International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2318 - 2321
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0342

AI Based Anomaly Detection in Endpoint Logs

Aromal Unni.A[1], A.L. Sriram[2], Akshay.R[3], Lakshmi.S[4]
[1,2,3] SRM University, Chennai, India
[4]Assistant Professor, SRM University, Chennai, India
Emails: *aa5193@srmist.edu.in[1], sa0507@srmist.edu.in[2], ar8252@srmist.edu.in[3] ,lakshmi9@srmist.edu.in[4]*

Abstract
Artificial Intelligence (AI) has achieved significant advancements in anomaly identification. Software systems frequently document critical runtime data in system logs for diagnostic purposes. The swift advancement of cybersecurity threats has rendered the detection of anomalies in endpoint logs essential for recognizing potential security breaches. Conventional rule-based detection techniques frequently inadequately identify complex and dynamic assault patterns. Explainable Artificial Intelligence (XAI) enhances the examination of system logs. It employs a white-box model to ensure transparency, comprehensibility, reliability, and dependability of Machine Learning (ML) and Deep Learning (DL) models. The methodology is corroborated using actual endpoint log datasets, exhibiting enhanced accuracy and diminished false positives relative to conventional techniques. The results underscore the capability of AI-driven anomaly detection to improve endpoint security through real-time threat intelligence and adaptive protection strategies.
Keywords: *Artificial Intelligence; Machine Learning; Deep Learning.*

1. Introduction
With the increasing sophistication of cyber threats, ensuring the security of endpoint devices has become crucial. Endpoint logs, which record system activities, user behavior, and network operations, play a vital role in identifying potential anomalies. However, traditional rule-based detection methods often fail to capture evolving and complex attack patterns, resulting in missed threats and false positives. This project proposes an AI-based anomaly detection system that analyzes endpoint logs in real time to improve detection accuracy and reduce errors. By leveraging Machine Learning (ML) and Explainable AI (XAI), the system enhances transparency, adaptability, and provides a more effective defense against emerging security risks. Traditional systems often rely on signature-based detection, which fails against unknown threats. Unsupervised learning models like Isolation Forest are effective but require fine-tuning for real-world applications. Hybrid models combining multiple techniques (e.g., behavior-based and statistical approaches) show better adaptability [1-2]

1.1.The Need for AI in Endpoint Security
As organizations increasingly rely on digital infrastructure, cyber threats continue to evolve in complexity and scale. Attackers exploit zero-day vulnerabilities, execute advanced persistent threats (APTs), and utilize sophisticated evasion techniques that can bypass conventional security solutions. Endpoint logs provide a vast amount of data that, if analyzed effectively, can uncover subtle patterns indicative of malicious activity. However, manually sifting through these logs is impractical, and conventional rule-based methods often struggle to distinguish between normal fluctuations in user behavior and actual threats. This is where AI-driven anomaly detection becomes essential, enabling proactive threat identification rather than reactive mitigation. [3-5]

1.2.Challenges with Traditional Detection Approaches
Each Traditional security solutions, such as signature-based detection (e.g., antivirus software and intrusion detection systems), rely on predefined attack patterns. While effective against known threats, these methods fail against novel or polymorphic attacks, where malicious actors continuously alter their tactics to evade detection. Moreover, rule-based systems generate a high number of false positives, overwhelming security

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2318 - 2321
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0342

teams with unnecessary alerts. This leads to alert fatigue, causing critical threats to be overlooked.

Another limitation is the lack of contextual awareness in traditional methods. Attackers often mimic legitimate user behavior, making it difficult to differentiate between a compromised insider threat and a genuine employee performing routine tasks. Behavior-based detection combined with statistical modeling offers a more dynamic approach by adapting to evolving attack patterns. [6]

1.3.Leveraging Machine Learning for Anomaly Detection

The Machine Learning (ML) has transformed cybersecurity by enabling automated pattern recognition and real-time anomaly detection. Instead of relying on static rules, ML models can learn from historical data, identify complex behavioral trends, and flag deviations that indicate potential threats. Unsupervised learning techniques, such as Isolation Forest, effectively isolate outliers without requiring labeled attack data. This approach is particularly valuable in cybersecurity, where new attack techniques emerge frequently. However, unsupervised models alone may lack interpretability and struggle with high-dimensional data. To address this, dimensionality reduction techniques like Principal Component Analysis (PCA) are employed to simplify complex feature sets while retaining crucial information. PCA allows security teams to visualize threat patterns, making it easier to distinguish between genuine anomalies and benign variations in user behavior. [7]

1.4.Explainable AI (XAI) for Transparent Security Decisions

One of the challenges of AI-driven security systems is the black-box nature of many ML models. Security analysts require clear explanations for why an action is flagged as an anomaly to make informed decisions. This project integrates Explainable AI (XAI) techniques, such as feature importance analysis and decision boundary visualization, to enhance transparency. By understanding why an anomaly is detected, security teams can validate threats more effectively, reducing investigation time and improving response efficiency.

1.5.Hybrid Models for Enhanced Accuracy

While Isolation Forest and PCA are effective at anomaly detection, they work best when combined with complementary techniques. A hybrid approach that integrates:

- Behavior-based analysis (detecting deviations in login patterns, file access, or network connections).
- Statistical modeling (examining frequency, duration, and correlation of events).
- Context-aware detection (considering user role, device type, and historical behavior).

This ensures higher detection accuracy while reducing false positives. Additionally, adaptive learning techniques can refine model performance over time by incorporating feedback loops from security analysts. [8]

1.6.Real-World Applications and Benefits

- Organizations across various industries—financial services, healthcare, and government sectors—can benefit from AI-powered anomaly detection in endpoint logs. This system can:
- Detect lateral movement within networks, preventing internal threats.
- Enhance insider threat detection by spotting unusual user activities.
- Reduce reliance on manual log analysis, improving operational efficiency.

By automating threat detection, organizations can strengthen their cybersecurity posture, reduce response times, and protect critical assets against emerging threats. [9]

2. Method

The proposed system uses a hybrid approach, combining behavior-based anomaly detection with context-aware AI. Key components include real-time data collection, feature engineering, dimensionality reduction, anomaly detection, and threat intelligence integration. The system integrates multiple detection techniques to enhance security coverage. The model continuously updates itself based on new threats and log patterns. The combination of feature reduction (PCA) and real-time monitoring ensures minimal latency. The architecture is designed to work across various

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2318 - 2321
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0342

infrastructures, including enterprise networks and IoT environments. [10]

3. Figures

The architecture consists of several interconnected stages. First, logs from endpoint devices and network packets are captured using Scapy. These logs undergo preprocessing where timestamps are cleaned and protocols are standardized to maintain consistency. Feature engineering involves extracting crucial attributes such as source IP, destination IP, protocol types, and timestamps, while also creating derived features like time buckets and behavioral patterns to provide a more comprehensive dataset. (Figure 1)

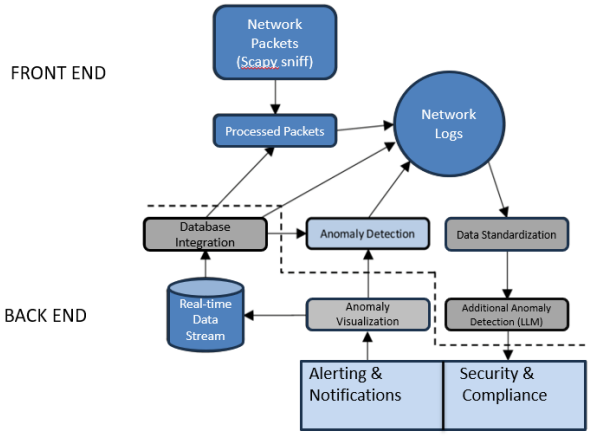

Figure 1 Architecture Diagram of the Anomaly Detection

4. Results and Discussion

4.1.Results

We visualized the detected anomalies using scatter plots and time-series charts. These visual aids highlight the clustering of anomalies, enabling better interpretation of anomalous patterns. The system also provides real-time alerts and graphical dashboards to monitor network activity efficiently. [11]

4.2.Discussion

The plot is titled "Anomaly Detection using Isolation Forest", indicating that it visualizes anomalies detected by the Isolation Forest algorithm. X-axis: Represents "PCA Component 1" (first principal component). Y-axis: Represents "PCA Component 2" (second principal component). These components come from Principal Component Analysis (PCA), which is used to reduce the dimensionality of the dataset while retaining most of the important information. The plot has two sets of points, shown in red and blue: Blue points likely represent normal data (inliers). Red points represent anomalies (outliers) detected by the Isolation Forest model. The Isolation Forest model is detecting patterns in the dataset. Points that deviate significantly from normal behavior (as captured by PCA) are flagged as anomalies (in red). The clustering of blue points suggests areas of "normal" activity, while isolated red points indicate "abnormal" or suspicious patterns. (Figure 2,3) [13-14]

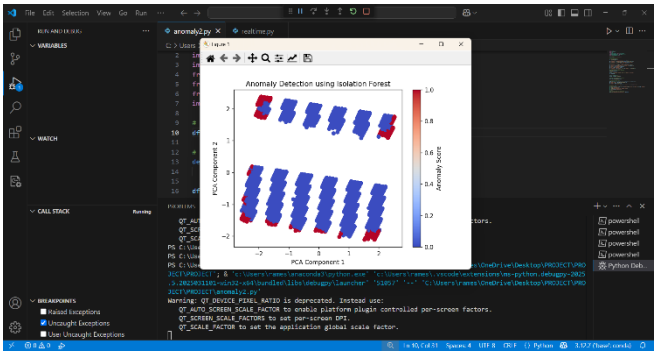

Figure 2 Graph of Anomaly has Detected

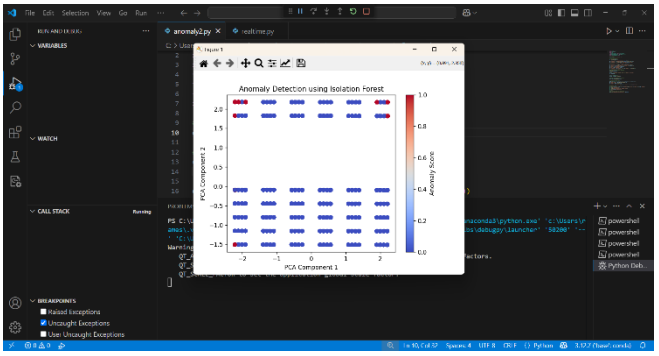

Figure 3 Graph of the Final Result

Conclusion

The proposed AI-based anomaly detection system effectively identifies irregular patterns in endpoint logs using the Isolation Forest algorithm. By leveraging features such as time difference, day of the week, and hourly access patterns, the system can detect deviations from normal behavior. The application of Principal Component Analysis (PCA) reduces the dimensionality of the dataset, enhancing computational efficiency and enabling clear

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2318 - 2321
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0342

visualization of normal and anomalous data points. The model, with a contamination rate of 5%, successfully flags potential anomalies, providing valuable insights for identifying suspicious activities like unauthorized access or unusual traffic patterns. This system is adaptable for real-time monitoring and can be extended to other domains requiring anomaly detection. Future improvements could include dynamic contamination adjustment and exploring advanced models like Autoencoders for better accuracy and performance. The resulting dataset with anomaly labels can be further analyzed for threat intelligence and proactive security measures. [15]

Acknowledgements

Place Acknowledgments, including information on the source of any financial support received for the work being published. Place Acknowledgments, including information on the source of any financial support received for the work being published.

References

[1]. L Zhao, C., Li, C., Feng, S., Su, N., & Li, W. A Spectral-Spatial Anomaly Target Detection Method Based on Fractional Fourier Transform and Saliency Weighted Collaborative Representation for Hyperspectral Images. 2020.

[2]. Zhang, L., Wang, Z., Gao, Q., Li, D., Liu, J., Wang, H., Yu, X., & Wang, Y. Insulator Anomaly Detection Method Based on Few-Shot Learning. 2021.

[3]. Sarda, K., Acernese, A., Noli, V., Manfredi, L., Greco, L., Glielmo, L., & Del Vecchio, C. A Multi-Step Anomaly Detection Strategy Based on Robust Distances for the Steel Industry. 2021

[4]. Yan, N., Zhu, L., Yang, H., Li, N., & Zhang, X. Online Yarn Breakage Detection: A Reflection-Based Anomaly Detection Method. 2021.

[5]. LAbdelrahman, O., & Keikhosrokiani, P. Assembly Line Anomaly Detection and Root Cause Analysis Using Machine Learning. 2020.

[6]. Hashemi, Tian, J., Ding, W., Wu, C., & Nam, K. W. A Generalized Approach for Anomaly Detection from the Internet of Moving Things. 2019.

[7]. Bao,L.,etal. Execution anomaly detection in large-scale systems through console log analysis 2018

[8]. Zhao, Z., Xu, C., & Li, B. A LSTM-Based Anomaly Detection Model for Log Analysis. Journal of Signal Processing Systems, 2021.

[9]. Guo, H., Yuan, S., & Wu, X. LogBERT: Log Anomaly Detection via BERT. arXiv preprint arXiv:2103.04475, 2021.

[10]. Bulusu, S., Kailkhura, B., Li, B., Varshney, P. K., & Song, D. Anomalous Example Detection in Deep Learning: A Survey. arXiv preprint arXiv:2003.06979, 2020.

[11]. Chalapathy, R., & Chawla, S. Deep Learning for Anomaly Detection: A Survey. preprint arXiv:1901.03407, 2019.

[12]. DeMedeiros, K., Hendawi, A., & Alvarez, M. A Survey of AI-Based Anomaly Detection in IoT and Sensor Networks. Sensors, 2023.

[13]. Kim, K.-S., Oh, S. J., Cho, H. B., & Myung, H. One-Class Classifier for Chest X-Ray Anomaly Detection via Contrastive Patch-Based Percentile. 2021.

[14]. Castellani, A., Del Pinto, L., Iacobone, F. D., & Ruggeri, A. Real-World Anomaly Detection by Using Digital Twin Systems and Weakly Supervised Learning. 2021.

[15]. Fang, X., Zhang, Q., Wang, W., & Wang, J. Sewer Pipeline Fault Identification Using Anomaly Detection Algorithms on Video Sequences. 2021.