NATIONAL BOARD FOR TECHNICAL EDUCATION
NATIONAL JOURNAL OF TECHNICAL EDUCATION
Volume 24 Nos. 1 2025
ISSN No. 2992-3522

Machine Learning Applications for Insider Threat Detection in Cloud Security: A Narrative Review

Oladimeji Ganiyu
Computer Science Dept., Moshood Abiola Polytechnic, Abeokuta.
oladimeji.ganiyu@mapoly.edu.ng

Lawal Olufunmilayo
Computer Science Dept., Moshood Abiola Polytechnic, Abeokuta.

Oyelowo Rofiat
Computer Science Dept., Moshood Abiola Polytechnic, Abeokuta.

Abstract

This narrative review aims at synthesizing recent literature that explores machine learning (ML) for addressing insider threats in cloud security environments. In 2023, a systematic search of the literature was performed for peer-reviewed papers, conference proceedings, and industry reports related to ML techniques, challenges, and real-world problems from 2018 onwards. Some of the supervised and unsupervised learning techniques used to detect malicious activities in this review are Radial Basis Function Neural Networks and Random Forests. We offer a new comparison of these techniques, examining how well each performs in various insider threat scenarios. ML improves threat detection, but challenges remain with data accessibility, feature selection, and ethical issues. The research demonstrates the need for a balanced approach leveraging both ML and human-centric strategies to achieve effective insider threat mitigation. Finally, we discuss future research directions, focusing on building transparency into AI-driven systems and exploring federated learning techniques to overcome current limitations.

Keywords: *Machine Learning; Cloud Security; Insider Threats; Anomaly Detection; User Behavior Analytics; Privacy.*

1. Introduction

In today's market, cloud computing has revolutionized data management by offering scalability and accessibility while being cost-effective. However, it has also introduced new security problems, particularly insider threats. These threats come from within, from individuals with

authorized access to an organization's assets, posing significant risks. This paper provides a narrative review of studies focused on detecting and mitigating insider threats in cloud computing using ML. We review relevant literature, evaluate existing methodologies, and identify challenges and best practices for implementing ML into cloud security frameworks.

This introductory section provides a brief overview of studies on detecting and mitigating insider threats in cloud computing using ML. The remaining part of the paper proceeds as follows: Section two considers both the sources and methods of study which include the database consulted, inclusion and exclusion criteria, the quality of studies and relevance to cloud security. The third section is concerned with the types of insider threats in cloud security. Section four discusses machine learning techniques for insider threat detection. The fifth section presents challenges and limitations, focusing on key themes and section six discussed the performance of ML models with case studies. Section seven proposed recommended practices, and, the purpose of the final section is to reflect on the extent to which this study highlighted the broad role of ML in improving insider threat detection in cloud environments and the way forward.

The section below reviews the literature related to Insider Threats in Cloud Security.

2. Types of Insider Threats in Cloud Security

According to an investigation by Shejin & Sudheer (2023), the study classifies cloud security insider threats into several categories: malicious insiders, negligent insiders, compromised accounts, privileged insiders, contractors, and third-party vendors. Understanding these types of threats is critical for developing efficient ML-based detection mechanisms.

Moving on now to consider machine learning techniques for insider threat detection in the next section.

3. Machine Learning Techniques for Insider Threat Detection

Previous studies have shown that anomaly detection builds baseline behavioral models from historical data. Deviations from normal patterns are flagged as potential insider threats (Namdev et al., 2023). For instance, Oliveira et al. (2023) proposed an ensemble learning approach combining Isolation Forest and One-Class Support Vector Machine (SVM), achieving a 95% detection rate for anomalous user behaviors in cloud environments. Threat classification - ML models distinguish between accidental and

malicious actions. Mehmood et al. (2023) applied this method to detect privilege escalation attacks, achieving 92% accuracy using Random Forest classifiers. User Behavior Analytics (UBA) evaluates user behavior to separate legitimate users from potential masqueraders. Research finding by Kambhampaty & Nygard (2019) demonstrated that a Long Short-Term Memory (LSTM) network could detect abnormal usage patterns with 88% accuracy, significantly outperforming rule-based systems. Deep learning techniques, including Random Forests and Radial Basis Function Neural Networks (RBFNN), show promise in detecting malicious activities. This becomes clear when one examines Attou et al. (2023) that compared these approaches, finding that RBFNN achieved a 97% detection rate for insider attacks, while Random Forests reached 94% accuracy.

This chapter has demonstrated that the techniques for insider threat detection can be efficient. it is now necessary to explain the challenges of insider threats.

4. Challenges and Limitations

According to a study published recently, it is difficult to obtain large training datasets of insider actions due to the sensitive nature of these activities and their low frequency (Wanyonyi et al., 2023). This often results in imbalanced datasets, severely affecting model performance as observed in Choubey, (2023). The multi-dimensional nature of insider threats makes feature selection for analysis a complex task. Mohammed (2022) proposed a hybrid feature selection approach, combining filter and wrapper methods to enhance detection accuracy by 7% compared to conventional techniques. Monitoring insider activity also raises privacy concerns as demonstrated in the work by Wanyonyi et al., (2023). Research findings by Nguyen (2023) found that 68% of employees were uncomfortable with continuous monitoring, underscoring the need for transparent policies and ethical AI implementation.

This section has reviewed the challenges in implementing ML for insider threats; the next section establishes the framework for research methodology.

5. Methodology

We conducted a comprehensive systematic review using databases such as Google Scholar, IEEE Xplore, ACM Digital Library, and ScienceDirect. Keywords included "insider threats," "cloud security," "machine learning," "anomaly detection," "privilege escalation," and "real-time threat

detection." We focused on peer-reviewed articles and scientific conference abstracts published between 2018 and 2023. Inclusion criteria comprised studies related to ML applications in cloud security, research on insider threat detection, and empirically supported or theoretically constructed frameworks. We excluded studies not applicable to cloud environments, those focused on traditional security measures, and non-English publications. The quality of studies was evaluated based on methodological rigor, sample size (for empirical studies), and relevance to cloud security. Data were extracted on ML techniques, challenges, performance metrics, and real-world applications.

The section (Result and Discussion) that follows takes up a number of case studies to support ML effectiveness.

6. Results and Discussion

Performance of ML Models and Case Studies

A major US bank used ML-based anomaly detection to identify insider trading. The system, a hybrid approach combining Isolation Forest and LSTM networks, uncovered multiple unauthorized trades by an executive, potentially saving $12 million in liabilities (Chen et al., 2018; Satyanarayana & Madhavi, 2023). In healthcare, a provider used a Random Forest classifier to detect data exfiltration attempts. The system identified an employee gradually stealing patient records, achieving 94% accuracy in distinguishing between normal and malicious data access patterns (Moraetes, 2018). A government agency employed a UBA system using LSTM networks to monitor user activities. The system detected a compromised account used to access classified information, preventing a potential breach. The LSTM model outperformed traditional rule-based systems by 23% in detection accuracy (Bhardwaj & Dave, 2023).

So far this section has focused on the case studies. In the next section, we present the recommended practices based on literature review and the case studies available.

7. Recommended Practices

Based on the literature review and case studies, we recommend the following best practices:

1) Hybrid Approach: Combine multiple ML techniques to detect a wide variety of threats.

2) Continuous Model Updates: Retrain models regularly with new data to adapt to evolving threats.

3) Explainable AI: Implement interpretable ML models to provide clear justifications for threat alerts.

4) Privacy-Preserving Technologies: Use federated learning or differential privacy to enhance data protection while maintaining model performance (Zhang et al., 2023).

5) Human-in-the-Loop Systems: Combine ML predictions with human expertise for final decision-making (Villarreal-Vasquez et al., 2023; Sarhan & Altwaijry, 2022).

In this section, recommended practices were highlighted, the next (last) section moves on to consider the way forward.

8. Conclusions

This paper highlights the broad role of ML in improving insider threat detection in cloud environments. However, challenges in data quality, privacy, and model interpretability remain. Future research should focus on developing privacy-preserving ML approaches for insider threat detection in cloud environments, exploring federated learning techniques to address data scarcity and privacy issues, investigating quantum machine learning tools to enhance threat detection capabilities, conducting large-scale longitudinal studies to assess long-term effectiveness, and creating standardized benchmarks and datasets for fair comparisons of ML techniques. By addressing these research directions, the field can move toward more robust, ethical, and effective ML-based insider threat detection systems for cloud security.

References

Attou, H., Mohy-Eddine, M., Guezzaz, A., Benkirane, S., Azrour, M., Alabdultif, A., & Almusallam, N. (2023). Towards an Intelligent Intrusion Detection System to Detect Malicious Activities in Cloud Computing. Applied Sciences. https://doi.org/10.3390/app13179588

Bhardwaj, S., & Dave, M. (2023). Integrating a Rule-Based Approach to Malware Detection with an LSTM-Based Feature Selection Technique. *SN Computer Science, 4*. https://doi.org/10.1007/s42979-023-02177-2

Chen, X., Zhang, L., Liu, Y., & Tang, C. (2018). Ensemble learning methods

for power system cyber-attack detection. 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 613-616. https://doi.org/10.1109/ICCCBDA.2018.8386588

Choubey, R. (2023). Machine Learning Algorithms for Cloud Computing Security: A Review. Tuijin Jishu/Journal of Propulsion Technology. https://doi.org/10.52783/tjjpt.v44.i4.2564

Homoliak, I., Toffalini, F., Guarnizo, J.D., Elovici, Y., & Ochoa, M. (2018). Insight into Insiders and IT. ACM Computing Surveys (CSUR), 52, 1 - 40. https://doi.org/10.1145/3303771

Kambhampaty, K., & Nygard, K.E. (2019). Identifying Insider and Masquerade attackers in Cloud Computing and IoT Devices.

Mayank Namdev, Dr. JayasundarS, Muhammad Babur, Dr. Deepak A. Vidhate, 5Santosh Yerasuri (2023). Enhancing Security in Cloud Computing with Anomaly Detection Using Machine Learning. Tuijin Jishu/Journal of Propulsion Technology.

Mehmood, M., Amin, R., Muslam, M.M., Xie, J., & Aldabbas, H. (2023). Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning. IEEE Access, 11, 46561-46576. https://doi.org/10.1109/ACCESS.2023.3273895

Mohammed, T.Y. (2022). Impact of Number of Features Selected and Size of Training Data on the Accuracy of Machine Learning Based Cloud Security Algorithms – An Empirical Analysis. SLU Journal of Science and Technology. https://doi.org/10.56471/slujst.v4i.279

Moraetes, G. (2018, December 10). The CISO's Guide to Managing Insider Threats. Security Intelligence. https://securityintelligence.com/the-cisos-guide-to-managing-insider-threats/

Nguyen T. (2023, January 10). What is Threat Detection and Response (TDR)? - CrowdStrike. crowdstrike.com. https://www.crowdstrike.com/cybers

ecurity-101/threat-detection-response-tdr/

Oliveira, J.M., Almeida, J., Macedo, D.F., & Nogueira, J.M. (2023). Comparative Analysis of Unsupervised Machine Learning Algorithms for Anomaly Detection in Network Data. *2023 IEEE Latin-American Conference on Communications (LATINCOM)*, 1-6. https://doi.org/10.1109/LATINCOM59467.2023.10361849

Sarhan, B. B., & Altwaijry, N. (2022). Insider threat detection using Machine learning approach. Applied Sciences, 13(1), 259. https://doi.org/10.3390/app13010259

Satyanarayana, S., & Madhavi, P. (2023). Big Data Analytics for Electrical Systems using Machine Learning Algorithms. 2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), 1-4. https://doi.org/10.1109/TEMSMET56707.2023.10149919

Shejin T. R., & Sudheer K. T. (2023). A Review on Major Cyber Threats and Recommended Counter Measures. International Journal for Research in Applied Science and Engineering Technology. https://doi.org/10.22214/ijraset.2023.49764

Villarreal-Vasquez M., Modelo-Howard G., Simant Dube, and Bhargava B. (2023) Hunting for Insider Threats Using LSTM-Based Anomaly Detection. EEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. 20, NO. 1, JANUARY/FEBRUARY 2023

Wanyonyi, E.N., Abeka, S.O., & Masinde, N. (2023). A Systematic Review on Machine Learning Insider Threat Detection Models, Datasets and Evaluation Metrics. International Journal of Network Security & Its Applications. https://doi.org/10.5121/ijnsa.2023.15603

Zhang, L., Zhu, T., Xiong, P., Zhou, W., & Yu, P.S. (2023). A Robust Game-Theoretical Federated Learning Framework with Joint Differential Privacy. *IEEE Transactions on Knowledge and Data Engineering, 35*, 3333-3346. https://doi.org/10.1109/TKDE.2021.3140131