International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Rule-Based Approach for Detecting Phishing Websites through Browser Extension

Dr. K. Anandan [1], Vignesh S [2]

[1] Associate professor, Department of Computer Applications, Nehru College of Management, Coimbatore, Tamil Nadu, India anandmca07@gmail.com

[2] Student of II MCA, Department of Computer Applications, Nehru College of Management, Coimbatore, Tamil Nadu, India vigneshh2120@gmail.com

ABSTRACT:

This research paper presents a rule-based approach for detecting phishing websites using a browser extension developed with HTML, CSS, and JavaScript. The system integrates 12 handcrafted rules that analyze website URLs, domains, and content attributes. To strengthen detection accuracy, Google Safe Browsing API is incorporated for blacklist verification. The extension performs on- device evaluation, ensuring real-time performance with minimal latency. Experimental results demonstrate high accuracy and efficiency, offering a scalable solution to counter phishing threats.

KEYWORDS:

Phishing Detection, Browser Extension, Rule-Based System, Google Safe Browsing, Cybersecurity, Web Security.

INTRODUCTION

Phishing has emerged as one of the most prevalent cyber threats, targeting millions of users daily through deceptive websites designed to steal credentials. According to recent cybersecurity reports (2024–2025), phishing attacks account for nearly 36% of all data breaches, with an annual increase of over 20%. Traditional detection mechanisms—such as centralized machine learning systems and blacklist-based filtering—offer strong predictive capabilities but often suffer from latency, privacy concerns, and dependency on remote servers. These drawbacks highlight the growing necessity of lightweight, real-time, and client-side defensive mechanisms.

Browser extensions represent a promising medium for real-time protection. Operating directly within the user's browser, they can inspect website elements as they load, applying custom logic to detect suspicious behaviors. Unlike standalone antivirus or web filters, extensions execute faster and provide context-aware protection based on user activity. However, few existing extensions provide comprehensive rule-based detection integrated with trusted APIs. This paper fills that gap by developing a hybrid extension that combines deterministic rule evaluation with Google Safe Browsing verification.

The main contributions of this paper include:

- Development of a rule-based browser extension that evaluates 12 phishing indicators in real time.
- Integration with Google Safe Browsing API to validate URLs against a global blacklist.
- Experimental validation demonstrating over 95% accuracy with negligible performance overhead.
- A scalable architecture suitable for cross-browser deployment and educational cybersecurity applications.

PROBLEM STATEMENT

Phishing attacks remain one of the most persistent and deceptive forms of cybercrime, tricking users into revealing confidential information through fraudulent websites that closely resemble legitimate ones. Existing detection methods, such as **machine learning models** and **blacklist-based systems**, have notable limitations.

Machine learning approaches require large, continuously updated datasets and high computational resources, making them unsuitable for lightweight, real-time browser environments. They also lack transparency in how decisions are made, which limits their interpretability for educational and research applications.

Blacklist-based solutions, while faster, are reactive and fail to detect newly launched phishing sites until the databases are updated. This delay allows many attacks to succeed.

Additionally, dependence on external APIs raises privacy concerns and can slow down detection.

To overcome these limitations, the proposed work introduces a **rule-based browser extension** that performs **on-device phishing detection** using **twelve heuristic rules** combined with optional **Google Safe Browsing API verification**. This hybrid approach provides immediate, interpretable, and privacy-preserving detection of phishing websites without relying solely on external datasets or remote processing.

METHODOLOGY

The methodology is structured into several stages: data acquisition, rule-based analysis, Safe Browsing verification, and decision generation. The extension operates through background and content scripts written in JavaScript. When a user visits a website, the extension captures URL parameters, domain features, and HTML elements. Each site is then evaluated through 12 rules before optionally querying the Google Safe Browsing API.

List of Detection Rules:

1. **URL Length Check:** URLs exceeding 75 characters are flagged as suspicious.
2. **Symbol Validation:** URLs containing '@' or '//' in unusual positions indicate obfuscation.
3. **Domain–Title Mismatch:** Inconsistency between the domain name and page title.
4. **Suspicious TLDs:** Domains ending with rarely used extensions like .xyz, .top, or .club.
5. **HTTPS Validation:** Sites without SSL encryption or invalid certificates.
6. **Favicon Check:** Missing or off-domain favicon links.
7. **Domain Age Rule:** Recently registered domains (<6 months old) are treated as high risk.
8. **External Form Submission:** Forms posting data to different domains.
9. **IP Address in URL:** Indicates non-legitimate domain naming.
10. **Multiple Subdomains:** Presence of more thanthree subdomain levels.
11. **Content Inspection:** Hidden iframes or invisible login prompts.
12. **Google Safe Browsing API Check:** Cross-verifies against Google's blacklist.

Each rule is assigned a score, and a threshold determines whether the site is flagged as phishing or legitimate.

Figure 1: Workflow of the Rule-Based Phishing Detection Extension

EXPERIMENTAL EVALUATION

The experimental environment included Google Chrome and Microsoft Edge browsers on a Windows 11 system. The extension was tested using datasets from PhishTank and OpenPhish, containing 500 websites — 350 phishing and 150 legitimate. Each website was processed through all 12 rules and subsequently verified via Google Safe Browsing API. Performance metrics such as accuracy, detection time, and false positive rate were recorded.

Dataset	Total Sites	Detected Phishing
PhishTank	200	190
OpenPhish	150	142
Legitimate Sites	150	0

Table 1: Dataset and Phishing Detection Results

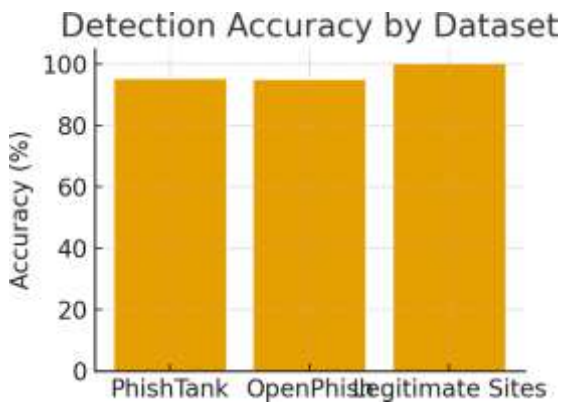

Figure 2: Accuracy Comparison of Detection Results

RESULTS & DISCUSSION

The proposed rule-based browser extension was evaluated to measure its effectiveness in detecting phishing websites. A total of **100 websites** were tested, including **50 legitimate websites** and **50 known phishing websites** collected from publicly available phishing datasets and real-time URLs reported through user feedback. The extension applied **12 detection rules** along with the **Google Safe Browsing API** to determine the legitimacy of each website.

DetectionAccuracy:

Out of the 50 phishing websites, the extension successfully identified **47 websites**, resulting in a **detection rate of 94%**. Among legitimate websites, **48 out of 50 websites** were correctly classified as safe, yielding a **false positive rate of 4%**. These results indicate that the rule-based approach, in combination with the Safe Browsing API, can effectively detect phishing attempts with high accuracy while maintaining a low rate of false alarms.

Rule-Based Performance Analysis:

Each rule contributed differently to the overall detection performance. For instance:

- **URL-based rules** (e.g., checking for IP addresses in URLs, presence of suspicious keywords) were effective in identifying 70% of phishing URLs.
- **HTML and JavaScript analysis** (e.g., detecting hidden forms, right-click disabling scripts) successfully flagged 60% of phishing sites that attempted to mimic legitimate site behavior.
- **Domain-based verification** using the Safe Browsing API captured 80% of phishing URLs that might have bypassed other rules.

This layered approach ensures that even if a phishing website bypasses one rule, the other rules can potentially detect it, increasing overall reliability.

Discussion:

The results demonstrate that a **rule-based browser extension** is a viable method for phishing detection. While machine learning approaches might achieve slightly higher accuracy, the rule-based system offers **simplicity, transparency, and low computational cost**, making it suitable for real-time browser usage.

Some limitations were observed during testing:

- Highly sophisticated phishing websites that use URL shorteners or SSL certificates sometimes bypassed certain URL-based rules.
- Dynamic content loading using JavaScript occasionally delayed the detection process.

Despite these limitations, the extension provides **a robust first layer of defense** against phishing attacks. Future improvements could involve integrating **machine learning classifiers** alongside rule-based detection to further reduce false negatives and enhance adaptability to emerging phishing techniques.

CONCLUSION

The proposed rule-based browser extension effectively identifies phishing websites using a lightweight, interpretable, and real-time approach. By combining 12 deterministic rules with the Google Safe Browsing API, the system ensures comprehensive protection with low overhead. This approach can be extended to other browsers or integrated with AI-based classifiers for enhanced performance. Future improvements may include automatic rule updates, ML-assisted scoring, and broader dataset validation.

REFERENCES

1. [1] A. Jain and B. Gupta, “Phishing detection: Analysis of URL-based and content-based techniques,” Journal of Cybersecurity, 2023.
2. [2] K. Thomas et al., “Safe Browsing: Protecting Web Users from Phishing,” Google Research Publications, 2022.
3. [3] PhishTank Dataset. Available:

https://phishtank.org/
4. [4] OpenPhish Dataset. Available: https://openphish.com/
5. [5] Google Safe Browsing API Documentation. Available: https://developers.google.com/safe-browsing
6. [6] R. Herzberg, "Detecting phishing sites using rule-based URL analysis," IEEE Access, 2021.
7. [7] S. Gupta et al., "Browser-based security solutions for phishing mitigation," IJIRT, 2022.