Project Report: LinkStop

Name: Samrudh M
Roll Number: 221CS148
Year: 3rd Year, B.Tech Computer Science and Engineering
Institute: National Institute of Technology Karnataka (NITK)
Under the guidance of: Dr. Shridhar Sanshi

Introduction

This project, titled *LinkStop Beta*, focuses on creating a secure mechanism to identify and handle malicious URLs. The primary goal is to address cybersecurity issues such as phishing and malicious links that pose significant threats, especially to less tech-savvy users. By integrating machine learning and network analysis techniques, the project aims to classify URLs, cross-reference them with known databases, and provide detailed insights about suspicious links.

Phase-wise Breakdown

Phase 1: Machine Learning Classifier Model

1. Download a dataset of URLs.
2. Preprocess the dataset and extract relevant features.
3. Implement various encoding methods.
4. Identify the best-fit neural network model.
5. Train, validate, and test the model.

Phase 2: Database Mapping

1. Check URLs against existing databases like AbuseIPDB and VirusTotal.
2. Flag/report malicious URLs or redirect safe ones.
3. Pass unknown URLs to the next phase.

Phase 3: Network Analysis

1. Perform WHOIS lookups, reverse IP searches, and geolocation analysis.
2. Identify the root owner of malicious URLs.
3. Generate reports, alert CERT-IN, and notify ISPs.
4. Display warnings in the user interface.

Phase 4: Integration

1. Develop an application for seamless URL verification.
2. Enable privileged access, where clicking a URL invokes the app for analysis(optional).

References:

[1] S. Marchal, J. Francois, R. State, T. Engel, "Detecting malicious web links and identifying their attack types", ResearchGate, 2014.