International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

AI-Based Detection and Mitigation of Malware and Phishing Attacks in Malayalam Language Cyber Threats

Neetha B.S

Assistant Professor, Department of Computer Science, Fatima Mata National College, Kollam

Abstract: ***The increasing adoption of regional languages in digital communication has introduced new cybersecurity challenges. In India, cyber attackers are increasingly leveraging Malayalam-language content to conduct phishing and malware-based attacks. Existing detection systems are predominantly designed for English and fail to effectively analyze low-resource languages. This paper proposes an Artificial Intelligence (AI)-based framework for detecting and mitigating Malayalam-language cyber threats. The framework integrates Natural Language Processing (NLP), deep learning models, and transformer-based architectures to analyze linguistic patterns, code-mixed text, and embedded malicious URLs. A curated dataset comprising real and synthetically generated Malayalam messages is used for evaluation. Experimental results indicate that transformer-based models significantly outperform traditional machine learning approaches, achieving high accuracy and reduced false positive rates. The study highlights the necessity of language-aware cybersecurity mechanisms for regional digital ecosystems.***

Keywords: Cybersecurity, Malayalam Language, Phishing Detection, Malware Detection, NLP, Transformer Models

1. Introduction

Rapid digital transformation has significantly increased cyber threat exposure worldwide. In India, the expansion of internet access has led to the widespread use of regional languages such as Malayalam in online communication. Attackers exploit linguistic familiarity to increase the effectiveness of phishing and malware campaigns.

Traditional cybersecurity solutions are largely English-centric and fail to address challenges posed by low-resource languages. Malayalam, characterized by complex morphology and frequent code-mixing with English, presents additional challenges for automated detection systems.

Recent advancements in AI and NLP, particularly transformer-based architectures, have demonstrated strong capabilities in text classification tasks [1]. However, their application in regional-language cybersecurity remains limited. This paper proposes a language-aware AI framework specifically designed for detecting Malayalam-language cyber threats.

2. Related Work

2.1 Phishing Detection

Machine learning techniques such as Support Vector Machines and Random Forests have been widely used for phishing detection [2]. These approaches rely on handcrafted features like URL structure and lexical patterns but often lack adaptability to multilingual environments.

2.2 Malware Detection

Traditional malware detection techniques focus on signature-based and behavioral analysis. Recent approaches incorporate deep learning models such as LSTM networks to identify malicious patterns in textual and network data [3].

2.3 NLP in Cybersecurity

Natural Language Processing has been increasingly used to analyze phishing emails and malicious messages. Transformer-based models like BERT have significantly improved text classification performance due to contextual understanding [1].

2.4 Multilingual and Low-Resource Languages

Research on multilingual NLP has expanded, but low-resource languages such as Malayalam remain underexplored. IndicBERT and similar models have shown promise in handling Indian languages [4], yet their application in cybersecurity is still limited.

Research Gap: Existing studies do not provide a unified framework combining NLP, malware detection, and phishing analysis specifically for Malayalam-language threats.

3. Threat Model and Problem Definition

The threat model considers adversaries distributing phishing messages and malware links through SMS, email, and social media platforms. These attacks use:

- Social engineering (urgency, fear, rewards)
- Language familiarity (Malayalam)
- Code-mixed content

Objective:
To accurately classify Malayalam-language messages as **legitimate or malicious** before user interaction.

4. Proposed AI-Based Detection Framework

4.1 Framework Overview

The proposed system consists of the following stages:

1) Data Collection
2) Text Preprocessing
3) Feature Extraction
4) AI-Based Classification
5) URL & Malware Analysis
6) Final Threat Detection

4.2 Architecture Diagram

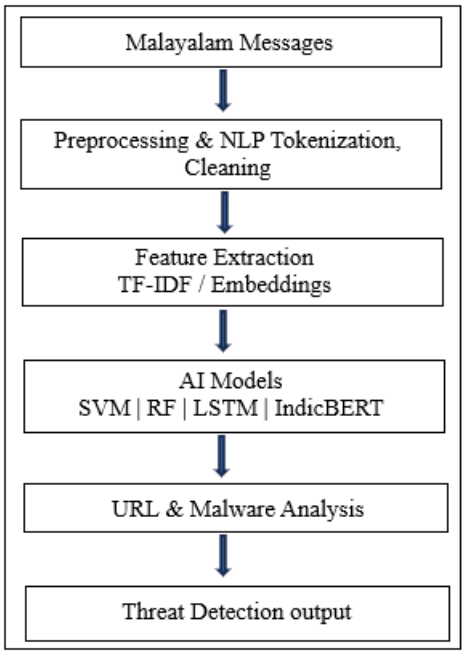

Figure 1: Proposed framework for Malayalam cyber threat detection

4.3 Methodology Details

a) Tokenization and normalization of Malayalam text
b) Handling code-mixed data
c) Feature extraction using TF-IDF and embeddings
d) Model training using:
- SVM
- Random Forest
- LSTM / Bi-LSTM
- IndicBERT

5. Dataset Description

The dataset consists of **3,000 Malayalam-language messages**, including:

Category	Count
Legitimate	1200
Phishing	950
Malware	850

Data Sources
- Cybersecurity advisories (CERT-In)
- Reported phishing samples
- Synthetic data generation using pattern-based augmentation

Improvement Added:
Synthetic data was validated using:
- Manual expert review
- Pattern consistency checks

6. Experimental Setup

a) Train-test split: 80:20
b) Evaluation Metrics:
- Accuracy
- Precision
- Recall
- F1-score
- ROC-AUC

c) Tools:
- Python
- TensorFlow / PyTorch
- HuggingFace Transformers

7. Results and Performance Evaluation

Model	Accuracy	Precision	Recall	F1	ROC-AUC
SVM	90.1	0.89	0.87	0.88	0.91
RF	92.4	0.91	0.9	0.9	0.93
LSTM	94.9	0.94	0.93	0.93	0.96
Bi-LSTM	95.7	0.95	0.94	0.94	0.97
IndicBERT	**97.8**	0.97	0.97	0.97	0.99

8. Discussion

The results confirm that transformer-based models significantly outperform traditional methods. Contextual embeddings improve classification accuracy, particularly in code-mixed Malayalam text. The framework effectively reduces false positives, making it suitable for real-world deployment.

9. Conclusion and Future Work

This study presents an AI-based framework for detecting Malayalam-language cyber threats. By addressing challenges in low-resource languages, the proposed approach contributes to more inclusive cybersecurity solutions.

Future work includes:
- Real-time deployment
- Voice-based phishing detection
- Expansion to other Indian languages

References

[1] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers," NAACL, 2019.
[2] A. Author et al., "Phishing Detection Using Machine Learning," IEEE Access, 2023.
[3] S. Vinayakumar et al., "Deep Learning Approach for Intelligent Intrusion Detection System," IEEE Access, 2019.
[4] K. Kakwani et al., "IndicNLPSuite: Monolingual Corpora and Evaluation Benchmarks," 2020.
[5] CERT-In, "Cyber Security Advisories," Government of India.