An Overview: Phishing Attack Detection and Mitigation Strategies

R.Padma Devi*[1], V.Khanaa[1]

[1]Department of Computer Science and Engineering, Bharath institute of higher education and research, Chennai, Tamilnadu, India
rpadmadevi@gmail.com

Abstract. Internet has also become very common and this has changed the daily lives of people because now people can easily shop online, communicate through the internet and access other government functions. This growth has also been accompanied by the fact that there is so much sensitive data being exchanged over the internet, and cybercriminals have taken advantage of this situation to carry out phishing in order to capitalize on such an opportunity. Such practice as phishing has also been among the most persistent cyber threats as any attacker constantly changes their methods, making malicious websites more plausible and simpler to implement. With the development of such tactics, it is becoming harder to detect and prevent phishing. This paper contains a comprehensive overview of various techniques of phishing attacks and how they are directed toward the online users. It also looks at the key security initiatives that have been created to combat such threats, their strength, and their weakness. It is on this knowledge that the study postulates better protection strategies aimed at strengthening the stability, dependability, and general security of anti-phishing systems.

Keywords: Phishing, Cybersecurity, Cyber Threats, AntiPhishing Techniques, Social media security, machine learning, Deep learning.

1. Introduction

Both the tremendous growth of digital information and the high rate of internet users have presented opportunities as well as challenges in the online world. The contemporary search engines are playing a crucial role as access points to extensive digital resources where people can find out the appropriate information in a few seconds. Nevertheless, even with significant technological progress, search engines are not always able to provide a way to filter dangerous or harmful information. Consequently, users are often subjected to malicious websites, such as phishing web pages in search results. The safety and reliability of information to the users is an issue that has become a significant concern in particular due to the fact that cyberattacks are becoming more advanced and prevalent. [1]

Phishing, initially recorded in 1996, is considered to be one of the most threatening and long-lasting cybercrime types. The name of the term is based on the word fishing; it describes the misleading procedure of how the attackers use false information to get the victims to share confidential data, posing as a respected company. These attacks are usually in the form of emails, websites, text messages or social media posts that resemble authentic messages. Phishing normally takes place in organized order: a deceptive bait, a scam hook in the form of a spoofed web page, and the final catch in which the victim freely and unconsciously discloses

S. P. Vijayaragavan et al. (eds.), *Proceedings of the Global Conference on Sustainable Energy Systems, Smart Electronics and Intelligent Computing (GCSESEIC 2025)*, Advances in Engineering Research 297,
https://doi.org/10.2991/978-94-6239-654-8_52

personal information. With the help of such attacks, cybercriminals can steal logins, financial data, and personal information, which can be used in identity theft, fraudulent transactions, and massive security breaches.

Despite the substantial amount of research carried out in the field of phishing detection and prevention, a lot of the literature available to date is specialised either in a single field (machine learning algorithms, URL-based classification, browser security tools, or user education). The only missing element is the wide and cohesive analysis of different strategies, their advantages and drawbacks, and the scrutiny of the emerging solutions that demonstrate the newest technological trends. As phishing remains among the biggest contributors to data breaching cases, it is increasing in magnitude. The 2024 IBM Cost of a Data Breach Report indicates that the average cost of a data breach in the world went up to USD 4.88 million in 2024 compared to USD 4.45 million in 2023, which is a reflection of the severity and economic cost of a data breach. [2]

With these issues, the paper at hand is expected to give an indepth review of the phishing attack mechanics, the contexts in which they thrive, and how the tactics of the attackers have been developed. Also, the research considers the traditional defense mechanisms and the advanced AI-based detection models and provides an analytical comparison to assess the efficiency of both. This review could be useful in supporting the existing endeavor to come up with more resilient and adaptive anti-phishing measures, particularly by pinpointing existing gaps as well as pointing at the promising avenues of the future.[3][4]

2. Literature Survey

Due to its persistence and development, phishing has become a popular topic of research in the field of cybersecurity. Initial research was mainly oriented to detecting the phishing attacks by means of heuristic and blacklist methods. These systems evaluated rules that were preset like suspicious URLs, mismatch in domain names, or irregular redirection patterns to identify fraud sites. Although it was effective, blacklistbased solutions were not able to keep up with the pace of new phishing sites and zero-day attacks, thus more adaptable
solutions were required. [5]

As machine learning (ML) progressed, scientists started addressing data-driven techniques of detection. Managed ML classifiers, including Decision Trees, Random Forests, Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) have been extensively used to distinguish legitimate and phishing websites with reference to URL structure, HTML characteristics, lexical structures and HTTP request patterns. Research has shown a significant increase in accuracy compared to conventional methods; but even the ML models have issues to deal with in terms of feature engineering and generalization to unseen phishing patterns.[6] The recent studies have moved more toward the methods of deep learning, which utilizes the models of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long-Short-Term Memory (LSTM), and Transformers. These models are automatic in

extracting the high level features of raw URLs, webpage screenshots and HTML code, which helps to eliminate the dependence on the manually created features. Image analysis, such as CNN, has been found to be useful in identification of visual similarities between phishing sites and legitimate sites. Equally, RNN and LSTM are applied to scan patterns of URLs and texts in order to identify the deceptive patterns. Even though the computational cost and datasets size are a major constraint, deep learning is highly accurate. [7]

The other primary research area is email phishing detection. Research in this field uses natural language processing (NLP) and text classification methods to detect deceptive email messages, grammar errors, calls to urgency or calls to social engineering. NLP and ML hybrid models or deep learning have recorded encouraging outcomes, particularly when identifying spear-phishing emails that are based on psychological manipulation and not on technical wonders. Some scientists have also focused on user-friendly innovations like awareness education, browser extensions, and graphic measures of security to minimize vulnerability to phishing. [8]

Although these techniques make users more vigilant, they are mostly dependent on the behavior of a user which may not be consistent. Social engineering attacks can be used to capitalize on human emotions and therefore behavioral interventions are not enough. More recent works attack the problem of adversarial machine learning, in which attackers intentionally corrupt input data to avoid detection mechanisms. With the growing sophistication of phishing schemes based on fast-flux hosting, HTTPS, AI-based content detection models should resist adversarial defenses. Research on this field suggests powerful learning systems, team learning, and anomaly-detection systems to improve
security.[9][10]

3. Techniques

Phishers use numerous approaches to distribution in order to make the most of their scam sites. The most widespread one is a fake emailing when attackers pose as well-known organizations like banks, financial services, social networks, or government bodies. The language such emails usually use is that of urgency, fake alerts or tempting offers in an attempt to exploit the emotional state of the users in order to make them act. These messages have hidden evil links or attachments that redirect the targeted victims to phishing sites to steal confidential data. In addition to the old-fashioned email-based strategies, the attackers are using various communication methods to expand their influence. Phishing links can be widely spread with the help of SMS messages (smishing), instant messaging systems, and social media because they utilize the popularity and fast communication nature of these channels. On social networks specifically, there are instances where visitors create false profiles, steal real accounts, or develop deceptive advertisements to make more people fall prey to accessing malicious links. The other new tactic is search engine poisoning, where the criminals are able to use search engine optimization (SEO) processes to improve the ranking of

the bad websites. Incorporating specific keywords and working around the vulnerabilities of algorithms effectively causes such malicious pages to be included in the list of search results that are legitimate, and this will raise the chances that the user will visit them by mistake.[11] Likewise, hijacked websites which are compromised may authorize pages which were taken over with a vulnerability may be used to host concealed phishing pages or redirect users to attacker-controlled domains against their knowledge. There are also advanced forms of distribution, including automated infrastructures, phishing kits and botnets, that facilitate the process of creating, copying and delivering phishing pages in large quantities. Phishing kits offer templates and scripts that need minimum technical skills and enable attackers to roll out a lot of campaigns in a short time. The botnets facilitate the propagation of the phishing URLs in internet, and the takedown is more difficult because the attack is distributed. The complexity and dynamism of these distribution strategies indicate how fast the threat environment is changing. With the attackers constantly moving to new channels and taking advantage of technological developments, the requirement of changing and intelligent detection mechanisms is becoming more and more important. [12][13] Fig. 1 shows the key methods that cybercriminals will employ to sell phishing sites and interact with target victims.

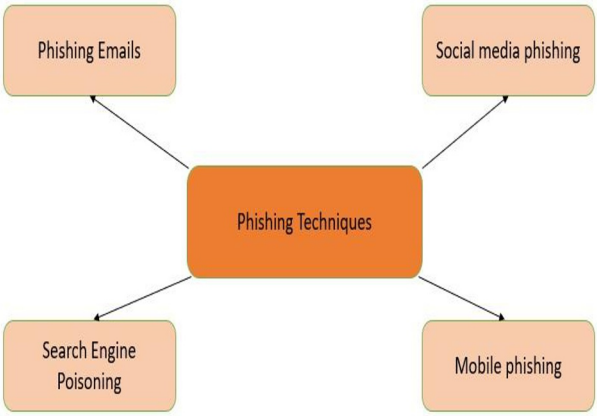

Fig.1. Phishing techniques

3.1 Phishing Emails

The most prevalent and ancient phishing method of the cybercriminals is via email. Attackers create deceptive mail which resembles the authoritative organization such as banks, e-commerce corporations, delivery service or government entities to influence the users to provide personal information. In these emails, different types of bait are used, such as giving free products, giving notifications of fake rewards, a verification need in the account, or a threat of a service interruption. The messages are designed so as to closely appear as a genuine communication by imitating branding, logos, writing styles, and identities of the senders.[14] The majority of phishing emails include harmful links that transfer the recipient to a fake webpage that would replicate

an authentic portal. As soon as the users provide their credentials or financial information to log in, the stolen data goes directly to the servers controlled by the attackers. Whereas phishing campaigns in the past were based on generic and widely broadcasted emails, present day attackers exploit more advanced means. Spear-phishing schemes focus on specific people by using personal information, whereas Business Email Compromise (BEC) attacks take advantage of the vulnerable systems used by organizations to authenticate users to anonymously act as an executive or employee, therefore, making the trick more believable and the success more probable.

3.2 Social Media Phishing

Since many people have embraced the use of social platforms, social media has been a very effective tool that cybercriminals are currently utilizing to perpetrate phishing activities. Fraudsters use bogus accounts that are set to impersonate a famous brand, a famous personality, or a service provider. These are rogue profiles that send phishing messages to the users either directly or in open posts. In the rest, hackers steal authorized accounts either by hacking the passwords, malware, or social engineering. After accessing an account, the attacker uses the inherent trust between the account holder and his/her network, and phishing becomes more believable. Social media phishing messages are usually similar to the usual email-based attacks. They usually contain enticing or threatening messages or warnings, and then bogus URLs redirecting the users into unsafe sites that are meant to steal confidential data. These connections may be sent directly, i.e., in such applications as Facebook Messenger, Instagram Direct, or a Twitter DM, or published publicly to attract more users. Leveraging of trusted social ties is one of the factors towards the fast development of the social mediadriven phishing schemes and their growing efficiency.
[15][16]

3.3 Mobile Phishing (Smishing and App-Based Phishing) The switch to mobile device usage has provided new opportunities to phishers as it has been changing very quickly. The SMS phishing, also referred to as smishing, is the process of sending a text message in which the author impersonates the bank, courier service, or an agency of the state. Phishing URLs are more frequently delivered in the popular messaging apps such as WhatsApp, Telegram, and Signal. These messages tend to replicate official messages, and they have malicious links like those of phishing emails. Once they are clicked, users will be redirected to fake websites that will collect personal details. The most recent studies show that cases of mobile-based phishing are increasing drastically. Research conducted by Lookout [17] has shown that phishing is being faced at an increasing pace by mobile users, and the number is growing at an annual rate of up to 85%. This explosion is informed by the fact that the users have a lot of trust in personal messaging applications and they are prone to reacting promptly to mobile alerts. The increased use of smishing and mobile-based phishing scams shows the necessity to build more effective mobile security systems, have better security on their devices, and educate
users to minimize the exposure to these new threats. [18]

3.4 Search Engine Poisoning

Search Engine Poisoning (SEP) is a methodology whereby attackers employ search engine algorithms to rank malicious webpages to the first page search ranking positions. SEP also contributes greatly to chances of victims visiting phishing sites because they usually trust the websites that are ranked top. The ways that cybercriminals use to gain top rankings include: keyword stuffing, manipulating links, and taking advantage of the weakness of the SEO. Particularly popular way of SEP is compromising legitimate and high-authority websites. Aggressors inject such sites with obscure keywords, rogue scripts or redirecting mechanisms. The hacked site is found in the top list when the users search in the related key words. The site is redirecting the user to a phishing site which is designed to steal sensitive information instead of loading the legitimate content expected. This legitimacy and high-visibility make SEP a very efficient and
dangerous phishing distribution technique**.**[19][20]

4. Solution

Various studies have been conducted with the aim of establishing methods that can protect users of the internet against phishing attacks. These approaches can broadly be classified into two large categories namely human-based solutions and software-based solutions Fig 2.

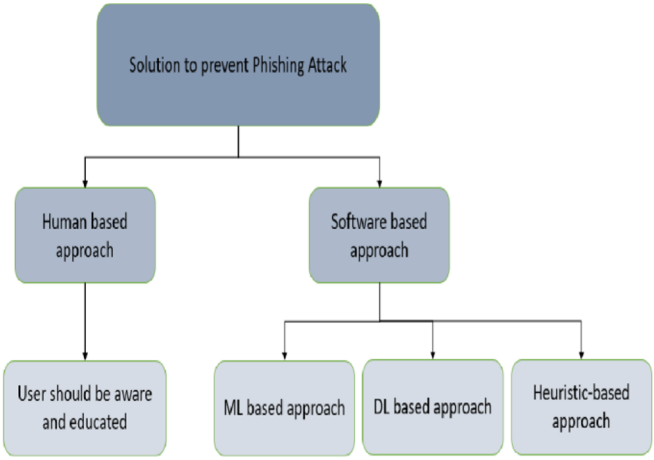

Fig.2. Solution to Phishing attack

4.1 Human-Based Approaches

The human-based methods are mostly focused on training the user and making him better at identifying suspicious or fake websites. Considering that most phishing attacks use social engineering to control user behavior, it has been suggested that raising user awareness is one of the most vital defense strategies. Such techniques

involve training courses, awareness activities, simulated phishing activities, browser cues, security tool bars, and visual messages that are meant to assist people to scrutinize the authenticity of websites. These methods equip users with decision-making empowerment skills that enable them to identify and prevent phishing more easily. Nevertheless, effective methods that rely on human intervention require regular interaction, awareness, and being computer savvy. [21]

$$Accuracy = TP + TNTP + TN + FP + FN \quad (1)$$

$$Precision = TPTP + FP \quad (2)$$

In the detection of phishing attacks, system performance is generally quantified through a set of four major criteria: Accuracy, which denotes the ratio of instances that were correctly classified (both phishing and legitimate) to the total number of cases Precision, which gives the amount of the sites labeled as phishing that is genuinely phishing Equation (1,2,3,4)thus minimizing the false alarms; Recall (Sensitivity), which shows the antiphishing systems ability to detect all the current phishing sites correctly and not to miss the threats;and F1, Score, which represents the trade, off between Precision and Recall by computing their harmonic mean, thereby taking into account both the detection capability and the reliability of the system. These evaluation metrics are based on the classification results of True Positives (phishing correctly detected), True Negatives (legitimate sites correctly classified), False Positives (legitimate sites misclassified as phishing), and False Negatives (phishing sites missed), thus together they offer a detailed measurement of the efficiency of anti, phishing systems.

$$Recall = TPTP + FN \quad (3)$$

$$F1 = 2 \cdot Precision \cdot RecallPrecision + Recall \quad (4)$$

4.2 Software-Based Approaches

The second one is automated systems which identify and block phishing sites without the participation of the user. These solutions are meant to categorize web pages into legitimate or malicious web pages with the help of various techniques, which include machine learning, deep learning, blacklists, whitelists, browser extensions, and real-time threat intelligence systems. After a potential phishing page has been detected, the software will block it, issue a warning or record the threat to be reviewed at a later point. There are also some tools, which show warnings or help messages and allow the user to realize what kind of threat it is and what precautions to take. The software-based techniques offer proactive and realtime protection, which is why the risk of human error is minimized and the defense can be displayed in a more homogeneous way to various groups of users.[22]

4.3 Machine Learning Approach

Researchers have also in recent years examined hybrid machine learning methods, wherein a group of classifiers are combined in order to improve the performance of phishing detection. Bagging, AdaBoost and Gradient Boosting are ensemble techniques that combine the benefits of multiple models in order to reduce variance, minimize false positives and identify more diverse phishing patterns. Besides, the current systems have behavioral and contextual capabilities such as the traffic pattern

of a website, the domain registration information, the SSL certificate information and visual similarity score to enhance robustness. Though these rich sets of features greatly enhance the accuracy in detection, they also add complexity in computation and they need to be constantly updated to be effective in overcoming evolving phishing techniques. The deep learning-based method has gained more popularity in phishing detection studies in an attempt to address the shortcomings of the traditional ML models that solely used manually crafted features. Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are models that can automatically acquire complex representations using raw data, i.e. URLs, HTML code, and webpage screenshots. The features allow DL models to detect advanced phishing websites that resemble valid websites or dynamically generated content. Nevertheless, deep learning methods also have such issues as the fact that they require large and high-quality datasets, require significant computational resources, and are open to adversarial manipulation. Consequently, scalable and real-time phishing detection is a research topic that is still under development.[23]

4.4 Deep Learning Approach

Besides these widely used architectures, scholars have started testing hybrid deep learning architectures or models that integrate several neural network architectures to enhance the strength of phishing detection. As an example, CNNLSTM and CNN-GRU hybrids are based on the use of both spatial and sequential feature representation, using which the models can be trained to detect the structure of URLs, HTML structure, and visual webpage features. Another type of attention-based mechanisms have also been applied to phishing detection pipelines, to emphasize the most pertinent portions of the input sequence, and enhance the classification accuracy of obfuscated or dynamically-generated phishing URLs. Such hybrid and attention-based frameworks are much more flexible but demand even more computational resources and attention to dataset curation to prevent overfitting. More recently, developments are in multimodal deep learning, in which data on the webpage, in the form of URL strings, webpage screenshots, DOM trees, and hosting metadata, is synthesized to build a complete representation of each webpage. This allows the detection systems to learn delicate visual, structural and behavioral information that may not be explicitly visualized by the traditional single-source methods. Nevertheless, the multimodal DL models create further changes including the issue of data synchronization, training complexity, and storage demands. Nevertheless, these constraints notwithstanding, multimodal and hybrid DL approaches are a promising avenue towards creating more resilient, more generalizable and scalable phishing detection systems that have the capacity to keep up with the faster changing cyber-attack methods.[24]

4.5 Heuristic Approach

Recently, heuristic-based detection systems have been extended to feature behavioral and contextual information, which allows them to gain a more dynamic insight into phishing actions. These features examine the interaction pattern of users, webpage loading behavior, script execution and redirection sequence to detect suspicious attributes that might be missed by the heuristics based on the stable URL. Also, the contemporary heuristics compare the domain registration data, the attributes of an issued certificate of the SSL, and page popularity statistics to

distinguish between the trusted and the untrusted websites. Though these enhanced heuristics are resistant to simple phishing attacks, they still fail in cases where attackers have advanced evasion techniques like fast URL changing, deception with legitimate hosting services or obfuscating a script to hide ill intent. To overcome these limitations modern studies combine heuristic techniques with machine learning and deep learning models to create new hybrid systems of detection that would provide the flexibility of data-driven models, as well as the interpretability of heuristic approaches [29]. In these methods, the heuristic features are used as inputs to the classifiers, enhancing the capability of the model to differentiate attacks that are of the zero-day nature and also analyze large amount of web data effectively. Although the hybrid approach will minimize false positives and enhance robustness, it is also associated with the difficulty of feature engineering, recurrent model re-training, and computational complexity. However, heuristic-improved hybrid systems can be viewed as a good way of developing more stable phishing detection systems, that can evolve in accordance with the current trends in cyber-attacks.[25]

5. Overcome Phishing Attack

In this section, the current research on the subject matter and the recent studies on web security and phishing detection methods are discussed and reviewed. The chosen scientific articles cover a variety of approaches and algorithms that would improve the accuracy and reliability of phishing detection, such as machine learning, deep learning, humanawareness programs, and visual similarity detection schemes. This review will give a crucial background on where our proposed system fits in the existing methods and the strengths and weaknesses of existing methods.[26][27][27][28]

Table. 1. Comparative strategies

Technique	Strengths	Weaknesses
Blacklist-based	Simple, fast, widely used	Cannot detect new/unknown phishing sites
Heuristic-based	Detects suspicious patterns	High false positives
Machine Learning (ML)	Learns complex patterns, adaptable	Requires large datasets, training cost
Deep Learning (DL)	High accuracy, automatic feature learning	Computationally expensive
Hybrid Approaches	Combines multiple methods for robustness	Complexity in integration

- Doe, Do et al. created a phishing detection system that categorizes the URLs as phishing or legitimate by effectively examining the URL structures by

incorporating both character-level and word-level embeddings Table 1. They used Multi-Head Self-Attention (MHSA) and Temporal Convolutional Networks (TCN) to resolve the shortcoming of Recurrent Neural Networks (RNNs) and standard Convolutional Neural Networks (CNNs). The MHSA mechanism allowed the system to be more discriminative with regard to features that are necessary to the URL and greatly enhanced the discriminative capabilities. Consequently, the model had an excellent accuracy of 98.78. Moreover, the use of TCN allowed to make the system more efficient to identify advanced phishing patterns, which proves to be highly effective against profile changes.

- Zhang and Liu et al. suggested a hybrid phishing detection system combining the visual similarity analysis with the deep learning-based URL category in 2024. They involved deriving structural and visual information on webpage screenshot- layout patterns, similarity of logos, and colour composition- until it was combined with lexical and host-based URL features. The visual data were processed using a dual-branch CNN and the sequential URL patterns were analyzed using the Bidirectional LSTM (BiLSTM). Combining these multimodal features the system recognized some 97.42 accurate, which was much better than the traditional URL-only detection techniques. The paper had observed that the effectiveness of text and visual characteristics in the combination to improve resistance to phishing websites that are developed to resemble legitimate websites.

- Alazab et al. proposed a phishing detection model based on Graph Neural Network (GNN) in 2023 and detected relational patterns among URLs, domains, IP addresses, and hosting infrastructures. The system considered phishing indicators as links with each other and the system could identify suspicious links, e.g. shared hosting environments, duplicate usage of malicious domain names, or abnormal redirection chains. The GNN model was able to identify and detect attacks with a high accuracy of 96.85% and it was particularly found to be efficient in detecting zero-day attacks. Even though the benefits are evident, the authors observed that the model was time-consuming to build the graph structure and that could be a hindrance to real-time implementation.

- Kumar and Singh created a phishing detection algorithm based on Transformer-based language models and specifically fine-tuned variants of BERT developed on large datasets of phishing URLs in 2025. They used semantics of URLs, dependencies of tokens, and subword interactions to identify obfuscation methods like homograph attacks and character substitutions. The model has a high accuracy of 99.12% with a high resistance to adversarial manipulation of URLs. It was also launched as a lightweight and browser-based version. The research however noted that
Transformers need periodic retraining to be effective since the URLs used in phishing are changing at a very fast rate, and large models do not come without computation cost**.**

6. Conclusion

Moreover, the analysis of the recent practices shows that the sphere of detecting phishing is also underdeveloped to meet new methods of attacks, which become more advanced. The methods that are traditionally based on blacklist and heuristic are still of some use in identifying known threats, but they cannot be changed to suit zero-day attacks. The methods based on machine learning have high detection rates due to their ability of pattern recognition but extensive data quality biasing (high-quality feature engineering and balanced datasets). Although more resistant to external shocks and able to extract features automatically, deep learning-based methods are more demanding in terms of computational resources and training data and therefore might be not applicable in a low-resource setting. These constraints signify that there has not been any single approach which can offer a full-scale resolution and this supports the application of hybrid and multi-layered defence mechanisms. To sum up, phishing attacks are becoming more sophisticated, and it will require a set of technical innovation, long-term threat intelligence, and user awareness. The new avenue of research should focus on devising scalable, real-time detection systems that combine various sources of data (URL structures, visual webpage content, and behavioral indicators) to enhance the system against new threats. Moreover, it is possible to further increase the transparency and reliability of systems with the implementation of adaptive learning models, adversarial-resistant architecture, and explainable AI methods. With the help of such developments, organizations will be able to develop more holistic and aggressive phishing defenses, which will eventually lead to a generally stronger cybersecurity resiliency in the constantly changing digital environment.

.

References

1. Y. Huang, Q. Yang, J. Qin, and W. Wen, “Phishing URL detection via CNN and attention-based hierarchical RNN,” in 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), IEEE, pp. 112–119, 2019.
2. H. Wang, L. Yu, S. Tian, Y. Peng, and X. Pei, “Bidirectional LSTM Malicious webpages detection algorithm based on convolutional neural network and independent recurrent neural network,” Applied Intelligence, vol. 49, pp. 3016–3026, 2019.
3. T. Feng and C. Yue, “Visualizing and interpreting rnn models in url-based phishing detection,” in Proceedings of the 25th ACM Symposium on Access Control Models and Technologies, 2020, pp. 13–24.

4. L. Yuan, Z. Zeng, Y. Lu, X. Ou, and T. Feng, "A character-level BiGRU-attention for phishing classification," in Information and Communications Security: 21st International Conference, ICICS 2019, Beijing, China, December 15–17, 2019, Revised Selected Papers 21, Springer, pp. 746– 762, 2020.

5. S. Al-Ahmadi, "PDMLP: phishing detection using multilayer perceptron," International Journal of Network Security & Its Applications (IJNSA) Vol, vol. 12, 2020.

6. N. Q. Do, A. Selamat, O. Krejcar, and H. Fujita, "Detection of malicious URLs using Temporal Convolutional Network and Multi-Head SelfAttention mechanism," Appl Soft Comput, vol. 169, p. 112540, 2025.

7. O. Sarker, A. Jayatilaka, S. Haggag, C. Liu, and M. A. Babar, "A Multi-vocal Literature Review on challenges and critical success factors of phishing education, training and awareness," Journal of Systems and Software, vol. 208, p. 111899, 2024.

8. D. Li, Q. Chen, and L. Wang, "Phishing Attacks: Detection and Prevention Techniques," Journal of Industrial Engineering and Applied Science, vol. 2, no. 4, pp. 48–53, 2024.

9. M. A. Adebowale, K. T. Lwin, and M. A. Hossain, "Intelligent phishing detection scheme using deep learning algorithms," Journal of Enterprise Information Management, vol. 36, no. 3, pp. 747– 766, 2023.

10. A. Mughaid, S. AlZu'bi, A. Hnaif, S. Taamneh, A. Alnajjar, and E. A. Elsoud, "An intelligent cyber security phishing detection system using deep learning techniques," Cluster Comput, vol. 25, no. 6, pp. 3819–3828, 2022.

11. Y. Lin et al., "Phishpedia: A hybrid deep learning based approach to visually identify phishing webpages," in 30th USENIX Security Symposium (USENIX Security 21), pp. 3793–3810, 2021.

12. R. S. Rao, T. Vaishnavi, and A. R. Pais, "CatchPhish: detection of phishing websites by inspecting URLs," J Ambient Intell Humaniz Comput, vol. 11, pp. 813– 825, 2020.

13. A. K. Jain and B. B. Gupta, "A machine learning based approach for phishing detection using hyperlinks information," J Ambient Intell Humaniz Comput, vol. 10, pp. 2015–2028, 2019.

14. Adewole, K. S., Akintola, A. G., Salihu, S. A., Faruk, N., and Jimoh, R. G. (2019). Hybrid rule-based model for phishing URLs detection. Lecture Notes Inst. Comput. Sci. Soc. Inf. Telecommun. Eng. 12, 119–135. doi: 10.1007/978-3-030-23943-5_9,2019.

15. Alabdan, R. (2020). Phishing attacks survey: types, vectors, and technical approaches. Fut. Int. 12:168. doi: 10.3390/fi12100168

16. Aljofey, A., Jiang, Q., Rasool, A., Chen, H., Liu,W., Qu, Q., et al. An effective detection approach for phishing websites using URL and HTML features. Sci. Rep.12:10841. doi: 10.1038/s41598-022-10841-5. (2022)

17. Jain, A. K., and Gupta, B. B. (2017). Phishing detection: analysis of visual similarity based approaches. Secur. Commun. Netw, 1–20. doi: 10.1155/2017/5421046. 2017

18. Anti-Phishing Working Group (APWG) (2024). Phishing Activity Trends Report, 3rd Quarter 2022

19. P. Yi, Y. Guan, F. Zou, Y. Yao, W. Wang, and T. Zhu, "Web phishing detection using a deep learning framework," Wirel Commun Mob Comput, vol. no. 1, p. 4678746,2018

20. H. Le, Q. Pham, D. Sahoo, and S. C. Hoi, "URLNet: Learning a URL representation with deep learning for malicious URL detection. arXiv 2018," arXiv preprint arXiv:1802.03162, 2018Ali M, Jung LT, Sodhro AH, Laghari AA, Belhaouari SB, Gillani Z. A Confidentiality-based data Classification-as-aService (C2aaS) for cloud security. Alex Eng J; 64(2): 749–760. 2023.

21. Balaji, A., Sathyasri, B., S, V.V.R., Indumathy, D., Krishnan, R., Vanaja, S.: Intruder Alert System in Smart Home based on IoT Technique. (2022). https://doi.org/10.1109/icpects56089.2022.10047243.

22. Aoudni Y, Donald C, Farouk A, Sahay KB, Babu DV, Tripathi V, Dhabliya D. Cloud security-based attack detection using transductive learning integrated with Hidden Markov Model. Pattern Recognit Lett; 157: 16–26. 2022.

23. Nadeem M, Arshad A, Riaz S, Zahra SW, Dutta AK, Al Moteri M, Almotairi S. An Efficient Technique to Prevent Data Misuse with Matrix Cipher Encryption Algorithms. Comput Mater Contin; 74(2): 4059–4079. 2022.

24. Upadhyay D, Zaman M, Joshi R, Sampalli S. An efficient key management and multi-layered security framework for SCADA systems. IEEE Trans Netw Service Manag; 19(1): 642–660. 2021.

25. Vanitha, V., Joe, S.B., Krishnan, R., Fletcher, A.S.A., Anju, M., Akila, V.: Cognitive Threats Detection Model using Nature Inspired Chimpanzee Optimization for IoT Networks (CCM-COM). In: Atlantis highlights in engineering/Atlantis Highlights in Engineering. pp. 629–637 (2025). https://doi.org/10.2991/978-94-6463-754-0_55.

26. N. Q. Do, T. C. Nguyen, and H. T. Nguyen, "Detection of malicious URLs using Temporal Convolutional Networks with Multi-Head SelfAttention," Journal of Network and Computer Applications, vol. 240, pp. 1–12, 2025.

27. Y. Zhang and H. Liu, "A hybrid phishing detection framework using visual similarity and deep learning-based URL classification," *IEEE Access*, vol. 12, pp. 145321–145334, 2024.

28. M. Alazab, S. Venkatraman, A. Alazab, and A. S. Alhyari, "Phishing URL detection using Graph Neural Networks (GNN): A relational learning

approach," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 5021–5033, 2023.

29. R. Kumar and S. Singh, "Transformer-based malicious URL detection using fine-tuned BERT models," *Computers & Security*, vol. 134, pp. 1–14, 2025.