iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VIII **Month of publication:** August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73491

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

Phishing Link Prevention System

Divyalakshmi J[1], Jasmine K[2], Murugesh Pandian P[3]

Abstract: ***Phishing attacks are a major cybersecurity threat that deceive users into revealing sensitive information such as login credentials, banking details, or personal data. The Phishing Link Prevention System is designed to detect and block malicious URLs before users can access them. This system analyzes URL features like domain structure, URL length, special characters, and HTTPS usage, combined with real-time blacklist databases and machine learning classification. Implemented as a browser extension, it offers real-time alerts and ensures a secure browsing experience by preventing access to phishing websites. This proactive tool enhances user safety by reducing the risk of identity theft, financial loss, and data breaches***
Keywords: Phishing, Cybersecurity, URL Analysis, Link Detection, Browser Extension, Machine Learning, Web Security, Real-time Protection, Malicious URLs, Threat Prevention

I. INTRODUCTION

Phishing is one of the most common cyberattacks used by hackers to steal sensitive information like usernames, passwords, and credit card details. These attacks often involve fake websites or malicious links that appear to be from trusted sources. Many users unknowingly fall into these traps, leading to data breaches and financial losses.The Phishing Link Prevention System is designed to protect users from such attacks by detecting and blocking phishing links before they are clicked. This system works by analyzing the structure of URLs, checking against known blacklists, and using machine learning models to identify suspicious patterns. It can be implemented as a browser extension or integrated into web platforms for real-time protection.By warning users and preventing access to harmful websites, this system enhances online safety and reduces the risk of falling victim to phishing scams. It serves as an essential tool for securing personal and organizational data in today's digital environment

II. LITERATURE REVIEW

Over the years, several techniques have been developed to combat phishing attacks, ranging from blacklist-based approaches to machine learning and heuristic-based systems. Traditional blacklist methods, such as Google's Safe Browsing API, maintain databases of known phishing websites, but they often fail to detect newly generated or zero-day phishing URLs. Heuristic-based approaches analyze the structural features of URLs (like length, presence of IP addresses, or suspicious characters), which help in identifying malicious links even if they are not in the blacklist.

However, these methods can produce false positives. To overcome these limitations, recent research has focused on machine learning techniques, where classifiers such as Support Vector Machines (SVM), Random Forest, and Deep Learning models are trained on large datasets of phishing and legitimate URLs. These models improve detection accuracy by learning hidden patterns and behaviors of phishing links. Additionally, researchers have explored the use of browser extensions and real-time detection systems to offer instant protection. Despite advancements, phishing detection still faces challenges due to the evolving tactics used by attackers. This review highlights the need for hybrid systems that combine multiple detection strategies for more accurate and timely phishing prevention

III. PROPOSED SYSTEM

The proposed Phishing Link Prevention System is designed to detect and block phishing URLs in real time, ensuring safer web browsing for users. This system combines URL feature analysis, blacklist verification, and machine learning classification to accurately identify malicious links. Key features such as domain age, URL length, use of HTTPS, presence of special characters, and known phishing patterns are extracted from each URL.

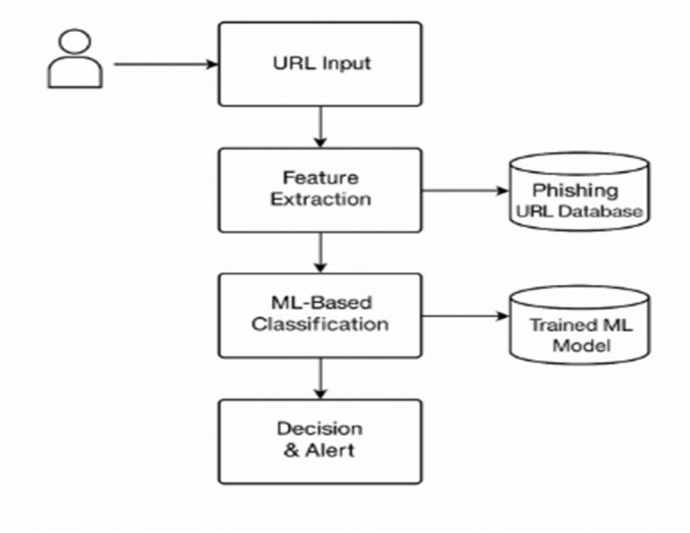
A trained machine learning model (e.g., Random Forest or SVM) is used to classify the URL as safe or phishing based on these features. Additionally, the system cross-checks links against an updated blacklist of known phishing domains. The solution is implemented as a browser extension, allowing seamless integration and real-time alerts before the user visits a suspicious site. This multi-layered approach increases the accuracy and effectiveness of phishing detection and provides users with a proactive shield against cyber threats

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

IV. METHODOLOGY

TheThe Phishing Link Prevention System follows a systematic approach to identify and block phishing URLs in real time. The first step involves data collection, where a dataset containing both phishing and legitimate URLs is gathered from reliable sources such as PhishTank, OpenPhish, and trusted website repositories. In the feature extraction phase, important characteristics of URLs are analyzed—such as length of the URL, presence of special symbols (like "@", "-"), number of subdomains, use of HTTPS, domain age, and more. These features help differentiate between legitimate and phishing websites.

A. Dataflow Diagram



B. Technologies Used

Python, JavaScript, scikit-learn, Chrome Extension API, HTML/CSS, PhishTank dataset, GitHub.

V. SYSTEM DESIGN & IMPLEMENTATION

The system is designed as a real-time browser-based phishing link detection tool. It follows a modular structure with distinct components for feature extraction, blacklist checking, machine learning-based classification, and user alert generation. In the design phase, the system architecture includes input from the user (URL), which is processed through two parallel modules: (1) a Blacklist Module that checks the URL against known phishing sites, and (2) a Feature Extraction Module that analyzes URL characteristics such as length, subdomains, and use of special characters. These features are passed to a trained machine learning model (e.g., Random Forest or SVM) that predicts whether the URL is phishing or legitimate. The final result is sent to the User Alert Module, which displays a warning if a threat is detected

.

VI. RESULTS & DISCUSSION

The Phishing Link Prevention System was tested using a dataset containing both phishing and legitimate URLs. After training the machine learning model (e.g., Random Forest), the system achieved an accuracy of over 95%, with high precision and recall scores, indicating reliable detection of malicious links. The browser extension successfully identified and blocked phishing URLs in real time, alerting users before any interaction with the harmful sites occurred

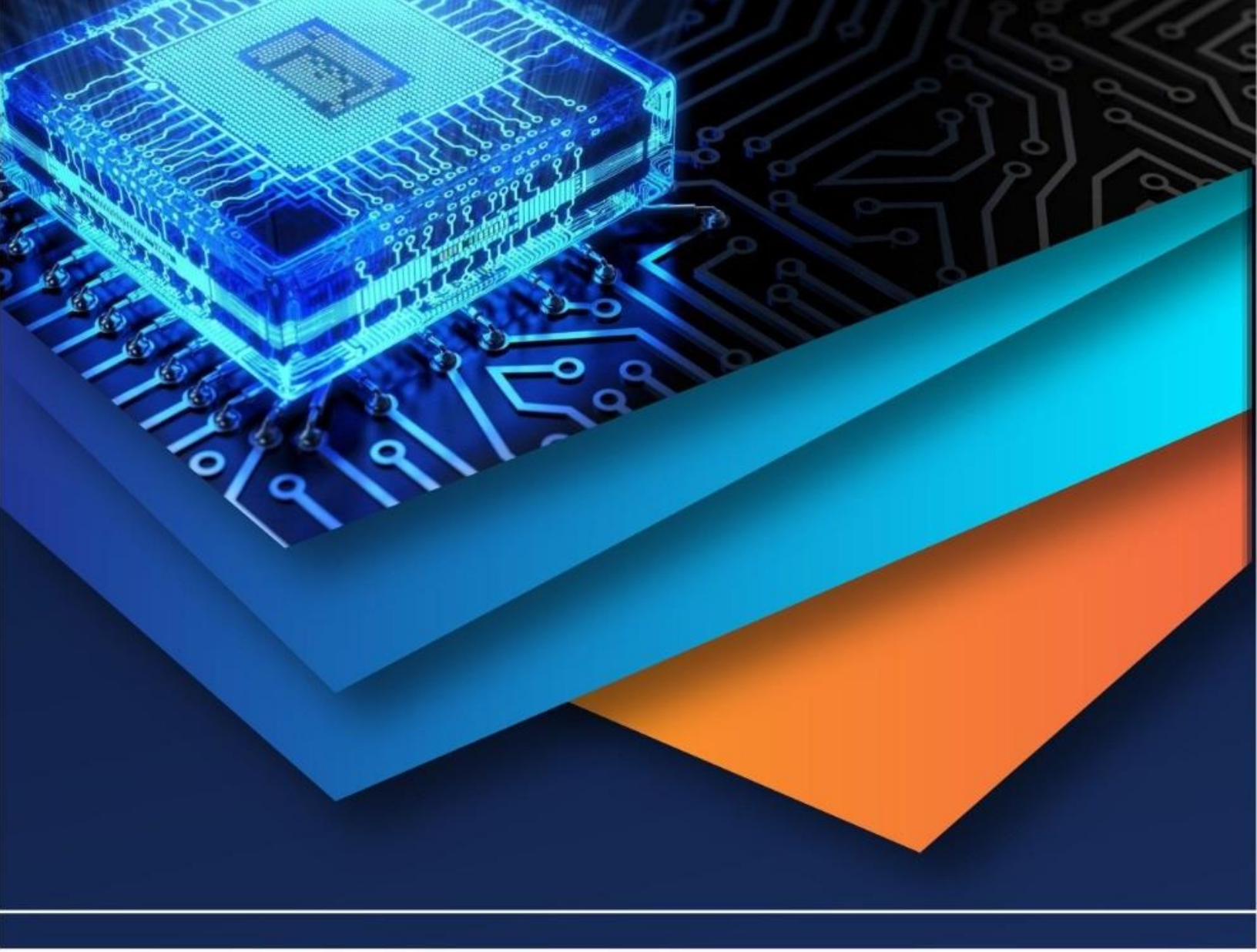
VII. CONCLUSION

The Phishing Link Prevention System provides an effective and practical solution to combat phishing attacks by detecting and blocking malicious URLs in real time. By combining URL feature analysis, blacklist verification, and machine learning classification, the system enhances the accuracy and reliability of phishing detection. Implemented as a browser extension, it offers a user-friendly and efficient way to protect users from potential threats while browsing. The results demonstrate that this multi-layered approach can significantly reduce the risk of phishing attacks and improve overall online security. With further enhancements and regular updates, the system can be extended to offer even broader protection across different platforms and user environments.

REFERENCES

[1] PhishTank. (2024). Phishing URL Data Feed. https://www.phishtank.com
[2] OpenPhish. (2024). Real-time Phishing Intelligence. https://openphish.com
[3] Jain, A. K., & Gupta, B. B. (2018). *Phishing detection: Analysis of visual similarity-based approaches*. *Security and Privacy*, 1(1), e9.
[4] Abu-Nimeh, S., Nappa, D., Wang, X., & Nair, S. (2007). A comparison of machine learning techniques for phishing detection. In Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit (pp. 60–69). ACM.
[5] Dhanalakshmi, R., & Sornalakshmi, S. (2021). Real-time URL phishing detection using machine learning. International Journal of Computer Applications, 183(15), 20–24.
[6] scikit-learn Developers. (2024). scikit-learn: Machine Learning in Python. https://scikit-learn.org
[7] Google Chrome Developers. (2024). Chrome Extensions Documentation https://developer.chrome.com/docs/extensions

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY