Intelligent Vulnerability Assessment and Exploitation Detection System: An AI-Enabled Autonomous Cybersecurity Response Framework

Nambiraj. S[1], Adhith K.R[2], Maadhula R[3]
[1,2]*III B.Sc Digital and Cyber Forensic Science, Department of Digital and Cyber Forensic Science, Nehru Arts and Science College, Coimbatore, Tamil Nadu, India*
[3]*Assistant Professor, Department of Digital and Cyber Forensic Science, Nehru Arts and Science College, Coimbatore, Tamil Nadu, India*

Abstract- The exponential growth of interconnected digital infrastructures has significantly increased organizational exposure to cyber threats. Traditional vulnerability assessment and incident response mechanisms rely on periodic scanning, signature-based detection, and manual intervention, which are inadequate against modern attack techniques such as zero-day exploits, advanced persistent threats, fileless malware, and automated exploitation frameworks. This research article proposes an Intelligent Vulnerability Assessment and Exploitation Detection System integrated with an AI-enabled autonomous cybersecurity response framework. The proposed model combines continuous vulnerability monitoring, behavioral analytics, machine learning-driven exploitation detection, and automated containment strategies to reduce detection latency and improve security posture. By leveraging supervised, unsupervised, and deep learning algorithms, the framework dynamically prioritizes vulnerabilities, detects real-time exploitation attempts, and initiates automated response actions based on risk confidence scores. The system aims to transition cybersecurity operations from reactive defense models to predictive and autonomous protection ecosystems.

Keywords: ***Artificial Intelligence, Vulnerability Assessment, Exploitation Detection, Risk Scoring.***

I. INTRODUCTION

Cybersecurity has evolved from perimeter-based defense strategies to complex, multi-layered protection models due to rapid advancements in digital transformation. Organizations now depend on cloud computing, virtualization, distributed networks, remote access systems, and application programming interfaces, which significantly expand the attack surface. Attackers exploit vulnerabilities in software, network configurations, authentication systems, and third-party integrations to gain unauthorized access and compromise critical data.

Traditional vulnerability scanners operate on periodic schedules and focus primarily on identifying known Common Vulnerabilities and Exposures (CVEs). While these tools are essential, they do not provide real-time exploitation detection or adaptive risk prioritization. Additionally, signature-based intrusion detection systems are ineffective against novel and obfuscated attack patterns. Manual incident response further delays mitigation efforts, allowing attackers to escalate privileges and move laterally across networks.

Artificial Intelligence (AI) and Machine Learning (ML) provide promising solutions for adaptive threat detection, behavioral modeling, predictive analytics, and automated response orchestration. Integrating intelligent vulnerability assessment with exploitation detection and autonomous response mechanisms can significantly enhance enterprise security resilience.

II. RELATED WORK

Previous research in vulnerability management has primarily focused on static risk scoring models such as CVSS-based prioritization. While these models provide standardized severity metrics, they do not account for real-time exploit trends, asset sensitivity, or threat actor behavior.

Behavior-based intrusion detection systems have introduced anomaly detection techniques, but many suffer from high false positive rates and limited contextual correlation with known vulnerabilities. Recent advancements in machine learning have demonstrated improved classification accuracy for intrusion detection; however, integration between vulnerability assessment, exploitation monitoring, and automated response remains limited in many existing frameworks.

This research bridges the gap by proposing a unified, AI-driven system that integrates continuous vulnerability monitoring, exploitation detection, and autonomous mitigation.

III. SYSTEM ARCHITECTURE

The proposed framework follows a multi-layered architecture designed to ensure scalability, flexibility, and real-time responsiveness.

The data collection layer aggregates telemetry from endpoints, servers, cloud platforms, firewalls, network sensors, application logs, and authentication systems. Endpoint agents monitor process creation, command-line execution, file modifications, privilege escalation events, and configuration changes. Network sensors capture metadata such as IP addresses, ports, protocols, packet size distribution, flow duration, and abnormal traffic patterns. Vulnerability databases and threat intelligence feeds provide contextual information regarding exploit availability and threat trends.

The preprocessing layer standardizes raw logs through normalization, timestamp alignment, duplicate removal, categorical encoding, and missing value handling. Cleaned and structured data is then passed to the feature engineering layer.

The feature engineering process extracts meaningful attributes such as abnormal process frequency, unusual parent-child process relationships, unauthorized registry changes, suspicious login attempts, geographic access anomalies, port scanning behavior, and exploit likelihood indicators. These features are transformed into numerical representations suitable for machine learning models.

The detection engine applies a hybrid learning approach. Supervised learning models such as Logistic Regression, Random Forest, and Support Vector Machine are trained on labeled datasets containing normal and malicious behavior samples. Random Forest enhances classification performance by combining multiple decision trees, while SVM is effective in high-dimensional anomaly detection scenarios.

Unsupervised models such as Isolation Forest identify rare behavioral anomalies without requiring labeled data. Deep learning architectures such as Long Short-Term Memory (LSTM) networks analyze sequential event patterns, enabling detection of multi-stage attack chains including reconnaissance, exploitation, privilege escalation, and data exfiltration.

The risk scoring module correlates vulnerability severity, exploit availability, asset criticality, and behavioral anomaly scores to generate a dynamic risk index. This contextualized scoring mechanism improves prioritization accuracy compared to static severity ratings.

The autonomous response module executes mitigation strategies based on predefined policies and model confidence levels.

IV. INTELLIGENT VULNERABILITY ASSESSMENT

The intelligent vulnerability assessment component continuously monitors system configurations and patch status rather than relying solely on scheduled scans. It tracks software version updates, open ports, exposed services, configuration drift, and cloud misconfigurations. By integrating threat intelligence feeds, the system evaluates whether a detected vulnerability is actively exploited in the wild.

Machine learning-based prioritization models analyze historical attack patterns, exploit kit references, and asset exposure levels to predict exploitation probability. This predictive prioritization enables security teams to allocate resources effectively and patch high-risk vulnerabilities promptly.

V. EXPLOITATION DETECTION MECHANISM

Exploitation detection focuses on identifying active misuse of vulnerabilities through behavioral analysis.

The system monitors abnormal process execution patterns, unauthorized privilege escalation attempts, suspicious command sequences, and unusual outbound network traffic.

Sequential pattern recognition using LSTM networks enables the system to detect subtle attack chains that may individually appear benign. For example, multiple failed login attempts followed by successful authentication from a new geographic location and sudden privilege escalation may indicate credential compromise.

Anomaly detection algorithms assign scores to events that deviate significantly from baseline behavior profiles. By combining anomaly scores with contextual vulnerability data, the system achieves higher detection accuracy while reducing false positives.

VI. AI-ENABLED AUTONOMOUS CYBERSECURITY RESPONSE

A critical innovation of the proposed framework is automated incident response. Traditional security operations require manual validation before containment, which increases response time. The AI-enabled response engine calculates a confidence score for each detected threat.

Low-risk events are logged for monitoring. Medium-risk events generate alerts for security analysts. High-risk events trigger automated containment actions such as terminating malicious processes, blocking suspicious IP addresses, disabling compromised accounts, isolating infected endpoints, enforcing password resets, and updating firewall rules.

This tiered response strategy minimizes operational disruption while ensuring rapid mitigation of critical threats. Continuous feedback from response outcomes is used to retrain machine learning models, enabling adaptive learning and performance improvement.

VII. IMPLEMENTATION FRAMEWORK

The system can be implemented using Python due to its extensive machine learning ecosystem. Libraries such as Scikit-learn, TensorFlow, and PyTorch support model development and deployment. Pandas and NumPy handle large-scale log analysis and data preprocessing.

Structured logs may be stored in relational databases such as PostgreSQL, while unstructured telemetry data can be stored in NoSQL databases like MongoDB. Containerization technologies such as Docker and orchestration platforms such as Kubernetes enable scalable deployment across distributed environments. Endpoint monitoring agents operate as lightweight background services to minimize system performance impact.

VIII. PERFORMANCE EVALUATION

The performance of the proposed system can be evaluated using classification metrics including accuracy, precision, recall, F1-score, and ROC-AUC. However, operational metrics such as Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR) provide a more realistic assessment of security effectiveness.

An optimized AI-driven system aims to achieve high recall to detect the majority of attacks while maintaining a low false positive rate to reduce alert fatigue. Experimental validation using benchmark intrusion detection datasets and enterprise telemetry data can demonstrate improved detection speed and prioritization accuracy compared to traditional systems.

IX. CHALLENGES AND LIMITATIONS

Despite its advantages, the proposed framework faces challenges including computational resource requirements, model drift, data privacy concerns, and adversarial machine learning threats. Attackers may attempt to evade detection by manipulating input data or injecting adversarial noise. Regular model retraining, secure data pipelines, encryption, and explainable AI mechanisms are necessary to maintain reliability and trust.

X. CONCLUSION

The increasing sophistication of cyber threats necessitates a shift from reactive security practices to proactive and autonomous defense strategies. The proposed Intelligent Vulnerability Assessment and Exploitation Detection System integrated with AI-enabled autonomous response provides a comprehensive solution for modern cybersecurity challenges. By combining continuous monitoring,

behavioral anomaly detection, predictive risk scoring, and automated mitigation, the framework enhances enterprise resilience and reduces response time.

As digital infrastructures continue to evolve, AI-driven autonomous cybersecurity systems will play a critical role in safeguarding critical assets, ensuring regulatory compliance, and maintaining organizational trust in an increasingly hostile cyber environment.

REFERENCES

[1] Mell, Peter, Scarfone, Karen, & Romanosky, Sasha (2006). *A Complete Guide to the Common Vulnerability Scoring System (CVSS)*.
[2] National Institute of Standards and Technology (2018). *Framework for Improving Critical Infrastructure Cybersecurity (NIST Cybersecurity Framework)*.
[3] National Institute of Standards and Technology (2012). *Guide for Conducting Risk Assessments (SP 800-30)*.
[4] Breiman, Leo (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
[5] Cortes, Corinna & Vapnik, Vladimir (1995). Support-Vector Networks. *Machine Learning*, 20, 273–297.
[6] Liu, Fei Tony, Ting, Kai Ming, & Zhou, Zhi-Hua (2008). Isolation Forest. *Proceedings of the IEEE International Conference on Data Mining (ICDM)*.
[7] Hochreiter, Sepp & Schmidhuber, Jürgen (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.
[8] Sommer, Robin & Paxson, Vern (2010). Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. *IEEE Symposium on Security and Privacy*.
[9] MITRE Corporation (2023). *MITRE ATT&CK Framework*.
[10] SANS Institute (2022). *Incident Handler's Handbook*.
[11] Open Web Application Security Project (2021). *OWASP Top 10: The Ten Most Critical Web Application Security Risks*.
[12] European Union Agency for Cybersecurity (2023). *ENISA Threat Landscape Report*.