International Journal of
Technology and Emerging Research
ISSN: 3068-1995 (Print) | 3068-109X (Online) | https://www.ioro.org/ijter

Identification Of URL-Based Attacks From IP Data

Cyril K S[1],Basil Varghese[2], Jaijith A[3], Yadhu Krishna[4], Dr Krishna Kumar P R[5]

[1,2,3,4]*Students , Dept of CSE , S.E.A College of engineering and Technology , India*

[5] *Faculty, Dept of CSE , S.E.A College of engineering and Technology, India*

Abstract:

With the rapid expansion of internet usage and web-based services, cyber threats such as phishing, malware distribution, and malicious URL attacks have significantly increased. Attackers often exploit IP-based patterns and URL structures to bypass traditional security mechanisms. This project focuses on identifying URL-based attacks using IP data analysis combined with machine learning techniques to improve detection accuracy and cybersecurity resilience.

The system analyzes URLs by extracting features such as IP address patterns, domain behavior, request frequency, and URL structure. By leveraging IP intelligence and classification models, the system can distinguish between legitimate and malicious URLs in real time. This approach enhances traditional URL filtering mechanisms by incorporating behavioral and network-level insights.

The proposed solution aims to provide a scalable, efficient, and automated detection system capable of preventing cyber threats before they reach end users. The system can be deployed as a web-based application or integrated into network security tools, contributing to safer browsing environments and improved threat intelligence systems.

Keywords: **URL Detection, Cybersecurity, IP Analysis, Phishing Detection, Machine Learning, Network Security**

1. INTRODUCTION

The increasing reliance on the internet for communication, transactions, and information sharing has led to a rise in cyber threats, particularly URL-based attacks. These attacks include phishing websites, malicious redirects, and malware distribution links that deceive users into revealing sensitive information or downloading harmful content.

Traditional security systems primarily rely on blacklists and signature-based detection methods, which are often ineffective against newly generated or obfuscated malicious URLs. Attackers continuously evolve their techniques, making it necessary to develop more advanced detection mechanisms.

IP data plays a crucial role in identifying malicious activities, as attackers often use suspicious or frequently changing IP addresses to host harmful content. By analyzing IP patterns, domain associations, and network behavior, it becomes possible to detect anomalies that indicate potential threats.

This project proposes a system that combines URL feature extraction with IP-based analysis and machine learning algorithms to accurately classify URLs as safe or malicious. The system is designed to be scalable, efficient, and adaptable to evolving cyber threats.

2. OBJECTIVES

The main objectives of this project are:

1. To develop a system for detecting malicious URLs using IP data analysis
2. To extract and analyze URL-based features for classification
3. To identify suspicious IP behavior associated with cyber attacks
4. To implement machine learning models for accurate detection
5. To improve detection of phishing and malware URLs
6. To create a real-time URL analysis system
7. To design a scalable and efficient architecture
8. To enhance cybersecurity awareness and protection

3. PROBLEM STATEMENT

URL-based attacks are one of the most common forms of cyber threats, targeting users through phishing links, fake websites, and malicious downloads. Existing systems often fail to detect newly generated or obfuscated URLs due to reliance on static blacklists.

Additionally, many systems do not fully utilize IP-level data, which can provide valuable insights into suspicious behavior. This creates a gap in detection accuracy and leaves users vulnerable to evolving cyber attacks.

4. EXISTING SYSTEM

Current URL detection systems include:

- Blacklist-based filtering
- Rule-based detection systems
- Browser security warnings

Limitations:

- Ineffective against new (zero-day) attacks
- Limited use of IP intelligence
- High false positives or false negatives
- Lack of real-time adaptability

5. PROPOSED SYSTEM

The proposed system introduces an intelligent URL attack detection model using IP data and machine learning.

Key Features:

- URL feature extraction (length, symbols, domain info)

- IP address analysis (geolocation, reputation, frequency)
- Machine learning classification (e.g., Random Forest, Logistic Regression)
- Real-time detection and response
- Scalable architecture

This system improves detection accuracy by combining URL structure analysis with network-level IP insights.

6. SYSTEM ARCHITECTURE

Input Module

Accepts URL from user or system

Feature Extraction Module

Extracts URL features (length, protocol, domain)

IP Analysis Module

Identifies IP address and analyzes behavior

Machine Learning Module

Classifies URL as safe or malicious

Database Module

Stores URL data and results

Output Module

Displays classification result

7. TECHNOLOGIES USED

Frontend

- HTML5, CSS3, JavaScript

Backend

- Python (Flask / Django)

Machine Learning

- Scikit-learn
- Pandas, NumPy

Database

- MySQL / MongoDB

Tools

- VS Code
- Jupyter Notebook
- Git & GitHub

8. METHODOLOGY

8.1. Requirement Analysis

- Identify system requirements such as URL input, feature extraction, and classification

8.2. Data Collection

- Gather dataset of malicious and benign URLs

8.3. Feature Extraction

- Extract features like:
  - URL length
  - Presence of IP in URL
  - Special characters
  - Domain age

8.4. Model Training

- Train ML models using labeled data

8.5. Testing

- Evaluate accuracy, precision, and recall

8.6. Deployment

- Deploy system as a web application

9. IMPLEMENTATION

URL Input Module

User enters URL

Feature Extraction

System processes and extracts relevant features

IP Analysis

Resolves domain to IP and analyzes patterns

Model Prediction

ML model predicts malicious or safe

Result Display

Output shown to user

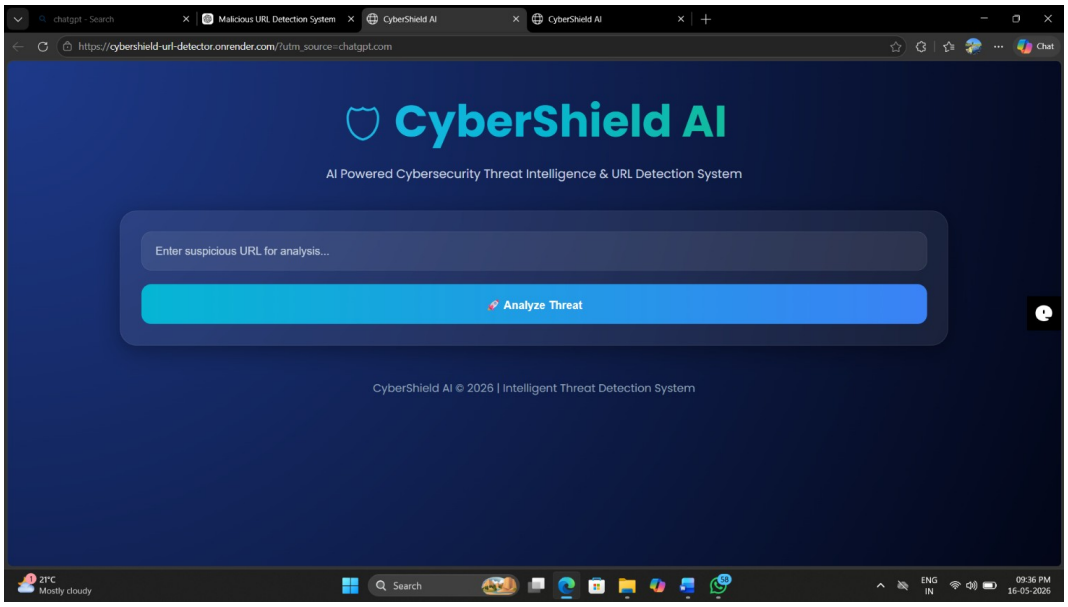
10. RESULTS

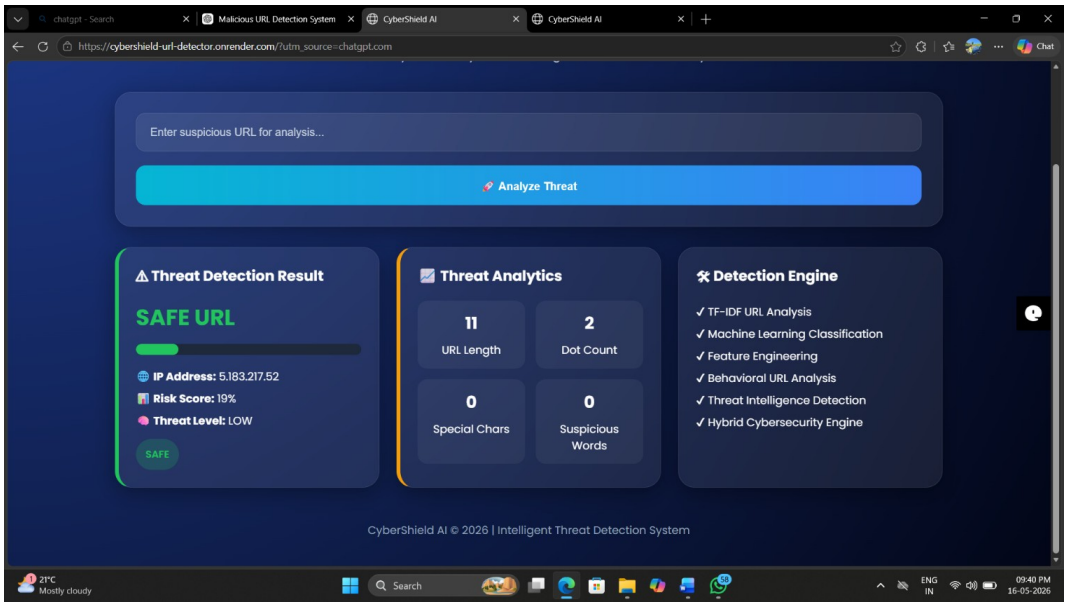
Functional Results

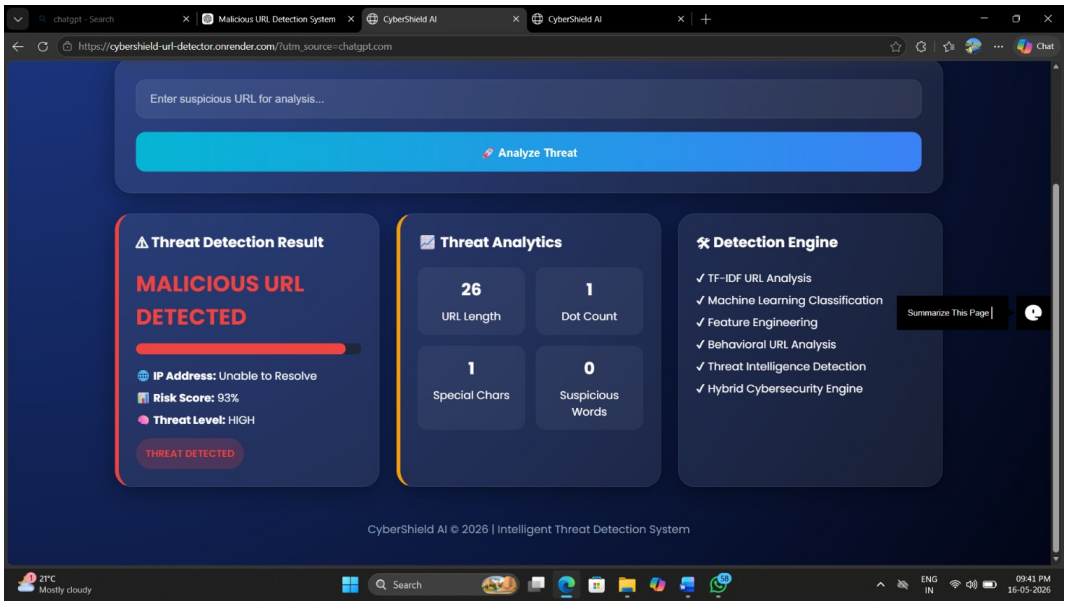
- Successful detection of malicious URLs
- Real-time analysis capability

Technical Results

- Improved detection accuracy
- Reduced false positives







11. FUTURE ENHANCEMENTS

- Integration with browser extensions
- Deep learning-based detection
- Real-time threat intelligence APIs
- Cloud-based deployment
- Automated blacklist updates

12. CONCLUSION

The project successfully demonstrates the identification of URL-based attacks using IP data and machine learning techniques. By combining URL feature extraction with IP analysis, the system provides a more accurate and efficient detection mechanism compared to traditional methods.

This approach enhances cybersecurity by enabling early detection of malicious links and protecting users from potential threats. The system can be further improved with advanced AI models and real-time data integration, making it a strong foundation for future cybersecurity applications.

13. REFERENCE

[1] Blum et al., “Phishing URL Detection Using Machine Learning,” ACM, 2010.
[2] Link: https://dl.acm.org
[3] J. Ma et al., “Beyond Blacklists: Detecting Malicious Websites,” ACM SIGKDD, 2009.
[4] Link: https://dl.acm.org
[5] M. Khonji et al., “Phishing Detection: A Literature Survey,” IEEE, 2013.
[6] Link: https://ieeexplore.ieee.org
[7] K. Thomas et al., “Real-Time URL Spam Detection,” IEEE Security & Privacy, 2011.
[8] Link: https://ieeexplore.ieee.org