INSTAGRAM MEDIA BASED MALICIOUS URL DETECTION

M.Sarukkesh
B.Sc Digital & Cyber Forensics Science
Rathinam College of Arts and Science, Coimbatore, India

Abstract

The rapid growth of social media platforms has increased the sharing of links through posts, comments, stories, direct messages, and bio sections. Among these platforms, Instagram has become a major target for cyber attackers who distribute malicious URLs to perform phishing, malware delivery, credential theft, and unauthorized redirection. Traditional URL detection methods often rely on static blacklist databases, which are not effective against newly generated or obfuscated malicious links.

This paper presents an **Instagram Media-Based Malicious URL Detection System**, designed to identify harmful URLs shared through Instagram-related media content using machine learning techniques. The system analyzes URL-based features such as length, domain structure, suspicious keywords, special characters, and redirection behavior to classify links as safe or malicious. It integrates data preprocessing, feature extraction, model training, and prediction modules to improve classification accuracy and reliability.

The proposed system supports both manual URL input and automated analysis workflows. Additionally, it incorporates visualization and evaluation mechanisms to measure model performance and improve detection quality. Experimental results demonstrate that the system can efficiently identify malicious URLs with good accuracy while maintaining low processing overhead.

Key Words: Malicious URL Detection, Instagram Security, Machine Learning, URL Analysis, Cyber Security, Phishing Detection, Random Forest, Support Vector Machine.

1. INTRODUCTION

With the increasing popularity of social media, users now frequently interact with links shared through captions, comments, direct messages, stories, and profile bios. Instagram, being one of the most widely used platforms, has become a common medium for spreading malicious URLs. These URLs may redirect users to phishing websites, trigger malware downloads, or steal sensitive user information such as usernames, passwords, and payment details.

Traditional security approaches often focus on browser warnings or blacklist-based filtering. However, these methods are not sufficient for handling rapidly changing malicious links. Attackers continuously generate new URLs, shorten links, or disguise domains to avoid detection. As a result, many harmful URLs remain undetected until users interact with them.

In real-world scenarios, malicious URL attacks commonly occur due to:

• Phishing links shared through messages

• Shortened or hidden malicious URLs

• Redirect links leading to unsafe websites

To address these challenges, the **Instagram Media-Based Malicious URL Detection System** is designed to provide an intelligent and automated solution. The system analyzes URL characteristics using machine learning models and predicts whether a link is safe or malicious. By doing so, it helps reduce cyber risks and improves the safety of social media users.

2. LIMITATIONS OF EXISTING SYSTEM

Existing malicious URL detection systems face several challenges that reduce their effectiveness in protecting users from cyber threats.

One major limitation is the heavy dependence on blacklist databases. These databases can only detect URLs that have already been identified as malicious. Newly created malicious URLs may bypass such

systems because they are not yet recorded.

Another limitation is the lack of real-time intelligent analysis. Many systems check only basic URL reputation and fail to analyze hidden patterns such as suspicious keywords, special characters, or abnormal domain structures.

Many existing tools are also not specifically designed for social media environments like Instagram. Since malicious links can be shared in captions, comments, and direct messages, traditional systems may miss contextual threats related to media-based distribution.

Another important issue is the high rate of false positives and false negatives. Genuine links may sometimes be flagged as malicious, while harmful links may remain undetected. This affects system reliability and user trust

.

3.PROPOSED SYSTEM

The Secure Backup and Recovery System is designed to overcome the limitations of existing solutions by integrating security, automation, and efficiency into a single framework.

The system ensures that data is securely backed up using encryption techniques and can be recovered quickly when required. It provides a structured approach to data protection by combining backup automation, secure storage, and controlled recovery.

The **Instagram Media-Based Malicious URL Detection System** is designed to overcome the limitations of existing approaches by integrating machine learning, automated feature extraction, and prediction modules into a single framework.

The system examines URLs shared through Instagram-related environments and classifies them as safe or malicious based on their structural and behavioral characteristics. It provides a systematic method for URL analysis by combining data collection, preprocessing, model training, and result generation.

3.1 Data Collection Module

This module is responsible for collecting safe and malicious URL datasets from trusted sources. The collected URLs are used for training and testing machine learning models.

3.2 Feature Extraction Module

This module extracts important URL features such as URL length, domain name, number of special characters, suspicious keywords, and protocol type. These features are used as input for classification.

3.3 Data Preprocessing Module

This component cleans the dataset by removing duplicates, handling null values, and converting categorical data into suitable numerical format. Proper preprocessing improves model performance.

3.4 Model Training Module

This module trains machine learning algorithms such as Linear Regression, Random Forest, and Support Vector Machine (SVM) using the prepared dataset.

3.5 Prediction Module

This module accepts a new URL and predicts whether it is safe or malicious based on the trained model.

3.6 Evaluation Module

This component measures system performance using metrics such as accuracy, precision, recall, F1-score, and confusion matrix.

3.7 Visualization and Reporting Module

All analysis results are shown through graphs, charts, and classification reports for easy interpretation and performance monitoring.

4. METHODOLOGY

The methodology of the Secure Backup and Recovery System ensures systematic data protection and recovery.

Workflow

User Input → Data Selection → Encryption → Backup Storage → Monitoring → Recovery → Verification

- Encryption Phase:

  Data is encrypted before being stored to ensure security.

- Backup Phase:

  Encrypted data is stored in selected storage locations.

- Monitoring Phase:

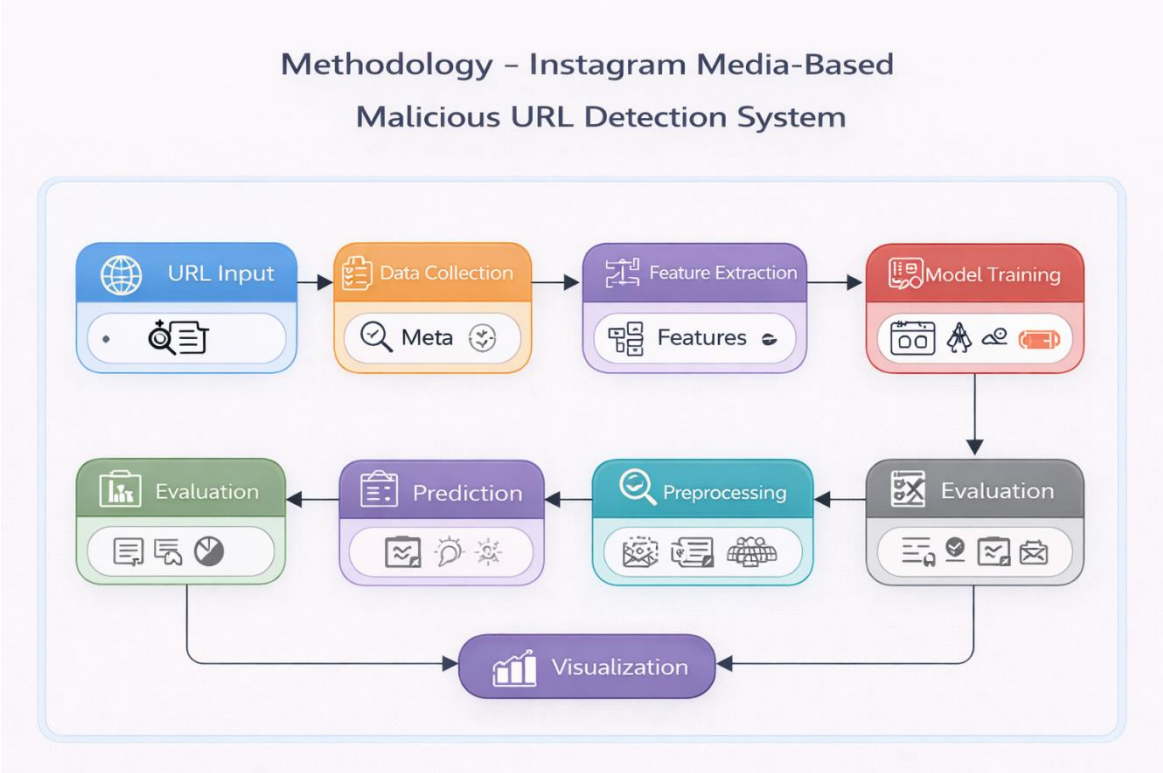
  The system continuously monitors backup status and logs activities.

- Recovery Phase:

  Data is restored when required by the user.

- Verification Phase:

  Ensures that recovered data is accurate and unaltered.



5. RESULTS AND ANALYSIS

The system was evaluated using different operational modes to simulate real-world backup and recovery scenarios.

Mode	Description	Observation
Linear Regression	Basic prediction model for URL-related scoring	Moderate accuracy
Random Forest	Ensemble learning model for URL classification	High accuracy and stability
SVM	Classification model for separating safe and malicious URLs process	Good accuracy and reliable detection

The system performance remained stable during testing and showed effective classification of malicious URLs with minimal processing delay.

Analysis

The system demonstrated the ability to identify harmful URLs based on extracted URL features. The preprocessing stage improved dataset quality by removing null values and duplicate entries. The feature extraction process helped the machine learning models recognize suspicious patterns more effectively.

Among the tested models, Random Forest and Support Vector Machine provided better performance than simpler approaches because they handled nonlinear URL patterns more efficiently. The confusion matrix and evaluation metrics confirmed that the proposed system can detect malicious links with good reliability.

The visualization module also helped interpret the results clearly by showing classification outputs, performance graphs, and analysis charts. Overall, the system proved to be suitable for enhancing cyber safety in Instagram-related environments.

6. CONCLUSION

The Instagram Media-Based Malicious URL Detection System provides an effective solution for protecting users against harmful links shared through social media. By integrating machine learning techniques, feature extraction, and evaluation mechanisms, the system improves the identification of malicious URLs beyond traditional blacklist-based approaches. The results indicate that using intelligent classification models significantly improves the detection of suspicious links and reduces the risk of phishing, malware infections, and unauthorized access. The system is lightweight, scalable, and adaptable to various digital security use cases.This work highlights the importance of combining automation, machine learning, and URL analysis to build a practical cybersecurity solution for modern social media environments..

7. FUTURE WORK

• Future improvements may include:

• Integration with live Instagram monitoring tools Advanced anomaly detection

• Browser extension for real-time URL scanning

• User-friendly dashboard for alerts and reports

8.REFERENCES

[1] NIST Cybersecurity Framework

[2] CIS Critical Security Controls

[3] OWASP Top 10

[4] Stallings, W., Network Security Essentials