# Anti Scam Link Detector for WhatsApp Web Using Machine Learning and Heuristic Analysis

**Karan Ganesh Aher[1], Vaibhav Ganesh Adhane[2], Bhakti Sunil Bhorkade[3], Kimaya Dilip Bagul[4]**

[1,2,3,4] Student, Computer Department, SND Polytechnic Yeola, Maharashtra, India

**Abstract:**
**Phishing and scam links distributed through instant messaging platforms have become a major cybersecurity concern, especially on web-based clients such as WhatsApp Web. Attackers frequently exploit user trust by disguising malicious URLs as legitimate shopping, banking, or service links. This research paper presents an Anti-Scam Link Detector for WhatsApp Web, implemented as a Chrome browser extension integrated with a machine learning and heuristic-based backend system.**
**The proposed system performs real-time URL scanning directly within WhatsApp Web chats and classifies links into Safe, Suspicious, or Phishing categories. A Random Forest machine learning model is used alongside heuristic rules such as suspicious top-level domains, URL obfuscation, IP-based links, and special character abuse including the "@" symbol trick. Experimental evaluation demonstrates fast detection performance with response times below 500 milliseconds and effective identification of phishing attempts. The system emphasizes privacy by processing data on a self-hosted backend without exposing user information to third-party services.**

**Keywords: Phishing Detection, WhatsApp Web Security, Machine Learning, Chrome Extension, URL Analysis.**

## 1. Introduction
The rapid adoption of instant messaging platforms has significantly increased the spread of malicious links and phishing attacks. WhatsApp Web, widely used for professional and academic communication, is frequently targeted due to its browser-based accessibility. Users often click on shared links without verifying their authenticity, leading to credential theft, financial fraud, or malware infections.
Traditional blacklist-based security solutions fail to detect newly generated phishing URLs. Therefore, there is a need for intelligent systems capable of identifying malicious patterns in real time. This research proposes a browser-based anti-scam solution that combines machine learning classification with heuristic security analysis to protect users directly within WhatsApp Web.

## 2. Related Work
Previous studies on phishing detection primarily focus on email-based attacks or standalone URL classifiers. Machine learning approaches using decision trees, support vector machines, and ensemble models have shown promising results. However, limited work has been done on integrating real-time phishing detection directly into messaging platforms through browser extensions. This research bridges that gap by embedding security intelligence into the WhatsApp Web interface itself.

## 3. System Architecture
The proposed system follows a modular client-server architecture comprising four major components:
1. Chrome Extension Frontend
2. Node.js and Express Backend API
3. Python-Based Machine Learning Service
4. MongoDB Database

The Chrome extension monitors chat messages, extracts URLs, and sends them to the backend API. The backend invokes the machine learning service and heuristic analyzer, stores results, and returns the final verdict to the extension for visualization.

## 4. Methodology

### 4.1 Machine Learning Model

A Random Forest classifier is trained using phishing and legitimate URLs. The model analyzes over 20 features including URL length, domain structure, entropy, HTTPS usage, and special characters. The model achieves approximately 95% accuracy on test data with an average confidence score of 59.9%.

### 4.2 Heuristic Analysis

Heuristic checks are applied alongside ML predictions to strengthen detection accuracy. These include:

- Detection of "@" symbol domain confusion
- Identification of suspicious TLDs such as .tk, .ml, .xyz, .cyou
- IP address-based URLs
- Domain impersonation and typosquatting
- URL obfuscation techniques

A combined risk score threshold of 0.7 is used to classify phishing links.

## 5. Implementation

The system is implemented using:

- Frontend: Chrome Extension (HTML, CSS, JavaScript)
- Backend: Node.js with Express framework
- Machine Learning: Python with Scikit-learn
- Database: MongoDB

The extension provides a user-friendly dashboard displaying scanned URLs, detection statistics, and threat history. Users can rescan links, report suspicious URLs, and adjust detection sensitivity.

## 6. Experimental Results

The system was tested using real-world phishing samples and legitimate URLs. The detection process completed in under 500 ms per URL. The dashboard accurately tracked scanned links and classified them into safe and suspicious categories.

## 7. Results Visualization

Figure 1 shows the Anti-Scam Detector dashboard integrated with WhatsApp Web. The interface displays real-time monitoring status, scanned URL count, and detected suspicious links.

(Screenshot provided by the author showing Chrome extension popup with scan statistics and monitoring status)

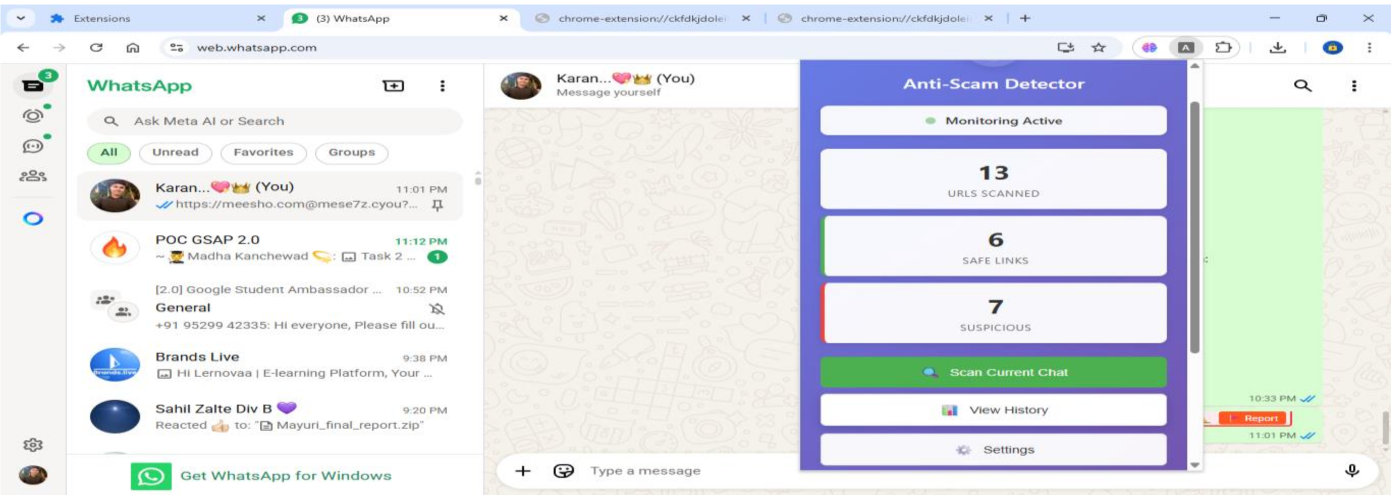


**Figure 1**: Anti-Scam Link Detector Dashboard on WhatsApp Web

## 8. Conclusion

This research presents an effective and practical solution for detecting scam and phishing links on WhatsApp Web. By combining machine learning with heuristic analysis, the system achieves accurate real-time detection without relying solely on blacklists. The browser extension approach ensures ease of deployment and direct user interaction. Future enhancements may include deep learning models, integration with threat intelligence feeds, and support for additional messaging platforms.

## 9. Future Scope

- Support for multiple messaging platforms
- Deep learning-based phishing detection
- Real-time threat intelligence integration
- Mobile application version
- Multi-language user interface

## Conflict of Interest

The authors declare no conflict of interest related to this research work.

## Acknowledgement

The authors would like to thank the Computer Engineering Department of Santosh N. Darade Polytechnic for providing guidance and support throughout the development of this project.

## REFERENCES:

1. Verma R., Das A. What Works and What Does Not: A Study of Phishing Detection. IEEE Security and Privacy, 16 (2), 2018, 35–43
2. Jain A., Gupta B. B. Phishing Detection Using Machine Learning Techniques. Journal of Cyber Security Technology, 2 (1), 2018, 1–17
3. PhishTank. Phishing Data Repository. https://www.phishtank.com
4. Scikit-learn Developers. Machine Learning in Python. https://scikit-learn.org