PHISHING DETECTION APPLICATION

SYNOPSIS ON MAJOR PROJECT

BACHELOR OF TECHNOLOGY

< Computer Science >

Submitted

To

Dr. A.P.J. ABDUL KALAM TECHNICAL UNIVERSITY LUCKNOW

Submitted By-

Uday Raj Singh	<2100100100182>
Zoya Ashiyam	<2100100100198>
Md Shabi Ahmed	<2100100100090>
Swati Kesarwani	<2200100109023>

UNDER THE SUPERVISION OF

Mr. Amit Roy

Assistant Professor

UNITED COLLEGE OF ENGINEERING & RESEARCH PRAYAGRAJ

2023-24

INTRODUCTION

The Phishing Detection System is designed to identify and mitigate phishing attempts by analyzing email content, URLs, or other suspicious communication channels. It uses machine learning techniques and rule-based systems to detect phishing attacks in real time, protecting users from fraudulent activities like identity theft, financial fraud, and sensitive information breaches.

Objective

The objective of this project is to develop an intelligent system that can analyze incoming emails or web traffic, detect potential phishing attempts using algorithms or machine learning techniques, and provide timely alerts or preventive actions to users or administrators.

Key Features

1. Data Collection: Implements modules to gather data from email servers, URLs, or web traffic logs for analysis.

2. Phishing Detection: Uses machine learning models and rule-based systems to detect phishing attempts based on URL patterns, email content analysis, sender reputation, or known phishing signatures.

3. User Interface: Develops a user-friendly interface where users or administrators can view phishing alerts, manage the system, and review suspicious communication.

4. Dashboard and Reporting: Includes features to visualize phishing trends, generate reports, and review historical phishing incidents.

5. Security and Compliance: Incorporates security measures such as encryption, secure email handling, and audit logs to ensure confidentiality and compliance with cybersecurity regulations.

Technologies Involved

- **Programming Languages:** Python.

- **Database Management**: MySQL, MongoDB, for storing and managing data.

- **Web Development:** HTML, CSS, JavaScript, and frameworks like React for the user interface**.**

- **Data Analysis:** Machine learning models for phishing detection.

- **Security:** Implementing encryption, secure communication, and compliance with cybersecurity protocols.

Potential Challenges

- **False Positives/Negatives**: Ensuring high accuracy in detecting phishing while minimizing false alerts.
- **Algorithm Efficiency**: Choosing and training appropriate models for real-time phishing detection.
- **Data Integration**: Efficiently integrating data from different email or traffic sources.
- **Real-time Processing**: Handling phishing detection in real time, especially in high-volume traffic.

Benefits

- **Protection Against Phishing**: Detects and mitigates phishing attacks before they compromise sensitive information.
- **Improves Security**: Strengthens the overall security posture by reducing exposure to phishing.
- **Real-Time Detection**: Provides instant alerts and blocks access to phishing threats in real time.
- **Automated Phishing Detection**: Reduces manual efforts by automating phishing detection and response.

FEASIBILITY STUDY

1. Technical Feasibility

-Technology Stack: Evaluates the availability of technologies and tools necessary for phishing detection, such as machine learning models, email processing tools, URL filtering mechanisms, and user interface development. Ensures the chosen tech stack aligns with the project requirements and team's expertise.

-Integration: Assesses the feasibility of integrating with email servers, web traffic logs, notification services, and security mechanisms. Considers compatibility issues, API availability for external phishing databases, and scalability requirements to handle large volumes of data.

-Performance: Determines if the system can process high volumes of emails and web traffic in real time without performance bottlenecks. Conducts performance testing to ensure scalability and system reliability under high loads.

2. Economic Feasibility

-Cost Analysis: Estimates the costs associated with developing the phishing detection system, including software development, hardware infrastructure, email/API services, personnel, and ongoing maintenance. Compare these costs to the expected benefits, such as reducing phishing threats and enhancing security.

-ROI Calculation: Evaluates ROI by considering factors such as cost savings from phishing prevention, reduced security breaches, operational efficiency, and potential commercial applications of the phishing detection system.

3. Operational Feasibility

-User Needs Assessment: Identifies key stakeholders, such as system administrators and end users, who will interact with the phishing detection system. Conducts surveys or interviews to understand their needs, expectations, and challenges regarding phishing threats.

-Workflow Analysis: Analyzes the workflow for detecting phishing attempts, alert generation, blocking access to phishing sites, and notifying users. Ensures the system integrates smoothly into existing workflows and enhances security measures.

-Change Management: Evaluates the organization's readiness to adopt the phishing detection system. Identifies potential resistance to change, and outlines training programs and user adoption strategies to ensure smooth integration.

4. Scheduling Feasibility

-Project Timeline: Creates a detailed project plan outlining milestones, development phases, testing, and deployment schedules. Estimates time requirements for each phase, including system updates and ongoing maintenance.

-Risk Assessment: Identifies project risks such as false positives, technical challenges, resource constraints, or changes in phishing tactics. Develops risk mitigation plans to handle potential disruptions, including regular system updates to combat evolving phishing techniques.

METHODOLOGY

1. Requirements Gathering

-Stakeholder Interviews: Conduct interviews with key stakeholders, including potential users, security experts, and IT managers, to gather requirements, understand their needs, and define the scope of the Phishing Detection System.

-Use Case Analysis: Identify and document use cases describing interactions between users and the system, including email analysis, URL inspection, phishing detection, alert generation, and user response actions.

2. System Design

-Architecture Design: Defines the system architecture, including modules for email/URL analysis, phishing detection algorithms, alert generation, notification services, user interfaces, and data storage.

-Database Design: Designs a database schema to store phishing logs, user profiles, system configurations, historical alerts, and system audit trails. Focuses on scalability, indexing, and performance optimization.

3. Development Phase

-Backend Development: Implements backend modules for email parsing, URL analysis, phishing detection algorithms, and alert triggers. Uses Python and relevant libraries suited for data processing and phishing detection.

-Frontend Development: Develops a user-friendly interface using HTML, CSS, JavaScript, and React, allowing users to configure phishing detection settings, view alerts, manage email/URL analysis preferences, and generate report.

4. Integration and Testing

-Integration Testing: Integrates all system components, including external APIs (e.g., for threat intelligence), email servers, and notification services. Ensures smooth data flow between components through integration testing.

-Functional Testing: Perform functional testing to validate phishing detection, alert generation, notification delivery, user interface interactions, and system performance under various phishing scenarios.

-**Security Testing**: Conduct security testing, such as penetration testing and phishing simulation, to identify and fix vulnerabilities. Ensure data encryption, secure email handling, and compliance with security standards.

5. Monitoring and Maintenance

-**Monitoring Tools:** Implements tools to monitor system performance, email/URL traffic, phishing alert volumes, user activity, and potential issues. Sets up alerts for critical system thresholds or anomalies.

-**Maintenance Plan:** Develops a maintenance plan for system updates, phishing signature updates, security patches, bug fixes, and scalability improvements. Establishes procedures for user feedback and system enhancements.

6. Deployment and User Training

-**Deployment Plan**: Prepares a deployment plan outlining steps for deploying the Phishing Detection System in a live environment. Includes version control, configuration management, and rollback procedures.

-**User Training**: Provide training sessions and documentation for users to understand how to configure phishing detection settings, interpret phishing alerts, take appropriate actions, and utilize reporting features.

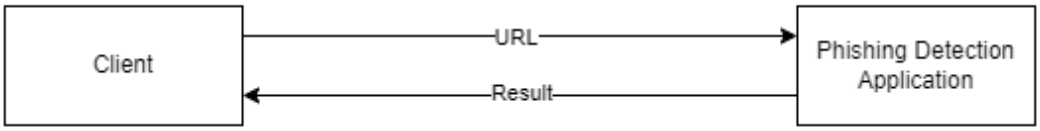
7. Continuous Improvement

-**Feedback Mechanisms**: Collect feedback from users, security experts, and system monitoring tools to assess system effectiveness, user satisfaction, and areas for improvement.

-**Iterative Development**: Uses an iterative approach (e.g., Agile, Scrum) to implement incremental updates, address user feedback, optimize performance, and adapt to evolving phishing tactics and technologies.

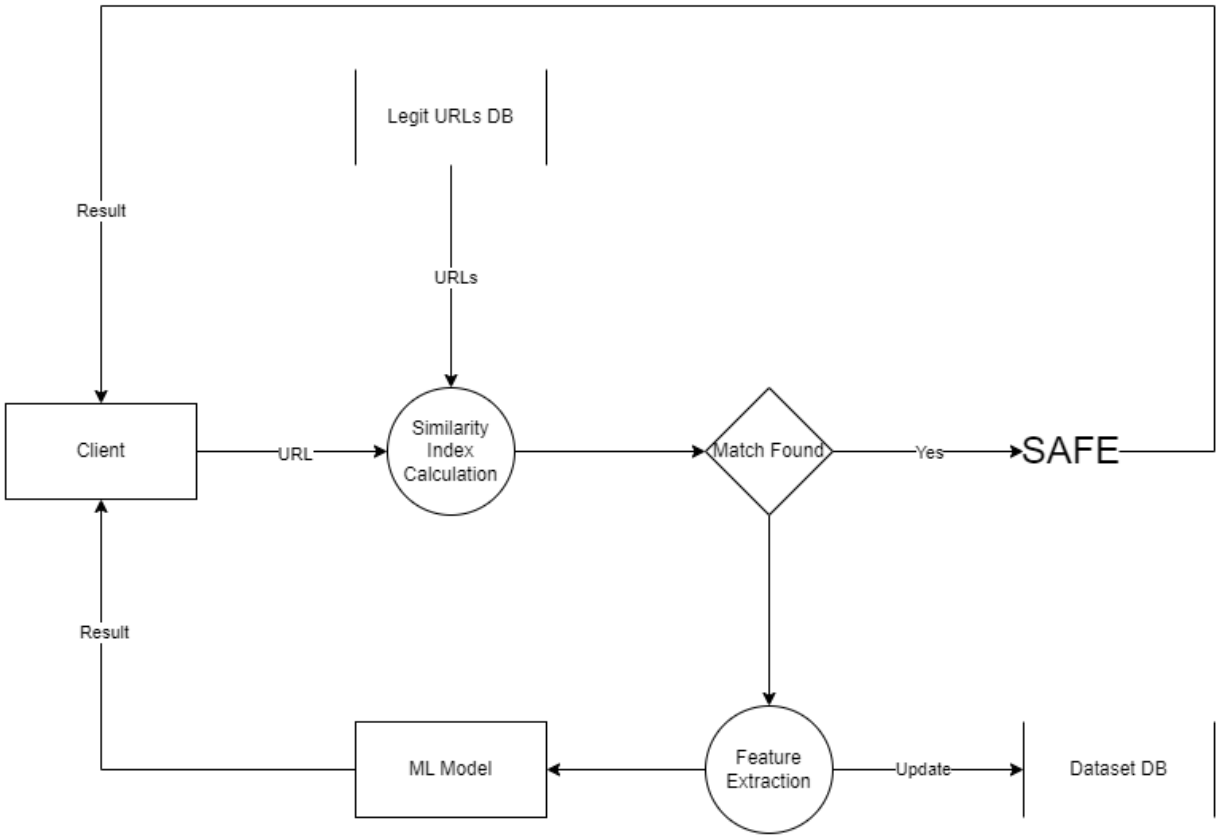
CONTEXT DIAGRAM

DFD 0



LEVEL 1 DFD

DFD 1



REFERENCES

Books

1. "Cybersecurity and Cyberwar: What Everyone Needs to Know" by P.W. Singer and Allan Friedman

2. "Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft" by Markus Jakobsson and Steven Myers

3. "Machine Learning and Security: Protecting Systems with Data and Algorithms" by Clarence Chio and David Freeman

Websites and Journals

1. IEEE Computer Society - https://www.computer.org/

2. ACM Digital Library - https://dl.acm.org/

3. Journal of Cybersecurity - https://academic.oup.com/cybersecurity

4. Information Systems Security Journal - https://www.tandfonline.com/loi/uiss20

BIBLIOGRAPHY

UDAY RAJ SINGH

Pursuing B,Tech (CS)

10th - 91.8%

12th – 91.8%

2nd Year - 70%

E-mail id – udayrajsingh288@gmail.com

Mobile No. 9695975787

ZOYA ASHIYAM

Pursuing B,Tech (CS)

10th - 92 %

12th - 84%

2nd Year -75 %

E-mail id – zazoya1234@gmail.com

Mobile no. 9555877325

MD SHABI AHMED

Pursuing B,Tech (CS)

10th - 90%

12th - 90%

2nd Year - 58%

E-mail id – mdshabi007@gmail.com

Mobile No. 8542035246

SWATI KESARWANI

Pursuing B,Tech (CS)

10th - 85.83%

12th - 71%

2nd Year – 72.10%

E-mail id -swatikesarwani2002@gmail.com

Mobile No. 9918080525