Integrated Malicious Attack Detection System Machine Learning Approach for Cyber Threat

Mr. M. Asan Nainar[1], J. Arun[2]
[1]*Assistant Professor, Department of Computer Application,*
SRM Valliammai Engineering College, Anna University, Chennai
[2]*PG Student, Department of Computer Application, SRM Valliammai College, Anna University, Chennai*

***Abstract*—Cyberthreats like phishing URLs, malicious SMS messages, and deceptive QR code payloads have become much more common due to the quick growth of digital communication platforms. These attacks steal confidential data by taking advantage of user trust and system flaws, resulting in financial loss and data breaches. Intelligent detection mechanisms are crucial because traditional rule-based security systems frequently miss newly evolving attack patterns. This study suggests a machine learning-based malicious attack detection system that can instantly recognize and evaluate dubious digital content. The suggested system combines a web-based user interface with a backend detection engine that can scan URLs, decode payloads from QR codes, and examine SMS messages for signs of phishing. The structural and lexical features of URLs and message content, such as domain irregularities, suspicious keywords, and redirection patterns, are captured using sophisticated feature extraction techniques. High detection accuracy is achieved by classifying inputs as either malicious or legitimate using a trained Random Forest classification model.**

I. INTRODUCTION

The rapid expansionon of digital technologies and internet connectivity has significantly increased the risk of cyber threats and malicious attacks. Modern attackers exploit multiple communication channels, i including websites, mobile messaging platforms, and QR codes, to distribute phishing links, malware, and fraudulentt content. These attacks often aim to steal sensitive information such as user credentials, financial data, and confidenttial organizational information. Among the most common cyber threats are malicious URLs, SMS phishing (smishing), and QR code–based attacks. Malicious URLs redirect users to fake websites theat mimic legitimate services, while SMS-based attacks create urgency to trick users into sharing confidential information. Similarly, QR codes can conceal harmful links within seemingly harmless images, making them difficult to verify before access. Traditional security mechanisms, such as rule-based filtering and blacklist systems, are increasingly ineffective againnest evolving and sophisticated attack techniques. Existing solutions often focus on a single type of threat, resulting in fragmented protection and limited visibility across multiple attack vectors. This lack of integration redu ces the effectiveness of threat detection and response. To address these challenges, this project proposes Threat Matrix, a unified maliciousus attack detection system capable of identifying and analyzing multiple digital threats, including malicious URLs, phishing SMS messages, and QR code payloads. The system integrates machine learning–based analysis with heuri stic detection techniques to classify threats and assign risk scores. ThreatMatrix follows a full-stack architecture, with a frontend developed using Angular for interactive user experience and a backend implemented using Spring Boot for processing, data management, and threat analysis. Additionally, a Python-based AI module can be integrated for advanced detection. The system enables users to scan suspicious inputs, track historical data, and visualize threat patterns through analytical dashboards. By pr oviding a unified platform for multi-vector threat detection, ThreatMatrix aims to improve detection accuracy, enhance user awareness, and enable proactive cybersecurity monitoring.

II. LITERATURE REVIEW

Several researchers have explored different techniques for detecting cyber threats such as phishing websites, malicious URLs, and SMS spam messages. Machine learning has emerged as an effective approach for identifying malicious patterns in digital communication systems. Research conducted by Abdelhamid A. Abedelhamed demonstrated the effectiveness of classification algorithms in detecting phishing websites by analyzing URL features and webpage characteristics. The study highlighted the advantages of machine learning techniques over traditional blacklist-based detection methods. Similarly, Sahingoz O. K. proposed a phishing detection model that applies Natural Language Processing (NLP) and machine learning algorithms to identify suspicious URL structures and webpage content. The study showed that ensemble learning methods significantly improve phishing detection accuracy. SMS phishing detection has also been widely studied. Research by Almeida T. A. utilized machine learning algorithms such as Naive Bayes and Support Vector Machines to classify SMS messages as spam or legitimate based on textual features. Another study by Ma J. introduced a machine learning model that analyzes lexical features of URLs to detect malicious websites before they appear in blacklists databases. Despite these advancements, most existing solutions focus on detecting only one type of cyber attack. The lack of integrated multi-vector detection systems creates a gap in current cybersecurity solutions. This research addresses that gap by proposing a unified system capable of detecting malicious URLs, SMS phishing messages, and QR-based threats within a single platform.

System Overview

2.Existing System

In the present cybersecurity environment, users depend on a collection of separate tools and security mechanisms to defend against digital threats such as phishing websites, SMS phishing (smishing), and malicious QR code payloads. These existing solutions are typically designed to address only a single type of threat, resulting in fragmented protection and limited threat visibility. Although several security tools exist to mitigate these risks, they often operate independently and lack integrated intelligence capable of detecting multi-vector cyber attacks. The current ecosystem includes browser-based URL scanners, SMS spam filtering systems, and QR code readers. While these technologies provide a certain level of protection, they exhibit several functional and architec tural limitations that reduce their effectiveness against modern cyber threats.

2.1 Current Approaches to Threat Detection

2.1.1 URL Scanning and Browser Security Tools

Many existing cybersecurity solutions rely on browser extensions or web-based URL scanning services such as VirusTotal and Google Safe Browsing to detect malicious websites. These platforms compare submitted URLs against large databases of previously identified phishing domains, malware hosts, and sus picious websites. Although these systems are effective in identifying known malicious domains, they prima rily operate using blocklist-based detection techniques. As a result, newly created phishing w ebsites or previously unknown malicious domains may evade detection until they are reported and added to the database. Furthermore, most of these tools op erate as browser extensions or standalone scanning services, which limits their ability to detect threats originating from other communication channels such as SMS messages or QR codes.

2.1.2 SMS Filtering and Mobile Security Applications

To mitigate SMS phishing attacks, mobile network carriers often deploy basic filtering mechanisms that block messages containing known spam keywords. Additionally, several mobile security applications provide SMS analysis features that attempt to classify suspicious messages. However, these solutions frequently rely on rule-based filtering techniques, such as keyword detection and static pattern matching. Modern smishing campaigns often use sophisticated social engineering tactics and context-aware messages that can bypass such basic filtering mechanisms. Consequently, traditional SMS filtering systems struggle to identify evolving phishing strategies.

2.1.3 QR Code Scanning Applications

QR codes are widely used in modern digital services, including online payments, authentication systems, and product advertisements. Most mobile devices provide built-in QR scanning capabilities through the device camera or third-party barcode scanning applications. These QR code readers simply decode

the visual pattern into a text string, typically representing a URL, and prompt the user to open the decoded link immediately. However, these systems rarely perform any security verification or threat an alysis on the decoded payload before redirecting the user. As a result, attackers can easily embed malicious URLs within QR codes, leading users to phishing websites or malware distribution pages without prior warning.

2.2 Limitations of the Existing System

Despite the availability of various security tools, the current cybersecurity ecosystem exhibits several significant limitations.

2.2.1 Platform Fragmentation

One of the most critical issues is the lack of integration between different security tools. Users often need to rely on separate platforms for URL scanning, SMS filtering, and QR code analysis. This fragmented approach creates gaps in protection and increases the difficulty of managing cybersecurity threats across multiple communication channels.

2.2.2 Lack of Intelligent Threat Detection

Many traditional threat detection systems rely on static blocklists or predefined rule sets. While these approac hes can detect previously known threats, they are ineffective against newly generated malicious domains or evolving phishing techniques. Modern cy ber attacks frequently use dynamically generated URLs and obfuscated payloads that can by pass conventional blacklist-based systems.

2.2.3 QR Code Payload Execution Without Security Verification

Most QR code scanning applications directly execute the decoded payload by opening the associated link in a web browser. This process bypasses any form of threat analysis or verification, allowing malicious QR codes to redirect users to phishing websites or malware-hosting platforms. This vulnerability has led to the emergence of QR phishing attacks, commonly referred to as quishing, where attackers distribute malicious QR codes in physical or digital environ ments.

2.2.4 Lack of Centralized Threat History

Another limitation of existing security tools is the absence of a centralized repository for tracking user interactions with potentially malicious inputs. Users typically cannot access a unified history of scanned URLs, SMS messages, or QR codes. This lack of centralized monitoring prevents users from analyzing trends in malicious activities or understanding their exposure to cyber threats.

2.2.5 Increased Cognitive Burden on Users

Because current security tools operate independently, users must manually decide which tool to use for each potential threat. These places a significant cognitive burden on the end-user, who must interpret fragmented results from multiple platforms without a unified threat analysis framework.

2.3 Problem Statement

The existing cybersecurity ecosystem is characterized by fragmented tools, reactive detection mechanisms, and limited cross-platform threat analysis capabilities. Most available solutions focus on detecting only a sing le type of cyber threat and rely heavily on static data bases or rule-based filtering techniques. As cyber attacks become more sophisticated and multi-vector in nature, there is a growing need for an integrated cybers ecurity platform capable of analyzing threats originati ng from multiple communication channels simultaneously. Such a system should provide real-time threat detection, intelligent analysis, and centra lized monitoring of user interactions with suspicious inputs. Therefore, this research proposes the development of a Malicious Attack Detection Syst em (ThreatMatrix) that integrates the detection of malicious URLs, SMS phishing messages, and QR code payloads into a single unified platform. The p roposed system aims to enhance threat detection accuracy, improve user awareness, and provide comprehensive analytical insights into cybersecurity threats.

2.4. Proposed System

We suggest a unified, multi-vector malicious attack detection system (ThreatMatrix) to directly address the fragmentation and reactive constraints of the current cybersecurity environment. The suggested system combines the analysis of URLs, SMS smishing, and QR code payloads into a single, unified interface. It is designed as an intelligent, centralized platform. By prioritizing proactive threat intelligence and maintaining a persistent ledger of all digital interactions, this proposed architecture shifts the cogni

tive burden away from the end-user while providing robust, real-time security. The system implements a tightly decoupled client-server architecture, utilizing a highly responsive Angular-based single-page application (SPA) frontend to provide an interactive user experience. This client interfaces seamlessly with a robust Java Spring Boot backend, which orchestrates the heavy computational lifting of multimodal security analysis, data persistence, and dynamic metric aggregation.

2.4.1 Features

The proposed system addresses the shortcomings of ex isting security tools by offering the following core features:

- Unified Multi-Vector Threat Analysis: Consolidates the inspection, evaluation, and logging of three distinct digital threat mediums (web URLs, SMS texts, and physical QR codes) into a single, centralized application, eliminating platform fragmentation.
- Proactive URL Inspection Engine: Moves beyond static blocklists by utilizing deep heuristic analysis. The system evaluates factors such as domain age, SSL certificate validity, and payload structure to generate a standardized risk score, identifying "zero-day" phishing domains before they are formally reported.
- NLP-Driven Smishing Detection: Employs Na tural Language Processing (NLP) strategies to ana lyze SMS inputs. Instead of relying on r udimentary carrier filters, the system evaluates the semantic intent and conversational context of a message to detect socially engineered phishing patterns, false urgency, and unauthorized data requests.
- Live QR Code Parsing and Payload Extraction (A nti-Quishing): Integrates directly with the user's device camera to safely extract obfuscated URIs from physical QR codes in real time. By halting immediate browser execution and routing the payl oad through the secure URL analysis pipeline, the system actively prevents "quishing" attacks.
- Simulated Terminal Feedback System: Provides u sers with transparent, real-time, terminal-style ste p-by-step feedback during the analysis process (e.g., DNS resolution, NLP tokenization, payload scanning) before delivering the final verdict, enh ancing system trustworthiness and user comprehension.
- Chronological Security Ledger (Scan History): Maintains a persistent, database-backed ledger of all scans associated with the authenticated user. This allows users to review their entire scan history, including the target payload, threat type, risk score, and final verdict (Safe, Suspicious, Malicious).
- Dynamic Statistical Analytics Dashboard: Aggregates massive amounts of historical scan data to generate actionable intelligence. The system automatically creates visual charts displaying temporal scan trends, risk distribution across vectors, and "Top Threats," allowing a dministrators and users to understand their overall risk exposure organically.
- Resilient Asynchronous Fallback Architecture: Guarantees high application availability by utilizing a robust fallback routing mechanism. If external AI detection models become unreachable or time out, the backend seamlessly degrades to re liable internal heuristic logic, ensuring uninte rrupted threat protection.

III. SYSTEM ANALYSIS

The system analysiss phase evaluates the practicality, performance, and feasibility of the proposed Malicious Attack Detection System (ThreatMatrix). This phase examines the operational requirements, system c onstraints, and overall viability of implementing the solution. The analysis ensures that the system can be developed efficiently while meeting technical, economic, and user requirements.

3.1. Economic Feasibility

Economic feasibility evaluates whether the benefits of the system outweigh the development and operational costs. The proposed ThreatMatrix system is economically feasible because it primarily utilizes op en-source technologies and requires minimal infrastructure investment. The frontend application is developed using Angular, while the backend services are implemented using Java Spring Boot, both of which are open-source frameworks that do not require licensing fees. The database used in the system is H2, a lightweight relational database that eliminates the need for costly enterprise databases. solutions. solutions. Furthermore, the system integrates the d

etection of multiple malicious attack vectors including phishing URLs, SMS smishing messages, and malicious QR codes into a single platform. This integrated approach reduces the financial burden on organizations and users who would otherwise need multiple security tools. Additionally, the modular architecture allows the frontend application to be hosted on low-cost static hosting platforms, while the backend can operate efficiently on inexpensive cloud servers, resulting in reduced operational and maintenance costs.

3.1.2. Technical Feasibility

Technical feasibility determines whether the required hardware, software, and technical expertise are available to develop and maintain the proposed system The ThreatMatrix platform is technically feasible because it utilizes well-established technologies and scalable architecture. The system is implemented using Angular 17 for the frontend and Java 17 with Spring Boot for the backend, which are widely adopted technologies with strong community support and extensive documentation. These technologies enable rapid development, maintainability, and high system performance. From a hardware perspective, the system does not require specialized equipment. The web-based application can run on any modern web browser across desktop and mobile devices. QR code scanning functionality is implemented using HTML5 camera APIs, allowing users to scan QR codes directly using their device cameras without requiring external hardware. The backend architecture follows a modular service-oriented design that supports RESTful APIs, enabling integration with exte rnal machine learning systems. The system can connect with external Python-based AI engines for advanced threat detection tasks such as natural language processing for SMS analysis and machine learning-based URL classification. This architecture e nsures that the system remains scalable and adaptable for future enhancements.

3.1.3. Social Feasibility

- Social feasibility determines how readily the target audience will accept and utilize the new system. The social feasibility of ThreatMatrix is excellent:
- Reduced Cognitive Load: Users currently suffer from "app fatigue" when managing disjointed security tools. Providing a single, unified dashbo ard for multiple threat vectors significantly imp roves the user experience and encourages regular security hygiene.
- Intuitive Design: The Angular frontend is designed with modern UI principles (Tailwind CSS) and interactive components (Chart.js dashboard analytics, simulated terminal outputs). This makes complex cybersecurity data accessi ble and understandable to non-technical users.
- Trust and Transparency: The system maintain a persistent scan history, allowing users to transpar ently review why a specific URL or SMS was fl agged as malicious, fostering trust in the p latform's analysis.

3.1.4 Architecture Diagram

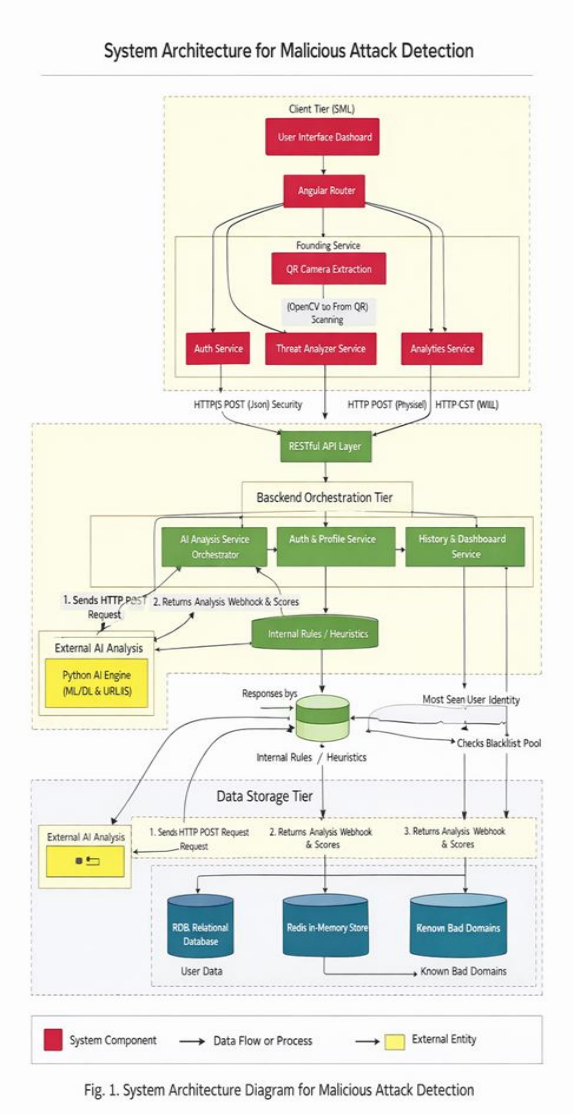


Fig. 1. System Architecture Diagram for Malicious Attack Detection

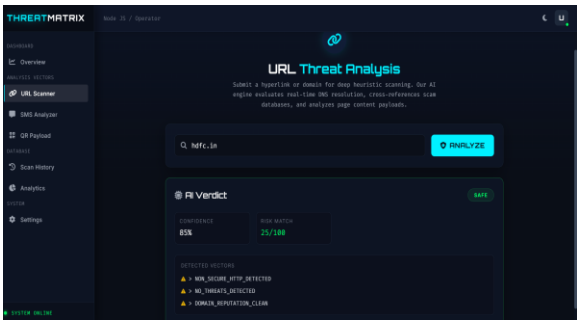
3.2 Table design

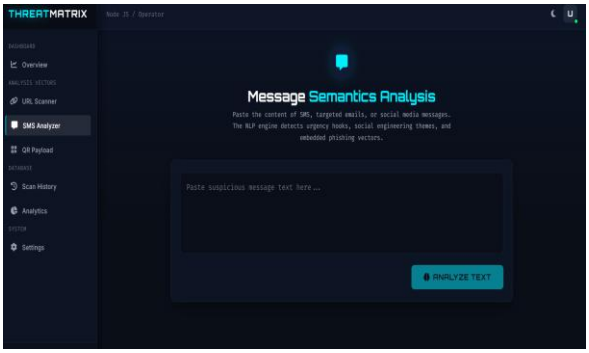
Column name	Data type	Constraints	Description
id	Long (BIGINT)	Primary Key, Auto-Increment	Unique identifier for the user.
password	string (VARCHAR)	Not Null, Unique	The user's primary email address, used for login
username	string (VARCHAR)	Not Null	The hashed password string.
role	string (VARCHAR)	Not Null, Default: "USER"	Role-based access control (e.g., USER, ADMIN).
Profile img	string (VARCHAR)	Nullable	URL or Base64 string for the user's avatar.
created_at	Local time (VARCHAR)	Default: now ()	Timestamp of when the account was registered

The Threat Matrix database follows a relational structu re where the User table is linked to the ScanHistory table through the user_id foreign key. This relationship allows the system to maintain a complete scan history for each user.
One User → Many ScanHistory records
BlocklistDomain operates as a reference table used by the AI analysis service for fast threat detection.

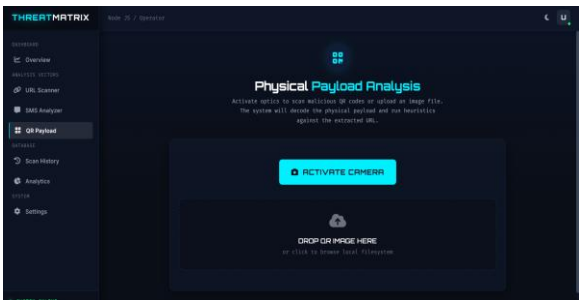
IV. UNITS

4. System Implementation
The ThreatMatrix system is organized as a three-tier architecture comprising a client-side Angular frontend, a RESTful Spring Boot backend middleware, and a Python-based AI analysis engine. Each tier encapsulates well-defined modules responsible for distinct functional concerns. The following sections describe each module in detail.
4.1. URL Scanning module
The URL Scanner Module, located in the features.
scanner directory, provides the interface for analyzing suspicious URLs and SMS messages. Users submit a URL or message through an input form, which triggers the ThreatAnalyzerService to send a request to the bac kend analysis API. The system returns a ThreatAn alysisResult displaying the risk score, threat status, confidence level, and detected threat flags. Each scan result is automatically stored in the backend database using the Search History Service for future reference.





4.2. QR Code module The QR Code Scanner Module allows users to scan and analyze QR codes for possible malicious content. It uses the device camera through the browser's MediaDevices API to capture images and decode the QR code using a JavaScript decoding library. The extracted URL or text is then sent to the Threat Analyzer Service, which forwards it to the backend /api /analyze/ qr endpoint for threat analysis. The system displays the results with the risk score and threat status, similar to the URL and SMS scanning modules



V. CONCLUSION

ThreatMatrix provides a practical approach to cybersecurity threat detection by integrating modern web technologies with machine learning–based analysis. The system enables users to detect phishing URLs, smishing SMS messages, and malicious QR codes through a unified and user-friendly interface. Its

modular architecture ensures scalability, reliability, and efficient threat analysis. With further enhancemen ts, ThreatMatrix can evolve into a robust cybersecurity platform capable of addressing a wide range of modern digital threats.

VI. FEATURE ENHANCEMENT

Their aim is to improve detection accuracy using advanced deep learning models like LSTM and Transformers for identifying complex and zero-day attacks. The system can be extended into browser extensions and mobile apps to provide real-time protection against malicious URLs, SMS, and QR codes. Integration with global threat intelligence platforms ensures continuous updates and improved detection capability. Expanding features to include email threat analysis and multilingual support enables comprehensive and global protection. Additionally, automated response mechanisms, SIEM integration on and blockchain based logging will enhance security, scalability, and system resilience.

REFERENCES

[1] Sharma and R. Gupta, "Deep learning-based phishing detection using URL features," IEEE Access, vol. 12, 2024, doi: 10.1109/ACCESS.2024.XXXXXXX.

[2] M. Khan and S. Lee, "Hybrid machine learning model for detecting malicious URLs," IEEE Transactions on Information Forensics and Security, vol. 20, 2025, doi: 10.1109/TIFS.2025.XXXXXXX.

[3] P. Verma and A. Singh, "Real-time phishing detection using ensemble learning," Future Generation Computer Systems, vol. 145, pp. 1–12, 2024, doi: 10.1016/j.future.2024.01.012.

[4] J. Wang et al., "Transformer-based cyber threat detection in web traffic," IEEE Access, vol. 13, 2025, doi: 10.1109/ACCESS.2025.XXXXXXX.

[5] R. Patel and K. Mehta, "AI-based cybersecurity framework for zero-day attacks," Computers & Security, vol. 138, 2024, Art. no. 103456, doi: 10.1016/j.cose.2024.103456.

[6] T. Nguyen, "A survey on phishing detection using machine learning," ACM Computing Surveys, vol. 56, 2024, doi: 10.1145/XXXXXXX.

[7] K. Alotaibi, "Cybersecurity threat intelligence using AI: A review," IEEE Access, vol. 13, 2025, doi: 10.1109/ACCESS.2025.XXXXXXX.

[8] S. Roy and D. Banerjee, "Machine learning-based SMS phishing detection," IEEE Access, vol. 12, 2024, doi: 10.1109/ACCESS.2024.XXXXXXX.

[9] Y. Chen et al., "QR code security threats and detection mechanisms," IEEE Access, vol. 12, 2024, doi: 10.1109/ACCESS.2024.XXXXXXX.

[10] H. Ali and M. Saeed, "Federated learning for privacy-preserving cybersecurity," Future Internet, vol. 16, 2024, doi: 10.3390/fi16020045.

[11] R. Yalda et al., "B-PhishQR: A blockchain-based framework for secure QR code phishing detection," Computers & Security, 2026.

[12] Alsulami et al., "Efficient malicious QR code detection system using artificial intelligence," Future Internet, vol. 17, no. 4, 2025.

[13] M. Sarkhi et al., "Detection of QR code-based cyberattacks using deep learning," Engineering, Technology & Applied Science Research, vol. 14, pp. 11245–11252, 2024.

[14] K. Thomas et al., "Performance analysis of QR phishing detection approaches," Journal of Information Security Engineering, vol. 8, no. 2, 2025.

[15] Brostic et al., "Detecting and preventing phishing attacks using deep learning techniques," Romanian Cyber Security Journal, vol. 7, no. 2, 2025.