Machine Learning Approaches for Malicious URL Detection: A Literature Survey

Meenu S Nair
Department of Computer Science
(Cyber Security)
Vimal Jyothi Engineering College
Chemperi, Kannur
Email:meenusnair692@gmail.com

Sanjay S Kumar
Department of Computer Science
(Cyber Security)
Vimal Jyothi Engineering College
Chemperi, Kannur
Email:sanjaysk.7704@gmail.com

Muhammed Hawaz PK
Department of Computer Science
(Cyber Security)
Vimal Jyothi Engineering College
Chemperi, Kannur
Email:muhammedhawaz1222@gmail.com

Akshay
Department of Computer Science
(Cyber Security)
Vimal Jyothi Engineering College
Chemperi, Kannur
Email:akshaysajeevan557@gmail.com

Mr.Rasheed Ahamed Azad V
Assistant Professor
Department of Computer Science
(Cyber Security)
Vimal Jyothi Engineering College
Chemperi, Kannur
Email:rasheedklpm@vjec.ac.in

***Abstract*—As cyber threats keep getting more advanced, malicious URLs have become a huge problem, causing data breaches, malware infections and financial losses. Hackers use tricky methods, like sneaking harmful links into emails, social media and even legit-looking websites. Traditional blacklist-based detection isn't very reliable since it needs constant updates and often misses new threats. To tackle this, machine learning provides a smarter approach by analyzing patterns and behaviors of malicious URLs. Algorithms like SVM, Random Forest, Logistic Regression and deep learning models like XGBoost and MLPs have shown really good accuracy in detecting harmful links. But there are still challenges, like dealing with imbalanced datasets, needing high-quality data and handling the high computational costs. This survey looks into different machine learning-based detection techniques, their strengths and weaknesses and why ongoing improvements are necessary to stay ahead of evolving cyber threats.**

***Index Terms*—Machine Learning , Malicious URL, Cybersecurity , Malware Detection , LSTM , Kaggle , Safe Browsing.**

I. Introduction

Cyber threats are increasing day after day and malicious URLs have become a huge problem. The attacker makes use of it for spreading malware, stealing personal data and phishing scams. As almost every person depends on the internet for communication and transactions, hackers have invented cunning ways to hide harmful links in emails, social media messages and websites that look totally legitimate. Traditional methods such as blacklists, signature-based systems and rule-based approaches are no longer sufficient. Blacklists become outdated quickly, leaving systems open to new threats and signature-based methods require constant updates but still fail to catch newly created malicious URLs.

Machine learning has been a game changer in detecting harmful URLs. Instead of simple string matching with a list, machine learning models analyze many factors like the URL's structure and domain details along with patterns in redirects to detect the threats. Different techniques of Random Forest, Decision Trees, Logistic Regression, K-Nearest Neighbors (KNN) and deep learning models like Multilayer Perceptrons (MLPs) and XGBoost have been tested.Each of these models has its own pros and cons,ensemble models like XGBoost and Random Forest are very accurate, while simpler models like Logistic Regression work faster, making them better for real-time detection.

However, machine learning-based detection isn't perfect. One major issue is class imbalance-there are way more normal URLs than malicious ones, which can make models less effective at spotting the malicious ones. The quality and variety of training data also matter a lot since cybercriminals keep changing their tactics to avoid detection. Some lightweight methods, like checking URL attributes (such as Top-Level Domains or dot count), are fast but might miss more advanced threats. Deep learning models are powerful, but they need a ton of computing power, making them harder to use in low-resource environments.

This literature survey looks at different research studies on machine learning-based malicious URL detection, discussing their methods, strengths, weaknesses and performance. It also explores various feature extraction techniques, model comparisons and new trends in the field, emphasizing the importance of constantly improving detection strategies to keep up with evolving cyber threats.

II. Literature Review

J. Khalife et al. proposed a new approach [1] to detect malicious URLs using a simple heuristic approach that focuses on only two key elements: The TLD (top-level domain) and the dot count in the URL[1]. The report highlights the growing threat of phishing and malicious links that can lead to financial losses,data theft and malware infections. While traditional machine learning techniques like random forests can effectively detect these threats, they often rely on analyzing multiple URL attributes, which makes the process more complex and time-consuming. To solve this problem, the plan uses specific selection criteria that identifies the most important features and uses simple rules for classification. This method resulted in a 95 percent improvement in allocation time and a 93 percent reduction in training time while keeping efficiency similar to existing methods. It is designed to be a quick and effective first step in identifying potentially problematic URLs before performing more advanced analysis. However, despite these improvements, it may still struggle against more advanced threats, as attackers can adapt and find ways to bypass the system.

M. D. Karajgar et al. [2] examined the use of machine learning techniques to improve cybersecurity by detecting harmful URLs before users interact with them. The study compares various models, like random forests, Naive Bayes, K-nearest neighbors and decision trees, to analyze over 700,000 URLs, aiming to differentiate between safe and malicious ones. One notable strength of this approach is the use of ensemble models, especially random forests and extra trees, which showed high accuracy of over a 89.7 percent in identifying harmful URLs.These machine learning models can analyze various kinds of URLs and identify patterns that indicate whether the URL is harmful or not. By recognizing these patterns, the models can more easily detect malicious links that might lead to phishing sites or spread malware. where there are way fewer malicious URLs in the dataset compared to normal ones. This makes it harder for the model to correctly spot the malicious URLs, lowering its precision and recall. The authors suggest that fixing this imbalance could make detection better. Another limitation identified in the research is the model's reduced performance when dealing with less common types of malicious URLs, which affects its overall accuracy.Even with these challenges, the study still highlights that tree-based ensemble models are pretty good at improving malicious URL detection and helping boost cybersecurity.

R. Stoleriu et al. presents a system [3] that uses machine learning for the detection of malicious short URLs, combined with threat intelligence services like VirusTotal and PhishTank. Since attackers often use short URL services to hide harmful links, it can be tough for users to spot them. To solve this, the system checks short URLs by following any redirects to their final destination and scanning them against over 70 Anti-Virus engines. If the URL isn't flagged as malicious, it's passed to an machine learning model that looks at 90 different features to classify the URL. Using the Random Forest algorithm, the system achieves an impressive 97 percent accuracy, providing real-time analysis with better accuracy than traditional blacklisting methods. However, the system does have some limitations, like relying on the quality of threat intelligence data, facing challenges with processing large datasets and struggling to detect new or unknown types of attacks that weren't included in the training data. In the future, the system plans to expand its capabilities for detecting malicious short URLs used in phishing attacks via SMS (smishing), helping protect users across different platforms.

R. R. K. Menon et al. [4] discusses the importance of detecting malicious URLs in order to protect web applications from threats like phishing, malware and other online dangers. Cybercriminals often exploit human mistakes, using tricks like phishing, pharming and social engineering, by hiding harmful content behind fake-looking URLs. To fight this, the paper suggests using machine learning to examine the behavior and characteristics of URLs, helping to identify the malicious ones. It also introduces a updated collection of URL features and actions that can make the detection process more accurate. The proposed method uses machine learning algorithms like Decision Tree , Random Forest , Support Vector Machine and K-Nearest Neighbour to analyze and categorize URLs.The results from the experiment shows that this method is effective in detecting harmful URLs and also offers a simple but still an efficient solution. Unlike traditional methods that rely on fixed rules or large datasets, this study uses unique URL features to detect malicious URLs while maintaining high precision. These findings could help improve the identification of malicious URLs in real-time security systems.But still, there are some downsides to this approach. It relies a lot on the quality of the dataset that the model was trained on, meaning that it may struggle to detect new types of malicious URLs that are absent in the training data. Additionally, when applied to a large number of URLs, the system may require more computational resources, which could greatly impact its performance.

U. S. B et al. [5] highlights the growing issue of malicious URLs, which are often used for phishing, spreading malware and causing data breaches-especially as the usage of mobile devices are increasing rapidly. To address this issue, the study explores logistic regression, a popular machine learning method, to separate harmful URLs from safe ones. The model looks at features like suspicious keywords, unusual character patterns and strange subdomains to identify potential threats. Logistic regression is easy to understand and efficient, making it a great choice for systems with limited resources. Despite this, there are some drawbacks to this method.Logistic regression works on the idea that the relationship between URL features and harmful behavior is straightforward, but this makes it less effective against more complex or tricky attacks used by cybercriminals. Additionally, gathering high-quality datasets, keeping models up to date with new threats and dealing with privacy concerns in data collection are all challenges that need to be addressed. While logistic regression is a good starting point and works well in many cases, integrating it with more advanced techniques, such as deep learning or ensemble models, could enhance detection accuracy.The study stresses the need to keep improving cybersecurity methods to stay ahead of rising threats.

Shantanu et al. [6] tackles the problem of detecting malicious URLs, which are a major cybersecurity threat responsible for financial losses, data breaches and malware infections. Traditional methods like blacklisting, regular expressions and signature matching struggle to keep up with the constantly changing patterns and the huge volume of data. To address this, the paper treats malicious URL detection as a binary classification problem and tests different machine learning models using a Kaggle dataset containing 450,000 URLs. Among all the models, Random Forest did the best, achieving the highest accuracy and F1 score. The model was then validated on OpenPhish data, proving its ability to effectively detect malicious URLs. While this machine learning approach improves detection and adapts well to new threats, it still faces some challenges like data imbalance, shifting attack patterns and the need for real-time scalability. One way to enhance its performance would be to train the model on more balanced datasets, which could help reduce false positives and improve detection accuracy.

S. Kinger et al. [7] looks at the rising issue of malicious websites that pose a serious threat to users by stealing personal information and spreading harmful content. It explores how machine learning models can be used to identify these dangerous sites by analyzing their static features to detect hidden vulnerabilities or exploit code. The study tested several models, with the Autoencoder Neural Network being the least effective, achieving only 61.5 percent accuracy. On the other hand, the Multilayer Perceptrons (MLPs) and XGBoost models performed much better, with accuracies of 85.5 percent and 85.35 percent, respectively. However, there are still some challenges: MLPs and XGBoost require significant computational resources, Autoencoders have trouble with accuracy and all of these models rely heavily on high-quality data to work effectively. The researchers used a large dataset of labeled URLs, splitting it into training and testing sets and extracted various features like lexical, network and host-based information. Models like Random Forest, Autoencoder, Decision Tree, MLPs, SVM and XGBoost were trained and tested. While MLPs and XGBoost proved to be the best at detecting malicious URLs, the paper highlights the ongoing challenges related to computational resources and the need for high-quality datasets to achieve the best results.

S. R. A et al. [8] addresses the growing issue of malicious URLs, which cybercriminals often use for phishing, spreading malware and stealing sensitive information. These harmful links are commonly shared through emails, text messages and social media, allowing cybercriminals to target a wide audience of unsuspecting users. Traditional methods like blacklists have trouble keeping up, especially with tactics like URL shortening that help evade security filters. The paper explores various ways to detect malicious URLs, including blacklist-based methods, rule-based systems, machine learning and deep learning techniques. It emphasizes the importance of extracting the right features from URLs, such as lexical, host-based and network-based data, to accurately identify malicious ones.As the usage of machine learning and deep learning is increasing rapidly, their ability to adapt to evolving threats is becoming more crucial. However, there are still challenges, like the need for specialized knowledge and the time-consuming nature of feature extraction. The paper concludes by stressing the importance of continuous innovation in detection systems to stay ahead of the ever-evolving tactics of cybercriminals.

B. R. Hanji and R. Kanagavalli [9] explores the role of machine learning techniques to detect malicious URLs, which are a huge threat to online security, leading to problems like phishing, malware and data theft. The main aim is to classify URLs as either harmful or safe using different machine learning models. Among the models tested, Random Forest performed exceptionally well with a 99.79 percent accuracy, along with perfect precision and recall. Logistic Regression also did great, with a 99.66 percent accuracy and a fast training time of just 3.3 seconds.On the other hand, Support Vector Machine (SVM) was accurate but took a lot longer to train (about 1 hour and 27 minutes). This research shows the important role that machine learning can play in protecting users by spotting harmful URLs before they can do any damage. The next steps for this work could involve expanding the datasets, refining feature selection and improving real-time detection. However, the study does have some limitations, like relying on existing datasets and dealing with class imbalance, which might affect its ability to detect new threats. Despite these challenges, the research emphasizes the significant capability of machine learning in enhancing online security.

A. Lakshmanarao et al. [10] focuses on usage of machine learning to detect malicious URLs, a growing cybersecurity issue due to threats like phishing, malware and data theft. As users browse the web, they often come across links that look safe but are designed to exploit vulnerabilities in systems. The aim is to analyze different URL features-such as domain names and path structures-and train a model that can tell the difference between harmful and safe links. We tested several machine learning models, like SVM, Random Forest and Logistic Regression and measured their performance in terms of precision, recall, accuracy and training time. The Random Forest model did the best, with an impressive 99.79 percent accuracy and near-perfect precision and recall. Logistic Regression was faster to train, making it a good option for real-time detection, while SVM had high accuracy but took longer to train. While the results are promising, there are some challenges. The performance of the model is influenced by the quality of the training data and issues like class imbalance can impact accuracy. The model may also struggle with identifying new or previously unseen malicious URLs. To improve the model, future work could focus on enhancing the dataset, refining feature selection and making the model more adaptable to new threats. Overall, this project highlights how machine learning can significantly improve online security and help protect users from malicious attacks.

TABLE I
Comparison Table

Reference	Description	Advantages	Disadvantages
[1]	A heuristic method uses machine learning to detect malicious URLs by analyzing attributes like TLD and DOTS for fast and efficient classification.	• Drastically reduces classification and training time. • Easy to implement with minimal computational resources.	• May miss sophisticated malicious URLs. • Performance is different across various datasets
[2]	The paper compares machine learning models (decision trees, random forests, KNN, naive Bayes) for detecting malicious URLs	• In-depth comparison of models for malicious URL detection. • High accuracy (over 93 percent) with ensemble models like random forests and extra trees.	• Class imbalance affects precision for detecting minority malicious types. • Limited to standard models, not exploring newer techniques like deep learning.
[3]	The paper presents a system that detects malicious short URLs using machine learning and threat intelligence, achieving 97 percent accuracy with 90 features.	• Achieves up to 97 percent accuracy in detecting malicious URLs. • Integrates threat intelligence for up-to-date URL analysis.	• Relies on external platforms that may not always be current. • Combining multiple techniques makes it complex.
[4]	The paper intoduces a machine learning approach for detecting malicious URLs, using algorithms like Decision Tree, KNN, SVM and Random Forest to improve cybersecurity by analyzing URL features.	• Analyzes new URL properties for better accuracy. • Evaluates various models for optimal detection	• Risk of benign URLs being flagged or malicious ones missed. • Results may not apply across all datasets or URL types.
[5]	Uses Logistic regression for detecting malicious URLs by analyzing features like keywords and character sequences	• Easy to understand classification decisions. • Uses relevant features for better accuracy.	• May miss complex patterns. • Struggles with advanced malicious URLs.
[6]	This paper explores how different machine learning classifiers can be used to detect malicious URLs, treating it as a binary classification problem	• Compares multiple classifiers to identify the most effective method for malicious URL detection. • Uses 450,000 URLs, ensuring more reliable and generalized results.	• Some classifiers, like Random Forest, may be too computationally expensive for real-time systems. • Excludes more advanced techniques like deep learning models, which could perform better.
[7]	This paper investigates various machine learning models (Decision Tree, Random Forest, Autoencoder, XGBoost, MLP, SVM) for detecting malicious URLs	• Evaluates various machine learning models, identifying the most efficient for malicious URL detection.	• Autoencoder showed low accuracy. • Continuous updates are required due to evolving cyber threats.
[8]	This paper reviews techniques for detecting malicious URLs, categorizing them into blacklist-based, rules-based, machine learning and deep learning methods.	• Offers a brief review of current malicious URL detection methods, beneficial for researchers. • In-depth look at features that help identify malicious URLs	• No case studies or real-world experiments to back up the methods. • Doesn’t provide enough details on how to implement the techniques in real-world scenarios.
[9]	The paper examines the growing threat of malicious URLs, focusing on the role of machine learning techniques to detect and prevent phishing, malware and data fraud.	• Focuses on a major issue, malicious URLs, affecting internet security. • Offers insights for developing better cybersecurity tools to protect users from malicious URLs.	• Models may not work well for all types of malicious URLs due to dataset limitations. • The fast evolution of cyber threats demands constant model updates and retraining.
[10]	This paper introduces a machine learning method for detecting malicious URLs,combining NLP techniques and different classifiers.Additionally, a Flask- based web application is created for real-time detec- tion.	• The integration of a Flask web app makes the model easily accessible, allowing users with minimal technical knowledge to interact with the system. • The use of multiple feature extraction techniques (Count Vectorizer, TF-IDF, Hashing) provides a well-rounded approach to handling various URL structures.	• Relying on a single Kaggle dataset may limit the model’s generalizability. • Despite the high accuracy, the model may still produce false positives or false negatives, which can lead to incorrect URL classifications and security risks.

III. Conclusion

With cyber threats constantly evolving, detecting malicious URLs has become more important than ever. This survey examined various machine learning techniques, ranging from basic models such as logistic regression to more complex ones like Random Forest, XGBoost, and deep learning. These models were proven to be more effective than traditional blacklist-based methods since they can recognize harmful links based on patterns rather than relying on outdated lists.However, there are still challenges, like class imbalance, the requirement for high-quality datasets and the fact that attackers keep coming up with new tricks to bypass detection. No single model is perfect, so combining different approaches and continuously updating detection systems is necessary. Future research should focus on improving accuracy, reducing false positives and making these models more efficient for real-time use.Overall, machine learning has made huge improvements in cybersecurity, but it needs to keep evolving to stay ahead of cybercriminals. By refining detection techniques and integrating real-time threat intelligence, we can build stronger defenses against malicious attacks and make the internet a safer place.

References

[1] J. Khalife, F. T. Hussain M Nassar and M. H. M KH Al Marri, ”New Heuristics Method for Malicious URLs Detection Using Machine Learning,” 2023 International Symposium on Networks, Computers and Communications (ISNCC), Doha, Qatar, 2023, pp. 1-6

[2] M. D. Karajgar et al., ”Comparison of Machine Learning Models for Identifying Malicious URLs,” 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), Bangalore, India, 2024, pp. 1-5

[3] R. Stoleriu, C. Negru, B. -C. Mocanu and F. Pop, ”Malicious Short URLs Detection Technique,” 2023 22nd RoEduNet Conference: Networking in Education and Research (RoEduNet), Craiova, Romania, 2023, pp. 1-6

[4] R. R. K. Menon and V. Anandhu, ”Machine Learning Supported Malicious URL Detection,” 2023 4th IEEE Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2023, pp. 1-5

[5] U. S. B, V. S. Bamla, V. Bandari, M. Bheemagani, C. Uppuganti and R. Akkenapally, ”Detection of Malicious URLs using Logistic Regression,” 2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2024, pp. 1127-1132

[6] Shantanu, B. Janet and R. Joshua Arul Kumar, ”Malicious URL Detection: A Comparative Study,” 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 1147-1151

[7] S. Kinger, P. Nirmal, A. Shrivastav, A. Sharma and S. Saindane, ”Malicious URL Detection Using Machine Learning,” 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 1062-1068

[8] S. R. A, M. R, R. N, S. L and A. N, ”Survey on Malicious URL Detection Techniques,” 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 778-781

[9] B. R. Hanji and R. Kanagavalli, ”Machine Learning for the Exposure of Malicious URLs,” 2023 International Conference on Recent Advances in Information Technology for Sustainable Development (ICRAIS), Manipal, India, 2023, pp. 83-88

[10] A. Lakshmanarao, M. R. Babu and M. M. Bala Krishna, ”Malicious URL Detection using NLP, Machine Learning and FLASK,” 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2021, pp. 1-4,