**Project Report**

**Phishing Link Detector**

Group Members :

Imtiaz Hussain (CY-24F-111)

M.Irtaza (CY-24F-149)

Sanaullah Lakho (CY-24F-131)

Taha Akbar (CY-24F-121)

1. **Introduction**

**Agenda**

Cybersecurity threats have been increasing, and phishing attacks are among the most common cybercrimes. Attackers use deceptive URLs to trick users into revealing sensitive information such as login credentials and financial details. This project aims to develop a Phishing Link Detector, a C++ program that evaluates URLs and determines whether they are potentially harmful.

**Background**

Phishing is a form of cyberattack where attackers create fake websites resembling legitimate ones to steal user credentials. Many phishing attacks exploit URL structures, domain names, and misleading words to deceive users. Detecting such fraudulent links manually is challenging, making automated detection essential.

**Objective**

The primary objectives of this project are:

* To develop a C++ program capable of detecting phishing links based on various indicators.
* To implement multiple checks such as domain analysis, special character detection, and suspicious words identification.
* To assign a Phishing Risk Score to URLs, categorizing them as safe, suspicious, or highly dangerous.

1. **Project Overview**

**Methodology**

This project follows a rule-based detection approach by analyzing different phishing indicators in URLs. The program evaluates URLs based on predefined conditions such as:

* Suspicious Domain Endings (e.g., .xyz, .tk)
* Misspelled Brand Names (e.g., g00gle, faceb00k)
* Excessive Use of Dashes or Special Characters
* Presence of ‘@’ Symbol (often used in phishing attacks)
* Use of IP Addresses Instead of Domains

A risk scoring system assigns points based on the severity of each factor, ultimately categorizing the URL as safe, suspicious, or highly dangerous.

**Flowchart**

Start

↓

Input URL

↓

Check for phishing domain endings

↓

Check for suspicious words

↓

Check for excessive dashes or special characters

↓

Check for ‘@’ symbol in URL

↓

Check if URL contains an IP address

↓

Calculate phishing risk score

↓

Display risk level

(Safe, Suspicious, or Dangerous)

↓

End

**Tools and Techniques**

* Programming Language: C++
* Development Environment: Visual Studio
* Libraries Used: <iostream>, <string>, <vector>, <regex>
* Detection Techniques: String Matching, Regular Expressions, and Pattern Recognition

1. **Implementation and Execution**

**Development Process**

The project was implemented as a command-line C++ program with the following key features:

* Data Collection: Lists of common phishing domains and misspelled brand names were compiled.
* URL Analysis Functions: Functions were implemented to detect phishing indicators, including domain endings, special characters, and suspicious words.
* Risk Scoring Mechanism: A risk score (0-10) was assigned based on multiple checks.
* User Interaction: The program accepts a URL from the user, analyzes it, and displays the risk level.

**Challenges Faced**

* Handling Different URL Formats: URLs can be written in various formats, making it necessary to account for different structures.
* Regex Pattern Matching: Implementing regex-based IP detection required proper testing to avoid false positives.
* Efficient String Matching: Optimizing string comparisons for better performance.

**Testing**

* Tested with real-world phishing URLs to check accuracy.
* Edge cases handled, such as incomplete URLs or special characters in unexpected places.

1. **Results and Achievements**

**Final Output**

The program successfully identifies phishing URLs based on predefined conditions and assigns a Phishing Risk Score (0-10).

**Key Achievements**

✔️ Successfully implemented multiple phishing detection techniques.

✔️ Developed a user-friendly interface for risk assessment.

✔️ Optimized performance with efficient string and regex matching.

1. **Conclusion**

**Summary**

This project demonstrates an effective phishing URL detection system using C++. The program evaluates URLs based on predefined conditions and assigns a risk score to determine whether a URL is safe or dangerous. By implementing string pattern matching and regular expressions, the program provides a simple but effective phishing detection mechanism.

**Lessons Learned**

* Regex is a powerful tool for pattern recognition in cybersecurity applications.
* Phishing detection requires multiple indicators to improve accuracy.
* String comparison optimization is crucial for better performance in security applications.

**Future Scope**

* Enhancing detection accuracy by incorporating machine learning models.
* Expanding the phishing database with real-world data sources.
* Integrating with a browser extension to provide real-time URL checks.

Final Thoughts

The Phishing Link Detector is a simple yet powerful cybersecurity tool that can help users identify potential phishing threats. While not a replacement for advanced security solutions, it serves as an educational and practical approach to understanding phishing detection techniques.