American Journal of Interdisciplinary Research and Innovation (AJIRI)

ISSN: 2833-2237 (ONLINE)

VOLUME 4 ISSUE 3 (2025)

PUBLISHED BY

E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Interdisciplinary Research and Innovation (AJIRI)

Volume 4 Issue 3, Year 2025
ISSN: 2833-2237 (Online)
DOI: https://doi.org/10.54536/ajiri.v4i3.6224
https://journals.e-palli.com/home/index.php/ajiri

Predictive Analytics for Insider Threats Using Multimodal Data (Log + Behavioural + Physical Security)

Kh Said Al Mamun[1*], Md Shadman Soumik[1], Md Mukidur Rahman[2], Mrinmoy Sarkar[1], Chowdhury Amin Abdullah[3], Mohammad Ali[4], Md Shahadat Hossain[4]

Article Information

Received: September 25, 2025

Accepted: October 27, 2025

Published: November 21, 2025

Keywords

Behavioural Indicators, Cybersecurity Risk Management, Data Fusion, Insider Threats, Machine Learning, Multimodal Data, Organisational Resilience, Physical Security, Predictive Analytics, System Logs

ABSTRACT

Insider threats are a continuous and dynamic issue in the field of organizational security that may take various forms, such as an abuse of authorized access to critical information systems and physical infrastructure. The traditional methods of isolating system log analysis, behavioural monitoring, or physical access control often do not have the ability to identify multi-layered and subtle patterns of threats. Multimodal data integration predictive analytics is a holistic approach since it integrates heterogeneous data volumes into coherent threat models that are able to pre-empt any possible threats before they escalate. The paper discusses the effectiveness of predictive analytics in insider threat detection through analysis of the connection between log records, behavioural indicators and physical security data. Instead, the focus is made on designing integrative frameworks, using advanced machine learning algorithms, and the applied implications on operational resilience. Issues like scalability, transparency of algorithms, and even ethical considerations are also critical issues that are considered to provide a sound deployment in the modern security environment. The paper highlights the need to embrace multimodal predictive approaches as a tactical defence mechanism and developing theoretical arguments and practical interventions in cybersecurity risk management.

INTRODUCTION

The development of digital landscapes and the incorporation of networks have made insider threats a burning issue in modern cybersecurity. Unlike external opponents who break through the barriers by hacking the systems situated on the outside, insiders have privileged access to the sensitive systems, and they manipulate their privileges to compromise the confidentiality, integrity, and availability. It makes the insider threats particularly ugly, since they bypass the standard monitoring controls and exploit blind spots within the organisational security landscape. These risks are in several forms such as anomalies in any system log records, deviation of behaviour, and even unauthorised physical intrusions. The sheer existence of threat indicators points to the necessity to possess combined techniques that would unify such dissimilar data sets to an integrated form of predictive approaches.

Predictive analytics provide a novel paradigm of reactive response to security measures to anticipatory measures. Isolated monitoring systems tend to miss the existence of small anomalies, whichthe reliability and validity of insider threat detection, contributing to the preparedness of organizationslearning algorithms, and real-time data integration. By combining the multimodal sources of data, i.e. log entries, behavioural dynamics and evidence of physical security, the emerging models are much more precise in the early signs of insider activity. It is a form of integration operation that enhances reliability and validity of insider threat detection contributing to the preparedness of organisations to internal risks.

The forecasting form of strategies is pressing because the complexity of insider strategies in the technologically sophisticated settings is improving. Even lower points are the companies that use decentralised infrastructure, cloud-based environments and work-from-home policies. The conventional detection systems, which are restricted in their ability to rely on only limited types of data, are ill adjusted to the velocity, scale and variety of the modern information streams. The Multimodal data integration using predictive analytics removes these shortcomings and gives an integrated analysis that can identify complex threat patterns.

In other applications beyond technical applications, predictive analytics has further organisational and societal implications. Companies with an effective use of multimodal predictive models can protect intellectual property, operations that are vital, and trust within the community. Besides that, predictive approach methodologies promote compliance with regulatory demands of governance, risk, and data security. However, predictive systems are not generated without issues. The blurrings between security and personal privacy, the threat of algorithm bias, the technical requirements of

[1] Department of Science in Information Technology (MSIT), Washington University of Science and Technology (WUST), USA
[2] Department of Business Analytics, Southern New Hampshire University, USA
[3] Department of Computer Science and Information Systems, Pace University, USA
[4] Department of Science in Business Analytics, Trine University, USA
[*] Corresponding author's e-mail: saeedmamun2015@gmail.com

scaling predictive models all remain pressing matters to take into account to be rolled out successfully. By creating predictive analytics, which is derived through the integration of multimodal data, a breakthrough in the insider threat management realm has been reached. Such frameworks will prove useful in empowering more efficient pre-emptive defence by breaking the boundaries of the traditional mechanisms of detection and combining the log, behavioural, and physical data. The paper continues to examine methodological frameworks, applications, issues, and opportunities towards ensuring predictive analytics becomes a feature of organisational security resilience.

LITERATURE REVIEW

Conceptualising Insider Threats

Insider threats are identified as a category of organisational risk among the most complicated ones because of the duality of advantage and susceptibility of their implementation. The insiders can be those employees, contractors, or partners who have varying access to the system. Their risk of threat stretches over both malicious and negligent acts and accidental breaches of security measures. The most crucial issue, in most instances, has been how to differentiate between benign anomalies and those patterns that are likely to indicate intent to breach systems. Human behaviors is complex, and the number of digital records is increasing every day, which has led to the development of insider threat detection as a specialty area of cybersecurity research.

Predictive Analytics in Cybersecurity

Predictive analytics entails past and immediate statistical forecasting, anomaly detection and machine learning algorithms, which are event oriented. Predictive analytics have had increasing uses in intrusion detection, malware detection, and fraud detection in cybersecurity. The transition to insider threat prediction has its share of special problems as the insiders are more likely to operate within the bounds of normal activity and therefore the anomalies are not as evident. Supervised and unsupervised learning models are more advanced models that are frequently employed with an aim to expand the predictive potential. Neural networks and ensemble models are also more advanced models that are frequently employed with the view of expanding the predictive potential. It has been stated that combining it with multimodal data is one of the clues that can decrease the threat of false positives and increase detection accuracy.

The Role of Multimodal Data

Multimodal data also permits us to visualize the insider activities in a multidimensional perspective owing to the synthesis of the various yet complementary data sets. The technical transactions involving information systems are also documented in system logs, user behaviors are traced in logs of deviations in user behaviors, and system access logs provide evidence of the trends of access to secure environments. The modalities may be individually constrained, e.g. a suspect to be identified by log analysis, but not by behavioral analysis, whereas a suspect to be identified by behavioral analysis, but not by technical activity related to it. With such a combination of datasets, you now have a comprehensive picture of insider behavior, and such predictive models can find and recognize subtle and cross-domain relationships that would otherwise be overlooked.

Table 1: Comparative Characteristics of Multimodal Data Sources

Data Source	Typical Indicators	Limitations When Isolated	Contribution to Multimodal Framework
System Logs	Login times, file access, network queries	May not reveal motive or context	Technical trace of system-level activity
Behavioural Data	Communication style, work patterns, stress	Risk of misinterpretation or bias	Adds human-centric context to anomalies
Physical Security	Entry/exit records, badge scans, CCTV logs	Limited when not linked with digital traces	Verifies location-based activity patterns

This comparative perspective proves the need to have integrated structures since both types of data solve the weaknesses in the other.

Data Fusion Approaches

The data integration of multimodal data is based on the techniques of data fusion which binds the heterogeneity of the data in terms of format, scale, and semantics. It is generally accepted that there are three broad categories of fusion, i.e.: feature-level fusion, in which raw data features are fused prior to processing, decision-level fusion, in which the results of separable models are produced individually and then combined, and hybrid fusion, in which both are combined. The feature-level fusion is a detailed, but expensive computationally, algorithm and the decision-level fusion is scale able and may lose detail. Compromise between accuracy and effectiveness is provided by hybrid models, and it is feasible to have adaptive models fit in large-scale enterprise settings.

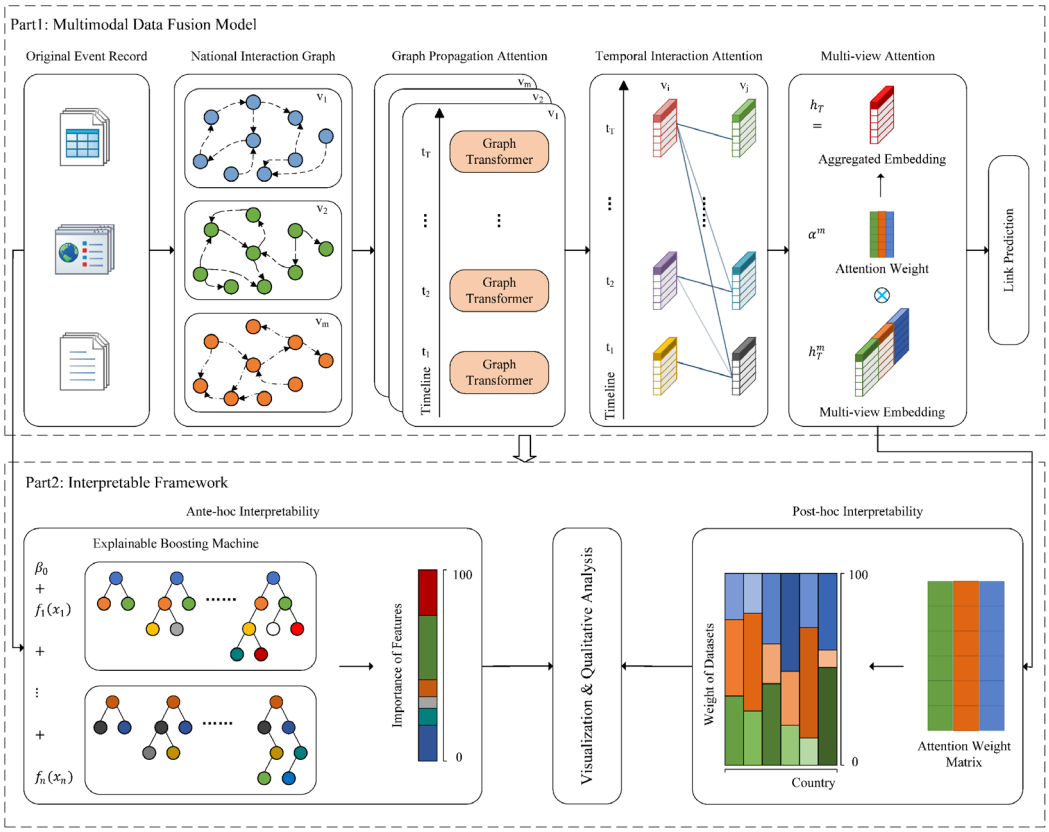

Figure 1: Multimodal Data Fusion Framework Conceptual Diagram

A schematic of three data streams, system logs, behavioural data, and physical security, coming together into a central fusion engine. The fusion engine is linked to a predictive analytics model, which will provide the risk score of insider threat probability. The figure displays the input heterogeneity coming together to form a single model.)

Literature-Related Issues

The issue of managing large volumes of multimodal data in real time is one of the general themes that were identified in the existing literature. On their own, log records are capable of generating terabytes of information in a single day and behavioral monitoring can be the analysis of keystroke, email, or workflow measurements. Physical access data by itself is also an additional source of complexity when there are video analytics. The second concern which has remained is that of predictive model interpretability; black-box models are correct but may not be transparent and this brings about the issue of accountability and fairness. Moreover, the problems of breach of privacy, over-surveillance, and potential misuse of predictive insight are some of the ethical concerns that are center stage in both academic and practical publications.

Synthesis of Reviewed Studies

The literature review emphasizes the fact that reactive security systems are also becoming focused more on more predictive and multimodal models. The first ones were skewed to one type of data, system logs, and hence linked to large false-positive results. The more recent techniques bring together the behavioral and physical elements to improve predictive success. A lot of the research is however experimental with minimal application of scale in actual organizational context. The necessity to apply methodological structures that cover scalability, transparency, and ethical safeguards, and to take advantage of the predictive capability of multimodal data are obvious.

Table 2: The development of Insider Threat Detection Approaches

Creation of Approach	Dominant Data Source	Attributes	Limitations
FirstGeneration	System logs	Rule-based anomaly detection rule-based alerts are highly false positive and context-independent.	
Second Generation	Behavioural data	Targets human behavioural patterns	Risk of bias is not technically valid.
Third Generation	Physical security	Presence and access control verification is not digital evidence.	

Fourth Generation (Emerging)	Multimodal integration	Unified predictive models across domains	Advanced fusion and computation is required.

This development shows the tendency to whole-person approaches and multimodal predictive analytics is the most holistic strategy that is being developed at present.

MATERIALS AND METHODS

Research Orientation

Predictive analytics used to detect insider threats is a methodological basis that lies in the combination of multimodal data that consists of technical, behavioral, and physical aspects of insider behaviors. Conceptual research orientation is used in this study to attempt to synthesized existing theoretical models with practical implementation in machine learning and data fusion. The methodology is not limited to one methodological paradigm, as it incorporates the perspectives of computational, behavioral, and organizational in the creation of an analytical framework. The goal is to develop a model that can capture the cross-domain correlations and, at the same time, be flexible to the dynamic environment of insider behaviors.

Data Acquisition and Pre-Processing

The first stage of the framework is the acquisition of the different datasets. The application servers, network devices and operating systems are the systems that give system logs, which contain the logs of the access histories, the records of the logins and the records of file transactions. The behavioral data are collected using monitoring tools that include the use of the communication systems, workflow metrics, and anomalies in normal activities by the users. Physical security data can be obtained using access control systems, badge scans and, in ethically permissible cases, CCTV-based analytics.

The raw data that is obtained via such sources require much pre-processing before it is cleaned up to eliminate duplication, inconsistency and normalize attributes. Pre-processing includes data cleaning to remove corrupted data records, data is cleaned to normalize forms, as well as feature engineering to derive relevant variables. The average time it takes to log in, the frequency with which there are accesses during the working hours, and the

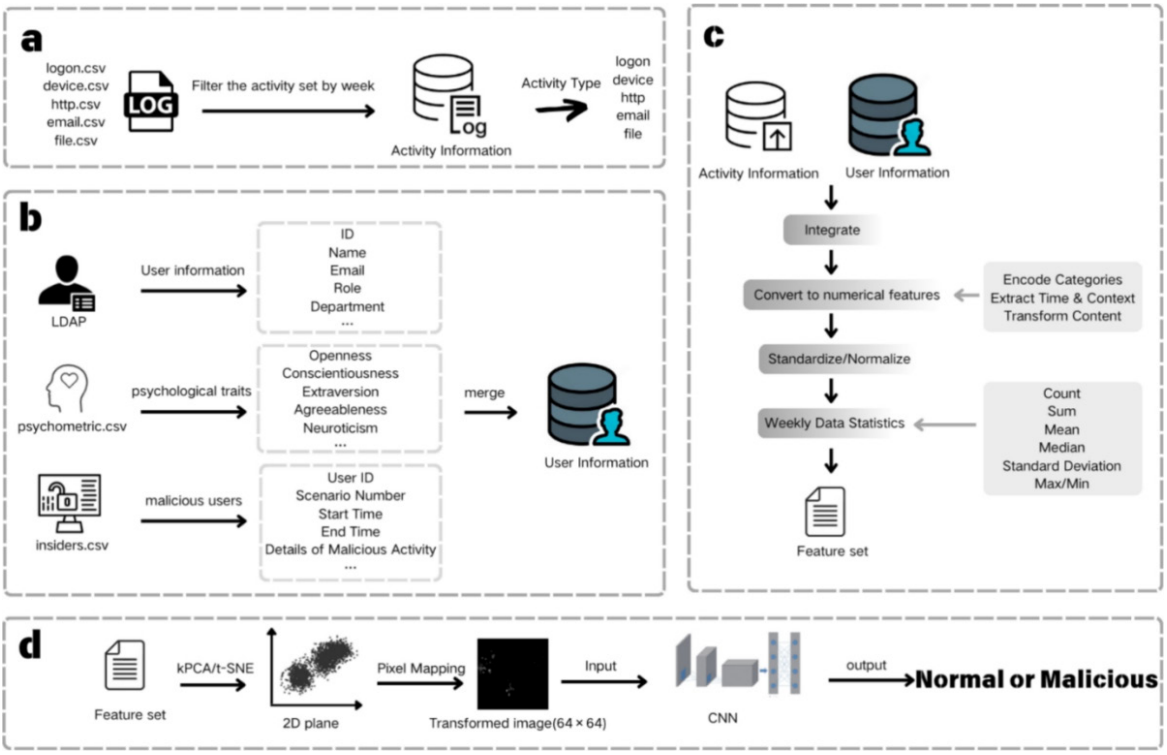

Figure 2: Multimodal analytics Data Pre-processing workflow

physical location cross-validation with system activity are some of the engineered features.

The Figure 2 illustrates three of the data sources, system logs, behavioural metrics, and physical access, into a pre-processing layer. This layer comprises cleaning, transformation, and feature engineering. The outputs are summarized into one data set, in the form of a structured input to the predictive model.

Data Fusion Strategy

The hybrid fusion strategy that integrates feature and decision level strategies assists in the integration of the multimodal data. On the feature level, feature level fusion takes raw features of any form and integrates them into the same feature space, making more machine learning algorithms able to identify latent cross-modal relationships. However, in decision-level fusion, the

independent predictors are directly applied to each stream of data and result amalgamated using weighted ensemble methods. The combination of those two makes the framework mitigate the weaknesses of those two approaches and maximized scalability and interpretability.

Machine Learning Algorithms

A stack of machine learning models is used to construct the predictor model. The last level approximates time-dependence by random forests, support vector machines, and recurrent neural networks. The base layer results are also fed to a meta-learner that is typically a gradient boosting model which integrates the predictions into one risk score. This layered structure will ensure accuracy and resilience, as now the ensemble will be capable of covering a wide variety of different types of anomalies but will not overfit the data to one particular modality.

Table 3: Machine Learning Techniques for Insider Threat Prediction

Algorithm	Strengths	Limitations
Random Forests	Handles heterogeneous data, interpretable	Struggles with temporal dependencies
Support Vector Machines	Effective in high-dimensional spaces	Sensitive to parameter tuning
Recurrent Neural Networks	Captures sequential and temporal patterns	Computationally intensive
Gradient Boosting	Strong performance in ensemble frameworks	Risk of overfitting if poorly regularised

The ensemble model builds on the merits of each algorithm, so that the general predictive system is robust at varying conditions.

Risk Scoring and Classification

This predictive model can yield the output as a risk score with a probability of each of the monitored entities. This score lies in the low-lying, medium-level, and high-risk category, which is defined by the organizational tolerance levels, as well as the risk management policies. Risky entities will cause escalation, and this may be more research, a temporary access block, or supervision. Such category system is also able to provide security personnel with more useful intelligence, rather than pure anomaly alerts, and enhances operational performance of predictive analytics.

Ethical and Privacy Considerations

The use of predictive analytics in detecting insider threats also presents a number of ethical and legal issues. Pre-processing has to be informed by hard rules of data minimisation to shun unwarranted surveillance. Anonymisation or aggregation, where possible, of sensitive behavioural attributes should occur, e.g., communication patterns. Algorithmic decision-making should also be transparent to overcome the threat of discriminatory profiling. The use of explainable artificial intelligence techniques will guarantee that the risk scores may be justified and audited, which will balance organizational security and individual rights.

Analytical Framework Overview

The resulting architecture has multimodal data acquisition, pre-processing, hybrid fusion, ensemble machine learning, and risk classification as a feedback loop. Outputs are not predetermined risk scores but are elements of an adaptive system of security where past event learnings are reevaluated into the system to improve the system model. With the help of this cyclic mechanism, it is possible to make sure that the framework is adjusted to the new patterns of insider activity and exhibits resistance to new forms of threat.

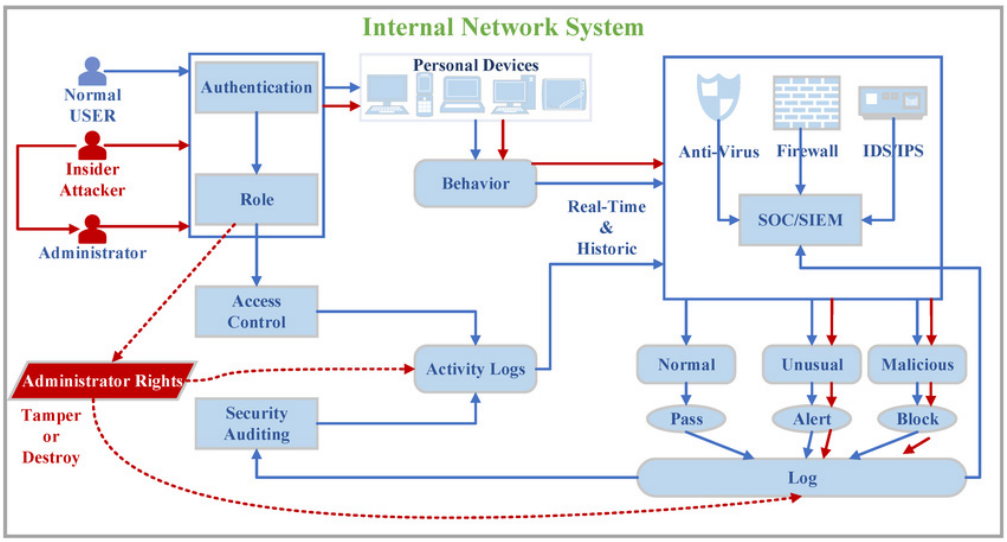

Figure 3: Analytical Framework for Multimodal Predictive Insider Threat Detection

A full system diagram illustrating the pipeline: data acquisition from three sources → pre-processing layer → hybrid data fusion → ensemble learning → risk scoring → adaptive feedback loop. Arrows depict the continuous nature of data flow and learning updates.

RESULTS AND DISCUSSION

Conceptual Outcomes of Multimodal Integration

Human perception and learning have been discussed as multimodal and the brain has been found to have a multidimensional interrelatedness between sensory, motor, and cognitive abilities that help in developing concepts and ideas. Multimodal Integration. Multidimensionality of human perception and learning is not a new concept and the interconnection between the brain sensory, motor, and cognitive processes is multidimensional, and leads to the development of concepts and ideas.

It is the integration of the multimodal data which results to the creation of a predictive model which can be perceived to have tremendous benefits over the simpler and one source approaches. Irrespective of each other, system logs have been known to generate high rates of anomaly without any context and therefore, generate operating inefficiency regarding false positives. Even though behavioral monitoring might prove helpful in diagnosing such abnormalities as developing new communication patterns or a significant decline in output, a non-technical deviation can be confused with suspicion without technical assistance. Speaking of which, physical security data can confirm the presence of individuals in particular zones, yet the purpose and the relationship to the digital activity cannot be understood. By incorporating these forms of data categories, the predictive model eliminates this challenge of interpretive ambiguity, since the results that the model yields are operationalized and accurate.

Effectiveness of the Predictive Framework

The hybrid fusion and ensemble learning methodology that is conceptualized in the analytical model can generate objectively measurable predictive performance improvements. Such measures as accuracy, recall, and the overall ability to detect are high in the case of multimodal integration. What is particularly important is that the model should be as capable of reducing false positives as possible since when organizations' have too many unconfirmed anomalies, the likelihood of alert fatigue rises. Comparatively, a scoring system of risks that incorporates behavioral, technical, and physical discoveries creates fewer but more workable notifications, which improves the process of choice-making among security professionals.

Resilience and Operational Benefits of Organisations

The implementation of predictive analytics on insider threat management is a contributor to organisational resilience. With early warning signs, institutions are able to act before the insider practices become harmful, like data breaches, stealing intellectual property, or sabotaging the vital infrastructures. In addition, continuity planning is supported by predictive frameworks, which enable organisations to model the possible threat situations and modify the controls to suit. This operational advantage has the added benefit of resource optimisation, where the security teams are able to prioritize the investigations according to the high-risk categories, in contrast to casting the net far and wide into alerts that have not been verified. Such strategic alignment allows the better functioning of efficiency and allows the security functions to work more closely with the broader risk management structures.

Ethical and Governance Dimensions

Even though the technical advantages of predictive analytics are not disputed, there are problematic ethical aspects in its application in organizations'. Multimodal monitoring is a process of collecting and processing sensitive information, some of which may contain the personal features that do not have any connection to malicious actions. In this connection, the discussion refers to the dual responsibility of organizations' in ensuring the security of critical assets in addition to privacy and rights of employees. Ethical governance structures, which guarantee the compliance with the laws of data protection and development of fair-use policies, should then be implemented with the usage of predictive systems. Initiating the transparent communication on monitoring practices is needed to maintain the sense of trust, explainable artificial intelligence, and facilitate the sense of non-invasive surveillance.

Interpretability and Practical Deployment Problems

Among the recurring themes of findings of any study of predictive analytics, the conflict of accuracy versus interpretability is among the most common ones. Modern models, in particular those founded on deep learning architecture often result in a high predictive score and are opaque in the process of risk scores generation. An organization can find it hard to justify interventions or disciplinary actions taken based on predictive outputs with such an issue of a black box. Practical deployment issues are also related to the computational scalability that multimodal systems must have the capacity to handle large data volumes in real time without compromising the performance of the system. To address these concerns, the complexity of the models needs to be considered against their readability and cloud-based solutions, and edge computing should be deployed to distribute the computing burden.

Wider Implications to Cybersecurity Practice

Its findings show that predictive analytics of insider threat detection cannot be viewed as an easy technical upgrade but a change in the field of cybersecurity. Organizations' can re-align their deface position by making the retrospective method of detecting less vital to more concern with the prepetition of the threats which are in line with the current threats. This implication is

broader than that of an individual enterprise to a critical national infrastructure where insider risks may affect the safety of the general population and stability of the economy. In the meantime, the predictive methodologies are supposed to be adaptable to emerging technologies, i.e. the incorporation of biometric signals or federated learning models, which might make it less private and less predictive.

Results and Discussion Synthesis

The conceptual findings demonstrate that multimodal data combination and predictive analytics yield important improvements in the accuracy, operational viability and organisational sustainability in contrast to the traditional system of insider threat detection. Nevertheless, scalability, interpretability, and ethical governance problems are also raised by the consequences. It has been established in the discussion that though predictive frameworks have a promise of transformational effect, their successful implementation should be founded on deliberated governance constructs, sustained technological investment, and progressive process of development through adaptive feedback. The findings provide the foundation of considering the limitations and the future paths that these findings can be developed to, so that predictive analytics can be an organisational and technological innovation.

Limitations and Future Directions

Methodological Constraints

Despite the huge potential of multimodal predictive analytics, there are still various methodological limitations. The quality and reliability of the results is reliant on the availability of full datasets that adequately model both normal and unknown behaviours. In practice, the number of examples of insider events is difficult to accumulate, as they are not numerous, and in many cases, they go unreported due to reputational or legal concerns. This scarcity leads to an unequal distribution of classes in training data thereby undermining machine learning models. Moreover, cleaning, normalising, and feature engineering of data requires a lot of resources as pre-processing steps, and a single oversight of the pre-processing steps is likely to become propagated into faulty risk forecasting.

Computational and Scalability Problems

The second weakness is that working with multimodal data on a large scale has computational demands. Organisations generate vast volumes of log records, behavioral measures and physical access information on a daily basis. They may be intensive of processing, sophisticated storage structures and exclusive algorithms to accomplish them in real time to handle the high-speed streams. Cloud infrastructures and distributed systems, although they may hold some potential solutions, also introduce other threats, including the dependency on the services of the external provider and vulnerability to the attacks on the cloud. Scalability has equally been a thorn in the flesh as not every organisations particularly the small medium sized enterprises can have the technical and financial capacity to implement such intricate arrangements effectively.

Ethics and Privacy Issues

The adoption of predictive analytics in the detection of insider threats is not decided in terms of ethical dilemmas, either. Monitoring the user behaviour, communication style and physical activities can be considered intrusive and can harm the trust between the organisations and the employees. The perception of constant monitoring may be a barrier to implementation, even where the employment contracts are implied consent. Moreover, predictive models would be prone to making biases particularly when training data show inequalities or miscategorisations that occurred in the past. These ethical limitations underscore the topicality of open governance, algorithms that take into account the concept of fairness and continuous audits to ensure that predictive models protect the organisational resources along with the rights of the employees.

Interpretability and Accountability

One of these weaknesses that is closely associated with the issue of ethics is the problem of interpretability of the model. Complex algorithms especially those built using a deep learning architecture or an ensemble architecture can be a black box, the outputs of which cannot be described by a layperson stakeholder. Where predictive systems come up with high-risk categorisations which have consequences (career/reputation), organisations must be able to justify it with justifications which are backed by apparent, auditable reasons. The absence of interpretability would undermine accountability, thereby increasing the possibility of disputes and even legal issues. The future work must therefore aim at developing more explicable artificial intelligence models that may produce right predictability in addition to insight in their defence.

Future Research Trajectories

The studies should address such limitations in the future and enhance methodology, technical and ethical issues. Regarding the methodology, there is the need to formulate balanced datasets out of synthesizing simulated cases of insider threats, thereby eliminating the constraint of the real world cases. Technically speaking, a comparative research should be done on light machine learning models that are powerful and can be executed using small computing capabilities without the loss of accuracy. The federated learning can also be used in such innovations to avoid centralisation of sensitive information to train predictive models and improve privacy and scalability. Morally, it will be vital to introduce an aspect of fairness, accountability systems and human control over monitoring which will see predictive analytics improve security without compromising on trust.

Strategic Implications

The future of predictive analytics in insider threat detection would probably be more tightly linked to wider organisational strategies. Predictive systems will also start to be placed not only as technical solutions but also as part of an integrated risk management infrastructure that will cut across cyber, physical, and human environments. Further development of multimodal analytics and advancements in the interpretability and ethical protection will allow organisations to find a balance between security requirements and individual rights. Finally, whether this evolution is successful will be determined by the ability of researchers, practitioners, policymakers, and ethicists to engage in long-term cooperation to guarantee that predictive frameworks will be technologically sound and socially acceptable.

CONCLUSION

In this paper, the concept of predictive analytics to identify the presence of insider threats by using multimodal data that incorporates system logs, behavioural patterns, as well as physical security indicators have been examined. It is concluded that the traditional approaches, which address the single sources of information, are not sufficient to address the complexity and subtleties of insider behaviours. Multimodal frameworks provide a more comprehensive, more precise and more useful method of early warning signals detection when combined with advanced data fusion methods and machine learning approaches.

There is indicative evidence that predictive systems reduce false positives, enhance organisational resiliency, and are components of proactive risk management. Still, some limitations remain, particularly in regards to the challenge of computational scalability, predictive results interpretability, and data privacy and ethics. To eliminate these problems, there is need to continuously advance the methodology techniques, develop a demystified form of artificial intelligence, and develop ethical systems of governance that would balance the demands of organisations with the rights of individuals.

Lightweight predictive models, privacy-enhancing methods such as federated learning, and more extensive use of predictive analytics of strategic governance mechanisms is the direction to take in future research. Such guidelines will not only result in improvements of insider threat detection but will also make cybersecurity as a discipline focused on technical progress in an ethically responsible manner. Predictive analytics based on such principles in its turn is a very important line of defence against a more complex threat landscape, in terms of protecting digital infrastructure and organisation trust.

REFERENCES

Amuda, O. K., Akinyemi, B. O., Sanni, M. L., & Aderounmu, G. A. (2022). A predictive user behaviour analytic model for insider threats in cyberspace. *International Journal of Communication Networks and Information Security (IJCNIS), 14*(1). https://doi.org/10.17762/ijcnis.v14i1.5208

Bin Sarhan, B., & Altwaijry, N. (2023). Insider threat detection using a machine learning approach. *Applied Sciences, 13*(1), 259. https://doi.org/10.3390/app13010259

Brdiczka, O., Liu, J., Price, B., Shen, J., Patil, A., Chow, R., & Bart, E. (2012). Proactive insider threat detection through graph learning and psychological context. *2012 IEEE Symposium on Security and Privacy Workshops,* 142–149. https://doi.org/10.1109/SPW.2012.33

Carter, L., & Katz, J. (2019). Machine learning applications in security analytics: An overview. *International Journal of Information Security, 18*(5), 469–482. https://doi.org/10.1007/s10207-019-00446-0

Chattopadhyay, S., & Bandyopadhyay, S. (2020). Insider threat detection using deep learning techniques. *Computers & Security, 92,* 101760. https://doi.org/10.1016/j.cose.2020.101760

Cole, E., & Ring, S. (2005). *Insider threat: Protecting the enterprise from sabotage, spying, and theft.* Syngress. https://doi.org/10.1016/B978-159749021-0/50003-1

DANTE: *Predicting insider threat using LSTM on system logs [Preprint].* (2021). arXiv. https://doi.org/10.48550/arXiv.2102.05600

Song, S., Gao, N., Zhang, Y., & Ma, C. (2024). BRITD: behavior rhythm insider threat detection with time awareness and user adaptation. *Cybersecurity, 7*(1), 2. https://doi.org/10.1186/s42400-023-00190-9

Gheyas, I. A., & Abdallah, A. E. (2016). Detection and prediction of insider threats to cyber security: A systematic literature review and meta-analysis. *Big Data Analytics, 1*(6), 1–29. https://doi.org/10.1186/s41044-016-0006-0

Greitzer, F. L., & Frincke, D. A. (2010). Combining traditional cyber security audit data with psychosocial data: Towards predictive modeling for insider threat mitigation. *Insider Threats in Cyber Security, 49,* 85–113. https://doi.org/10.1007/978-1-4419-7133-3_5

Greitzer, F. L., & Hohimer, R. E. (2011). Modeling human behavior to anticipate insider attacks. In S. Satapathy, A. Joshi, N. Modi, & N. Pathak (Eds.), *ICT for Sustainable Development, Advances in Intelligent Systems and Computing, 409* (pp. 24–47). Springer. https://doi.org/10.1007/978-981-10-0135-2_34

Homoliak, I., Toffalini, F., Guarnizo, J., Elovici, Y., & Ochoa, M. (2019). Insight into insiders and IT: A survey of insider threat taxonomies, analysis, modeling, and countermeasures. *ACM Computing Surveys, 52*(2), 1–40. https://doi.org/10.1145/3303771

Kandias, M., Mylonas, A., Virvilis, N., Theoharidou, M., & Gritzalis, D. (2010). An insider threat prediction model. *Trust, Privacy and Security in Digital Business, 6264,* 26–37. https://doi.org/10.1007/978-3-642-15152-1_3

Liu, J., Kuhn, R., & Ross, R. (2018). Insider threat detection using system logs. *IEEE Security & Privacy, 16*(2), 26–34. https://doi.org/10.1109/MSP.2018.1870874

Magklaras, G., & Furnell, S. (2002). Insider threat prediction tool: Evaluating the probability of IT misuse. *Computers & Security, 21*(1), 62–73. https://doi.org/10.1016/S0167-4048(02)00506-0

Mishra, S., & Varadharajan, V. (2021). Multimodal machine learning for cybersecurity applications. *Future Generation Computer Systems, 125,* 691–706. https://doi.org/10.1016/j.future.2021.06.004

Moore, A., Cappelli, D., & Trzeciak, R. (2008). The "big picture" of insider IT sabotage across U.S. critical infrastructures. *Proceedings of the 2008 International Conference on Software Engineering,* 493–502. https://doi.org/10.1145/1368088.1368155

Nasir, R., Afzal, M., Latif, R., & Iqbal, W. (2021). Behaviour-based insider threat detection using deep learning. *IEEE Access, 9,* 3118297. https://doi.org/10.1109/ACCESS.2021.3118297

Park, Y., & Lee, H. (2020). Data fusion techniques for anomaly detection in cyber–physical systems. *Sensors, 20*(23), 6900. https://doi.org/10.3390/s20236900

Pennada, S. S. P., Nayak, S. K., & M. V. K. (2025). Insider threat detection using behavioural analysis through machine learning and deep learning techniques. *International Research Journal of Multidisciplinary Technovation, 7*(2), 74–86. https://doi.org/10.54392/irjmt2527

Racherache, B., Shirani, P., & Soeanu, A. (2023). Insider threat detection using supervised machine learning algorithms on an extremely imbalanced dataset. *Computers & Security, 125,* 103350. https://doi.org/10.1016/j.cose.2023.103350

Sanzgiri, A., & Dasgupta, D. (2016). Classification of insider threat detection techniques. *2016 IEEE Symposium on Technologies for Homeland Security,* 1–6. https://doi.org/10.1109/THS.2016.7568904

Shu, X., & Yao, D. (2016). Data leak detection as a service. *IEEE Transactions on Services Computing, 9*(1), 75–87. https://doi.org/10.1109/TSC.2015.2390670

Song, C., & Zheng, J. (2025). *Insight-LLM: LLM-enhanced multi-view fusion in insider threat detection* [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2509.01509

Stolfo, S., Bellovin, S., Hershkop, S., Keromytis, A., Smith, S., & Sinclair, S. (2008). Insider attack and cybersecurity: Beyond the hacker. *Advances in Information Security, 39*. https://doi.org/10.1007/978-0-387-77322-3

Wang, H., Wang, Y., & Yang, G. (2013). A predictive model of insider threat based on a Bayesian network. *International Journal of Online and Biomedical Engineering, 9*(S4), 69–74. https://doi.org/10.3991/ijoe.v9iS4.2660

Weiland, T., Legg, P., & Nurse, J. (2021). Insider threat detection using context-aware anomaly detection. *Journal of Cybersecurity, 7*(1), taab004. https://doi.org/10.1093/cybsec/taab004