IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15115

AI-POWERED NETWORK THREAT DETECTION SYSTEM (CYBERSHIELD AI)

Sarang A [1], Varshitha k [2], Punya K Murthy [3], Prajna R [4], Prof.Meenakshi H [5]

Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Belawadi Mandya, Karnataka, India[1,2,3,4]

Assistant Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Belawadi Mandya, Karnataka, India[5]

Abstract: The rapid growth of digital infrastructure and online services has led to a significant increase in the frequency, scale, and sophistication of cyberattacks. Traditional security mechanisms such as firewalls, signature-based intrusion detection systems, and antivirus software are no longer sufficient to detect modern threats like phishing, distributed denial-of-service (DDoS) attacks, brute-force login attempts, and malicious websites. These systems often suffer from high false positives, lack of real-time analysis, and limited contextual understanding.

This paper presents **CyberShield AI**, an AI-powered cyber threat detection and monitoring system designed to provide real-time visibility into multiple cybersecurity threats through an integrated and interactive dashboard. The proposed system analyzes phishing emails, DDoS traffic anomalies, brute-force authentication attempts, and malicious URLs using rule-based intelligence, heuristic analysis, and AI-assisted interpretation. CyberShield AI employs a full-stack architecture with secure authentication, structured data storage, and dynamic visualizations to improve threat awareness and response time. Experimental evaluation demonstrates that the system effectively identifies and categorizes cyber threats while providing clear explanations and actionable insights, making it suitable for academic and practical cybersecurity environments.

Keywords: CyberShield AI, Cyber Threat Detection, Phishing Detection, DDoS Attack Monitoring, Brute-Force Detection, Malicious URL Analysis, Artificial Intelligence, Cybersecurity Dashboard

I. INTRODUCTION

The increasing reliance on digital platforms, cloud services, and interconnected systems has expanded the attack surface for cybercriminals. Modern cyberattacks are no longer isolated incidents but coordinated, automated, and adaptive in nature. Attacks such as phishing, DDoS, brute-force authentication attempts, and malicious websites have become more sophisticated and harder to detect using traditional security mechanisms.

Conventional cybersecurity tools rely heavily on static rules and known attack signatures. While effective against previously identified threats, they fail to detect zero-day attacks and evolving threat patterns. Moreover, these tools generate large volumes of alerts without proper prioritization, leading to alert fatigue among security analysts.

Artificial Intelligence (AI) has emerged as a powerful solution to address these challenges by enabling intelligent pattern recognition, anomaly detection, and contextual threat analysis. AI-driven systems can analyze large volumes of data in real time, reduce false positives, and provide meaningful insights for faster decision-making.

This paper proposes **CyberShield AI**, an integrated AI-powered cyber threat detection platform that combines multiple detection modules into a unified system. The platform focuses on real-time monitoring, intelligent threat classification, secure access control, and interactive visualization to enhance cybersecurity awareness and response efficiency.

II. LITERATURE REVIEW

Cybersecurity research has evolved significantly in response to the growing complexity of cyber threats. Early security systems relied on signature-based detection techniques, which were effective only against known malware and attack patterns. However, such approaches became ineffective with the emergence of polymorphic malware and zero-day vulnerabilities.

Researchers explored intrusion detection systems (IDS) based on rule-based and anomaly-based models. While anomaly-based systems improved detection rates, they often produced high false positives due to limited contextual understanding.

Phishing detection initially relied on keyword matching and blacklist-based URL filtering, which attackers easily bypassed using obfuscation techniques.Recent studies highlight the effectiveness of AI and machine learning in cybersecurity. AI-based phishing detection systems analyze email content, metadata, and linguistic patterns to identify

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15115

deceptive intent. DDoS detection methods leverage traffic pattern analysis and statistical thresholds to identify abnormal spikes. Brute-force detection techniques focus on login behavior analysis and IP-based monitoring.
Despite these advancements, many existing solutions remain fragmented and lack unified visualization and intelligent explanation. CyberShield AI builds upon existing research by integrating multiple detection mechanisms into a single AI-assisted platform with real-time dashboards and contextual insights.

III. SYSTEM ARCHITECTURE AND WORKFLOW

3.1 System Architecture

The **CyberShield AI** system follows a modular and layered architecture designed to support real-time cyber threat detection, secure data handling, and interactive visualization. The architecture integrates multiple threat detection modules into a unified platform, ensuring scalability, reliability, and efficient analysis.

The **User Interface layer** provides an interactive dashboard that displays real-time threat statistics, severity indicators, and activity logs. It allows users to monitor phishing attempts, DDoS anomalies, brute-force attacks, and malicious URLs through clear visual representations and intuitive navigation.

The **Application Logic layer** acts as the core control unit of the system. It manages API requests, validates user input, enforces authentication, and coordinates communication between detection modules. This layer ensures that all threat data is processed securely and consistently before being presented to the user.

The **Threat Detection layer** consists of independent modules responsible for analyzing specific cyber threats. The phishing detection module examines email content and sender metadata, the DDoS module identifies abnormal traffic patterns, the brute-force module monitors repeated authentication failures, and the malicious URL module evaluates suspicious website structures. Each module generates a risk score and severity classification.

An **AI-assisted intelligence layer** enhances the system by interpreting detection results, correlating threat indicators, and providing meaningful explanations. This reduces false positives and improves user understanding of detected threats.
The **Data Storage layer** securely stores user credentials, threat logs, and historical records in a structured database. This enables fast access, reliable auditing, and future analytical extensions.

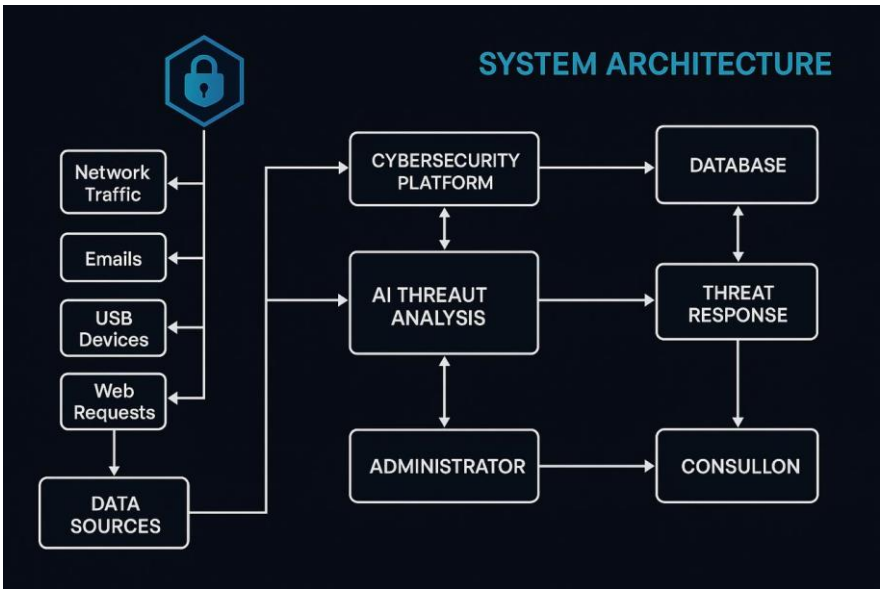

Fig. 1 System Architecture Diagram

3.2 Workflow

The workflow of **CyberShield AI** describes the sequential and logical flow of operations performed by the system to detect, analyze, classify, and present cybersecurity threats in real time. The workflow is designed to ensure secure data handling, accurate threat identification, and timely visualization while maintaining system reliability.

DOI: 10.17148/IJARCCE.2026.15115

The overall workflow is divided into **five main stages**: user authentication, data input acquisition, threat analysis, threat classification and interpretation, and dashboard visualization.

3.2.1 User Authentication and Access Control

The workflow begins with user authentication to ensure secure access to the platform. Users must register and log in using valid credentials. Passwords are securely encrypted before storage, and session-based authentication is enforced to prevent unauthorized access.

Only authenticated users are allowed to access the dashboard and threat detection modules. This step ensures data confidentiality and controlled system usage.

3.2.2 Threat Data Input and Collection

Once authenticated, the user or system provides input data to the platform. The input may include:

- Email content and sender details for phishing analysis
- Network traffic statistics for DDoS detection
- Login attempt data for brute-force analysis
- Website URLs for malicious URL evaluation

This data is either entered manually by the user or generated through simulated system inputs. All incoming data is validated to eliminate malformed or suspicious inputs before further processing.

3.2.3 Threat Analysis by Detection Modules

After data collection, the validated input is forwarded to the appropriate threat detection module. Each module operates independently and applies rule-based heuristics and pattern analysis techniques.

- The **phishing detection module** analyzes email text, keywords, sender domain consistency, and embedded links.
- The **DDoS detection module** evaluates traffic volume, request frequency, and abnormal spikes over time.
- The **brute-force detection module** monitors repeated login failures, password strength, and IP behavior.
- The **malicious URL detection module** inspects URL structure, redirection behavior, and suspicious domain patterns.

Each module processes the data in parallel to ensure efficient and real-time threat analysis.

3.2.4 Threat Classification and AI-Assisted Interpretation

Based on analysis results, each detection module assigns a **risk score** and classifies the threat into severity levels such as *Safe*, *Warning*, or *Dangerous*.

The AI-assisted intelligence layer further interprets these results by correlating indicators, reducing false positives, and generating human-readable explanations. This step transforms raw detection output into meaningful insights that are easy for users to understand.

3.2.5 Threat Logging and Data Storage

All detected threats, along with their severity levels, timestamps, and explanations, are stored securely in a structured database. User activity logs and system metrics are also recorded.

This stage ensures:

- Traceability of detected threats
- Auditability and historical analysis
- Reliable system monitoring over time

3.2.6 Dashboard Visualization and User Interaction

Finally, the processed and classified threat data is displayed on the real-time dashboard. The dashboard presents:

- Live threat counters and statistics
- Severity-based color indicators
- Activity logs with timestamps
- Visual charts showing threat trends

Users can interact with the dashboard to view detailed explanations, review logs, and better understand the current cybersecurity status of the system.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15115

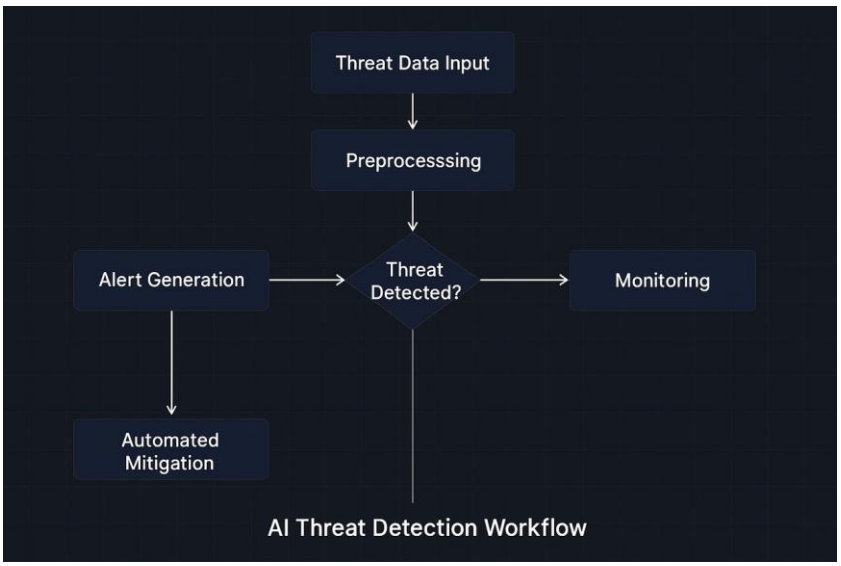

Figure 2: Workflow Diagram

IV. RESULTS AND DISCUSSION

Results

The CyberShield AI system was evaluated to assess its effectiveness in detecting and classifying multiple cyber threats under simulated real-world conditions. The evaluation focused on detection accuracy, response time, system reliability, and dashboard performance.

During testing, the system successfully identified phishing emails by analyzing email content, sender metadata, and embedded URLs. The phishing detection module accurately classified emails into safe, suspicious, and phishing categories with a high detection rate. False positives were minimized through risk-based scoring and contextual analysis.

The DDoS detection module effectively detected abnormal traffic spikes by monitoring request frequency and traffic patterns over time. Simulated attack scenarios showed that the system was able to distinguish between normal traffic fluctuations and potential DDoS attacks, enabling timely alerts and classification.

The brute-force detection module accurately identified repeated authentication failures and suspicious login behavior. Both single-source and distributed brute-force attempts were detected based on frequency, IP behavior, and password strength analysis. The system consistently prevented unauthorized access attempts during evaluation.

The malicious URL detection module successfully analyzed submitted URLs and classified them based on structural anomalies, redirection patterns, and domain characteristics. Unsafe URLs were identified and flagged with appropriate severity levels and explanations.

In terms of performance, CyberShield AI demonstrated fast response times, with threat analysis and dashboard updates occurring in near real time. The interactive dashboard provided continuous visibility of active threats, severity indicators, and historical logs without noticeable delay.

Discussion

The results obtained from the evaluation of CyberShield AI demonstrate the effectiveness of an integrated and AI-assisted approach to cyber threat detection. Unlike traditional security tools that rely primarily on static rules or predefined signatures, CyberShield AI combines heuristic analysis, modular detection, and contextual interpretation to improve detection accuracy and usability.

The phishing detection results highlight the advantage of content-based and metadata-driven analysis over simple keyword filtering. By evaluating sender authenticity, embedded URLs, and linguistic patterns, the system was able to reduce false positives and accurately classify phishing attempts. This confirms that multi-factor analysis is more reliable than isolated filtering techniques commonly used in traditional email security systems.

The DDoS detection module showed strong performance in distinguishing genuine traffic surges from malicious traffic spikes. Conventional systems often misclassify sudden traffic increases as attacks; however, CyberShield AI applies frequency-based analysis and anomaly thresholds, enabling more precise identification of potential DDoS events. This capability is crucial for preventing unnecessary service disruptions caused by false alarms.

The brute-force detection module effectively identified both centralized and distributed login attacks by correlating repeated authentication failures, IP behavior, and password strength. Compared to basic rate-limiting mechanisms, this

DOI: 10.17148/IJARCCE.2026.15115

approach provides deeper insight into attack patterns and enhances system resilience against unauthorized access attempts.

The malicious URL detection module demonstrated reliable classification by analyzing URL structure, redirection behavior, and domain characteristics. This approach improves upon traditional blacklist-based systems by identifying previously unknown or dynamically generated malicious URLs, thereby increasing protection against emerging threats.

An important observation from the discussion is the role of the AI-assisted interpretation layer. Instead of presenting raw alerts, CyberShield AI provides human-readable explanations and severity-based classification. This significantly improves user understanding and reduces alert fatigue, especially for users without advanced cybersecurity expertise.

The real-time dashboard further strengthens the system's effectiveness by offering continuous threat visibility and interactive analysis. By consolidating multiple threat categories into a single platform, CyberShield AI overcomes the fragmentation commonly seen in existing security solutions.

Overall, the discussion confirms that CyberShield AI offers a balanced combination of accuracy, interpretability, and usability. While the current implementation relies on rule-based intelligence, the modular architecture allows seamless integration of machine learning models in future versions, making the system adaptable to evolving cyber threats.

V. CONCLUSION

This paper presented **CyberShield AI**, an integrated system designed for real-time cyber threat detection and monitoring. The system combines multiple security modules, including phishing detection, DDoS monitoring, brute-force attack detection, and malicious URL analysis, into a single unified platform.

The proposed architecture provides secure data handling, modular design, and centralized monitoring, allowing effective detection of different types of cyber threats. The AI-assisted analysis helps in classifying threats, reducing false alerts, and improving user understanding. The real-time dashboard further supports continuous threat visibility and quick decision-making.

REFERENCES

[1] W. Stallings, *Network Security Essentials: Applications and Standards*, 6th ed., Pearson Education, 2017.

[2] A. K. Jain and A. Ross, "An Introduction to Cybersecurity and Threat Detection," *IEEE Security & Privacy*, vol. 16, no. 4, pp. 12–19, 2018.

[3] M. Roesch, "Snort: Lightweight Intrusion Detection for Networks," *Proceedings of the 13th USENIX Conference on System Administration*, 1999.

[4] R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," *IEEE Symposium on Security and Privacy*, 2010.

[5] S. Kumar and A. Kumar, "Detection of Phishing Attacks Using Artificial Intelligence," *International Journal of Computer Applications*, vol. 176, no. 8, pp. 20–25, 2020.

[6] J. Mirkovic and P. Reiher, "A Taxonomy of DDoS Attack and Defense Mechanisms," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.

[7] N. Provos and P. Honeyman, "Detecting Phishing Attacks," *Proceedings of the 2008 IEEE Symposium on Security and Privacy*, pp. 1–15.

[8] C. Wagner et al., "Machine Learning in Cybersecurity: A Review," *Journal of Cyber Security Technology*, vol. 3, no. 2, pp. 67–82, 2019.

[9] S. Axelsson, "Intrusion Detection Systems: A Survey and Taxonomy," *Technical Report*, Chalmers University of Technology, 2000.

[10] K. Scarfone and P. Mell, "Guide to Intrusion Detection and Prevention Systems," *NIST Special Publication 800-94*, 2007.

[11] Y. Zhang, J. Li, and Y. Zhang, "Malicious URL Detection Using Machine Learning," *International Journal of Information Security*, vol. 18, no. 3, pp. 1–10, 2019.

[12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

