International Journal of Multidisciplinary Research and Scientific Innovations (IJMRSI) > Volume 01, Issue 01, March 2026 | PP: 01-07 > E-ISSN: [Applied For] | Website: www.ijmrsi.com

AI BASED MALICIOUS URL DETECTION SYSTEM ENHANCING WEB SECURITY THROUGH INTELLIGENT ANALYSIS

Shamix Subaitha. S
Department of Computer Science Sri Krishna Arts and Science College Coimbatore,India

Shalini.M
Department of Computer Science Sri Krishna Arts and Science College Coimbatore,India

Tejashvi.V
Department of Computer Science Sri Krishna Arts and Science College Coimbatore,India

Farheena Naaz.Z
Department of Computer Science Sri Krishna Arts and Science College Coimbatore,India

Abstract - People interact with numerous web links daily through emails, websites, and online platforms. The widespread adoption of the internet and digital However, many of these links are malicious traps communication technologies has transformed how created by attackers to steal sensitive information, individuals and organizations access information and install malware, or redirect users to fraudulent services. Users routinely interact with web links websites. Such malicious URLs can cause serious obtained through emails, social media platforms, security consequences without the user's awareness. search engines, messaging applications, and online Therefore, an effective mechanism to verify the safety advertisements. While these links provide convenience of links before accessing them is essential ,this presents and connectivity, they also serve as a major attack the design of a Malware URL Scanner developed using vector for cybercriminals. Malicious URLs are React.js for the frontend and Supabase for backend commonly used to launch phishing attacks, distribute services. The system identifies suspicious URLs using malware, steal sensitive information, and redirect users an AI-based analysis powered by the Claude API. It to fraudulent or compromised websites, resulting in evaluates links by comparing them with known threat serious security and privacy risks **[1], [5]**.

Keywords - Malware URL Scanner, Artificial Intelligence, Cybersecurity, URL Analysis, Phishing Detection.

INTRODUCTION

links to conceal the true destination of malicious content. Due to these deceptive techniques, users frequently fail to recognize harmful links before accessing them, leading to unintentional exposure to cyber threats. The impact of such attacks can range from personal data theft and financial loss to large- scale organizational breaches and system compromise **[1]**.

signature-based detection mechanisms. These of cyberattacks. Modern malware URL scanners aim to approaches identify threats by comparing URLs provide fast, accurate, and real-time threat detection. against predefined databases of known malicious links. This proactive approach enhances online safety for While effective in detecting previously identified individuals and organizations.

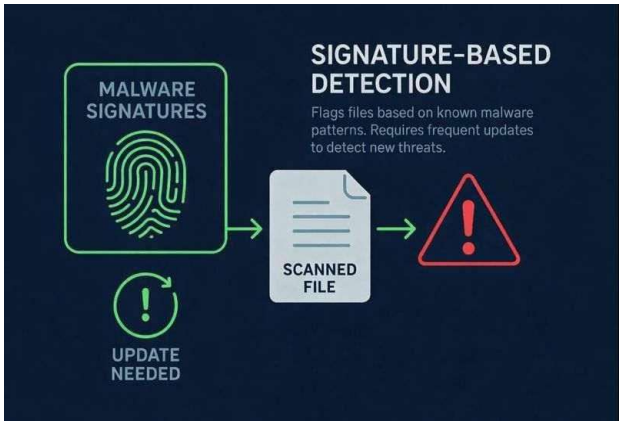
threats, blacklist-based systems are inherently reactive. Newly created malicious URLs or zero-

day threats often bypass detection until they are reported and added to threat databases. This delay leaves users vulnerable during the critical window between threat emergence and database updates **[2], [6]**.

In addition to limited detection capabilities, many existing systems provide only binary outputs such as “safe” or “unsafe,” without offering sufficient explanation or transparency regarding the decision- making process. The lack of contextual information makes it difficult for users, especially non-technical individuals, to understand the nature of the threat or assess the associated risk. Furthermore, inconsistent results across different security engines can reduce user trust in automated scanning tools and hinder effective decision-making **[2]**.To overcome these limitations, modern cybersecurity solutions increasingly incorporate artificial intelligence and machine learning techniques. AI-based systems are capable of identifying complex patterns, detecting anomalies and adapting to evolving threat landscapes in real time. By analyzing both structural and behavioral characteristics of URLs,intelligent scanners can identify suspicious activity even when a link does not exist in known blacklists. Such capabilities are particularly important for detecting phishing campaigns and zero-day attacks that rely on previously unseen URL patterns **[5],[7].** This is used against the cyber attacks and stealing of data from the people for using their personal data , their feed to display ads for other apps or promotion and marketing of their product ,u may not know the phishing activities behind the url and we can’t detect with the naked eye so this site is used for checking the level of clean and detection of phishing and also keep track of the history. Here is a **5-line introductory paragraph** you can use for a **Malware URL Scanner**:A malware URL scanner is a crucial cybersecurity tool designed to identify and prevent access to malicious websites. It analyzes web links to detect threats such as phishing, malware distribution, and fraudulent websites.The lack of adaptive learning mechanisms prevents these systems from systemsvulnerable toAnother issue is inconsistent detection results advanced attacks.

II . PROBLEM STATEMENT

Existing URL scanning tools primarily rely on traditional techniques such as blacklists and signature-based detection. While these approaches can identify previously known threats, they often fail to keep pace with rapidly evolving cyberattacks, particularly zero- day threats. The outputs from such tools can be inconsistent and delayed, as different security engines may produce conflicting results. Moreover, most of these solutions provide only basic link verification and lack real-time protection. They rarely leverage advanced artificial intelligence to analyze features such as unusual patterns, special characters, or numerical sequences that could indicate malicious behavior. Some systems also depend on community reports, which may be incomplete or slow to update. These limitations highlight the need for a more intelligent, efficient, and reliable URL scanning system capable of providing timely and accurate threat detection. Furthermore, traditional URL scanners often struggle to detect sophisticated phishing attacks that mimic legitimate websites and fraudulent activities before users interact with them. Additionally, many detection methods leaves systems vulnerable to advanced attacks.



III . EXISTING SYSTEM

VirusTotal is a widely used online platform that analyzes URLs, files, and domains for potential threats. It aggregates results from multiple antivirus engines and URL scanning tools to provide users with an overall assessment of malicious activity. While VirusTotal is effective in detecting known threats, it primarily relies on signature-based and heuristic methods, which can delay the identification of new or zero-day attacks. Additionally, the system provides binary or aggregated outputs without offering detailed explanations of the underlying risk factors, which may limit user understanding of specific threats. The platform also depends on updates from its integrated engines and community reports, which may not always be timely or comprehensive.

Traditional malware scanning platforms like VirusTotal face several limitations when it comes to detecting modern threats. They often rely heavily on blacklists and known virus signatures, which means that newly emerging or previously unseen threats can go undetected. Additionally, threat databases and virus definitions are not always updated in real time, leaving users exposed to the latest risks. The platform also tends to provide only brief or generic explanations about why a particular link or file has been flagged, which can cause confusion for non- technical users. Moreover, there is no deep AI- driven analysis of URLs; features such as structural patterns, unusual symbols. across different antivirus engines, making it difficult to confidently assess whether a file or URL is actually safe.

IV . PROPOSED SYSTEM

The proposed system is an AI-based malware URL scanning solution designed to provide real-time and accurate threat detection. Unlike traditional systems, it integrates multiple security checks including live DNS resolution, SSL certificate validation, domain age analysis, and response time monitoring to assess the authenticity and behavior of URLs. An advanced AI engine analyzes URL structures, suspicious patterns, and deceptive characteristics to identify phishing and malware-related threats. Based on this analysis, URLs are classified into high, medium, or low risk levels with clear explanations. The system is scalable, adaptive, and capable of detecting both known and previously unseen threats, making it suitable for individual and enterprise-level cybersecurity applications.

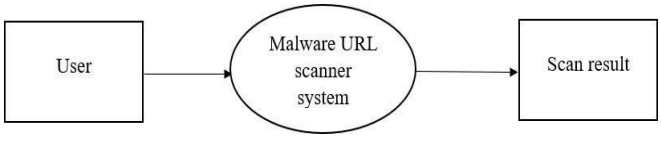
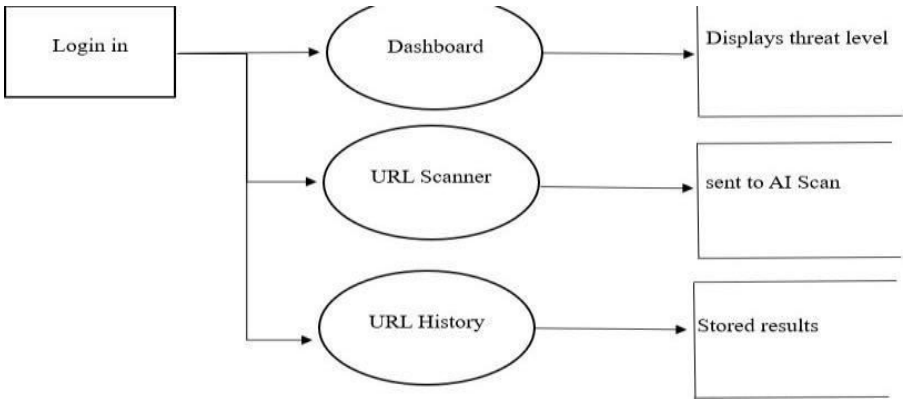


Fig.1 Block Diagram

The block diagram illustrates the basic working structure of the proposed Malware URL Scanner system. The process begins with the **User**, who submits a website link for analysis. The input URL is forwarded to the **Malware URL Scanner System**, which serves as the core processing unit. This system performs multiple security checks, including URL preprocessing, domain and network verification, and AI-based analysis to evaluate the legitimacy and risk associated with the submitted link. After completing the analysis, the system generates the **Scan Result**, which indicates the threat level of the URL. The result is presented to the user in a clear and understandable format, enabling informed decision-making regarding the safety of the link.The block diagram that is shown here is most simple and first step of the url scanner , a representation of user , intermediate and server this can be discussed more.

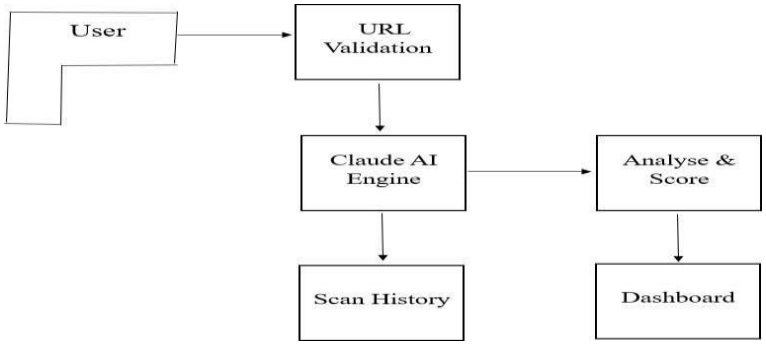
V . WORKING FLOW

The working flow of the proposed AI-based malware URL scanner begins when a user submits a URL through the application interface. The system first performs URL preprocessing to normalize the input and extract basic lexical features. Next, live DNS resolution and SSL certificate verification are carried out to validate domain authenticity and secure communication. Domain age analysis is then applied to identify newly registered or suspicious domains. Based on the AI-driven analysis, the system classifies the URL into a corresponding threat level such as high, medium, or low. The evaluated results are displayed to the user on the dashboard with clear risk information. Simultaneously, the scanned URL and its threat assessment are securely stored in the URL history module for future reference. This sequential flow ensures secure access, real-time analysis, transparent reporting, and effective monitoring of malicious URLs.



From the table, it can be understood that the proposed AI-based malware URL scanner provides stronger and more practical protection than the existing system.

Parameter	Existing system	Proposed system
Detection Method	Signature- and heuristic-based scanning	AI-driven multi-layer analysis
Real-Time Analysis	Limited depends on periodic updates	Live DNS, SSL, and AI analysis
Zero-Day Threat Detection	Reactive limited and	Proactive AI- based detection
URL Feature Analysis	Basic inspection	Advanced structural and lexical analysis
Domain Verification	Partial reputation checks	Domain and age live validation DNS
SSL Certificate Validation	Not explicitly verified	SSL authenticity verification
Threat Classification	Binary aggregated or results	High, Medium, Low risk scoring

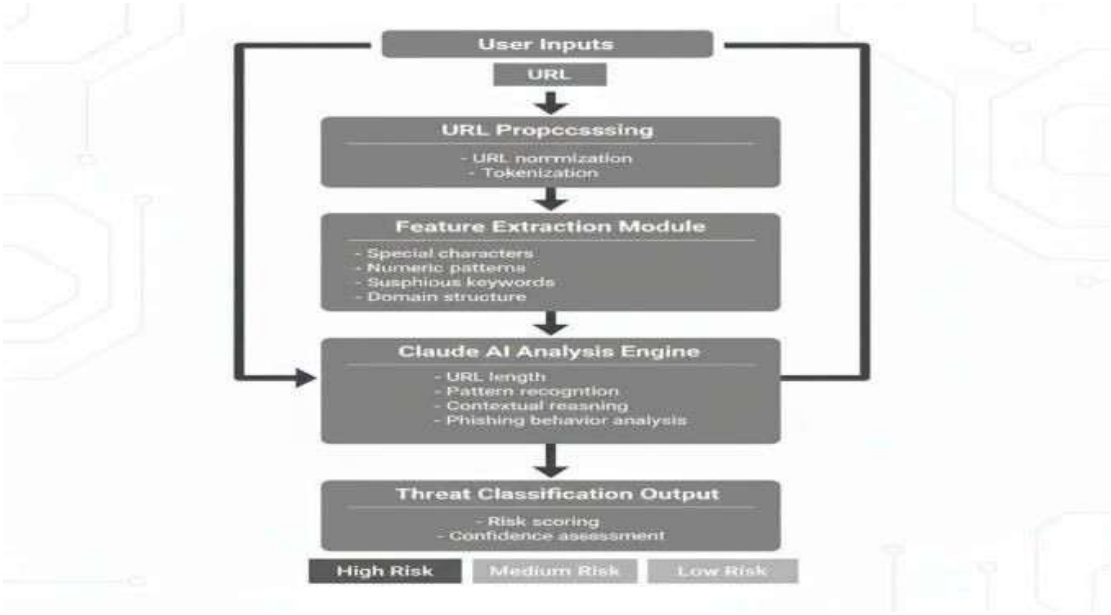


Fg . 3 Block Diagram

From fig 3 the proposed architecture illustrates a URL analysis and scoring workflow. The process begins when the user submits a URL, which is first passed through a URL validation module to ensure correctness and safety.The dashboard serves as the central interface of the proposed AI-based malware URL scanner. It provides users with access to core system functionalities, including URL submission, threat level visualization, and scan history review. The dashboard displays analyzed results in a structured and user- friendly format, enabling users to quickly interpret risk levels.

The system workflow begins with the end user accessing the application through a web interface

hosted on Netlify, which deploys the React-based frontend.This step ensures that the URL is in a consistent format before further evaluation.The processed URL is then passed to the feature extraction module, which derives relevant lexical and structural characteristics such as character patterns, keywords, and domain structure. These extracted features are analyzed by the Claude AI API, which applies intelligent reasoning and pattern recognition to detect potential indicators of malicious or phishing behavior. Based on the AI analysis, the system performs threat evaluation and assigns a classified risk level . *Claude and Its Role* - Claude is an advanced artificial intelligence model designed to analyze and understand textual patterns with high accuracy. It is capable of The image illustrates the internal analysis workflow of identifying subtle irregularities, contextual clues, and the proposed AI-based malware URL scanning system. deceptive structures that may not be easily detected It highlights how a submitted URL is processed in through traditional rule-based or signature-based multiple structured stages before a final decision is methods. Claude uses intelligent reasoning to made. Each block in the diagram represents a distinct recognize suspicious content by examining patterns, functional layer, ensuring that the URL is thoroughly language usage, and structural characteristics, making examined rather than being evaluated through a single, it effective for detecting phishing and malicious shallow check. This process enables the system to behavior. examine individual elements of the URL in detail.



Tokenization helps in identifying suspicious segments, unusual patterns, or deceptive naming techniques that are commonly used in phishing and malware-based URLs. By analyzing both structural and contextual characteristics, the AI engine is capable of detecting malicious intent.

VI . CONCLUSION

The proposed AI-based malware URL scanner has a significant impact on improving web security by enabling faster and more accurate detection of malicious links. By integrating real-time analysis with artificial intelligence, the system reduces user exposure to phishing attacks, malware distribution, and fraudulent websites. Its transparent threat classificationenhances user awareness and supports informednames," in *Proc. RAID*, Gothenburg, Sweden, 2014, decision-making

REFERENCES

[1] A. K. Jain and B. B. Gupta, "Phishing detection: Analysis of visual similarity based approaches," *Security and Communication Networks*, vol. 9, no. 17, pp. 386–404, 2016.

[2] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: A literature survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2091–2121, 2013.

[3] T. Banks, "Modern web application development using React.js," *International Journal of Web Engineering*, vol. 8, no. 2, pp. 45–52, 2020.

[4] Supabase Inc., "Supabase: Open source Firebase alternative," [Online]. Available: https://supabase.comdoesn't: A study of phishing detection," in *Proc. IEEE Conference on Technologies for Homeland Security*, 2017.

[12] K. Liao, X. Zhao, A. Doupe, and G. Ahn, "Behind closed doors: Measurement and analysis of CryptoLocker ransomware attacks," in *Proc. APWG eCrime Researchers Summit*, 2016.

[13] M. Garera, N. Provos, M. Chew, and A. Rubin, "A framework for detection and