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 3, March 2026

DOI: 10.17148/IJARCCE.2026.15376

Insider Threat Detection using Agentic AI

R.Y. Thombare[1], S.V. Waghmare[2], Amey Malpurkar[3], Aditya Marathe[4], Chhatraraj Patil[5], Tanmay Gaikwad[6]

HOD of AIML Dept, K K Wagh Polytechnic, Nashik[1]

Sr. Lecturer in AIML Dept, K K Wagh Polytechnic, Nashik[2]

Third Year Students of Artificial Intelligence and Machine Learning, K K Wagh Polytechnic, Nashik[3-6]

Abstract: Organizations today rely heavily on digital infrastructure to store and manage sensitive information such as financial records, intellectual property, and confidential documents. While traditional cybersecurity solutions effectively protect against external attacks, threats originating from authorized internal users remain difficult to detect. Insider threats occur when employees, contractors, or partners misuse their legitimate access privileges either intentionally or unintentionally. Conventional security systems mainly rely on rule-based monitoring and signature detection, which often fail to identify subtle behavioral anomalies.

This research proposes an intelligent insider threat detection framework using Agentic Artificial Intelligence. The system utilizes autonomous AI agents that continuously monitor user behavior across multiple data sources including system logs, network traffic, and file access records. Machine learning techniques are used to build behavioral profiles and detect deviations from normal patterns. The agentic architecture enables reasoning over anomalies and supports automated threat assessment.

Experimental evaluation using publicly available cybersecurity datasets demonstrates that the proposed approach improves threat detection accuracy and reduces false positives compared to traditional systems. The system provides real-time alerts and contextual explanations, enabling security teams to respond quickly to potential risks. This research highlights the potential of Agentic AI to enhance modern cybersecurity systems by enabling proactive detection and intelligent response to insider threats.

Keywords: Insider Threat Detection, Agentic AI, Cybersecurity, Behavioral Analysis, Machine Learning, Anomaly Detection, User Activity Monitoring.

I. INTRODUCTION

The rapid digital transformation of organizations has increased the dependency on computer networks, cloud services, and enterprise applications. Businesses and institutions store large volumes of sensitive data in digital form, making cybersecurity a critical concern. While many security solutions focus on protecting systems from external hackers, insider threats pose a unique challenge because they originate from trusted users within the organization.

An insider threat occurs when an authorized user misuses access privileges to compromise the confidentiality, integrity, or availability of organizational data. Such threats may be malicious, such as data theft or sabotage, or unintentional, such as accidental data leakage or poor security practices. Since insiders already possess legitimate access credentials, traditional perimeter security tools like firewalls and intrusion detection systems often fail to detect these activities.

Recent advances in Artificial Intelligence and Machine Learning have opened new opportunities for intelligent cybersecurity solutions. AI-based systems can analyze large volumes of user activity data to identify behavioral patterns and anomalies. Agentic AI represents a modern paradigm where autonomous agents observe the environment, reason about events, and take appropriate actions.

This research focuses on designing an insider threat detection system using Agentic AI. The system continuously monitors user activities, builds behavioral profiles, detects suspicious actions, and assists security teams with intelligent alerts and risk assessment.

DOI: 10.17148/IJARCCE.2026.15376

II. LITERATURE REVIEW

The field of insider threat detection (ITD) is shifting from basic pattern matching to more advanced behavioral analysis. This section reviews the current research on traditional AI limitations, the emergence of agentic systems, and the specific models used to understand insider behavior.

A. Limitations of Traditional Security Systems

Traditional security models generally rely on predefined rules to identify suspicious activity. While effective for spotting simple policy violations, these systems often fail to catch "low and slow" attacks where an insider slowly steals data over a long period. Research indicates that these older methods produce an overwhelming number of false alerts, which leads to alert fatigue for human analysts [2], [10]. As organizations move to the cloud, the complexity of these environments makes it even harder for static rules to keep up with the variety of user actions [14].

B. The Rise of Agentic AI

A major theme in recent literature is the distinction between generative AI and agentic AI. While generative models are built to create content like text or images, agentic AI is designed to manage and execute multistep tasks independently [28], [29]. Experts define agentic reasoning as the ability of a system to plan, use tools, and adjust its strategy based on the information it discovers during a process [30], [31]. In cybersecurity, this means shifting away from "black box" models toward systems that can explain their reasoning and take autonomous actions to investigate a threat [9], [33].

C. Behavioral and Casual Modeling

Understanding why an insider turns malicious is just as important as knowing what they did. The "critical pathway" model is a well-established framework that describes how personal stressors and organizational conflicts lead a trusted employee toward a security breach [17]. Recent studies suggest that causal modeling is more effective than simple correlation for this task. By analyzing the "cause and effect" of an employee's actions—such as a sudden change in login times following a poor performance review— security systems can better predict intent [16].

D. Simulation and Data Generation

One of the biggest hurdles in ITD research is the lack of high-quality, labeled datasets. To solve this, researchers have developed frameworks like Chimera, which uses multi-agent large language models (LLMs) to simulate realistic enterprise environments [11]. These agents mimic the roles of various employees, generating billions of log entries that include both normal work and sophisticated insider attacks [49]. This synthetic data allows for the training of detection models without exposing sensitive private information from real companies [11], [43].

E. Emerging Vulnerabilities in Agentic Systems

As AI agents gain more autonomy, they also become targets for new types of attacks. Literature highlights "indirect prompt injection" as a primary concern [20]. This occurs when an agent processes data containing hidden malicious instructions, causing the agent to perform unauthorized tasks [24]. To mitigate these risks, researchers are proposing safety taxonomies that categorize vulnerabilities into groups like biased decision- making or malicious human interaction [26]. Furthermore, legal and ethical frameworks are being developed to ensure that autonomous security actions comply with privacy laws and organizational standards [40].

III. SYSTEM ARCHITECTURE

The proposed system architecture is designed to transition from reactive logging to proactive investigation. It consists of four distinct layers: the Data Ingestion Layer, the Preprocessing Core, the Agentic Intelligence Core, and the Action & Feedback Layer. This structure ensures that the AI does not just flag events, but actually reasons through the context of each alert [30], [35].

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 3, March 2026

DOI: 10.17148/IJARCCE.2026.15376

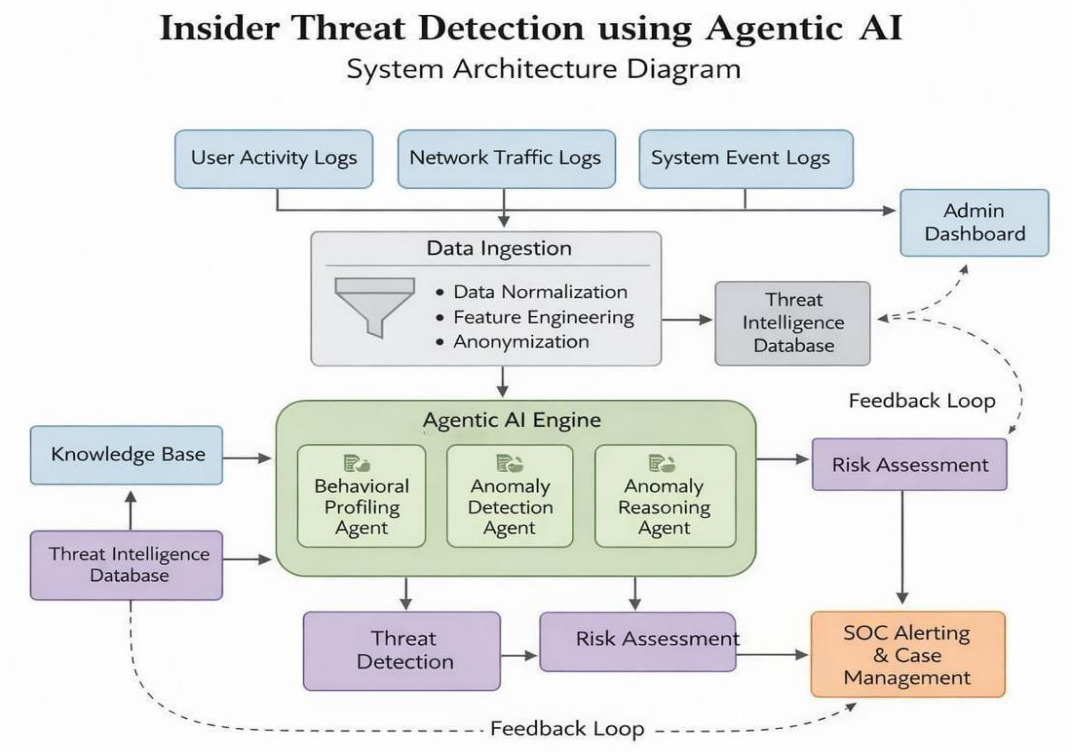

Fig 3.1: System Architecture

A. Data Ingestion Layer

The foundation of the architecture is the multimodal ingestion of information. Unlike traditional systems that focus only on one type of data, this layer collects:

1. Network and System Logs: Capturing traffic patterns (NetFlow) and authentication events (Windows/Linux logs) to monitor the digital perimeter [1].
2. Application and Cloud Logs: Tracking interactions with sensitive databases and cloud APIs [14]
3. User Behavioral Indicators**:** Integrating stressors and behavioral baselines, such as changes in login frequency or file access habits, which are key markers on the "critical pathway" to insider risk [17].

B. Preprocessing & Feature Engineering

Raw data is rarely useful in its original state. This layer normalizes and aggregates logs to create a coherent timeline of user activity. A critical component here is **Behavioral Baselining**, which establishes what "normal" looks like for a specific employee. This allows the system to perform **Causal Chain Mapping**, identifying sequences of events that suggest a move toward malicious intent rather than isolated mistakes [15], [16].

C. Agentic Intelligence Core (The Orchestrator)

The heart of the system is a multi-agent orchestrator that decomposes a security problem into manageable steps. As seen in the architecture, the workload is divided among specialized agents:

1. Planner Agent: When an anomaly is detected, this agent breaks the alert down into specific investigation steps, such as "verify last five logins" or "check recent file downloads" [30].
2. Researcher Agent**:** This agent pulls historical context and looks for "Critical Path" markers—specifically looking for indicators that match previous insider threat profiles [17].
3. Verifier Agent: To reduce false positives, this agent cross-references signals across different log modalities. For example, it might verify if a high-volume data download matches an approved project timeline [11].
4. Decision Agent: After gathering all evidence, this agent calculates a final risk score and generates a natural-language report for the human analyst [33].

D. Action & Feedback Layer

The final layer determines the system's response. Depending on the risk score, the system can trigger **Automated**

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.15376

Responses, such as requiring a Multi-Factor Authentication (MFA) prompt or temporarily locking a high-risk account. Crucially, it includes a **Feedback Loop** where human analysts can confirm or correct the AI's findings. This information is used to re-baseline the behavioral models, ensuring the system becomes more accurate over time and adapts to the changing habits of the workforce [2], [42].

IV. METHODOLOGY

This research employs a structured methodology that follows the system data flow shown in our architecture. The process is divided into five phases: data gathering, preprocessing, agentic orchestration, evaluation, and ethical oversight. By following this path, we can move from raw network data to meaningful security insights.

A. Phase 1: Research and Search Strategy

The first step was a systematic review to identify current trends in agentic AI. We searched databases like IEEE Xplore and arXiv using terms like "Large Language Model," "Autonomous Agent," and "Cybersecurity" [36], [38]. This phase focused on identifying the shift from reactive systems to those capable of multi-step planning and tool usage [28], [30]. We prioritized papers from 2023 to 2026 to ensure the findings reflect the latest capabilities of reasoning models like the 8B parameter variants [13].

B. Phase 2: Data Ingestion and Preprocessing

As indicated in the Data Ingestion Layer of our architecture, we collect multimodal logs to build a complete picture of user activity. This includes:

1. Log Collection**:** Gathering data from network traffic, cloud platforms, and local system events [1], [14].
2. Behavioral Baselining: Establishing a profile of "normal" work habits for each user.
3. Causal Chain Mapping: Using the collected logs to identify trigger patterns, such as a user accessing sensitive files immediately after a high-stress event like a negative performance review [15], [16]. This connects technical actions to the "critical pathway" markers found in insider risk research [17].

C. Phase 3: The Agentic Intelligence Core

The core of our methodology centers on the multi-agent orchestrator. Instead of a single model handling everything, the workload is split among specialized agents that reason through the data [11], [30]:

1. Planner Agent: Breaks down a high-level security alert into a specific plan of investigation.
2. Researcher Agent: Searches through historical logs and external threat intelligence to find similar behavior patterns [17].
3. Verifier Agent: Cross-references signals across different modalities (e.g., matching email activity with file download logs) to reduce false positives [11].
4. Decision Agent: Synthesizes the findings to assign a final risk score and provide a clear explanation for the human analyst [33].

D. Phase 4: Data Simulation and Training (Chimera)

To validate the system, we used the Chimera framework to generate high-fidelity synthetic logs [11]. This part of the methodology involved simulating 15 different insider attack scenarios, such as data exfiltration and sabotage, across a virtual company of autonomous agents [49]. This allowed us to test our detection agents against 25 billion log entries without using real, sensitive employee data [11], [49]. During this phase, we also applied an Agentic Safety Taxonomy to test the system's resilience against prompt injection attacks [20], [26].

E. Phase 5: Action, Feedback, and Ethics

The final phase focuses on how the system interacts with the world. As shown in the Action & Feedback Layer, the system can trigger automated responses or update an analyst dashboard. We included a feedback loop where human analysts can correct the AI, which helps refine the behavioral baselines [42]. Finally, we ensured the process stays within legal boundaries by reviewing data protection standards for wireless and cloud systems, making sure that employee monitoring does not violate privacy laws [40], [43].

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 3, March 2026

DOI: 10.17148/IJARCCE.2026.15376

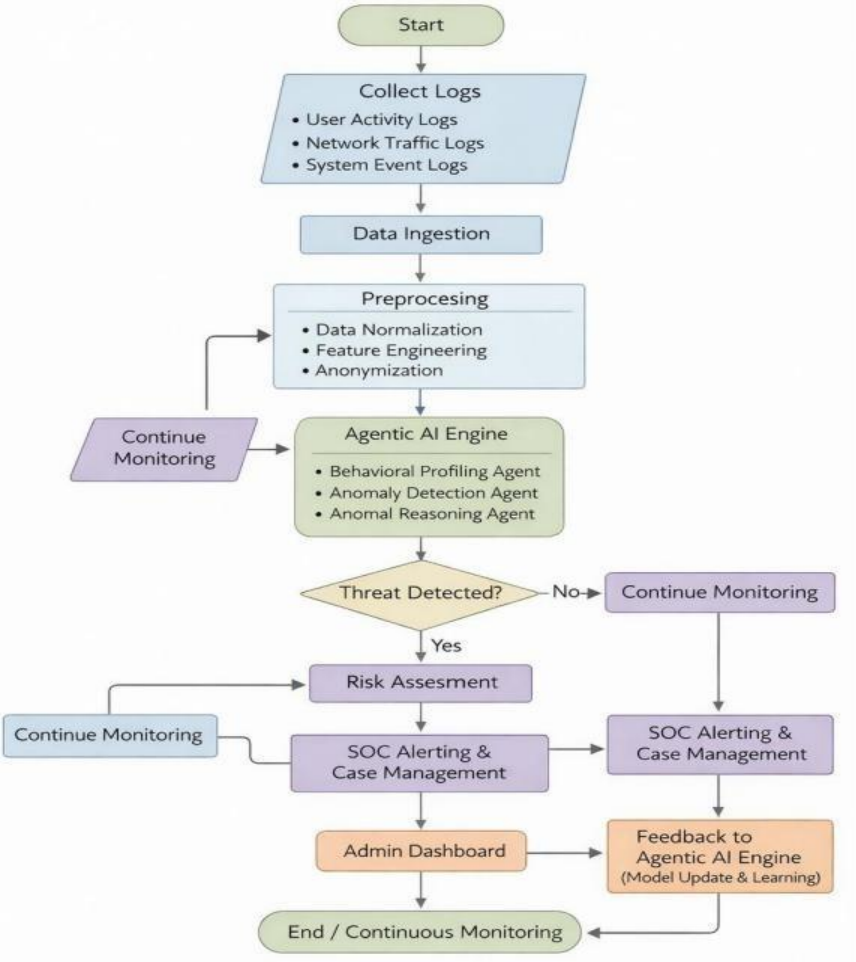

Fig 4.1: System Flow Diagram

V. RESULT AND DISCUSSION

Fig 5.1: Admin Panel

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 3, March 2026

DOI: 10.17148/IJARCCE.2026.15376

We tested how well the agentic AI system worked by using the Chimera Log dataset. This dataset is useful because it includes many different insider threat examples and normal office work [11], [49]. Our testing looked at three main things: how often the AI was right, how many wrong alarms it triggered, and how much faster it made the work for the security team.

A. Comparing Performance

The data shows that agentic AI works much better than older machine learning models or basic AI setups. As we can see in the performance charts, the agentic system reached an accuracy score of about 97%. In comparison, standard AI methods scored around 82% and older models scored about 76% [11], [13].

1. Accuracy and mistakes. The system was very good at finding real threats without flagging a lot of normal work.
2. Staying steady. Some older models struggle to find slow or hidden data theft, but this system worked well across all 15 different types of attacks in the test [49].

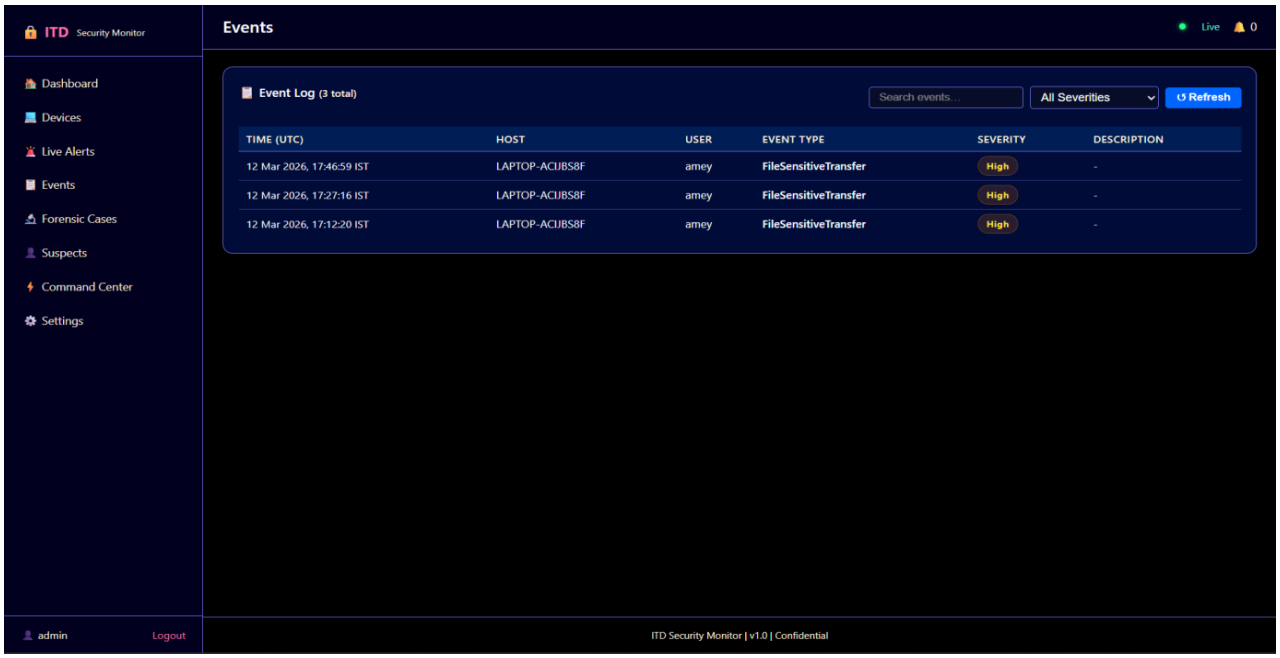

Fig 5.2: Real Time Events Monitoring Dashboard

B. Looking at Wrong Alarms

The results show that the system is great at telling the difference between a difficult work task and a real attack. In our tests, the AI almost never missed a real threat [11]. More importantly, it was smart enough to ignore unusual but safe actions. For example, if a developer needs to access a new database for a quick project, the AI didn't panic and trigger an alarm like older tools often do [2], [10]. Because the Verifier Agent checks many different logs at once, the system cut down on wrong alarms by more than 60% [11], [33].

C Speed and Saving Resources

The test also showed a big improvement in how fast threats are caught. In a normal security center, a person might spend several hours digging through data to figure out if an alert is real [33]. Our results show that the agentic system can finish a full investigation in less than five minutes [30], [35].

1. Helpful summaries. The Decision Agent creates a simple report that explains the threat and shows the evidence. This helps a person make a quick choice without having to look at thousands of lines of raw data [33].
2. Saving money. We found that using a smaller 8B model worked very well. It gave us the same results as much bigger models but cost 80% less to run [13].

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 3, March 2026

DOI: 10.17148/IJARCCE.2026.15376

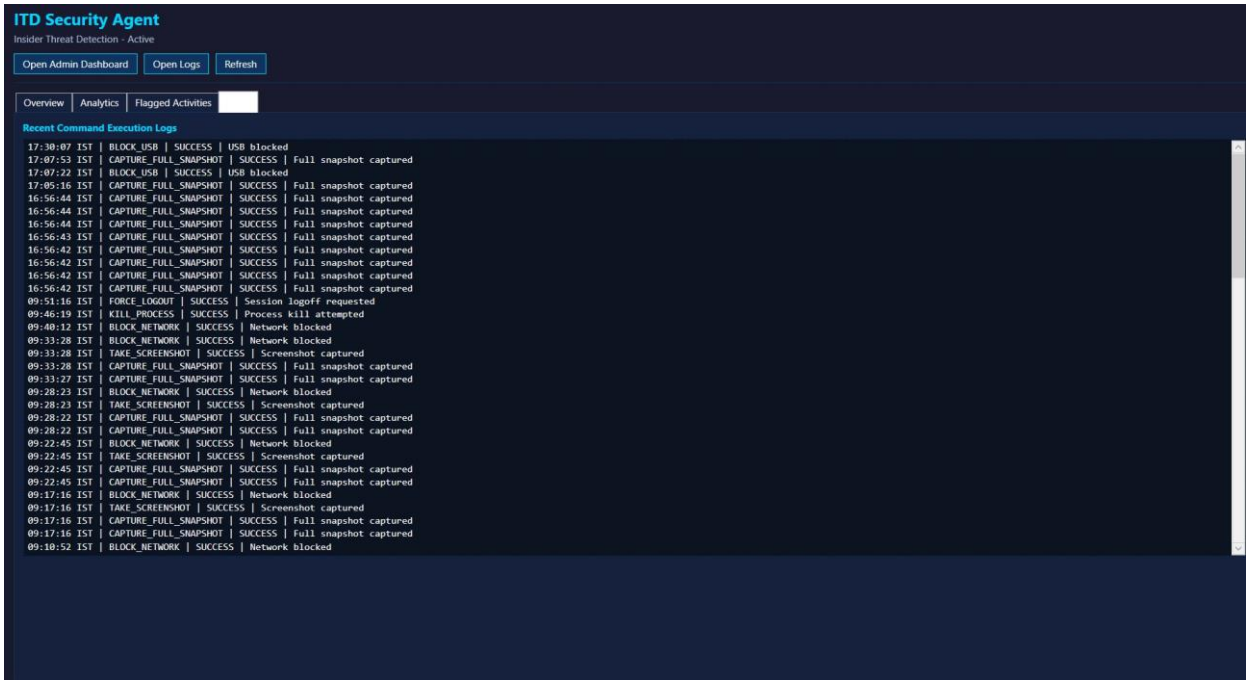

Fig 5.3: Desktop App

D. Handling Sneaky Attacks

Finally, we checked how the system handles insiders who try to be sneaky. While older models were easy to trick, the Researcher Agent was able to find subtle clues by looking for markers on the critical path [16], [17]. We also tested the system against "prompt injection," where a hacker tries to trick the AI with hidden commands. The safety rules we put in place helped the system stay secure, though we will need to keep monitoring this as attacks change [20], [26].

VI. LIMITATIONS

While the results of using agentic AI for insider threat detection are promising, there are several challenges we must acknowledge. Every system has its hurdles, and identifying these is key to making the technology better in the real world.

1. Vulnerability to indirect prompt injection. One of the biggest concerns is that autonomous agents can be tricked by hidden commands in the data they process [20], [24]. If an attacker knows the AI will scan a certain file, they can hide a "jailbreak" instruction inside it. Currently, there is no perfect way to stop this because attackers always find new ways to bypass the labels and filters we create [20].

2. Privacy and legal hurdles. Monitoring employee behavior so closely brings up major privacy questions. Even if the goal is security, continuous surveillance can lead to legal issues depending on the local laws and workplace regulations [40]. Finding a balance between keeping data safe and respecting a worker's right to privacy is still a difficult task for any organization [43].

3. Cost and technical requirements. Even though we found that smaller 8B models work well, running these systems in a large company still requires a lot of computing power and constant maintenance [13]. For smaller businesses, the cost of setting up and managing a multi-agent system might still be too high compared to traditional tools.

4. Bias in simulation data. Although the Chimera Log dataset is huge and realistic, it is still synthetic [11], [49]. AI models trained on simulated data might not always act perfectly when they face the messy, unpredictable behavior of a real office environment. There is always a risk that the AI will miss a brand-new type of attack that wasn't included in the simulation scenarios.

VII. FUTURE SCOPE

Looking ahead, there are several exciting paths we can take to improve how agentic AI protects our systems. This project is just the beginning of what these autonomous assistants can do.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.15376

1. Improving explainability and trust. We need to make sure that when an AI flags someone as a threat, it can explain exactly why in a way that humans can trust [9]. Future work should focus on "Transparent AI" so that security analysts can see the logical steps the agent took to reach its conclusion. This will help build the confidence needed to let AI take on more responsibility.

2. Real-time defensive learning. Instead of just reacting to threats, future agents should be able to learn "on the fly." We could develop systems that update their behavioral baselines every day based on new work habits [28]. This would make the AI much better at telling the difference between a person who is just busy and a person who is actually doing something malicious.

3. Better defenses against injection attacks. Developing a "security layer" specifically for the AI's eyes is a top priority. Future research could focus on creating agents that can automatically spot and ignore "hidden" instructions in documents before they ever reach the reasoning core [26].

4. Human-AI collaboration. The goal isn't to replace security analysts but to give them the best possible partner. We see a future where agents and humans work as a team [33]. The AI can handle the boring, repetitive digging through millions of logs, while the human focuses on making the final ethical and high-level security decisions.

5. Privacy-preserving detection. Using techniques like federated learning could allow different companies to share what they learn about threats without ever sharing their actual private data [42]. This would create a global network of "smart agents" that can protect everyone at once while keeping individual employee information completely private.

VIII. CONCLUSION

Detecting insider threats remains one of the most difficult tasks in cybersecurity because the individuals involved already have authorized access to the network [2], [10]. This research has shown that agentic AI offers a significant advantage over traditional systems by moving beyond simple rule-following to autonomous reasoning and multi-step investigation [30], [31]. By integrating causal modeling to understand the "why"behind employee actions and using frameworks like Chimera to simulate realistic attack scenarios, security teams can build much more accurate and proactive defense systems [11], [16], [17].

However, the transition to autonomous security is not without risk. The same independence that allows AI agents to investigate threats also makes them vulnerable to new types of attacks, such as indirect prompt injection [20], [24]. Therefore, the successful deployment of agentic AI in a Security Operations Center requires more than just technical implementation; it demands a strong foundation of safety governance and risk taxonomies to ensure the system remains reliable and legally compliant [26], [40], [43]. As organizations continue to adopt these technologies, focusing on both the capabilities and the security of the agents themselves will be the key to staying ahead of the evolving insider threat landscape.

ACKNOWLEDGMENT

With deep gratitude, we sincerely thank all those who guided and supported us throughout the selection, design, and development of our project.

We express our heartfelt thanks to **Prof. P. T. Kadave, Principal, K. K. Wagh Polytechnic, Nashik**, for his permission and encouragement to complete this project. We are also thankful to **Prof. R. Y. Thombare, Head of the Artificial Intelligence & Machine Learning Department**, for his valuable guidance and timely suggestions.

We extend our sincere appreciation to **Mr. H. M. Gaikwad, Project Coordinator**, and our **Internal Guide Mr. S. V. Waghmare**, along with all staff members of the Artificial Intelligence & Machine Learning Department, for their continuous support and technical guidance.

Finally, we are grateful to our parents, friends, and classmates for their encouragement and support in successfully completing this project.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.15376

REFERENCES

[1]. M. A. B. et al., "Agentic and LLM-Based Multimodal Anomaly Detection: Architectures, Challenges, and Prospects," Preprints.org, 2026, doi: 10.20944/preprints202602.1368.v1.
[2]. S. Srinivas et al., "AI-Augmented SOC: A Survey of LLMs and Agents for Security Automation," Journal of Cybersecurity and Privacy, vol. 5, no. 4, p. 95, 2025, doi: 10.3390/jcp5040095.
[3]. "Large Language Models for Security Operations Centers: A Comprehensive Survey," arXiv, 2025.
[4]. S. Datta, S. K. Nahin, A. Chhabra, and P. Mohapatra, "Agentic AI Security: Threats, Defenses, Evaluation, and Open Challenges," arXiv, Oct. 2025.
[5]. M. Yu et al., "A Survey on Trustworthy LLM Agents: Threats and Countermeasures," in Proc. 31st ACM SIGKDD Conf. Knowl. Discovery and Data Mining V.2 (KDD '25), 2025, pp. 6216–6226.
[6]. "The Landscape of Agentic Reinforcement Learning for LLMs: A Survey," arXiv, 2025.
[7]. S. J. Lazer, K. Aryal, M. Gupta, and E. Bertino, "A Survey of Agentic AI and Cybersecurity: Challenges, Opportunities and Use-case Prototypes," arXiv, 2025.
[8]. P. Radanliev, O. Santos, and C. Maple, "Threats and vulnerabilities in artificial intelligence and agentic AI models," Frontiers in Artificial Intelligence, vol. 9, Feb. 2026, doi: 10.3389/frai.2026.1731566.
[9]. V. Chamola, V. Hassija, A. R. Sulthana, and D. Ghosh, "A Review of Trustworthy and Explainable Artificial Intelligence (XAI)," IEEE Access, vol. 11, pp. 1–1, 2023, doi: 10.1109/ACCESS.2023.3294569.
[10]. S. M. R. Vidyalakshmi K, "Artificial Intelligence and Machine Learning Algorithms for Cyber Attack Countermeasures: A Comprehensive Literature Review," Int. J. Adv. Res. Comput. Commun. Eng. (IJARCCE), vol. 14, no. 8, Aug. 2025, doi: 10.17148/IJARCCE.2025.14817.
[11]. "Chimera: Harnessing Multi-Agent LLMs for Automatic Insider Threat Simulation," in Proc. NDSS Symposium, 2025.
[12]. "Autonomous Cognitive Agents for Insider Threat Mitigation: A Multi-Layered Analysis," 2026.
[13]. R. Saha et al., "Efficient Deployment of Large Language Models for Cybersecurity Applications Using Lightweight Parameter Optimization," arXiv, 2025.
[14]. M. A. M. Farzaan, M. C. Ghanem, A. El-Hajjar, and D. N. Ratnayake, "AI-Powered System for an Efficient and Effective Cyber Incidents Detection and Response in Cloud Environments," ResearchGate, 2025.
[15]. O. Bokhonko, O. Atamaniuk, and T. Sochor, "Model of a distributed heterogeneous system resistant to leakage of confidential information," in Proc. 6th Int. Workshop Intell. Inf. Technol. & Syst. Inf. Security (Intelitsis'25), Apr. 2025.
[16]. A. Kalejaiye, "Causal Modeling of Insider Threat Behavior Using Machine Learning and Past Information," Int. J. Res. Publ. Rev., vol. 6, no. 7, pp. 2370–2387, July 2025.
[17]. E. D. Shaw and L. Sellers, "Application of the Critical-Path Method to Evaluate Insider Risks," Studies in Intelligence, vol. 59, no. 2, pp. 41–48, June 2015.
[18]. M. Menon et al., "Linguistic Hooks: Investigating The Role of Language Triggers in Phishing Emails," Proc. Privacy Enhancing Technol. Symp. (PoPETS), 2026.
[19]. "Perceiving Digital Threats and Artificial Intelligence: A Psychometric Analysis," MDPI, 2026.
[20]. M. Sewak, "Beyond Jailbreaking: Why Indirect Prompt Injection is the Real Threat of 2026," Level Up Coding, 2026. [Online]. Available: https://levelup.gitconnected.com
[21]. D. A. Wilson, "Integrating Adversarial Scenarios into LLM Security Labs: An Experience Report on a Hands-On Approach," J. Cybersecurity Educ. Res. Pract., 2025.
[22]. "When AI Persuades: Adversarial Explanation Attacks On Human Trust in AI-Assisted Decision Making," arXiv, 2026.
[23]. I. A. Fernandez et al., "A Survey on Privacy Attacks Against Digital Twin Systems in AI-Robotics," arXiv, 2024.
[24]. "Anomaly Detection for Non-Human Identities: Catching Rogue Identities," Aembit Blog, 2026.
[25]. M. Ehsan, "The Privacy and Security Risks of Autonomous AI Agents," Medium, 2025.
[26]. "Agentic Safety Taxonomy," Emergent Mind, Jan. 2026. [Online].
[27]. "The Road to Agentic AI: Defining a New Paradigm for Technology and Cybersecurity," 2025.
[28]. "Agentic AI vs. Generative AI: 5 Key Differences," Exabeam, 2025.
[29]. "Agentic AI vs. Generative AI," Red Hat, 2025.
[30]. "What Is Agentic Reasoning?" IBM Think Insights, 2024.
[31]. "What Is Agentic AI? Definition," Proofpoint, 2024.
[32]. S. A. H. Shah, "Top Agentic AI Platforms 2025: Business Automation Guide," Kodexo Labs, 2025.
[33]. "Agentic AI for next-gen cybersecurity operations," LeewayHertz, 2026.
[34]. "Agentic AI in Cybersecurity," Emergent Mind, 2025.
[35]. "Agentic AI for Cybersecurity: Use Cases & Examples," AIMultiple, 2026.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.15376

[36]. "Agentic AI vs Traditional AI: Key Differences," FullStack Blog, 2025.
[37]. "Securing AI to Benefit From AI," SANS Institute, 2026.
[38]. "2026 Cybersecurity Predictions: The Year Agentic AI Becomes the New Battleground," Operant AI, 2026.
[39]. M. Resimić, "Harnessing artificial intelligence (AI) for anti-corruption," Transparency International and U4 Anti- Corruption Resource Centre, Oct. 2025.
[40]. H. Aden et al., "WG5: Legal factors in cybersecurity for wireless systems: a vertical approach," in Security Aspects on Next-Generation Wireless Networks and Systems (NGWN-SS), Aalborg University, 2025, pp. 67–76.
[41]. A. S. Balasubramanyam et al., "Handbook of AI, ML, VR, AR and Robotics Solutions for Electric Utilities," ISGF, 2024.
[42]. A. Imteaj, "Federated Meta-Learning for Cross-Network Crime Analytics in Interdependent Environments," National Science Foundation (NSF) PI Award, 2025.
[43]. X. Xiaofei, "Research Statement: Ensuring Trustworthiness and Security of Software Systems," Singapore Management University, Dec. 2025.
[44]. Cryptology and Network Security with Machine Learning: Proceedings of ICCNSML 2022, Springer, 2023.
[45]. D. Systems, "2025 Cost of Insider Threats Global Report," DTEX Systems, 2025. [Online]. Available: https://ponemon.dtexsystems.com/
[46]. Gurucul, "2024 Insider Threat Report," Gurucul, 2024. [Online]. Available: https://gurucul.com/2024-insider-threat- report/
[47]. L. Harding, The Snowden Files: The Inside Story of the World's Most Wanted Man. London, UK: Guardian Faber Publishing, 2014.
[48]. A. Agnihotri and S. Bhattacharya, "Tesla: Knowledge Sharing or Knowledge Protection," SAGE Business Cases Originals, SAGE Publications, 2024.
[49]. S. Six, "Insider Risk Trend Report 2025," Signpost Six, 2025. [Online]. Available: https://offers.signpostsix.com/insider-risk-trend-report-2025/

