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 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103

Insider Threat Detection System Using Machine Learning

Mr. M.V. Prabhakaran[1], Naveen A[2], Saran R[3]

HOD, Computer Science Engineering (Cyber Security),

Dhanalakshmi Srinivasan College of Engineering and Technology, Chennai, Tamil Nadu - 603 104[1]

Computer Science Engineering (Cyber Security),

Dhanalakshmi Srinivasan College of Engineering and Technology, Chennai, Tamil Nadu - 603 104[2,3]

Abstract: Insider threats are one of the major challenges in securing an organizational network since the insider possesses proper access authority, making it hard to trace their malicious activities using traditional security measures. Currently, most insider threat detection systems are based on supervised learning methodologies, which demand a lot of labeled data, most of which tends to be imbalanced. To tackle these problems, this research work will employ a hybrid insider threat detection model that combines Isolation Forest with temporal behavior profiling and with random forest algorithm for classification. The proposed solution is based on simulating normal user activity and identifying irregularities, which could be associated with insider attacks. As opposed to conventional solutions, which rely solely on static variables, this solution uses temporal behavioral variables, including access rate, session time, abnormal system activity during offline sessions, and sudden changes in user activity patterns. The Isolation Forest algorithm is leveraged to identify abnormal activity at the algorithmic level without relying on resampling, thus mitigating overfitting issues and distorting the data. In this paper, the proposed approach has been tested using the CERT insider threat dataset, known to be an extreme class imbalance. Results show that the hybrid approach of the model, which combines the benefits of temporal profiling with the strengths of anomaly-based approaches, greatly increases the accuracy level while at the same time ensuring low false positives. This indicates that the model can be said to be quite robust.

Keywords: Insider threat detection, anomaly detection, Isolation Forest, Random forest, behavioral analysis, temporal profiling, imbalanced dataset, framework.

INTRODUCTION

Currently available insider threat detection models are mostly based on supervised machine learning models that need a lot of data to be properly labeled as insider threats and normal activities. But properly labeling insider threat data is not only quite expensive but also difficult to implement in practical settings. Also, insider threat data is typically imbalanced, as malicious events usually happen very rarely as opposed to normal activities. It is found that the performance of traditional machine learning models is significantly harmed when dealing with imbalanced data, resulting in higher rates of false positives. To deal with these, there has been an increasing interest in behavior-driven and anomaly-based methods that are capable of capturing the normal user behavior and pinpointing the unusual patterns, which might point to potential danger. Specifically, unsupervised learning algorithms, such as the Isolation Forest, have shown promising results in identifying anomalies within high-dimensional and imbalanced datasets without the need for resampling and intense labeling (Liu et al., 2008). Nonetheless, the methods used are still highly dependent on static features, disregarding the temporal patterns that are essential for the identification of sophisticated malicious activities by insiders. With these considerations in mind, this paper advances a hybrid model of insider threat detection, combining Isolation Forest and Random Forest with temporal behavior profiling. The proposed approach uses temporal characteristics of user behavior, including frequency of accesses, duration of user sessions, unusual system usage by an individual outside working hours, and sudden changes in user behavior, to model user behavior that varies with time. The Isolation Forest technique can be used for detecting unusual user behavior by deriving user behavior probability distributions, while Random Forest can be employed to classify identified threats into particular categories of insider threats.The performance of the new approach has been analyzed by employing the CERT insider threat dataset, which has been acknowledged for its extreme class imbalance problem and accuracy in simulating real-life insider threat behavior. The experimental analysis has revealed that by employing temporal behavior profiling and anomaly detection techniques, the detection accuracy has been significantly improved while maintaining a low false positive ratio. This indicates that the hybrid approach provides a robust and efficient way for detecting insider threats in real-life organization setups.

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 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103

LITERATURE REVIEW

Alketbi & Mehmood (2025): This survey review Explainable AI techniques are applied to insider threat detection and highlight their role in improving transparency and analyst trust. It identifies challenges such as the lack of standardized explainabi lity metrics and accuracy–interpretability trade-offs. **Dong et al. (2025) in this study** DDCC combines diffusion models with curriculum learning to capture insider behavior from simple to complex temporal patterns. The approach improves detection of both short-term anomalies and long-term malicious activities on CERT datasets. **Xiao et al. (2025): This author introduced** SENTINEL models insider behavior using multi-timescale interaction graphs and graph neural networks Notably, it can recognize abrupt and gradual threats even without labels and scales well. Kim et al. introduced a method for combining log analysis via natural language processing and reinforcement learning for adaptive host-based insider attack detection, Kim et al. (2025). It performs effectively in real-time environments with evolving attack patterns. **Kong et al. (2025) analysis of this review paper** DPI-ITD proposes a lightweight insider threat detection framework for IoT systems using symbolic tagging and GloVe embeddings. It achieves high accuracy with low computational overhead in resource-constrained environments. **Alzaabi & Mehmood (2024): This** survey reviews classical ML, deep learning, NLP, and LLM-based approaches for insider threat detection. It highlights challenges such as class imbalance, limited labels, and evolving user behavior. **Al-Shehari et al. (2024) t**his study applies a Density-Based Local Outlier Factor approach to insider threat detection without data resampling. High F1-scores on CERT datasets demonstrate its effectiveness in detecting rare malicious behaviors. **Zhu et al. (2024): This author describes AUTH, which** uses adversarial autoencoders with temporal convolution and LSTM to detect insider threats in an unsupervised manner. Adversarial learning improves robustness and detection of stealthy attacks. **Wang et al. (2024)** FedITD combines federated learning with large language models to enable privacy-preserving insider threat detection. It achieves performance comparable to centralized models without sharing raw data.

Al-Shehari et al. (2023) This work demonstrates the effectiveness of Isolation Forest for insider threat detection on highly imbalanced datasets. High accuracy is achieved without requiring labeled attack data. **Wang & El Saddik (2023)** This framework models user behavior using digital twins and transformer architectures for long-sequence analysis. NLP-based augmentation mitigates data imbalance while maintaining high detection accuracy. **Mehmood et al. (2023)** This study detects insider privilege escalation in cloud systems using supervised ensemble classifiers. LightGBM outperforms other models in accuracy and robustness. **Rahman et al. (2022)** This paper focuses on unintentional insider threats using a NARX-based predictive model. It enables proactive risk assessment by forecasting risky behavior patternsAhmadi-Assalemi et al. (2022) In this study, the use of an ensemble model of one-class classifiers for anomaly detection has been proposed. This model has been highly accurate with less computational complexity.**Nasir et al. (2021)** This work applies deep learning to capture temporal user behavior patterns for insider threat detection. It outperforms traditional machine learning approaches in detection accuracy. **Le et al. (2020)** This study shows that fine-grained user activity data significantly improves early insider threat detection. Temporal behavior monitoring reduces false positives and enables faster response.

The overall survey introduced the below techniques and algorithm such as state-of-the-art works on insider threats use diverse approaches, some of which are traditional machine learning models such as RF, LightGBM, unsupervised anomaly detection methods such as Isolation Forest, density-based outlier detection algorithms, deep models using LSTMs, Transform models, and diffusion models, graph neural networks to design user-resource interaction models, reinforcement learning to design dynamic models for detection rules, federated learning models to promote collaboration in insider threats in an organization without exchanging data, explainable AI models such as SHAP values, LIME to enhance model trustworthiness, behavior models based on time to design models for insider threats that vary in time, ensemble models to design models for imbalanced datasets related to insider threats.

Research gaps in existing insider attack detection solutions primarily target anomaly detection, failing to offer attack type information that would be useful for incident handling. Most sophisticated solutions are not explainable, deployable, or computationally complex. Existing solutions also do not efficiently address behavior drift for longer periods of attack scenarios or minimize false positives. Nonetheless, the extensive utilization of synthetically produced CERT datasets is raising concerns over generalization capability for real-world scenarios, implying the need for a lightweight, explanations-providing, and attack-aware insider attack detection solution.

Contribution of the paper:

The proposed model implements a hybrid insider threat detection by fusing unsupervised anomaly detection with supervised attack classification. First, temporal behavior profiling is utilized to model changing user activities and detect anomalies without reliance on labeled data. Next, the outliers that are detected will be classified into various categories of insider attacks by means of a Random Forest classifier, which has been selected considering its resistance to the large

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 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103

dimensional nature of the behavior features. Provide a module for explainable results by means of SHAP, which can help in understanding the inference process. Incorporate an adaptive mechanism for suppressing false positives.

METHODOLOGY

SYSTEM PREMILINARIES:

1. Insider Threat

An insider threat occurs when a legitimate user misuses their authorized access, either intentionally or unintentionally, to harm an organization. Since insiders already have valid credentials, their malicious actions are difficult to detect using traditional security tools.

2. System Assumptions

The proposed system assumes that the user activity logs are being collected continuously, and most of the user activity is normal, while insider attacks are rare. Insider activity patterns vary over time, so temporal patterns are important for insider attack detection. There is a lack of labeled attack data, particularly in a real-world setting.

3. Dataset Used

The proposed framework is tested using the CERT Insider Threat Dataset, which models real-world enterprise user behavior such as login, file access, computer use, and email activity. The dataset is very imbalanced and represents a real insider threat attack scenario.

4. Behavioral Features

User behavior is modeled using time-series features like access frequency, session time, off-hours access, anomalous file access, and abrupt behavior changes. These features are useful in modeling the dynamics of user behavior over time.

5. Anomaly Detection

Isolation Forest is employed to detect anomalous user behavior without requiring labeled data. It is very effective on imbalanced datasets and quickly points out the suspicious regions of user behavior that may be indicative of insider attacks.

6. Threat Classification

After detecting anomalies, a Random Forest classifier is employed to classify them into various types of insider attacks. This allows security analysts to better understand the nature of the threat and act accordingly.

7. Explainability

To increase the interpretability of the system, SHAP is employed to provide explanations for why a particular user is marked as suspicious. It points out the most important behavioral features, which helps analysts build trust in the system.

8. Evaluation Metrics

The proposed system is evaluated based on accuracy, precision, recall, F1-score, AUC, and false positive rate, with a focus on minimizing false alarms.

9. System Flow

The proposed system consists of five major steps: data collection, profiling of temporal behavior, anomaly detection, threat classification, and explainable analysis.

SYSTEM ARCHITECTURE

The proposed insider threat detection solution is conceptualized as a multi-layer system that constantly tracks user behavior and detects insider threats with high accuracy and low false positives.

The initial part of the process begins with the gathering of raw data logs of user behaviour from the enterprise's resources. The sources of the raw log files are login attempts, access to files, device usage and network communication, and user activities collected in the course of a single day as they interact with the enterprise resource system.

After collecting the raw log files from the various enterprise resources, the data will be cleaned and pre-processed to remove the noise, missing data and inconsistencies. Feature sets will be extracted from the log files and presented in time windows so that we can monitor the behaviour of the users over time, rather than monitoring them by each individual action.

Once the normal behavior profiles of users have been built, temporal behavior profiling (stage 3) uses that information to identify a user's usual patterns of conduct on the system (e.g., hours worked, frequency of logins, length of login sessions, and types of things done within the system). Any activity that deviates from these usual patterns will be considered unusual and therefore suspicious if it occurs.

The Isolation Forest module utilizes behavior profiling to identify anomalous behavior. Because relatively few insider attacks have occurred, and labeled data is extremely limited, an unsupervised model is ideal for detecting abnormal user behavior and does not rely on knowing about previous intrusions.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

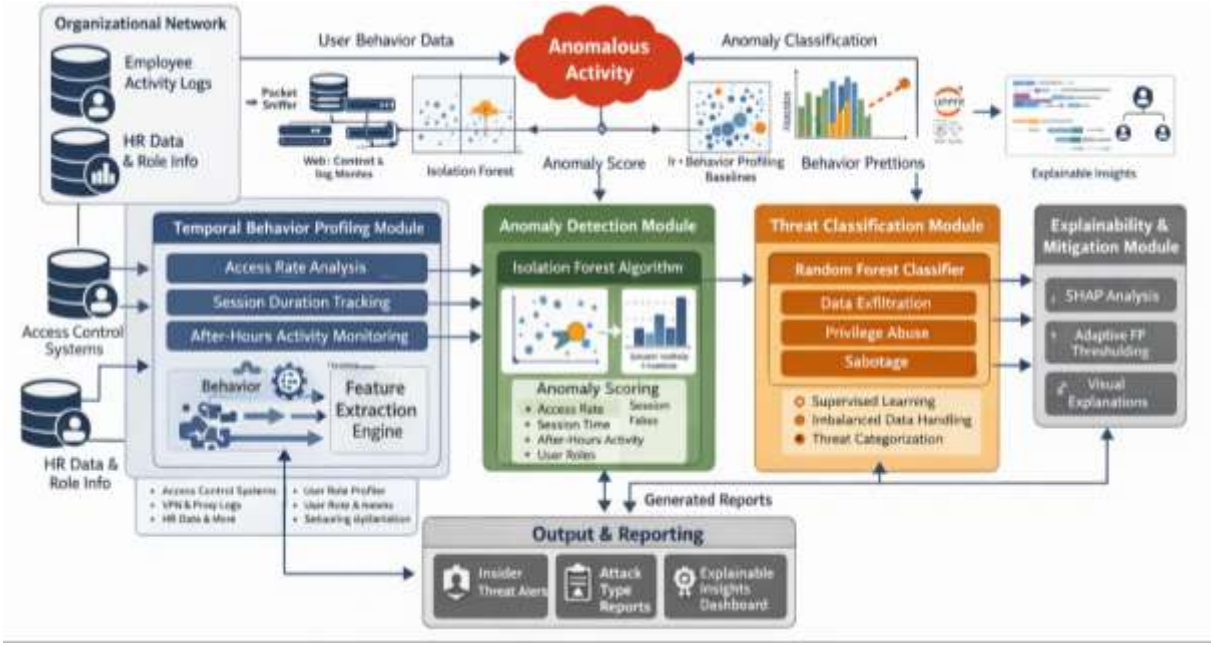
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103

After the Isolation Forest module detects anomalous behavior, it then passes that data on to the Random Forest Classification module so that each potentially suspicious indicator can be assigned to categorizations of insider threats. This step enables security personnel not only to know that a potential attack has taken place, but also what type of attack occurred.

An explainability layer based on SHAP in the architecture aims to create greater transparency in the process of flagging users or actions, thus improving trust and usability; this is accomplished through presentation of the most influential behavioural features, allowing users to understand why a user or action has been flagged.

Finally, the False-Positive suppressor is adaptive in nature; it will insert historical data/decisions into the context as well as the current state of affairs to help reduce the number of false-positive alerts. Therefore, the system can create alerts and reports that will help security administrators provide a more timely and informed response to incidents.



IMPLEMENTATION OF PROPOSED WORK

DEPLOYING THIS WORK

We follow a defined workflow when compiling user activity logs into relevant security analysis within our current project methodology. The workflow is made up of five primary steps: Data Prep, Modeling Temporal Patterns of User Behavior, Anomaly Detection/Classification of Threats, Classification of Threat Levels, and Provide an Explanation for the Results of Each of the Above Classes.

1. Data Acquisition and Preprocessing

Data is first acquired from the user's activity logs which contain logon/logoff records, file access events, device usage and web activity within the organization. Cleaning these logs includes removing missing values, duplicates and irrelevant fields.

2. Temporal Behavior Profile

Temporal features for each user were developed over time in order to define the temporal evolution of user behaviour over time.

3. Isolation Forest performs Anomaly Detection.

Model Isolation Forest takes the profiles of the temporal behaviour and classifies them as anomalous or normal using certain features.

4. Insider Threat Classifier uses Random Forest Classification.

The anomalous instances classified by Isolation Forest are passed to a Random Forest Classifier, which identifies the type of attack associated with each of the anomalous instances.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155103

This step enables the system to distinguish between different insider attack types rather than only flagging anomalies.

5. Explanation via SHAP.
To create greater transparency to model users, the System uses SHAP to provide justification for its models' decisions. SHAP outputs the contribution of each model feature to the model's end prediction.

6. Suppressing False Positives
Adaptive thresholds along with historical context will help minimize alert fatigue. For example, if similar activity has already been verified as safe, that would result in suppression of alerts generated from similar types of activity that occur in the future to reduce the occurrence of false positives.

7. Performance Evaluation
The system performance is evaluated using standard metrics:

8.Implementation Summary
To sum up, the proposed implementation combines behaviour modelling over time, with unfactored anomaly detection, plus factored classification. This hybrid methodology is effective at detecting rare incidents involving insiders; identifying their nature; and providing results that are explainable for use in the real world.

EXPERIMENTAL SETUP

Algorithm 1: Hybrid Detection of Insider Threats through Temporal Profiling
Input:
User activity logs (L) (logon, file access, device usage, web usage)
Time window size (W)
Isolation forest parameters
Random forest parameters
Output:
Detected insider threats with attack type and rationale
Step 1: Log Collection
Collect and store all raw user activity logs in a central repository.
Step 2: Data Preprocessing
At this stage, all raw logs will be cleaned, by removing missing values, duplicate records, and irrelevant fields. Categorical fields will be converted to numerical values and timestamps will be normalized.
Step 3: Temporal Window Creation
User activities will be divided into fixed time periods (W) (daily or weekly) so the evolution of each individual's behavior can be captured over time.
Step 4: Extract features
For each user (u) for each time period (t) extract all applicable temporal behavior features. That is:
Login frequency.
Session length.
Number of times files were accessed.
Ratio of off hours activity.
Behavior deviations from the previous period.
Form a behavior vector: Bu,t=[f1,f2,…,fn]
Step 6: Anomaly Detection - Isolation Forest
Training the Isolation Forest model using the profiles of normal behavior.
Compute anomaly score s(x) for each behavior instance.
If $s(x)>\theta$: Mark instance as **anomalous**
Step 7: Anomaly Filtering
All instances identified as anomalous should be collected and passed from this module to the classification module.
Step 8: Threat Classification - Random Forest
Train a Random Forest classifier on instances of labeled anomalous behavior. Classify each anomalous instance into an insider threat category using majority voting.
Step 9: Explainability - SHAP
Utilize SHAP on the Random Forest model to identify the most significant behavioral features contributing to each prediction made by the classifier.
Step 10: Suppressing False Positives

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 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103

Utilize adaptive thresholds and the historical context of the alert to suppress duplicate or lower confidence alerts.

Step11: Generating Alerts

Make a final alert that includes the following pieces of information:

- user ID
- type of threat detected
- anomaly score
- features explaining the prediction

Step 12: Performance Evaluation

To evaluate the system performance, use metrics such as precision, recall, F1-score, AUC, and false positive rate.

End of Algorithm

RESULTS AND DISCUSSION

Detection Performance

The Hybrid Model (Temporal Profiling + Isolation Forest + Random Forest) provided better performance than either the supervised or unsupervised system performed alone.

Detection Accuracy for Each Model:

Method	Accuracy	F1-Score	AUC
Supervised Model	0.78	0.65	0.74
Isolation Forest	0 .81	0 .72	0.81
Hybrid Model	0.89	0.80	0.92

Model Comparison Analysis:

1) The Hybrid Model's detection accuracy of 89% is the best of the three models.
2) The F1-Score also improved significantly with the use of the Hybrid Model, which indicates a better balance between precision and recall.
3) The Hybrid Model had the highest AUC value of .92 among all three models, which indicates that the Hybrid Model was able to create good separation between normal and malicious behavior.
4) Thus, combining anomaly detection with supervised classification provides a more robust overall detection system.

ROC Curve Analysis

The ROC Curve shows that The Hybrid Model had higher True Positive Rates (TPR) at every False Positive Rate (FPR) compared to the other models.

Key Observations:

The supervised model is limited by the amount of labeled attack data.

Isolation Forest was able to deal with the imbalance better than the other models.

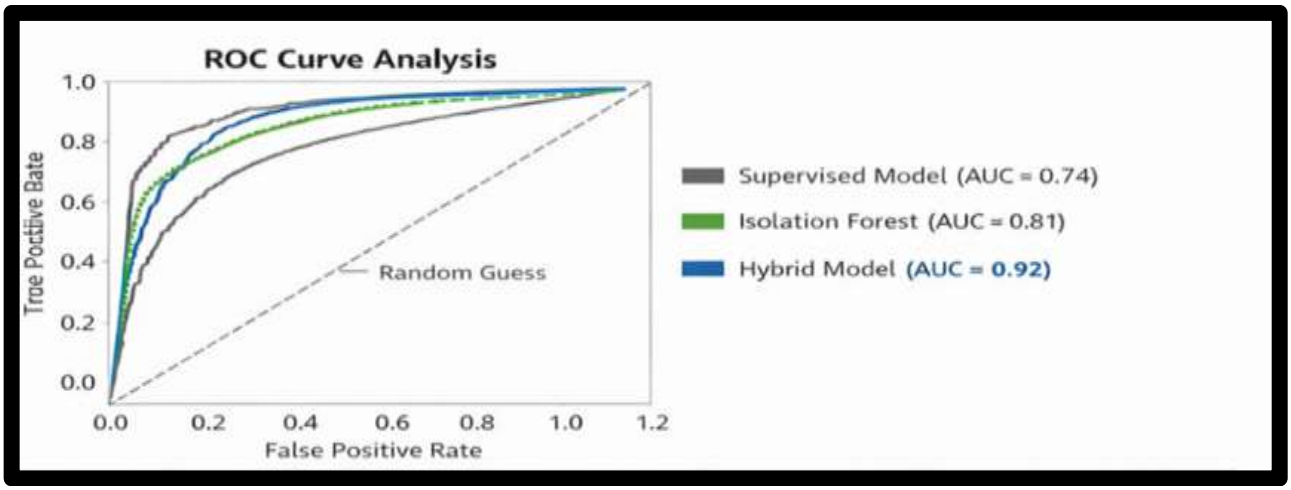
The Hybrid Model had the highest AUC value of .92, which demonstrates that it had the best discrimination ability.

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 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103



3. The Effectiveness of Temporal Behavioral Assessment

Experiments were conducted to demonstrate the contribution of temporal factors to the relevance of temporal assessments by using scoring methods both including and excluding temporal factors in assessing the significance of temporal factors.

Configuration	F1-Score
Excluding Temporal Factors	0.55
Including Temporal Factors	0.82

Summary:

The F1 score demonstrated a significant increase when including the use of temporal behavioral factors. The following temporal behavioral features facilitated the identification of subtle insider activity where features do not capture.

Analysis:

- Off-hours accessed
- Short-term behavioral changes
- Session duration length/sudden spikes

This demonstrates that insider threats will continue to be determined by behavioral patterns over time.

4. Reduction of False Positive Alerts

Before applying adaptive suppression.

- False Positive Rate was 12.4%.

After applying adaptive suppression.

- False Positive Rate was 4.1%.

Summary:

The use of the adaptive suppression alert filtering process demonstrated a reduction in false positive alerts by nearly a factor of 3.This will be essential in large organizational security operations where excessive false positive alerts lead to security analysts feeling overwhelmed by signal debilitation from focus while performing their duties.

A reduction in the presence of false positive alerts will produce:

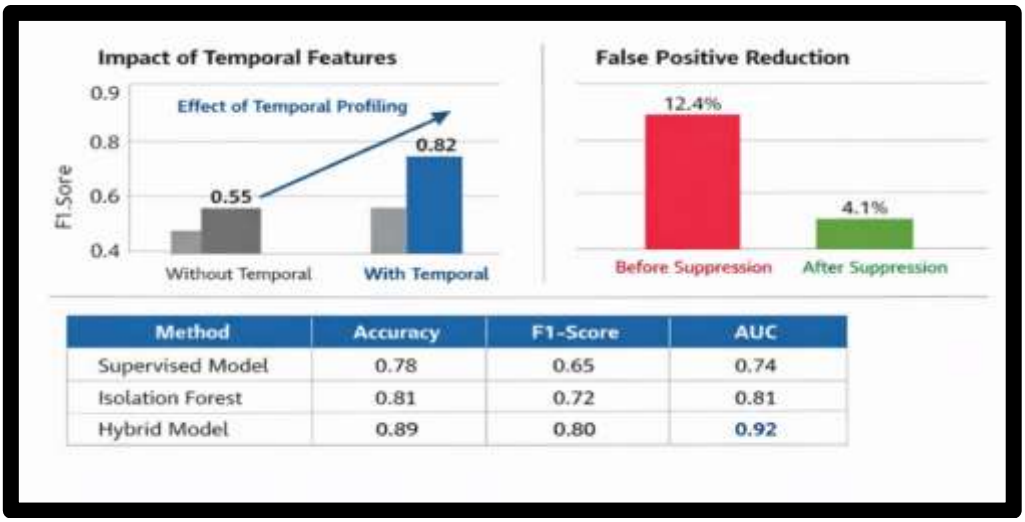
- Faster response times.
- Increased trust from security analysts.
- Greater operational efficiency.

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 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103



Method	Accuracy	F1-Score	AUC
Supervised Model	0.78	0.65	0.74
Isolation Forest	0.81	0.72	0.81
Hybrid Model	0.89	0.80	0.92

5. SHAP Explainability Analysis

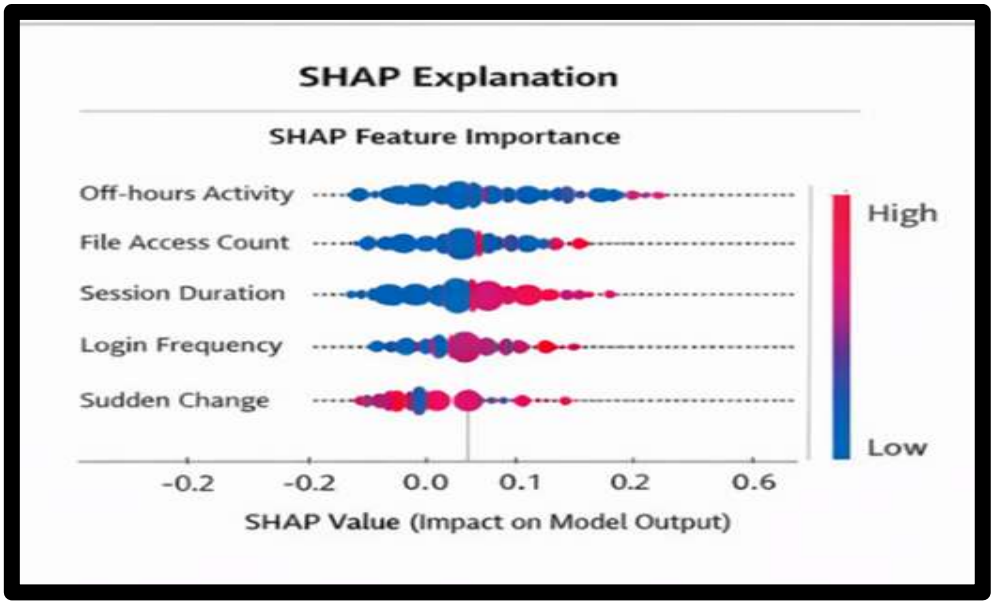
The SHAP analysis highlights the most dominant features that contribute to identifying insider threats.
Most Influential Features are as follows:

- Activity that occurs outside of business hours
- Number of file accesses
- Length of existing sessions
- Frequency at which someone logs into a system
- Sudden changes in behavior

An analysis of the SHAP values demonstrated that the most significant indicators of intended maliciousness were abnormal activity after-hours and unusual patterns of file accessing.
This will provide improvements in:

- Transparency of the model
- Confidence of the analyst
- Accountability for decision-making

Explainability is a crucial factor in deploying AI-based systems in environments where security is critical.



6. Comparative Discussion:

The results of the study show:

- That traditional supervised models have issues caused by the unbalancing of class instances.

DOI: 10.17148/IJARCCE.2026.155103

- That the pure anomaly detection method can detect regular behaviour and abnormal behaviour but does not provide the attack type classification.
- That combining both methods provide a more robust solution.

The components used in the hybrid method as follows:

- Temporal modeling
- Isolation Forest anomaly detection method
- Random Forest classification method
- SHAP explainability method
- False-positive reduction
  Makes the system more practical for enterprise deployment.

7.Key Points

- Detecting an insider threat is aided by temporally based attributes.
- Isolation forests work well when dealing with unbalanced datasets.
- Using a hybrid approach yields a higher detection rate and F1 score.
- Customer usability benefits from suppressing false positives.
- Analysts' levels of trust in the AI and its processes are bolstered by the use of explainable AI.

8.Limitations and Future Considerations

While the system did exceedingly well against the CERT data set, the information was synthetically created by CERT and does not support real-world enterprise validation. The next steps involve optimizing for real-time use.

Some examples of future work could include:

- Federated learning that allows for privacy preserving deployment.
- Long sequence modeling through a transform based model.
- Implementation of real-time streaming capabilities.

CONCLUSION

The creation of a mixed insider threat detection architecture, which includes temporal characteristics of behaviour, outlier detection using Isolation Forest, and classification of outliers using Random Forest for a high-security and explainable use case, has been presented in this paper. Furthermore, insider threat detection to be considered when building a system is the class imbalance of instances, the limited availability of labelled data, high false positive rates, and lack of interpretability.

When using the CERT insider threat data set to conduct experiments, the results show that incorporating temporal behaviour features into the detection process allows for the most accurate detection of insider threats. The hybrid detector achieves better overall detection performance, measured by accuracy, F1-score, and Area Under Curve (AUC), when compared with the use of only one of the supervised or unsupervised methods for detecting insider threats.

The Isolation Forest is an effective method for handling imbalanced data sets as it does not require a large number of labelled instances to detect anomalous behaviour patterns, while the Random Forest classifier is effective in classifying detected anomalies into various attack types. Furthermore, the use of SHAP provides an explanation for the decision made by the hybrid detector, which improves transparency and, therefore, increases analyst confidence in detection results.

Additionally, the adaptive mechanism for suppressing the false positive detections of the hybrid detector reduces the number of unnecessary alerts and makes the high-security hybrid insider threat detection framework more suitable for use in a real-world environment.

The overall framework presented in this paper provides a robust, scalable, and explainable solution for insider threat detection in today's complex and rapidly changing organisational environments.

REFERENCES

[1] F. R. Alzaabi and A. Mehmood, "A Review of Recent Advances, Challenges, and Opportunities in Malicious Insider Threat Detection Using Machine Learning Methods," in IEEE Access, vol. 12, pp. 30907-30927, 2024, doi: 10.1109/ACCESS.2024.3369906.

DOI: 10.17148/IJARCCE.2026.155103

[2] D. C. Le, N. Zincir-Heywood and M. I. Heywood, "Analyzing Data Granularity Levels for Insider Threat Detection Using Machine Learning," in IEEE Transactions on Network and Service Management, vol. 17, no. 1, pp. 30-44, March 2020, doi: 10.1109/TNSM.2020.296772

[3] T. A. Al-Shehari et al., "Enhancing Insider Threat Detection in Imbalanced Cybersecurity Settings Using the Density-Based Local Outlier Factor Algorithm," in IEEE Access, vol. 12, pp. 34820-34834, 2024, doi: 10.1109/ACCESS.2024.3373694.

[4] K. Saeed Alketbi and A. Mehmood, "A Comprehensive Survey of Explainable Artificial Intelligence Techniques for Malicious Insider Threat Detection," in IEEE Access, vol. 13, pp. 121772-121798, 2025, doi: 10.1109/ACCESS.2025.3587114.

[5] T. Al-Shehari, M. Al-Razgan, T. Alfakih, R. A. Alsowail and S. Pandiaraj, "Insider Threat Detection Model Using Anomaly-Based Isolation Forest Algorithm," in IEEE Access, vol. 11, pp. 118170-118185, 2023, doi: 10.1109/ACCESS.2023.3326750.

[6] M. M. H. Rahman, M. A. A. Naeem and A. Abubakar, "Threats From Unintentional Insiders: An Assessment of an Organization's Readiness Using Machine Learning," in IEEE Access, vol. 10, pp. 110294-110308, 2022, doi: 10.1109/ACCESS.2022.3214819.

[7] J. Dong et al., "DDCC: Synergizing Denoising Diffusion Probabilistic Models and Curriculum-Based Complexity Control for Insider Threat Detection," in IEEE Transactions on Industrial Informatics, vol. 21, no. 11, pp. 8351-8361, Nov. 2025, doi: 10.1109/TII.2025.3574417.

[8] R. B. Peccatiello, J. J. C. Gondim and L. P. F. Garcia, "Applying One-Class Algorithms for Data Stream-Based Insider Threat Detection," in IEEE Access, vol. 11, pp. 70560-70573, 2023, doi: 10.1109/ACCESS.2023.3293825.

[9] X. Zhu et al., "AUTH: An Adversarial Autoencoder Based Unsupervised Insider Threat Detection Scheme for Multisource Logs," in IEEE Transactions on Industrial Informatics, vol. 20, no. 9, pp. 10954-10965, Sept. 2024, doi: 10.1109/TII.2024.3393491

[10] F. Xiao, S. Chen, S. Chen, Y. Ma, H. He and J. Yang, "SENTINEL: Insider Threat Detection Based on Multi-Timescale User Behavior Interaction Graph Learning," in IEEE Transactions on Network Science and Engineering, vol. 12, no. 2, pp. 774-790, March-April 2025, doi: 10.1109/TNSE.2024.3519155.

[11] Z. Q. Wang and A. El Saddik, "DTITD: An Intelligent Insider Threat Detection Framework Based on Digital Twin and Self-Attention Based Deep Learning Models," in IEEE Access, vol. 11, pp. 114013-114030, 2023, doi: 10.1109/ACCESS.2023.3324371.

[12] M. Mehmood, R. Amin, M. M. A. Muslam, J. Xie and H. Aldabbas, "Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning," in IEEE Access, vol. 11, pp. 46561-46576, 2023, doi: 10.1109/ACCESS.2023.3273895.

[13] A. Azaria, A. Richardson, S. Kraus and V. S. Subrahmanian, "Behavioral Analysis of Insider Threat: A Survey and Bootstrapped Prediction in Imbalanced Data," in IEEE Transactions on Computational Social Systems, vol. 1, no. 2, pp. 135-155, June 2014, doi: 10.1109/TCSS.2014.2377811.

[14] L. Liu, O. De Vel, Q. -L. Han, J. Zhang and Y. Xiang, "Detecting and Preventing Cyber Insider Threats: A Survey," in IEEE Communications Surveys & Tutorials, vol. 20, no. 2, pp. 1397-1417, Secondquarter 2018, doi: 10.1109/COMST.2018.2800740.

[15] U. Rauf, F. Mohsen and Z. Wei, "A Taxonomic Classification of Insider Threats: Existing Techniques, Future Directions & Recommendations," in Journal of Cyber Security and Mobility, vol. 12, no. 2, pp. 221-252, March 2023, doi: 10.13052/jcsm2245-1439.1225.

[16] Z. Qiang Wang, H. Wang and A. El Saddik, "FedITD: A Federated Parameter-Efficient Tuning With Pre-Trained Large Language Models and Transfer Learning Framework for Insider Threat Detection," in IEEE Access, vol. 12, pp. 160396-160417, 2024, doi: 10.1109/ACCESS.2024.3482988.

[17] K. Kong, X. Jin, D. Liu, S. Xu, Z. Liu and G. Geng, "DPI-ITD: A Dual-Perspective Information-Driven Framework for Insider Threat Detection in IoT Systems," in IEEE Internet of Things Journal, vol. 12, no. 19, pp. 40731-40749, 1 Oct.1, 2025, doi: 10.1109/JIOT.2025.3589636.

[18] R. Nasir, M. Afzal, R. Latif and W. Iqbal, "Behavioral Based Insider Threat Detection Using Deep Learning," in IEEE Access, vol. 9, pp. 143266-143274, 2021, doi: 10.1109/ACCESS.2021.3118297.

[19] R. A. Alsowail and T. Al-Shehari, "Empirical Detection Techniques of Insider Threat Incidents," in IEEE Access, vol. 8, pp. 78385-78402, 2020, doi: 10.1109/ACCESS.2020.2989739.

[20] B. Böse, B. Avasarala, S. Tirthapura, Y. -Y. Chung and D. Steiner, "Detecting Insider Threats Using RADISH: A System for Real-Time Anomaly Detection in Heterogeneous Data Streams," in IEEE Systems Journal, vol. 11, no. 2, pp. 471-482, June 2017, doi: 10.1109/JSYST.2016.255850

[21] L. Liu, C. Chen, J. Zhang, O. De Vel and Y. Xiang, "Insider Threat Identification Using the Simultaneous Neural Learning of Multi-Source Logs," in IEEE Access, vol. 7, pp. 183162-183176, 2019, doi: 10.1109/ACCESS.2019.2957055.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155103

[22] A. Kim, J. Oh, J. Ryu and K. Lee, "A Review of Insider Threat Detection Approaches With IoT Perspective," in IEEE Access, vol. 8, pp. 78847-78867, 2020, doi: 10.1109/ACCESS.2020.2990195

[23] G. Ahmadi-Assalemi, H. Al-Khateeb, G. Epiphaniou and A. Aggoun, "Super Learner Ensemble for Anomaly Detection and Cyber-Risk Quantification in Industrial Control Systems," in IEEE Internet of Things Journal, vol. 9, no. 15, pp. 13279-13297, 1 Aug.1, 2022, doi: 10.1109/JIOT.2022.31441127

[24] A. Y. Al Hammadi et al., "Novel EEG Sensor-Based Risk Framework for the Detection of Insider Threats in Safety Critical Industrial Infrastructure," in IEEE Access, vol. 8, pp. 206222-206234, 2020, doi: 10.1109/ACCESS.2020.3037979.

[25] S. Walker-Roberts, M. Hammoudeh and A. Dehghantanha, "A Systematic Review of the Availability and Efficacy of Countermeasures to Internal Threats in Healthcare Critical Infrastructure," in IEEE Access, vol. 6, pp. 25167-25177, 2018, doi: 10.1109/ACCESS.2018.2817560.

[26] Y. Kim, S. -Y. Hong, S. Park and H. K. Kim, "Reinforcement Learning-Based Generative Security Framework for Host Intrusion Detection," in IEEE Access, vol. 13, pp. 15346-15362, 2025, doi: 10.1109/ACCESS.2025.3532353.

