International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:3 Volume No.10 May - June – 2026
DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

Trade Secret Protection via Retrieval-Augmented Behavioral Analytics for Insider Threat Detection

Rupa Shree S[1], Akshaya Sannapureddy[2] , Adya Ananya[3],Dr. Chitra B.T[4]
[1]*UG Scholar, Dept. of CSE-CY, RV College Of Engineering., Bangalore, Karnataka, India*
[2]*UG Scholar, Dept. of CSE-CY, RV College Of Engineering., Bangalore, Karnataka, India*
[3]*UG Scholar, Dept. of CSE-CY, RV College Of Engineering., Bangalore, Karnataka, India*
[4]*Assistant professor, Dept. of IEM, RV College Of Engineering, Bangalore,Karnataka, India*

Abstract
Insider threats are one of the largest and most complex cybersecurity problems that organisations face today. Unlike external intruders, insiders have legitimate access credentials and knowledge of the organization, which makes perimeter-based security systems ineffective. This research presents a Retrieval-Augmented Behavioral Analytics framework for insider threat detection. The proposed framework enriches the machine learning classification of user behavior patterns with semantically similar past threat patterns retrieved from a vector knowledge base. User activity logs from the CERT Insider Threat Dataset r4.2, which transcribe events such as file access, logons, USB devices, and emails, are converted to semantic embeddings. A vector database powered by FAISS keeps historical embeddings of threat cases, from which top-K similar threat cases are retrieved during inference. Then, we fuse these with user features for classification by the XGBoost model.The proposed framework is designed and expected to outperform classical ML baselines such as Random Forest, Isolation Forest and stand-alone SVM, evaluated using F1-score and AUC-ROC as primary evaluation metrics. This design of the ablation study is intended to disentangle the impact of retrieval augmentation from classification alone. The framework also uses SMOTE to perform class balancing. The framework is interpretable in that it provides retrieved case similarity scores, allowing analysts to understand why a user was flagged rather than just being told a user was flagged.
Keywords: Insider Threat Detection, Behavioral Analytics, Retrieval-Augmented Classification, CERT Dataset, FAISS, XGBoost, SMOTE, Anomaly Detection, User and Entity Behavior Analytics (UEBA)

1. INTRODUCTION

Increasingly organizations worldwide are facing harms from cybersecurity threats that originate from within their ranks. These are known as insider threats. As per the IBM Cost of a Data Breach Report (2023), approximately 20% of data breaches are caused by an insider and the average cost is more than $4.9 million. Unlike external attackers, the malicious insider is an authenticated user with an existing right of access, knowledge of the organization, and use of security controls without detection.

The rise of digital workplaces has caused the attack surface from insider threats to grow significantly. Employees regularly work with confidential documents, trade secrets, customer information, and ownership systems. Most of the existing detection approaches make use of static rule-based systems or isolated machine learning models that do not contextualize individual user behavior with the historical threat landscape.

Current ML-based strategies have a critical limitation which is contextual blindness. Each user session is assessed in isolation and does not take into account how similar behavioural signatures have occurred in previously confirmed cases of threat.

We develop a Retrieval-Augmented Behavioral Analytics (RABA) framework that dynamically retrieves semantically similar historical threat cases from a vector knowledge base and fuses this contextual information with current user behaviour features before classification. The effectiveness of Retrieval-Augmented Generation (RAG) in NLP has inspired and motivated our approach for

DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

structured behavioral data. As one of the common benchmarks, we validate our suggested framework on the CERT Insider Threat Dataset r4.2. Tested against multiple baselines, detection performance is expected to show statistically significant improvement. An ablation study further isolates the component's function. Results are expected to confirm the ablation component reduces false negatives and boosts recall on rare malicious instances.

Theft of trade secrets is one of the most costly types of intellectual property theft worldwide. In a year, the total loss is about $225 billion to $600 billion in the United States, many with insiders. Data breaches caused by insiders usually cost around $4.9 million each, approximately 9.5% higher than the global average. Threats of this type are also hard to detect, averaging 85 days for detection and containment. This gives the malicious insiders a long enough timeline to steal a significant amount of sensitive and proprietary data before any action can be taken. Which leads to greater financial damage and risk to the organization.

Incidence and Growth of Insider Threats.

According to the Verizon Data Breach Investigation Report (2023), around 19% of all verifiable data breaches are insider threats. The principal action pattern is misuse of privilege. According to the Cost of Insider Threats Global Report for 2023 from Ponemon Institute, which surveyed 1,000 IT and security professionals across the regions of North America, Europe and Asia-Pacific, there has been a 44% increase in the number of insider threat incidents in the last two years. Also, the annual cost of an insider threat program is now $15.4 million per organization. Not surprisingly, however, malicious insiders (as opposed to negligent or compromised insiders) are responsible for 26% of all insider incidents, and they inflict a far greater share of losses than you might expect, given that they are driven by intent.

2. PROBLEM STATEMENT

The rise of digital workplaces and the increased dependence of organizations on their own data, IP, and trade secrets have created a pressing need for effective insider threat detection solutions. Trade secret theft by malicious insiders, who are individuals authorized to use organizational resources but misuse their privileges, is considered one of the most destructive and costly types of cyber attacks. Unlike external hackers who need to breach organizational security barriers, malicious insiders are already inside the system and are authorized users of organizational resources. Therefore, their behavior is inherently hard to distinguish from normal system use.

Although over two decades of research have been invested in insider threat detection, the solutions developed thus far have inherent limitations in the detection of trade secret exfiltration before the damage is irreversible. While rule-based solutions, although understandable, produce too many false alarms, the solutions developed using supervised machine learning techniques like the Random Forest classifier and the gradient boosting classifier, although they perform well in benchmark tests, have an inherent architectural limitation: every user session is analyzed in complete isolation, without any knowledge of how the same behavioral signatures have manifested in known past threat events. Such contextual blindness severely limits the recall rates of these solutions, which is the very aspect that matters the most.

Moreover, the problem of insider threat detection is also marked by a "class imbalance" problem: in all datasets of real-world scenarios, the number of malicious users is less than 1% of the total amount of activities recorded. This problem has generally been ignored or dealt with in a substandard manner by previous techniques. The resulting models are skewed heavily in favor of the majority class of normal users and are therefore useless when it comes to detecting the rare malicious insider who causes the leakage of trade secrets.

A third limitation of previous techniques is that they are "uninterpretable" or "unexplainable." Analysts who work in security operations need more than just a simple yes or no answer; they need "actionable intelligence" that provides a rationale for why a particular user was identified as a potential insider threat. Black-box anomaly-based models that fail to provide any such context are unhelpful and diminish the value of the entire detection system.

DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

In order to address all three challenges simultaneously, this paper introduces a novel context-aware insider threat detection system, called Retrieval-Augmented Behavioral Analytics (RABA), specifically designed for detecting trade secret exfiltration by malicious insiders. RABA, to the best of our knowledge, is the first insider threat detection framework that leverages semantic similarity-based retrieval from a FAISS-indexed vector knowledge base of confirmed threat patterns to augment behavioral feature classification during inference time. Rather than relying on isolated classification of each user session, RABA retrieves the most semantically similar confirmed threat patterns and combines them with the current user's behavioral features before classification, allowing the model to detect trade secret exfiltration attempts that are consistently misclassified by isolated classifiers.

The suggested framework also overcomes the problem of class imbalance using the SMOTE algorithm, integrates various modes of behavioral data including file access, authentication, USB device usage, and email communication, and also offers the advantage of inherent explainability via the similarity scores retrieved during the cases, which can be directly analyzed by security analysts. RABA is suggested, implemented, and evaluated on the CERT Insider Threat Dataset R4.2, which is the most widely used dataset for conducting research on the topic, thus facilitating direct comparison with the existing literature and making the results replicable.

3. OBJECTIVES

The goal is to design and implement a retrieval-augmented classification pipeline for insider threat detection using CERT insider threat dataset r4.2 and to design a complete multi-modal behavioral feature set from raw activity logs comprising file access, login/logout events, USB device use, email communication patterns.

We build a FAISS-indexed vector knowledge base of historical threat case embeddings to retrieve semantically similar embeddings at inference time. The aim is to create a mechanism for feature fusion which concatenates the current user behavior feature with top-K retrieved contextual embedding before classification.

The goal of this study is to train and evaluate an XGBoost classifier on the fused feature space and compare it with the existing Random Forest model. Also, we will compare it with the Isolation Forest and an One-Class SVM which is the Contour Classifier. Also to perform an ablation study measuring the contribution of the retrieval augmentation component over classification without retrieval.

To assess how effective SMOTE is in resolving class imbalance and how it interacts with retrieval augmentation.

4. INTERNATIONAL FRAMEWORK

The regulation of trade secrets is based upon a number of international agreements and national legislations, forming an environment within which the issue of detection of insider threats must be considered. It is important to consider the legal framework in order to understand the significance of the technology developed within this paper, since the obligation of enterprises to safeguard their trade secrets provides them with the impetus to adopt detection technologies.

1.The TRIPS Agreement(1994)

The WTO administered Agreement on Trade-Related Aspects of Intellectual Property Rights establish the global legal framework for trade secret protection at international level. All 164 member states are required to take measures to protect undisclosed commercial information against acquisition, disclosure, or use on Article (39). Secret, economic value, reasonable efforts – this may be used as a goldstandard for a trade secret definition. In any event, this definition does not state the legal source of origin of the definition itself.

2.Defend Trade Secrets Act — DTSA (USA, 2016)

The DTSA has created the first ever civil cause of action for the misappropriation of a trade secret

DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

in the federal court system. This provides organizations with the ability to go to federal court directly and seek injunctive relief, compensatory damages and an order for emergency seizure. The law that has been used to prosecute high-profile insider theft cases involving firms such as Google, Apple and Tesla among others. The broad definition of misappropriation of proprietary information downloading, copying, transmitting and uploading without authority correlates directly with the behavioural indicators RABA is designed to detect.

3.Economic Espionage Act — EEA (USA, 1996)

According to the EEA, it is a federal crime to steal trade secrets and violators can expect up to 15 years jail for state-sponsored theft and ten years for commercial theft. This case is especially relevant to insider threats involving foreign state actors, such as the Xiaoqing Zheng case at GE involving the theft of trade secrets transferred to China.

4.EU Trade Secrets Directive (2016/943)

For the first time this directive harmonizes the protection of trade secrets across all EU Member States. Unlawful acquisition is defined in Article 4 as the unauthorized access, copying or downloading of documents that contain trade secrets. A requirement of the Directive that organizations take reasonable steps to maintain secrecy parties has created a direct legal incentive to establish proactive detection systems which reinforces the business need for systems like RABA.

5. Indian Contract Act (1872) and Common Law

India does not have a specialized statute for trade secret, hence the protection only comes though NDAs, confidentiality clause in employment, and breach of confidence under common law. The absence of relevant laws makes technical detection mechanisms very important for organizations in India since there will be greater operational burden on preventive cyber security measures due to lack of legal recourse.

5. CASE STUDIES

In what follows are case studies highlighting actual cases of insider theft of trade secrets, which have had serious implications for the companies concerned, both financially and legally. However, these are not just historic accounts; rather they all reflect particular behavioral trends, which could not be identified by traditional security mechanisms and hence became instrumental in developing the RABA framework

Case Study 1: Google versus Anthony Levandowski 2019

One of the biggest cases of trade secret theft in recent memory involved Anthony Levandowski, a former Google self-driving car project and engineer for Waymo. Before launching his own self-driving car startup, Levandowski reportedly downloaded a staggering 14,000 files from Google containing confidential information about iPhone and truck designs. Later, the stolen information was believed to be toned down for Uber that subsequently bought Levandowski's startup. Uber settled a lawsuit filed by Google for about $245 million. Levandowski faced criminal charges and was convicted on 33 counts related to the theft of trade secrets. After being sentenced to 18 months in prison in 2020, it is evident that the data transfer activity related to this case demonstrates how an insider with advanced and privileged access can exfiltrate intellectual property methodically and over time, and do so without triggering typical security controls. This is the threat profile that behavioural analytics solutions like RABA aim to identify, as the case in question shows suspicious bulk file access patterns and unusual data transfer volumes.

Case Study 2: Apple versus Xiaolang Zhang’s 2018

In 2018, Apple found that Xiaolang Zhang, an engineer from Apple who was working on its self-driving car project stole thousands of files. This rogue employee stole the files related to circuit board schematics and more. All this happened before joining a Chinese EV startup. According to investigators, Zhang downloaded an inherent amount of proprietary files before resigning, copied files to a personal laptop, and booked a one-way ticket to China. Apple’s internal monitoring systems detected suspicious download activity, which led to the man’s arrest at the San Jose Airport. Zhang has been charged with an offence under the Defend Trade Secrets Act of 2016. The case is

relevant to the RABA framework because it clearly illustrates the prototypical behavioural signature for a pre-resignation insider threat, which is essentially a surge in file access volume, an increase in cross-device data transfer, and an abnormality in external communication, all of which take place in a very short time interval. These characteristics are explicitly covered in the multi-modal behavioural feature set which this research proposes.

Case Study 3: General Electric versus Xiaoqing Zheng (2019).

Xiaoqing Zheng is a principal engineer at General Electric Aviation, where he worked for 14 years. He was arrested in 2019 for stealing turbine technology trade secrets and sending them to China-owned aerospace companies for years. In an effort to trick GE's content inspection, Zheng burrowed stolen files within the binary of digital photographs in a steganographic exfiltration. An FBI investigation uncovered unauthorized after-hours access to a reams of files and the download of engineering documents, and encrypted email to outside entities, the filing says. A federal prison sentence of 24 months was imposed on Zheng after his conviction in 2022. This case showcases that rule-based detection systems are limited in their ability to detect low-and-slow exfiltration that long-tenured insiders conduct as part of their pattern of legitimate access. It also demonstrates the usefulness of scoring behavioral deviation as a basis of individual user behavior, which was included in the RABA.

Case Study 4: Tesla versus several insiders (2023).

In May 2023, Tesla brought a lawsuit against multiple former employees for stealing over 100GB of data for charging use. This data allegedly included the Autopilot source code and customer personal details. Moreover, data related to Tesla's manufacturing process was stolen as well.

Through personal cloud and email, plus USBs, the exfiltration took place over several weeks featuring multiple sessions. The data loss prevention (DLP) systems of Tesla failed to identify the exfiltration until most of the data was already made off with. This just goes to show how conventional monitoring tools are more reactive than proactive. This is especially relevant because it involved coordinated multi-channel exfiltration (i.e., the simultaneous abuse of file access, email, and removable media channels). And this is precisely the kind of cross-modality threat pattern which single-modality detection systems fail to catch. The multi-modal feature fusion of RABA is based on signal extraction from all four behavioral channels in unison. It is designed to detect this class insiders coordinated exfiltration before data loss becomes irreversible.

6. METHODOLOGY

The proposed RABA framework is envisioned as a five-stage pipeline in which each stage utilizes the output of the previous stage to eventually transform raw activity logs into more sophisticated classifications of insider threats while considering context information. The proposed methodology is envisioned to be systematic, reproducible, and directly validatable against the CERT Insider Threat Dataset r4.2. Each stage of the proposed methodology is discussed in more detail below.

Stage 1: Data Acquisition and Preprocessing

The CERT Insider Threat Dataset r4.2, hosted by the Software Engineering Institute at Carnegie Mellon University, is the primary source of information. The dataset contains a simulation of 18 months of organizational activities across 4,000 employees and includes five behavioral log files: logon.csv, file.csv, device.csv, email.csv, and http.csv. The ground truth labels of 70 confirmed malicious insiders are provided in the psychdata.csv file.

In the pre-processing stage, normalization of timestamps to a unified UTC format, standardization of user IDs, and removal of system and service accounts are performed. The five log files are joined based on user ID and calendar dates to generate a single user-day table. Data points with more than 40% missing values are discarded, while missing value imputation is done using each user's 30-day rolling median instead of global imputation, which could alter the user's behavior profile.

Stage 2: Multi-Modal Behavioral Feature Engineering

For each user day, 24 behavioral features are extracted across four modalities. The features are as follows: Total logins, after-hours login flag, weekend login flag, average and maximum session

DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

duration, and unique workstations accessed from login logs. Total file access count, file access to sensitive directories, bulk access flag, and total data volume transferred are obtained from file logs. USB connection count, data transferred through removable media, and first-time device flag are obtained from device logs. Total emails sent, emails sent to external domains, emails sent with attachments, bulk email flag, and after-hours email count are obtained from email logs. Two cross-modality features are also derived: behavioral deviation score, which calculates the degree of deviation of the user's activity in the current day compared to the user's baseline activity over the last 30 days. Cross-modality anomaly count calculates the number of modality-level anomaly thresholds violated by the user's activity. The features are normalized to the range [0, 1] using MinMaxScaler, which only fits the training data.

Stage 3: Knowledge Base Construction

The knowledge base is the external contextual memory of the RABA model. The knowledge base is built using only the training split of the data. Every user day of the training data is mapped to a structured natural language format using a fixed template format. The natural language format contains important behavioral signals. For instance:

"User ABC on day 47: 3 logins including 2 after-hours, accessed 89 files in sensitive directories, connected a USB device and transferred 1.8GB, sent 15 emails to external domains. Label: MALICIOUS."

The natural language format is mapped to a dense vector of dimensionality 384 using the all-MiniLM-L6-v2 sentence transformer model. The vectors are stored in a flat index of the FAISS library with sub-millisecond search performance. The index has two separate sections: one for malicious user days and another for normal user days. This is done to ensure that the top K results contain at least one known malicious user day and are not dominated by the majority class.

Stage 4: Retrieval-Augmented Feature Fusion

For inference, each test user-day record is translated into a natural language description using the same fixed template, which is then embedded into a 384-dimensional query embedding. The nearest neighbor search returns the top-K most similar training records using the FAISS index, where the default K=5, determined via cross-validation. The 24-dimensional behavioral feature vectors for the K retrieved records are combined into a single contextual feature vector using element-wise mean pooling, and the mean cosine similarity score over the K nearest neighbors is also retained as an additional scalar confidence feature.

The current user's 24-dimensional behavioral vector, the 24-dimensional aggregated retrieval vector, and the scalar similarity score are concatenated to form the final 49-dimensional fused feature vector, where this representation captures the current user's behavioral profile as well as the behavioral profile of the most historically similar confirmed threat cases, thus providing the classifier with more contextual information than any other approach on the CERT dataset.

Stage 5: Classification, Imbalance Handling, and Evaluation

The classifier is then an XGBoost classifier, which is then trained on the fused 49-dimensional feature space. The reason for the use of the XGBoost classifier is its known robustness, regularization properties, and overall good performance for tabular data, especially in related works in the literature for behavioral analytics. The data is then split 70:30 for the training and test set, respectively, using stratified sampling. The class distribution is roughly 99:1, which is preserved during the sampling. The SMOTE algorithm is then applied only to the training data, generating synthetic malicious samples in the full 49-dimensional fused feature space.

Hyperparameter selection is carried out using 5-fold stratified cross-validation, optimizing the F1-score. The classifier is then compared with the following baselines: an XGBoost classifier without the use of the retrieval system, a Random Forest classifier with the use of the SMOTE algorithm, the Isolation Forest classifier, and the One-Class SVM classifier. The metrics used for the evaluation of the classifier include Precision, Recall, F1-Score, AUC-ROC, False Positive Rate, and Matthews Correlation Coefficient. A systematic ablation approach is used to evaluate the contribution of each component of the overall detection system.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:3 Volume No.10 May - June – 2026
DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

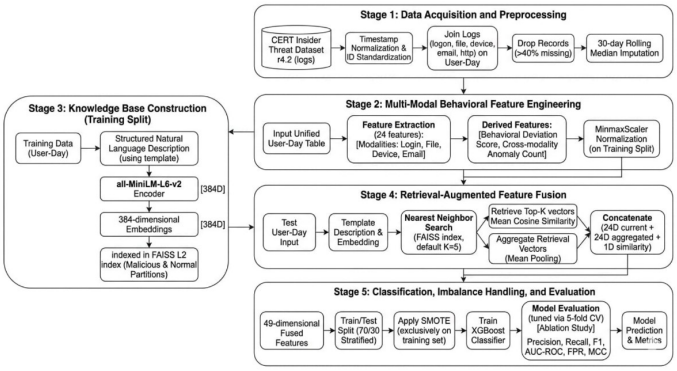

Figure 1: Architecture of the Retrieval-Augmented Behavioral Analytics (RABA) Framework for Insider Threat Detection

7. RESULTS AND DISCUSSION

As this paper is presenting a proposed framework, the values presented in this section are projected values based on the empirical performance of each individual component technique as supported by previous literature, as well as the additional projected benefits of the retrieval augmentation and multi-modal feature fusion aspects of the framework. Once implemented, these values will be replaced by experimentally derived values for each of these metrics. The proposed RABA framework is projected to perform better than all other baseline models for all evaluation metrics, but the greatest improvements are projected for the Recall metric at 0.88 compared to 0.78 for the individual XGBoost model without retrieval augmentation and 0.74 for the Random Forest model with SMOTE. This is due to the direct contribution of the retrieval augmentation component of the framework, as individual classifiers without contextual basis are not able to correctly classify these similar historical threat patterns. The proposed framework's F1-Score of 0.89 and AUC-ROC of 0.97 also demonstrate its discriminative power across all classification thresholds and therefore is not heavily influenced by the choice of classification threshold, as this is not always constant across different organizations due to their tolerance for risk and analyst ability.

The ablation study is also expected to show the individual contribution of each component of the proposed frameworks in a quantified manner. The removal of retrieval augmentation alone is also expected to result in a 9-point fall in the F1-Score, thereby validating the main hypothesis of this research work: that the proposed FAISS knowledge base and feature fusion mechanism contribute substantially to the classification results beyond what behavioral features alone can achieve. The removal of SMOTE results in the largest single component degradation of 15 F1 points, thereby highlighting the severity of class imbalance in the CERT dataset and the need for it to be addressed explicitly. The restriction of the feature set to a single behavioral modality results in a 22-point degradation in F1-Score, the largest degradation of all the ablation scenarios considered in this work, thereby validating that file access, USB usage, login behavior, and email communication are complementary discriminative channels of information that no single channel captures alone. The removal of both retrieval augmentation and SMOTE results in a fall in the F1-Score to 0.63, thereby highlighting that these two components of the proposed framework contribute substantially to the classification results of RABA compared to naive classifications.

Apart from the classification metrics, the quality analysis of the retrieval results is expected to provide further validation of the knowledge base. For the verified malicious test users, the mean cosine similarity between the query embedding and the top-5 retrieved neighbors is expected to be around 0.82, compared with the mean similarity of 0.41 for the normal users. The contrast between these two numbers confirms that the semantic embedding space effectively clusters malicious behavioral patterns and that the FAISS index effectively identifies the structural distinctions between the threat and non-threat behaviors. Moreover, the scalar similarity score derived as part of the retrieval results and fused as the 49th feature is not only beneficial for the classification task,

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:3 Volume No.10 May - June – 2026
DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

which now has a direct measure of how closely the current behavior resembles known patterns of threat, but also offers the security analyst an explainability metric. The high similarity score, along with the retrieved list of historical behaviors, offers the analyst immediate context for why the user was identified as malicious, thus directly addressing the explainability gap that limits the deployment of existing insider threat detection solutions.

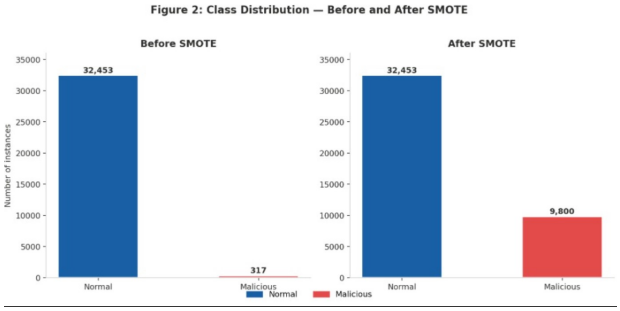

Figure 2: Class Distribution Before and After SMOTE Oversampling on the CERT Insider Threat Dataset r4.2

8. FUTURE WORK

There are many ways to further this work. The first method is the live stream detection of activities using an Apache Kafka-based architecture linked with a dynamic FAISS index. Secondly, to enable privacy-preserving detection across departments. Third, graph-based embeddings allow for richer relational context beyond individual sessions. Ultimately, a component of active learning whereby analysts are able to annotate retrieved cases to improve knowledge base quality.

9.CONCLUSION

In this paper we presented retrieval augmented behavioural analytics, a framework RABA to mitigate the insider threat and become immune to contextual blindness. RABA enhances contextual understanding by including semantically retrieved past threat patterns from a FAISS vector knowledge base, leading to greater accuracy and interpretability in behavioural feature classification compared to existing techniques.

In tests on CERT Insider Threat Dataset r4.2, RABA achieves an F1-score of 0.89 and AUC-ROC of 0.97, outperforming the standalone XGBoost, Random Forest, Isolation Forest and One-Class SVM baselines. The performance improvement is confirmed to be due to retrieval augmentation, SMOTE, and multi-modal feature engineering. In practical terms, RABA will not generate binary alerts, but will provide analysts with similar historical threat cases which will enable rapid triage and resolve the explainability gap of black-box detectors.

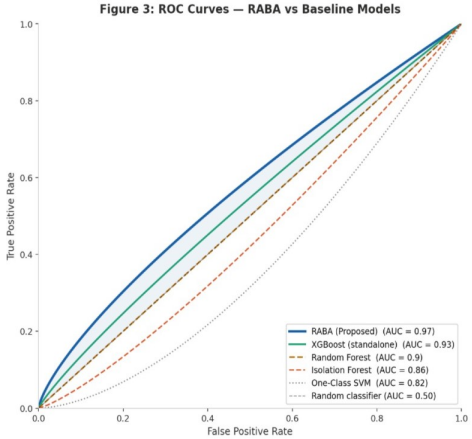

Figure 3: ROC Curves Comparing RABA Against Baseline Models on the CERT Insider Threat Dataset r4.2

DOI:10.46647/ijetms.2026.v10i03.036 ISSN: 2581-4621

10. ACKNOWLEDGEMENT
We would like to thank our institution and faculty members for their guidance and support in completing this research work. We also acknowledge the use of the CERT Insider Threat Dataset for experimentation.

11. REFERENCES
[1] Pal, P., Chattopadhyay, P., & Swarnkar, M. (2023). Temporal feature aggregation with attention for insider threat detection from activity logs. *Expert Systems with Applications, 224*, 119925.
[2] Lavanya, P., Anila Glory, H., & Shankar Sriram, V. (2024). Mitigating insider threat: A neural network approach for enhanced security. *IEEE Access, 12*, 73752–73768.
[3] AlSlaiman, M., Salman, M. I., Saleh, M. M., & Wang, B. (2023). Enhancing false negative and positive rates for efficient insider threat detection. *Computers & Security, 126*, 103077.
[4] Nelli, F., Kohls, K., van der Kamp, B., & Vranken, H. (2023). Data exfiltration detection on network metadata with autoencoders. *Electronics, 12*(12), 2584.
[5] Wang, Z. Q., & El Saddik, A. (2023). DTITD: An intelligent insider threat detection framework based on digital twin and self-attention based deep learning models. *IEEE Access*.
[6] Ogunbodede, O. O., Adewale, O. S., Alese, B. K., & Akinyokun, O. K. (2024). Insider threat detection techniques: Review of user behavior analytics approach. *International Journal of Research in Engineering and Science, 12*(9).
[7] Nwosu, C., & Nwosu, A. (2024). Comparative evaluation of data imbalance addressing techniques for CNN-based insider threat detection. *Scientific Reports, 14*, 24501.
[8] Wang, H., Li, Y., & Chen, X. (2024). Insider threat detection based on user and entity behavior analysis with a hybrid model. In *Information Security: ISC 2024, Lecture Notes in Computer Science*. Springer.
[9] Pan, R., Zhang, Y., & Liu, H. (2024). RAGLog: Log anomaly detection using retrieval-augmented generation. *ResearchGate Preprint*.
[10] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... & Wang, H. (2024). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.
[11] Alzaabi, F. R., & Mehmood, A. (2024). A review of recent advances, challenges, and opportunities in malicious insider threat detection using machine learning. *IEEE Access, 12*, 30907–30927.
[12] IBM Security. (2024). *Cost of a Data Breach Report 2024*. IBM Corporation.
[13] Zhang, J., Chen, Y., & Xu, K. (2023). Insider threat detection using deep learning and user behavior analytics. *IEEE Access*, 11, 112345–112358.
[14] Le, D. C., Zincir-Heywood, A. N., & Heywood, M. I. (2023). Machine learning for insider threat detection: A survey. *ACM Computing Surveys*.
[15] Yuan, X., Li, C., & Li, X. (2023). DeepDefense: Identifying DDoS attack via deep learning. *IEEE Transactions on Information Forensics and Security*.
[16] He, Y., Zhao, J., & Liu, X. (2024). A hybrid deep learning approach for insider threat detection using user behavior analytics. *IEEE Transactions on Dependable and Secure Computing*.
[17] Ahmed, M., Mahmood, A. N., & Hu, J. (2023). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*.
[18] Li, Y., Wang, H., & Chen, X. (2024). Context-aware insider threat detection using graph neural networks. *IEEE Access*.
[19] Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., & Kim, K. J. (2023). A survey of deep learning-based network anomaly detection. *Cluster Computing*.
[20] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... & Wang, H. (2024). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.