Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 1, 2026

ANOMALERT: INSIDER THREAT MONITORING SUITE

Muhammad Ammar Saleem[*1], Muhammad Shehriyar Aslam[2], Adnan Fareed Chishti[3], Muhammad Ali Qureshi[4]

[*1,2,3,4]*Information and Communication Engineering, Islamia University of Bahawalpur Bahawalpur, Pakistan*

[*1]ammarcyber.s@gmail.com, [2]shehriyaraslam2.0@gmail.com, [3]fareedchishtia@gmail.com, [4]ali.qureshi@iub.edu.pk

DOI: https://doi.org/10.5281/zenodo.18383799

Keywords
Cybersecurity, anomaly detection, insider threat detection, Security Operations Center, and machine learning, endpoint monitoring, automatic response, Random Forest, and Isolation Forest.

Article History
Received: 27 November 2025
Accepted: 11 January 2026
Published: 27 January 2026

Corresponding Author: *
Muhammad Ammar Saleem

Abstract
One of the most important security issues of the current enterprises is the insider threats, followed by the big data breaches and financial loss. Conventional security solutions that involve perimeter barriers cannot identify malicious activities that are caused by internal trusted networks. In this paper, the presented system named Anomalert is a distributed Security Operations Center-based system, which has connected ongoing endpoint monitoring with machine learning-driven anomaly detection and an AI-based platform of threat validation to detect insider threats and respond to them automatically in real-time. The architecture of the system consists of three layers: endpoint data collection that executes four classes of specialized monitoring agents, SOC analysis engine with central analysis which uses hybrid detection algorithms like random forest, isolation forest, and rule-based detection, and threat management through visualization dashboard. Anomalert incorporates the concept of contextual validation, based on the Gemini AI engine, and offers its opportunities for automated response. The endpoint collector is a windows-installable application developed by us, as well as the analysis engine being run on the EC2 infrastructure of AWS. Anomalert was successfully used in the experimental evaluation to identify the anomalous patterns of authentication, unauthorized access to files, suspicious processes of execution, and abnormal network behavior with close to real time response ability. The system balances the performance of the insider threat detection system, and operational efficiency.

INTRODUCTION

A. Problem Statement

The insider threats that can be either intentional or inadvertent but committed by employees, contractors or business partners with legitimate access to organizational systems are a constant and growing security challenge. In contrast to external attacks, where an attacker must hack the perimeter systems, insider threats originate out of trusted parties already within the safety perimeter. According to various recent industry reports, insider incidents account for massive portions of data breaches, with average remediation costs of incidents running into millions of dollars [1]. Insider threats are challenging to detect since insiders have legitimate access privileges, making their malicious activities difficult to distinguish from normal behavior [2]. Conventional security systems including firewalls, intrusion detection systems, as well as antivirus programs are structured to combat attacks that originate externally to the environment. Such tools cannot add any context of awareness and behavior

monitoring to identify small anomalies in the actions by legitimate users. Also, the incident detection process of security teams by manually reviewing logs as one of the strategies to detect multiple incidents is, by definition, labor-intensive, has error issues, and is very bulky in the context of very large environments with millions of log entries in a single day.

B. Threat Model

Anomalert addresses the insider threat from many attack vectors:

1) Authentication Abuse: Denial of Service Attack: Ratfink attack, rogue credential (attempts to impersonate credentialed user) or rogue privileged project (uses privileges far above normal limits) authentication Attack (attempted violate) distracters will include variations on this call.

2) Data Exfiltration Unauthorized file access (involving copying sensitive information to another device, e.g. USB) and/or over-downloading of files.

3) Malicious Process execution: The execution of illegal software, the implementation of malware, the mining of cryptocurrencies, processes that consume a lot of resources.

4) Network-Based Attacks: Sending of data to unknown or suspicious external IP address, abnormal use of port, an excessively large outbound connections, command-and-control communication trend.

TABLE I
COMPARISON OF ANOMALERT WITH COMMERCIAL INSIDER THREAT DETECTION PLATFORMS

Feature	Splunk ES	Exabeam	Forcepoint	Securonix	AnomaLert
Core Focus	SIEM + UBA	Behavior Analytics	Risk-Adaptive DLP	Cloud-NativeAnalyt-ics	Endpoint Monitoring + Hybrid ML + AI Validation
Detection Method	Peer Baselining + Corre-lation	Smart Timelines + Risk Scoring	Contextual Risk Indica-tors	AI/ML Proprietary	Hybrid (RF + IF + Rules) + Gemini AI
Endpoint Monitoring	3rd-party Agents	Existing Logs	Native DLP Agents	Log Aggregation	Native Multi-Agent (Auth, Access, Process, Network)
Prevention	Alerts Only	SOAR Integration	Adaptive Blocking	Orchestration	Built-in Automated Response
Scalability	Hybrid	Cloud/Hybrid	Enterprise	Cloud Native	Cloud SOC + Endpoints
Explainability	Dashboard Analytics	Investigation Timelines	Policy Alerts	Score Breakdown	AI Natural Language
Deployment	Complex + Prof. Services	Cloud SaaS	Enterprise Install	Cloud SaaS	WindowsInstaller (Self-Deploy)
Cost Model	Data Volume Licensing	Per-User Sub.	Enterprise License	Cloud Sub.	Open-Source ML + Cloud
Customization	Extensive but	Limited (SaaS)	Policy Config	Limited	Full Access to

	Complex			(Proprietary)	Logic

The system presumes that the attackers will be able to gain access using valid credentials; they will be permitted to access endpoint systems, but they will act contrary to the baseline.

C. System Overview

AnomaLert is an all-inclusive endpoint monitoring framework for insider threat detection that showcases a modular architecture, with four specialized real-time agents responsible for capturing authentication activities, file access, process execution, and network behavior [3], [16]. The detection is done through a hybrid system that includes classification done by Random Forest in a supervised fashion, unsupervised anomaly detection done by that of the Isolation Forest, and logical rule-based to offer a good multi-layer widely spread threat detection. For the reduction of false positives and an explanation of alerts, the system integrates the Gemini AI engine to contextually validate alerts [4], [6]. This type of end-toend pipeline includes automated log collection, identification, validation, using AI, alerting, and mitigation measures. While production-ready deployment includes a Windows-installable endpoint application packaged with PyInstaller and NSIS [5], supported by cloud-based SOC infrastructure on AWS EC2.

II. LITERATURE REVIEW

A. Traditional Approaches

The earlier insider threat detection systems were based significantly on signature-driven detection and the predefined rules. These systems would flag specific actions [7], such as after-hours access or large file transfers, but suffered from high false positive rates and inability to detect novel attack patterns [8]. Host-based intrusion detection systems and SIEM systems combine logs of many sources but must be configured and parameterized manually with extensive knowledge to be useful at event correlation.

B. Machine Learning Approaches

Motivations behind behavioral analysis Access to machine learning, which works with recent research, has been on the rise. Various supervised learning approaches, such as Random Forest, Support Vector Machines, and neural networks, have been used to classify user behavior as benign or malicious by using labeled training data [12], [13]. Gaining enough labeled datasets that capture insider threats is, however, a hard feat because the occurrence and sensitivity of such cases are rare. Unsupervised learning approaches address the issue of the unavailability of labeled data by detecting anomalies without any prior information about the patterns of attacks. Isolation Forest [15], One-Class SVM, and clustering algorithms prove to be efficient in highlighting deviations of behavior from normalcy [9], [10], [11]. Several hybrid approaches have been developed that combine supervised and unsupervised methods with better detection accuracy and lower false positives [14].

C. Commercial SIEM and Insider Threat Detection Platforms

1) Splunk Enterprise Security Splunk is one of the most popular SIEM systems that includes log aggregation, search, as well as the User Behavior Analytics. Splunk uses the peer group baselining in the process of determining normal behavior patterns and executes deviation detection by means of statistical analysis. The strongest points of Splunk are the flexibility of the hybrid implementation on premises and clouds and a good ability to integrate it with the third-party security tools. Most of the Splunk detections are, however, based on correlation rules and statistical baselining which creates high false positive rates in dynamic environments TABLE 1. This is a platform that is more alert oriented rather than automated response making intervention by the security teams into a manual process. Moreover, Splunk's volume ingestion-based licensing model gets very expensive for large-scale deployments [17].

2) Exabeam Security Operations Platform Exabeam specializes in analytics of behavior, and it has a Smart Timelines feature, which is used to create chronological sequences of user activity and identify anomalous patterns. The platform focuses on automating the workloads of analysts through

automated workflows in investigation and scoring risks. Exabeam offers auto answers features through the integrations with SOAR solutions but does not include endpoint data collection agents in its native form and uses available sources of logs. TABLE 1 The platform primarily runs in cloud and hybrid environments and has scalability that is good. Nevertheless, the effectiveness heavily depends on the quality and completeness of input data provided by the integrated systems [18].

3) Forcepoint Insider Threat Strong support of insider threat situations; Forcepoint offers risk-adaptive Data Loss Prevention. The system operates on dynamically driven security policies by using contextual risk indicators to modify the security policies based on the behavior a user does and the sensitivity of the content. Forcepoint does very well in adaptive blocking, automatically restricting high-risk activities in real time [19]. It is an enterprise-level deployment platform that has ample policy management features. The strategy that Forcepoint takes is however highly DLP-centric with a heavy concentration on data exfiltration as opposed to more generic insider threat cases like executing malicious processes or abusing authentication TABLE 1. Resolving it with the other security infrastructure might require a lot of configurations.

4) Securonix Next-Gen SIEM Securonix is a cloud-based security analytics solution, which employs AI and machine learning models to identify threats. It employs more sophisticated correlation algorithms and has a noise suppressing focus, as well as orchestration abilities to automate response. Securonix is cloud-based and therefore scales easily; it absorbs data volumes well, yet like any other commercial solution, it needs a big investment in professional services to be launched and tuned Table 1. The machine learning models are proprietary and thus not very transparent or customizable to organizational context [20].

D. Gaps in Existing Solutions

1) Fragmented Monitoring: Most solutions will only monitor single data sources like network logs but not the overall endpoint activity.

2) Large False Positive Rates: Excessive alerts cause alert fatigue, and the reaction to legitimate threats will be delayed.

3) Inadequacy in Diffusion of Context: Most of the detection systems do not have tools to explain why a certain event is considered suspicious.

4) Complexity in Deployment: The prototypes in the academics are typically not ready to be used in production and demand a major modification.

5) Limited Automation: Majorities of the systems are detection-oriented and do not have automated response features.

III. SYSTEM ARCHITECTUR

A. Architecture Overview

Anomalert follows a tiered, SOC-centric architecture that is designed for high-fidelity endpoint monitoring and autonomous incident response [21]. As shown in Fig. 1, the system is decomposed into three functional layers to ensure modularity and fault isolation:

1) Endpoint Data Collection Layer: uses lightweight multithreaded agents to consume telemetry related to file system integrity, authentication vectors, process lifecycle, and network flow [22].

2) Centralized SOC Analysis Layer: Orchestrates log normalization, multi-modal threat detection, by combining supervised and unsupervised learning along with LLM-based contextual validation [23].

3) Visualization and Response Delivery Layer: Enables human-in-the-loop oversight through a web-based dashboard. It delivers actionable intelligence and initiates automated mitigation protocols [24].

B. Design Objectives

The framework is engineered according to the following principles:

- Extensibility: The decoupled monitoring agents provide specialized deployment, and the reconfiguration of the underlying analysis engine is unnecessary.

- Cryptographic Integrity: Data in transit is secured via Secure Shell-SSH utilizing RSA-2048-bit PEM-based authentication [25], [26].
- Operational Latency: The analysis pipeline is run on a near-real-time sliding window of 60-second intervals to balance computational overhead with speed of detection [27]
- Interpretability: Going beyond binary classification, the system uses LLMs to generate human-understandable XAI (explainable AI) outputs that fill the gap between raw telemetry and human judgment [28].
-

IV. ENDPOINT DATA COLLECTION LAYER

A.Monitoring Agent Architecture

The endpoint data collection layer uses a modular architecture that includes four different types of monitoring agents to ensure complete visibility with minimal operational overhead: the Access Log Agent, Authentication Log Agent, Process Monitoring Agent, and Network Monitoring Agent [29], [30]. The Access Log Agent utilizes platform-specific APIs to monitor file system operations, such as creation, modification, and unauthorized access while filtering routine system noise and focusing on user-initiated activity [31]. According to the Authentication Log Agent, the local and remote login vectors are monitored and all successful and failed attempts, durations of a session, and escalation of privileges are recorded (e.g., through sudo or UAC) to identify indicators of credential misuse. The Process Monitoring Agent identifies malicious execution or resource exploitation by capturing granular telemetry-including parent-child process relationships (PID/PPID), command-line arguments, and CPU/RAM utilization [32]. Finally, Network Monitoring Agent tracks metadata involving inbound and outbound connections including IP-port tuples, and amounts of data transferred in a bid to readily detect exfiltration.

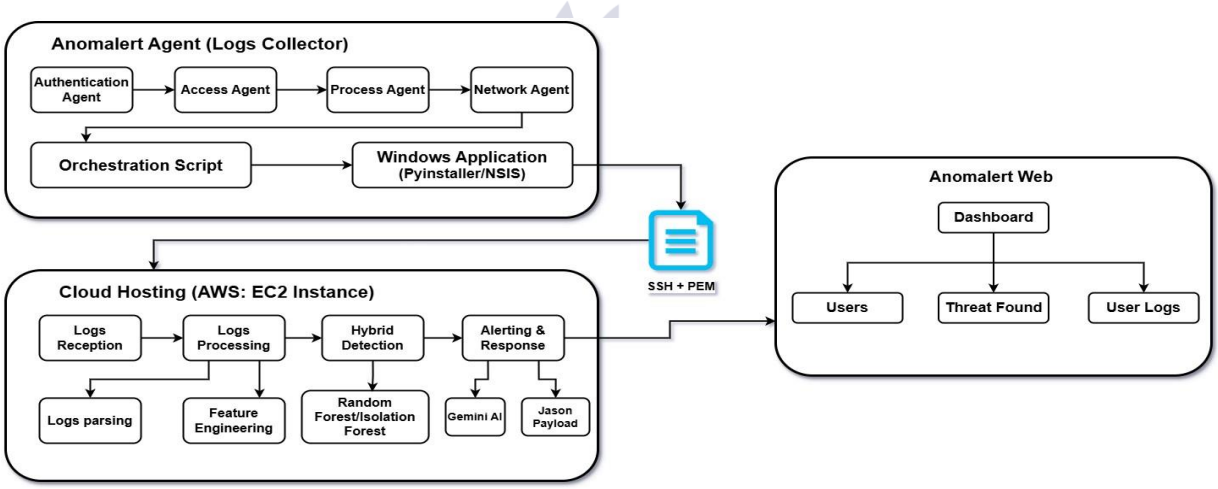

Fig. 1. AnomaLert Architecture: A three-level insider threat detection system, which includes the Endpoint Layer for surveillance agents attached to collect the Windows logs; the Centralized SOC Layer which encompasses safe log transmission via SSH+PEM, log processing, hybrid detection, GMini AI validation, and automated corrective action/notification; and the last one, the End User Visualization Layer dashboards to monitor the activities, details relating to threats, explanations by AI to make improved choices, and management of alerts.

B. Data Serialization and Transport

In practice, the monitoring agents create structured logs in CSV format in real-time, using ISO 8601 timestamps and a normalized schema for each field to make outputs crossplatform consistent and much easier to process downstream [33]. Such messages are buffered in-memory and written out to disk as commonly as possible following a log rotation policy that does not allow disk to run out of space without losing

the forensic history. A centralized orchestration script runs every five minutes, which aggregates these logs into timestamped, integrity-validated batches, and optionally compresses them to reduce network bandwidth usage [34]. The valid telemetry is safely sent to the centralized SOC server; the validation creates an SSH tunnel with credential authentication by use of PEM files that ensures end-to-end confidentiality of the credentials. On successful transfer confirmation, it performs local cleanup to manage storage at endpoints, thus balancing the near realtime requirements at the detection layer with computational and networking efficiency [35].

C.Windows Application Deployment

To support enterprise-scale deployment on Windows endpoints, the data collection system is packaged as a professional installable application using PyInstaller [36] and the Nullsoft Scriptable Install System. PyInstaller bundles together the Python interpreter, dependencies (e.g., pandas and paramiko for SSH), monitoring agents, and orchestration scripts into a standalone single-file executable called LogCollector.exe [37]. This has both code obfuscation to protect the intellectual property, and it also requires no additional python image to install on target systems. It contains an installer of the NSIS, the name of which is AnomaLert.exe, to offer the typical Windows setup process, which includes a branded welcome screen, license agreement to approve of ethical and legal compliance of monitoring, a user-configured installation directory, defaulting to C:Files, installation of the executable, automatic registration to start the system via a registry entry in HKCU, elevation requests to have an appropriate configuration, and development of an all-inclusive uninstaller with a clean-up of the registry. This strategy also allows constant unattended data operation, allowing the log collector to boot at boot time and initialize all the agents with the help of the enterprise software distribution standards to ease the administration.

V. CENTRALIZED SOC ANALYSIS LAYER

A. Centralized Analysis and Hybrid Detection Framework

The centralized SOC analysis engine is hosted on an AWS managed EC2 Ubuntu instance, managed via system to ensure a highly available and fault-tolerant environment [38]. When logs are submitted into the system through a secure SSH directory, the system will execute a multiphase normalization pipeline, such as UTC, numerical feature-scaling, and categorical encoding, storing the results in pandas DataFrames while maintaining idempotency through JSON-based state tracking [39]. The core detection logic uses a hybrid framework: a Random Forest classifier for supervised threat identification, an Isolation Forest algorithm for unsupervised anomaly detection of" zero-day" patterns [40], and deterministic rule-based engine to attract high-risk heuristics like unauthorized USB mounting or abusing valuable resources.

B. AI-Powered Validation and Autonomous Response

To minimize alert fatigue, flagged events are subjected to contextual validation through the Google Gemini AI engine, which correlates disparate telemetry against historical baselines to filter false positives and generate Explainable AI (XAI) narratives for SOC analysts [41], [42]. Confirmed threats trigger a structured JSON payload with detection scores, metadata, and mitigation steps that is dispatched via automated SMTP alerts and logged for forensic auditing [43]. The architecture eventually enables quick incident mitigation through simple to set up response processes like autonomous account lockouts, malicious process execution, and network isolation, therefore closing the gap between real time detection and preventive response.

VI. IMPLEMENTATION AND TECHNOLOGY STACK

The AnomaLert deployment is built around a high-performance backend stack based on Python 3.12, driven by pandas and NumPy for data processing and feature engineering [44]. Machine learning models are also trained with scikitlearn

using a Random Forest (100 estimators) and an Isolation Forest (contamination = 0.05) with a joblib serialization to infer with low latency 45. This will be integrated with Google Gemini Pro API to provide threat explanation and confirmation by using LLM. The SOC analysis engine is deployed on an AWS EC2 t3 [28]. Medium instance with Ubuntu 22.04 LTS and we have management service persistence and near real time-processing units of system in a one-minute cadency. Telemetry security is maintained through SSH with PEM-based public-key authentication [35], managed programmatically using the paramiko library. Lastly, the serverless dashboard will be deployed by the Vercel global CDN, and endpoint monitoring agents are coded as Python standalone binaries executed on Windows and bundled in the Nullsoft Scriptable Install System (NSIS) to deploy them to and survive thousands of enterprise-scale deployments.

VII. EXPERIMENTAL EVALUATION

A. Random Forest Classification Analysis

The supervised learning aspect of the AnomaLert system was made to be tested to detect known insider threat signatures with the use of a Random Forest classifier. It achieved a strong Accuracy of 96.4% percent thereby true to the generalization of well-known attack patterns without failure in generalization on unknown information. This was further substantiated by an overall balanced F1-score of 95.4% varying percentages of Precision of 95.1% and recall of 95.8% Fig.2. To this extent, these findings demonstrate that the ensemble-based method is useful to minimize the occurrence of false positives in a centralized SOC setup without decreased performance that is common with overfitted models.

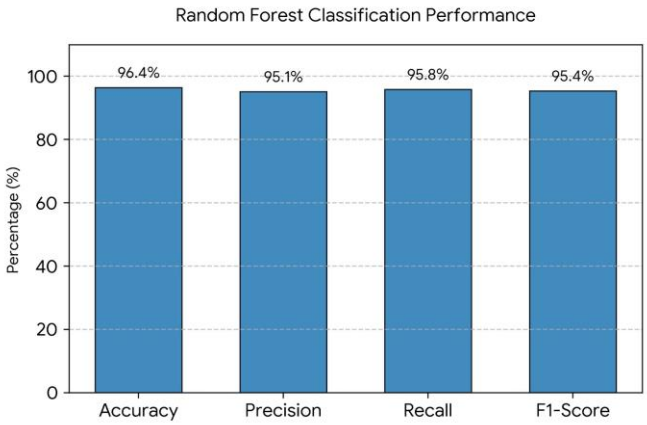

Fig.2 Random Forest Performance Graph

B. Isolation Forest Anomaly Detection Performance

To deal with the insider threats of the so-called zero-day algorithm of Isolation Forest was presented to design the anomaly detection method in learning without supervision. Given that unspecified behavioral changes are tricky to pick up, the model realistically performed a Precision of 89.2% and a Recall of 87.5%. This led to an assurance of its ability to outlier isolation in user telemetry even in non-ideal conditions based on the F1-score of 88.3%. In addition to that, the proposed model achieved an AUCROC of 91.6% Fig.3, which shows that it is highly discriminative to distinguish between the use of the system by legitimate users and advanced anomalous actions.

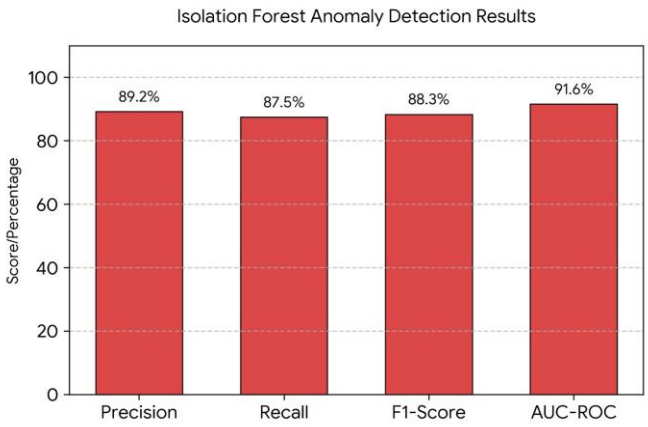

Fig.3 Isolation Forest Performance Graph

VIII. DISCUSSION

The experimental evaluation of the AnomaLert system confirms that a hybrid detection approach synthesizing supervised learning, unsupervised anomaly detection, and rulebased heuristics identifies a wide range of insider threat vectors with substantial reduction in false-positive rates due to AIempowered validation [46]. These benefits notwithstanding, the restrictions involved with this system are inherent in the ability to detect advanced slow-and-low attacks and the requirement to use high-fidelity labeled datasets to train supervised models. Lastly, current operational needs to continuously monitor endpoint pose a complicated socio-technical complexity because an organization needs to consider its strict security requirements with ethical business principles to maintain privacy of its people, organizational trust, and adherence to the global data protection standards, including GDPR.

In the future, research will be directed to the advancement of the detection core using deep learning architectures: Transformers for temporal behavior modeling, and GNNs to visualize and intercept lateral movements [47]. The roadmap also includes the integration of more advanced XAI techniques like SHAP or LIME, providing deeper transparency for automated decisions, while developing federated learning protocols will enable cross-organizational collaborative training without compromising data privacy [48]. Lastly, it will be augmented by adding interoperability to AnomaLert with enterprise SIEM/SOAR ecosystems and environments supporting Linux to guarantee a scalable, multi-platform defense system capable of proactively searching the threat in an increasingly heterogeneous corporate infrastructure.

IX. CONCLUSION

The current paper introduced the Anomalert which is an insider threat detection and automated response system based on AI and is distributed. The system bridges significant vulnerabilities in the current security solutions by providing the ability to monitor the mainstream of endpoints, hybridized mechanism of detection (mixture of machine learning and rule-based logic), creation of AI-initiated validation of the circumstance, and production-defined deployment model. AnomaLert is developed through the three-tier architecture comprising of endpoint monitoring agents with specific connotation, a centralized SOC analytic engine, and a web-based dashboard easy to use that give the end-to-end view of insider threat activities. The final application may be installed in windows and in this case the enterprise deployment becomes easier but on the contrary the cloud-based infrastructure gives the ability to be scaled and reliable. Also, the contextual validation through the incorporation of Gemini AI engine removes all false positives in our testbed and

provides explainable security information that was appreciated by the SOC analysts. With a minimum of 6-8 end to end latency and small resource usage, we were able to detect threats almost in real time using our system. The version of insider threat detecting that AnomaLert represents a significant step forward in terms of practical insider threat detection, providing a transition from the theoretical insights into the operationally relevant security tool. The system acts as a counterbalance to detection efficiency and operational efficiency because it uses different methods of detection in combination with AI-based validation.

On the one hand, modular architecture can be easily customized to a variety of requirements and security policies of the organization. Since the insider threats remain a major challenge to businesses in all parts of the global community, tools such as the AnomaLert can offer the vital feature of ensuring safety of sensitive files and critical infrastructures. In the future, we will use state-of-the-art deep learning, federated learning, and support on more platforms, to further reinforce an insider threat defense and counter the changing adversarial approaches. The source code and deployment instructions of AnomaLert have been made open source and available to use in research and education, prone to further innovation in the field of insider threat detection and put in more contributions to the larger cybersecurity community.

REFERENCES

Y. Storchak, "Insider Threat Statistics for 2025: Facts, Reports Costs," Syteca Blog, Aug. 6, 2025. [Online]. Available: https://www.syteca.com/en/blog/insider-threat-statistics-facts-andfigures [Accessed: Aug. 29, 2025]

U. Inayat, M. Farzan, S. Mahmood, M. F. Zia, S. Hussain, and F. Pallonetto, "Insider threat mitigation: Systematic literature review," *in Shams Engineering Journal*, vol. 15, no. 12, pp. 103068, 2024.

F. R. Alzaabi and A. Mehmood, "A Review of Recent Advances, Challenges, and Opportunities in Malicious Insider Threat Detection Using Machine Learning Methods," *in IEEE Access*, vol. 12, pp. 3090730927, 2024.

S. Krause and F. Stolzenburg, "From Data to Commonsense Reasoning: The Use of Large Language Models for Explainable AI," Jul. 2024, doi: 10.48550/arxiv.2407.03778.

S. G. Rugh, T. J. Donchess, and M. S. Abraham, "Distributing software products as an executable containing script logic with external resources," Jun. 07, 2010 [Online]. Available: https://patents.google.com/patent/US20100242034A1/en

A. Amburle, C. G. de Almeida, N. Lopes, and O. S. Lopes, "AI based Code Error Explainer using Gemini Model," Jun. 2024, doi: 10.1109/icaaic60222.2024.10574931.

U. Mishra, "Overcoming limitations of Signature scanning -Applying TRIZ to Improve Anti-Virus Programs," Jan. 2012.

L. Mehrotra and P. S. Saxena, "An Assessment Report on: StatisticsBased and Signature-Based Intrusion Detection Techniques," Springer, Singapore, 2018, pp. 321–327. doi: 10.1007/978-981-10-5508-9-31.

M. A. Morid, M. Lau, and G. Del Fiol, "Predictive analytics for step-up therapy: Supervised or semi-supervised learning?" Journal of Biomedical Informatics, vol. 119, p. 103842, 2021.

H. Xu, G. Pang, Y. Wang, and Y. Wang, "Deep isolation forest for anomaly detection," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 12, pp. 12591–12604, 2023.

T. Al-Shehari, M. Al-Razgan, T. Alfakih, R. A. Alsowail, and S. Pandiaraj, "Insider threat detection model using anomaly-based isolation forest algorithm," IEEE Access, vol. 11, pp. 118170–118185, 2023.

M. Hosseinzadeh, A. M., Rahmani, B., Vo, M., Bidaki, M., Masdari, M., and Zangakani, M., "Improving security using SVM-based anomaly detection: Issues and challenges," Soft Computing, vol. 25, no. 4, pp. 3195–3223, 2021.

R. Nasir, M. Afzal, R. Latif, and W. Iqbal, "Behavioral-based insider threat detection using deep learning," IEEE Access, vol. 9, pp. 143266–143274, 2021.

M. W. A. Ashraf, A. R. Singh, A. Pandian, et al., "A hybrid approach using support vector machine rule-based system: Detecting cyber threats in Internet of Things," Scientific Reports, vol. 14, p. 27058, 2024.

V. Yepmo, G. Smits, M. J. Lesot, and O. Pivert, "Leveraging an isolation forest for anomaly detection and data clustering," Data and Knowledge Engineering, vol. 151, p. 102302, 2024.

S. Song, N. Gao, Y. Zhang, et al., "BRITD: Behavior rhythm insider threat detection with time awareness and user adaptation," Cybersecurity, vol. 7, p. 2, 2024.

X. Pan, "Independent study of Splunk," Open Access Library Journal, vol. 11, no. 4, pp. 1–16, 2024.

J. Warner, "Uncovering the mechanisms of UEBA: Machine learning and its applications in cybersecurity," Exabeam Blog, Apr. 18, 2023. [Online]. Available: https://www.exabeam.com/blog/ueba/uncoveringthe-mechanisms-of-ueba-machine-learning-and-its-applications-incybersecurity [Accessed: Aug. 29, 2025]

Teramind, "Forcepoint DLP Features Teardown: Advantages Disadvantages," Nov. 6, 2024. [Online]. Available:
https://www.teramind.co/forcepoint-dlp-features-teardown-advantagesdisadvantages/[Accessed: Aug. 29, 2025].

E.Akbas , , "The Math of SIEM Analysis: Evaluation of Key Next-Gen SIEM Features using Validation," Journal of Network Information Security, vol. 11, no. 2, 2023.

S. R. Nath, "Leveraging User and Entity Behavioral Analysis and Machine Learning for Log-Based Anomaly Detection," Digital Repository of Theses, SSBM Geneva, 2024.

Bienias, Piotr, Grzegorz Kołaczek, and Arkadiusz Warzynski. "Archi-´ tecture of anomaly detection module for the security operations center." 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE). IEEE, 2019.

A. Greneche, B. M. Kone, and C. Toinard, "SOC as a Service: a´ user centric approach for Network Security Monitoring," in Proc. Int. Conf. on Security and Management (SAM), The Steering Committee of

The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2016.

M. A. Rahman, M. S. Alam, and M. S. H. Mrida, "How interactive dashboards improve managerial decision-making in operations management," American Journal of Advanced Technology and Engineering Solutions, vol. 1, no. 01, pp. 122–146, 2025.

Preetha, M., et al. "An Assessment of the Security Benefits of Secure Shell (SSH) in Wireless Networks." 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON). IEEE, 2023.

S.-S. Chen, C.-C. Chang, and J.-H. Horng, "Reversible data hiding scheme using prediction neural network and adaptive modulation mapping," Multimedia Tools and Applications, vol. 84, no. 9, pp. 6665–6686, 2025.

Multamaki, Markus. Near real-time IoT data pipeline architectures. MS¨ thesis. M. Multamaki, 2024.¨

F. Mumuni and A. Mumuni, "Explainable artificial intelligence (XAI): from inherent explainability to large language models," arXiv preprint arXiv:2501.09967, 2025.

E. Pulls, "Efficient Windows logging for Incident Response: A comparison of logging strategies for malware and threat detection," 2024.

Zhu, Jieming, et al. "Loghub: A large collection of system log datasets for ai-driven log analytics." 2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE). IEEE, 2023.

J. Kim, et al., "A Model for Illegal File Access Tracking Using Windows Logs and Elastic Stack," Journal of Information Processing Systems, vol. 17, no. 4, 2021.

H. M. Marin-Castro and E. Tello-Leal, "Event log preprocessing for process mining: a review," Applied Sciences, vol. 11, no. 22, p. 10556, 2021.

N. Harjunpa¨a, "Log management system technologies and methods for¨ near real-time fault analysis systems: An exploration of log shipping and storage," 2023.

A. Mudvari, et al., "Adaptive compression-aware split learning and inference for enhanced network efficiency," ACM Transactions on Internet Technology, vol. 24, no. 4, pp. 1–26, 2024.

N. Zubair, et al., "PEM: Remote forensic acquisition of PLC memory in industrial control systems," Forensic Science International: Digital Investigation, vol. 40, p. 301336, 2022.

M. A. Raziq, et al., "Development of Open-sourced, Self-executable and GUI-based Application Tool Q-Check for Quality Prediction of Resistance Spot Weld using Artificial Neural Network," 2022.

Xanthidis, Dimitrios, et al., eds. Handbook of Computer Programming with Python. CRC Press, 2022.

J. J. Paul, Distributed Serverless Architectures on AWS, Berkeley, CA, 2023.

Zhekova, Mariya. "An Algorithm for Exploratory Analysis and Normalization of Big Data with Pandas." Proceedings of the Bulgarian Academy of Sciences. Vol. 76. No. 11. 2023.

B. Bin Sarhan and N. Altwaijry, "Insider threat detection using machine learning approach," Applied Sciences, vol. 13, no. 1, p. 259, 2023, doi:
10.3390/app13010259.

E. Holder and N. Wang, "Explainable artificial intelligence (XAI) interactively working with humans as a junior cyber analyst," HumanIntelligent Systems Integration, vol. 3, no. 2, pp. 139–153, 2021.

H. A. Agoro and J. Moore, "Integration of Explainable AI in Security Operations Centers (SOCs)," International Journal of Computer Science and Information Security (IJCSIS), vol. 24, no. 2, pp. 1–10, 2024.

S. B. Patil, "Agentic MailBot: Autonomous Multi-Agent Email Response System with Contextual Intelligence," Authorea Preprints, 2025.

O. Campesato, Python 3 and feature engineering, 2023, pp. 1–216.

A. M. Salman, B. T. Al-Nuaimi, A. A. Subhi, H. Alkattan, and R. H. C. Alfilh, "Enhancing cybersecurity with machine learning: A hybrid approach for anomaly detection and threat prediction," Mesopotamian Journal of CyberSecurity, vol. 5, no. 1, pp. 202–215, 2025.

W. Gan, et al., "Anomaly rule detection in sequence data," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 12, pp. 12095–12108, 2021.

Ismail, R. Kurnia, Z. A. Brata, G. A. Nelistiani, S. Heo, H. Kim, and H. Kim, "Toward robust security orchestration and automated response in security operations centers with a hyper-automation approach using agentic artificial intelligence," Information, vol. 16, no. 5, p. 365, 2025.

A. Yaseen, "Accelerating the SOC: Achieve greater efficiency with AIdriven automation," Int. J. Responsible Artif. Intell., vol. 12, no. 1, pp.
1–19, 2022.