Detection Challenges

5. FEATURE ENGINEERING FOR PHISHING DETECTION

The success of deep learning and machine learning models used for phishing detection largely depends on feature engineering. The model's ability to generalize over different phishing campaigns, false-positive rate, and detection accuracy are all aspects that heavily influence the selection of robust and discriminative features. Effective features are

Phishing Type	Primary Vector	Detection Difficulty	Key Challenge
Email Phishing	Email	Medium	High volume and similarity to legitimate emails
Spear Phishing	Email	High	Personalized content bypasses filters
Whaling	Email	Very High	Targets executives with tailored attacks
Smishing	SMS	High	Limited URL visibility on mobile devices
Vishing	Voice	Very High	Real-time human interaction
Pharming	DNS	High	No malicious link interaction required
AITM	Web Proxy	Very High	MFA bypass and session hijacking

5.1. URL Related
Lexical Analysis - String Length, Entropy, Suspicious Tokens.
WHOIS domain registration age and history records.
Redirection chains and closeness to trusted domains.

5.2. Content Level
HTML DOM Patterns.
Keywords indicating an urgent transaction or phishing.
Visual similarity metrics to identify brand impersonation.

5.3. Behavioral Features
Click patterns and session analysis.
Geolocation Anomalies and Device Fingerprint
Feature sets are often combined with real-time intelligence to develop robust detection resistance to false alarms.

6. PHISHING DETECTION TECHNIQUES

The methods used to detect phishing have evolved significantly over time because of the increasing complexity of the attack methods.We can categorize existing methods into four: blacklist-based, heuristic-based, machine learning-based, and deep learning-based techniques. Each category has its own advantages and disadvantages, therefore they can be applied to different scenarios.

6.1 Blacklist-Based Detection

Blacklist-based phishing detection methods rely on a list of URLs, domains, or IP addresses known to be malicious. If a URL or message matches one of the entries on the list, it is either blocked or marked as phishing. Such methods are very fast and simple to execute, hence their widespread usage in web browsers and email gateways. At the same time, blacklists have a couple of very serious limitations when it comes to phishing detection. They are almost unable to detect zero-day phishing attacks and the short lifespan of malicious URLs. Most phishing attacks today generate their URLs dynamically, or a single-use URL is utilized. Hence, blacklists are often updated only after the attack, which greatly diminishes their detection capability.

6.2 Heuristic-Based Detection

Heuristic detection mechanisms review manually created rules and patterns to identify phishing attempts. These rules are derived from various observable traits such as URL length or usage of suspicious keywords, special characters, strange HTML structures, and the similarity of the phishing site to the real site. Heuristics can also find novel phishing attacks; however, in order to maintain effectiveness, constantly updating the rules is necessary. Besides that, very strict rules can lead to increased false positives, thereby negatively impacting the user experience. As phishing attacks evolve rapidly, it becomes more and more challenging to continuously update the heuristic rules.

6.3 Machine Learning-Based Detection

Machine learning (ML) methods interpret phishing detection as a classification task. The models are taught using labeled datasets to distinguish between legitimate and phishing instances. Some of the commonly used ML algorithms are Logistic Regression, Naïve Bayes, Support Vector Machines, Decision Trees, and Random Forest.
Ensemble methods like Random Forest and Gradient Boosting demonstrate greater robustness and higher accuracy. They are capable of uncovering complex, nonlinear feature relationships. Although ML methods outperform traditional methods substantially, the success of the former depends largely on the feature engineering quality and the access to labeled data. Moreover, ML models are vulnerable to concept drift, meaning the changes in phishing tactics result in the models' accuracy dropping over time.

6.4 Deep Learning-Based Detection

Deep learning (DL) models not only tackle the drawbacks of traditional machine learning (ML) methods but also allow for automatic feature extraction from raw inputs such as URLs, webpage content, and email text. People have experimented with Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and CNN and

LSTM hybrid models for phishing detection.The CNN-based model is capable of determining the spatial features of a URL and the layout of a website, whereas the LSTM model is good at capturing the time dependencies. While DL strategies could raise the detection accuracy at a higher level, they also come to light with issues such as their high complexity, long training times, lack of interpretability, and restrictions on real-time deployment.

Table 2: Comparison of Phishing Detection Techniques

Detection Technique	Key Approach	Advantages	Limitations
Blacklist-Based	Matches URLs/domains with known malicious lists	Low computation cost, fast response	Ineffective against zero-day and single-use URLs
Heuristic-Based	Rule-based analysis of URLs and webpage content	Detects previously unseen attacks	High false positives, requires frequent rule updates
Machine Learning-Based	Classification using handcrafted features	Improved accuracy and adaptability	Feature dependency, concept drift
Deep Learning-Based	Automatic feature extraction using neural networks	High detection accuracy for complex attacks	High computational cost, lack of interpretability
Hybrid Approaches	Combines intelligence + ML/DL	Balanced performance, reduced false positives	Increased system complexity

7. COMPARATIVE ANALYSIS OF PHISHING DETECTION TECHNIQUES

Researchers have carried out a study comparing various phishing detection methods, demonstrating that each method has a limited set of attacks that it can effectively defend against and a set of attacks that it can defend against to a lesser extent. Phishing detection methods using blacklists are extremely efficient in detecting known threats, but they fail when the attacker uses a new way. Heuristics can be less of a problem and more easily adapt to new attack methods, however, the use of heuristics is often associated with higher maintenance costs, more false positives, and a less semantically clear classification of the outcomes. Machine learning approaches can result in higher detection accuracies and can be made to work in different situations through different feature extraction methods such as deep learning. Nevertheless, it is necessary to retrain these methods regularly and also selection of features is a very important matter. Due to their superior performance in recognizing complex attacks, deep learning methods also face difficulties with scalability, explainability, and resource utilization. Therefore, the implementation of hybrid detection systems which combine different methods are being increasingly suggested as a way to strike a balance between the accuracy, scalability, and real-time performance of the systems.

8. ADVANCED AND EMERGING PHISHING TECHNIQUES

Nowadays, phishing attacks often utilize fancy evasion techniques to escape detection. A single-use phishing URL is one of a few tricks. Once a victim has clicked on this phishing link, it will no longer work and the threat of detection via blacklists is therefore greatly reduced. Adversary-in-the-Middle (AITM) attacks, where perpetrators insert themselves between the communication of a user and a genuine service, thus being able to steal credentials and session tokens, frequently bypass the multi-factor authentication mechanism. Compromised domain attacks, on the other hand, are when attackers manage to host phishing pages on legitimate yet compromised websites. These kinds of attacks present the biggest difficulty in detection since visitors are conditioned to trust those sites.

9. CHALLENGES AND OPEN RESEARCH PROBLEMS

Phishing detection has come a long way, but here are some challenges that still have to be cracked:

- Zero-day phishing detection: It is still challenging to identify attack types at the moment when they appear.
- Concept drift: Model performance deteriorates as a result of the rapid transformation of phishing tactics.
- Explainability: the lack of understanding of deep learning models restricts users from trusting and thus utilizing them.

- Data imbalance: Non-phishing cases outnumber phishing samples to a large extent.
- Real-time constraints: Selection of the rightful balance point between detection accuracy and latecy.
- Resolving these issues is what really defines powerful anti-phishing tools.

Table 3: Challenges and Future Research Directions

Challenge	Description	Future Research Direction
Zero-Day Phishing	No prior signature exists	Zero-second detection systems
Single-Use URLs	URL expires after one use	Real-time intelligence feeds
AITM Attacks	Session hijacking	Token-aware detection
Explainability	Black-box DL models	Explainable AI (XAI)
Privacy	User data inspection	Federated learning

10. FUTURE RESEARCH DIRECTIONS

The spotlight of forthcoming research should be on creating intelligent phishing detection systems. These are some of the main areas of focus:

- Explainable Artificial Intelligence (XAI): Enhancing the understanding and confidence of users in model decisions.
- Multimodal detection architectures: Integrating the characteristics of URLs, content, images, and user behaviors.
- Federated learning: Enabling a system where private model training is accomplished across several organizations without revealing data.
- Zero-second detection: Detect phishing attacks right at the very first point of contact with the user.
- Integration with zero-trust security models: Steady and resilient enterprises.

11. CONCLUSION

From time to time, phishing goes standalone, changes very quickly, and always lurks on the internet. It is broadly based on taking advantage of technological vulnerabilities and human behavior for stealing sensitive information. This article attempts to provide a comprehensive overview of phishing attacks through their evolution, classification, feature strategies, detection methods, and new challenges. The findings of the study indicate that before deep learning was introduced, traditional methods based on blacklists and heuristics could only offer a limited level of protection. However, with the adoption of machine learning and deep learning approaches, detection has become more accurate and versatile in dealing with sophisticated phishing campaigns. Nonetheless, the paper also argues that a single detection method is unsuitable for a full-fledged fight against the diverse and rapidly changing nature of phishing attacks. Nowadays, sophisticated attacks such as Spotlight on these research areas will be key to the development of scalable, robust, and reliable phishing detection systems that can effectively take on the challenge of increasingly complex and automated phishing campaigns in the wild. Single-use phishing URLs, adversary-in-the-middle attacks, and domain compromise techniques pose a great threat to existing defense systems. Thus, it is of utmost necessity to have a multi-layered defense system that integrates real-time threat intelligence, adaptable learning models, and transparent decision-making processes to efficiently mitigate phishing risks. Moreover, the outcomes of this survey illuminate the direction that future research should be steered towards, such as zero-second detection, explainable artificial intelligence, multimodal detection configurations, and privacy-preserving learning methods, which are equally important. Concentrating on these research domains will even be more indispensable in developing scalable, robust, and dependable phishing detection systems capable of effectively counteracting complex and automated phishing campaigns in the wild.

REFERENCE

[1] Anti-Phishing Working Group (APWG), Phishing Activity Trends Report, APWG, 2023.
[2] R. Dhamija, J. D. Tygar, and M. Hearst, Why phishing works, *Proc. SIGCHI*, 2006.
[3] M. Jakobsson and S. Myers, Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft, Wiley, 2007.
[4] C. Whittaker, B. Ryner, and M. Nazif, Large-scale automatic classification of phishing pages, *Proc. WWW*, 2010.
[5] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, Beyond blacklists: Learning to detect malicious web sites from suspicious URLs, *Proc. ACM SIGKDD*, 2009.

[6] M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, Intelligent phishing detection system for e-banking using fuzzy data mining.
[7] S. Marchal, J. François, R. State, and T. Engel, Off-the-hook: An efficient and usable client-side phishing detection approach, *Proc. NDSS*, 2014.
[8] M. Khonji, Y. Iraqi, and A. Jones, Phishing detection: A literature survey, *IEEE Commun. Surveys Tuts.*, 2013.
[9] Phishing website detection via deep learning, *IEEE Access*, 2019.
[10] O.K. Sahingoz et al., Machine learning based phishing detection from URLs, *Proc. IEEE ICMLA*, 2019.

[11] S. Rao and A. Pais, Detection of phishing websites using an efficient CNN–LSTM model.
[12] A. Oest et al., Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis, *Proc. USENIX Security*, 2018.
[13] Proofpoint, *State of the Phish Report*, Proofpoint Inc., 2024.
[14] Microsoft Security, Adversary-in-the-Middle Phishing Attacks, Microsoft Threat Intelligence, 2023.
[15] Zvelo, *Phishing Detection In-Depth*, Zvelo Inc., 2024. Available: https://zvelo.com/phishing-detection-in-depth/
[16] Kaspersky Lab, Phishing page lifecycle analysis, *SecureList*, 2023.
[17] ENISA, Phishing: Threat Landscape, European Union Agency for Cybersecurity, 2023.
[18] "Phishing attack life cycle," ResearchGate. Available: https://www.researchgate.net/figure/Phishing-attack-life-cycle_fig2_358678073