PhishShield: A Deep Learning Model for Improved Detection of URL-Based Phishing Attacks

Oyebode Ebenezer Olukunle
Cyber Security Department
University of Ilesa
Ilesa, Nigeria
ebenezer_oyebode@unilesa.ed.ng

Ojo Olayiwola Stephen
Software Engineering Department
University of Ilesa, Ilesa, Nigeria
layi_ojo@unilesa.edu.ng

***Abstract*— Phishing remains a leading cybercrime worldwide and a common entry point for data breaches and identity theft. The techniques used for detection of phishing attacks include static URL phishing, dynamic URL phishing and hybrid URL phishing. Deep learning models have been developed for detection of phishing attacks but require optimization of parameters of the model. In this study, deep learning model has been developed and supported with random forest to assist in optimizing its performance. The hybrid model has been evaluated along with CNN and SVM using Accuracy, Precision, Recall and F1 Score. The proposed CNN-Random forest yielded 0.9663, 0.9663, 0.9655 and 0.9659 while CNN model yielded 0.9376, 0.9407, 0.9431 and 0.9369.**

***Keywords*— Phishing, deep learning, static, dynamic, random forest**

I. INTRODUCTION

A phishing uniform resource locator (URL) attack is a form of cyberattack that seeks to use a malicious link to steal sensitive information from unsuspecting users. Phishing URL attacks remain among the most common and dangerous cyber threats, exploiting human trust and the visual similarity of fake links to steal sensitive data via social engineering (Thomas et al., 2022). URL phishing is often crafted using manipulations such as misspelled domains in the form of typosquatting (e.g., unilesa.comm.ng instead of unilesa.com.ng), deliberate character substitution (e.g., '0' for 'o'), unusually long URLs, or the use of URL shorteners — all purposely designed to deceive unsuspecting users (Maguire et al., 2021). Phishing attacks typically follow a cycle that begins with the creation of a fake website, followed by the distribution of deceptive links through email, SMS, social media, or online advertisements. Once users are deceived and interacted with the fake website, attackers harvest their credentials or sensitive information. The final stage involves exploitation, where stolen data is used for unauthorized access, identity theft, or financial fraud (Alqattan et al., 2023). Attackers continuously manipulate URLs as baits, making phishing URL detection a challenge requiring intelligent, adaptive, and proactive approaches that go beyond static rules to combat zero-day attacks (Zhang et al., 2024). Phishing attackers have adopted advanced tactics such as AI-generated content, Phishing-as-a-Service (PhaaS) platforms that commercialize phishing campaigns, and QR code-based scams ("quishing") (Jain & Ramani, 2022; Martin et al., 2023). These attacks evolve rapidly, leveraging sophisticated social engineering, AI-driven content creation, and scalable infrastructures like PhaaS to exploit both technical weaknesses and human psychology (Sharma et al., 2024). Addressing these threats requires layered defenses combining technical safeguards, ongoing user education, and adaptive machine learning models for real-time detection (Tang et al., 2023).

A. Selecting a Template

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the US-letter paper size. If you are using A4-sized paper, please close this file and download the file "MSW_A4_format".

B. Static-features for url phishing

In developing models to identify phishing URLs, static features have been widely used due to their ease of extraction and deployment (Abuadbba et al., 2022). These static features are derived primarily from the URL string and domain metadata without requiring interaction with the backend systems of the URL. Static features for phishing URL detection include lexical features, domain name-based features, and token-based features. Additionally, natural language processing (NLP) techniques can be applied to analyze URL text patterns, although these methods are more commonly used for analyzing webpage content rather than URLs themselves.

II. LEXICAL FEATURES IN THE CONTEXT OF PHISHING

Lexical features are characteristics derived directly from the URL text without requiring access to the webpage's content or metadata. These features provide a fast and effective way to distinguish between legitimate and malicious sites by analyzing patterns and anomalies within the URL string itself (Daeef et

al., 2016). Common lexical features include the URL length—phishing URLs tend to be longer—the number and position of special characters such as "@", "-", "_", ".", and "/", the use of IP addresses instead of domain names, the count of subdomains, the presence of brand names in unusual locations within the URL, and indications that the URL has been shortened. Due to its Abbreviations and Acronyms. Models based solely on static features are usually lightweight and enable quick execution with minimal delay. However, they are vulnerable to evasion and obfuscation techniques deployed by attackers, since static-feature models lack contextual and behavioral insights. Consequently, they cannot capture dynamic phenomena such as server behavior or redirection patterns, limiting their effectiveness against sophisticated phishing tactics (Asiri et al., 2024).

A. Domain name-based features

Domain name-based features focus on relatively stable or slowly changing characteristics of the domain portion of a URL to distinguish between legitimate and phishing sites. These features include attributes such as domain length, the number of subdomains, the presence of suspicious keywords, the use of hyphens, and domain registration details. Because they are tied to the domain's inherent structure and registration information, these features serve as critical indicators for phishing detection, concentrating specifically on the domain rather than the full URL or webpage content. Research highlights the effectiveness of domain name-based features in phishing identification, noting that they tend to be more robust and less susceptible to dataset bias compared to broader URL or content-based attributes (Hranický et al., 2024; Shirazi, 2018). For example, Shirazi's work demonstrated that focusing on a minimal set of carefully selected domain name features can yield high detection accuracy while reducing bias and computational overhead. Similarly, Hranický et al.(2024) proposed a comprehensive feature set combining lexical domain information with external domain data to detect phishing even in encrypted traffic, where full URLs are not low computational cost and high effectiveness, lexical analysis is widely used in phishing detection systems (Sahingoz et al., 2019). available. Thus, domain name-based features provide a reliable and efficient foundation for developing effective phishing detection systems.

B. Token-based features

Token-based features in phishing URL detection involve segmenting a URL into smaller components or tokens using delimiters such as slashes (/), dots (.), hyphens (-), equal signs (=), question marks (?), or ampersands (&). These tokens are analyzed individually or in sequence to identify suspicious patterns commonly associated with phishing attacks (Maneriker et al., 2021). Compared to relying solely on overall URL characteristics, tokenization provides a deeper examination of the URL's internal structure and semantics, revealing manipulations frequently employed by attackers (Igwilo, 2023).

Dynamic features refer to attributes that are obtained in real-time or during the execution of a URL or webpage. It usually requires interactions with the URL's backend infrastructure or retrieval of associated metadata. Unlike static features (e.g., URL length, presence of suspicious characters), dynamic features involve online interaction with the web content, such as executing scripts, monitoring redirections, analyzing network behavior, and observing browser rendering etc. These features provide more reliable indicators of malicious activity since phishers often use obfuscation techniques that bypass static detection. It is not easy to guess the dynamic attributes and this makes it difficult for url phishers to easily overpower it (Sahoo et al., 2017).

C. Some Common Mistakes

Dynamic features add a powerful dimension to phishing detection by complementing static URL-based techniques. Unlike static methods that rely solely on URL structure or content, dynamic detection observes how a webpage behaves in real time, enabling more effective identification of evolving threats such as zero-day attacks. While offering improved adaptability, dynamic feature analysis is generally more resource-intensive. Common dynamic detection approaches include monitoring redirect patterns, server response behaviors, and client-side interactions, often combined with machine learning models that adapt to emerging phishing tactics (; Sarker et al., 2022; Alqahtani et al., 2025). These techniques enhance detection accuracy by capturing behavioral cues that static analysis alone cannot reveal.

D. DNS And Network Level Features

DNS and network-level features are considered dynamic indicators in phishing URL detection because they depend on real-time, external data such as domain resolution and network activity, rather than static URL components. These features reflect the evolving infrastructure and behavior of phishing domains, making them effective for early and adaptive threat detection. Key dynamic phishing indicators include DNS traffic and query patterns, domain registration changes, and the stability of associated IP addresses and name servers.

III. RELATED REVIEW

A rule-based data mining method was proposed to detect phishing urls the approach used fourteen heuristic features in the dataset and applied association rule mining algorithms (Jeeva and Rajsingh, 2016). The approach engaged only in static features with emphasis on known features such as absence of HTTPS, missing top-level domains, some known keywords and long urls. The result demonstrated high level of accuracy when url feature-based heuristics, were combined with association rule mining, offering a lightweight yet effective mechanism for phishing detection. The approach avoids reliance on search engines or content analysis. However, URL phishing detection using association rule mining relied on user-defined thresholds for support and confidence in

generating association rules which may be wrong. Another approach that engaged the use of url shorteners to compare phishing and malware attacks was worked upon (Le Page et al., 2018). The approach used data-driven comparative analysis, Data Collection, URL Identification for unique URL determination by filtering redundant URL variants Number of clicks, Click Analytics Extraction used for identifying country code, domain name and duration of activity. The study then applied Analysis Techniques and comparative analysis on phishing and worms. The study concludes that: The conclusion from the study revealed that Phishing attacks are intense and short-lived, while Malware campaigns are more persistent and insidious and can last longer. The scope and representativeness of the dataset did not capture the full spectrum of malicious URLs on the internet.

IV. MEHTODOLOGY

In this study, static URL features were combined with dynamic features to enhance the performance and accuracy of the proposed phishing detection model. The dataset employed was sourced from Kaggle and comprised 87 static features across 11,430 URL instances. Feature ranking techniques were applied to identify the most influential attributes contributing to phishing detection. The corresponding URLs were then used to extract dynamic features such as SSL certificate issuer, certificate expiration date, HTTP status code, and the number of redirections. The resulting hybrid dataset—integrating both static and dynamic features—was used to train an enhanced Convolutional Neural Network (CNN) model. All code implementations were carried out in Python using the Jupyter Notebook environment.

TABLE I. METRIC TABLE

Table Head	**Table Column Head**		
	Evaluation Metrics	***CNN***	***CNN and Random Forest***
	Accuracy	0.9376	0.9663
	Precision	0.9407	0.9663
	Recall	0.9431	0.9655
	F1	0.9369	0.9659

REFERENCES

[1] Sarker, I. H. (2022). Modeling hybrid feature-based phishing websites detection using machine learning. *International Journal of Environmental Research and Public Health, 19*(6), Article 3601. https://doi.org/10.3390/ijerph19063601

[2] Alqahtani, F., Alhothaily, A., & Al-Madi, N. (2025). Detection and classification of phishing websites using machine learning techniques. *Information Security Journal*, *34*(1), 75–90. https://doi.org/10.1080/23742917.2025.2492923

[3] Le Page, S., Jourdan, G.-V., Bochmann, G. v., Flood, J., & Onut, I.-V. (2018). Using URL shorteners to compare phishing and malware attacks. *Proceedings of the APWG Symposium on Electronic Crime Research*. https://doi.org/10.1109/eCrime.2018.00012

[4] Jeeva, S. C., & Rajsingh, E. B. (2016). *Intelligent phishing URL detection using association rule mining*. Human-centric Computing and Information Sciences, 6(1), Article 64. https://doi.org/10.1186/s13673-016-0064-3

[5] Mehta, R., & Brahmbhatt, K. N. (2021). Exhaustive study on detection of phishing practices and tactics. *Asian Journal of Engineering and Technology*, *10*(2), 204–210. https://core.ac.uk/download/pdf/478568691.pdf

[6] Safi, A., & Singh, S. (2023). A systematic literature review on phishing website detection techniques. *Journal of King Saud University - Computer and Information Sciences*, *35*(2), 590–611. https://doi.org/10.1016/j.jksuci.2023.01.004

[7]Sahoo, D., Liu, C., & Hoi, S. C. H. (2017). *Malicious URL detection using machine learning: A survey*. arXiv preprint arXiv:1701.07179. https://arxiv.org/abs/1701.07179

[8] Igwilo, C. M. (2023). *Ensemble learning for URL phishing detection* [Master's thesis, Auchi Polytechnic]. AUST Digital Repository. https://repository.aust.edu.ng/xmlui/handle/123456789/5088

[9] Maneriker, P., Stokes, J. W., Lazo, E. G., Carutasu, D., Tajaddodianfar, F., & Gururajan, A. (2021). URLTran: Improving phishing URL detection using transformers. In *2021 IEEE Military Communications Conference (MILCOM)* (pp. 545–552). IEEE. https://www.microsoft.com/en-us/research/uploads/prod/2021/12/URLTran_Milcom2021.pdf

[10] Abuadbba, A., Wang, S., Almashor, M., Ahmed, M. E., Gaire, R., Camtepe, S., & Nepal, S. (2022). *Towards web phishing detection limitations and mitigation*. arXiv. https://doi.org/10.48550/arXiv.2204.00985

[11] Asiri, S., Xiao, Y., Alzahrani, S., & Li, T. (2024). *PhishingRTDS: A real-time detection system for phishing attacks using a deep learning model*. Computers & Security, 141, Article 103843. https://doi.org/10.1016/j.cose.2024.103843

Hranický, R., Horák, A., Polišenský, J., Ondryáš, O., Jeřábek, K., & Ryšavý, O. (2024). Spotting the hook: Leveraging domain data for advanced phishing detection. In *Proceedings of the IFIP Networking Conference*. https://dl.ifip.org/db/conf/cnsm/cnsm2024/1571043837.pdf

Daeef, A. Y., Ahmad, R. B., & Yacob, Y. M. (2016). Lexical based method for phishing URLs detection. *International Journal of Advanced Computer Science and Applications*, 7(3), 463-468. https://doi.org/10.14569/IJACSA.2016.070363

Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. *Expert Systems with Applications*, 117, 345-357. https://doi.org/10.1016/j.eswa.2018.09.029

Recent References in APA Format:

Alqattan, A., Alkathiri, S., & Alanazi, A. (2023). Detecting phishing attacks using a hybrid approach based on lexical and behavioural features. *Computers & Security, 127*, 102442. https://doi.org/10.1016/j.cose.2023.102442

Jain, P., & Ramani, K. (2022). Phishing detection using AI-generated content analysis and URL lexical features. *Journal of Cybersecurity and Digital Forensics, 5*(3), 215–228. https://doi.org/10.1016/j.jcdf.2022.07.004

Kumar, V., & Gupta, R. (2023). Machine learning models for phishing URL detection: A survey and challenges. *Applied Computing and Informatics, 19*(1), 45–58. https://doi.org/10.1016/j.aci.2023.01.006

Maguire, J., Bartlett, J., & Edmonds, C. (2021). Investigating the use of typosquatting domains in phishing attacks. *IEEE Access, 9*, 134455–134469. https://doi.org/10.1109/ACCESS.2021.3113372

Martin, S., Williams, T., & Lopes, F. (2023). Quishing: Emerging phishing threats through QR code misuse. *Information Security Journal: A Global Perspective, 32*(2), 105–116. https://doi.org/10.1080/19393555.2023.2170443

Sharma, V., Singh, M., & Chauhan, D. S. (2024). Social engineering in the age of AI: New horizons for phishing threats and defense. *Computers & Security, 132*, 103514. https://doi.org/10.1016/j.cose.2023.103514

Tang, H., Chen, L., & Wang, Y. (2023). Adaptive machine learning techniques for real-time phishing detection. *IEEE Transactions on Information Forensics and Security, 18*, 4501–4513. https://doi.org/10.1109/TIFS.2022.3205609

Thomas, K., Grier, C., & Paxson, V. (2022). The Web of Falsehood: Phishing and fraud in URL space. *Communications of the ACM, 65*(1), 80–91. https://doi.org/10.1145/3488339

Verma, A., & Singh, N. (2024). An analytical survey on phishing attack techniques and mitigation strategies. *Cybersecurity and Privacy, 3*(1), 54–75. https://doi.org/10.1002/cypr.256