American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

NEURAL NETWORK-BASED INSIDER THREAT DETECTION FOR ENHANCED CYBERSECURITY

DR. P. DILEEP KUMAR REDDY[1], S. SINDHU PRIYA[2], M. VAMSHI KRISHNA[3], P. SHIVA PRASAD[4], G. SRIRAJ REDDY[5]

PROFESSOR[1], UG SCHOLAR[2,3,4&5]

DEPARTMENT OF AI&ML, NARSIMHA REDDY ENGINEERING COLLEGE (UGC- AUTONOMOUS) MAISAMMAGUDA (V), KOMPALLY, SECUNDERABAD, TELANGANA-500100

PROBLEM STATEMENT

Insider threats pose a significant risk to modern cybersecurity systems because they originate from individuals who have authorized access to organizational resources, such as employees, contractors, or partners. These insiders may intentionally or unintentionally misuse their access privileges to steal sensitive data, disrupt services, or compromise system security. Traditional security mechanisms such as firewalls, access control systems, and signature-based intrusion detection systems are primarily designed to detect external attacks and often fail to identify malicious activities performed by trusted users.

Moreover, insider threats are difficult to detect because malicious actions may appear similar to normal user behavior. With the increasing volume of user activity logs and network data in organizations, manual monitoring and rule-based systems become inefficient and incapable of detecting subtle behavioral anomalies. Therefore, there is a need for an intelligent and automated system capable of analyzing large volumes of data and identifying suspicious patterns in user behavior.

To address this issue, neural network-based approaches can be utilized to learn complex behavioral patterns and detect anomalies in real time. By applying machine learning techniques to user activity data, organizations can improve their ability to detect insider threats early and enhance overall cybersecurity.

OBJECTIVES

1. To develop a neural network-based system for detecting insider threats in cybersecurity environments.
2. To analyze user behavior patterns using machine learning techniques to identify abnormal or suspicious activities.
3. To design an automated threat detection model capable of processing large volumes of system logs and network data.
4. To improve the accuracy of insider threat detection by utilizing deep learning and neural network algorithms.
5. To provide early warning alerts when potential insider attacks or suspicious activities are detected.
6. To enhance organizational cybersecurity by preventing data breaches and unauthorized access caused by internal users.
7. To evaluate the performance of the proposed model using metrics such as accuracy, precision, recall, and F1-score.

ABSTRACT

Insider threats have become one of the most critical challenges in modern cybersecurity, as malicious or negligent insiders can exploit their authorized access to compromise sensitive organizational data and systems. Traditional security mechanisms such as rule-based monitoring and signature-based detection are often ineffective in identifying insider attacks because these threats originate from trusted users and may follow normal usage patterns. To address this challenge, advanced artificial intelligence techniques, particularly neural networks, have been explored for detecting anomalous behaviors and potential insider threats in real time.This study presents a Neural Network-Based Insider Threat Detection System designed to enhance cybersecurity by analyzing user behavior patterns within organizational networks. The proposed system utilizes machine learning models to monitor user activities such as login patterns, file access behavior, system usage, and network interactions. By learning normal user behavior, the neural network can identify deviations that may indicate malicious insider activities, including data theft, unauthorized access, or privilege misuse.The system employs data preprocessing, feature extraction, and neural network-based classification to accurately detect suspicious activities. Experimental evaluation demonstrates that the proposed approach improves detection accuracy and reduces false alarms compared to traditional security mechanisms. This intelligent system provides organizations with a proactive solution for protecting sensitive information and strengthening cybersecurity infrastructure.

1.INTRODUCTION

With the rapid growth of digital technologies and interconnected systems, cybersecurity threats have become increasingly sophisticated and difficult to manage. While many security systems focus on protecting networks from external attacks, insider threats remain a significant concern for organizations worldwide. Insider threats occur when employees, contractors, or other trusted individuals misuse their authorized access to compromise sensitive data, disrupt system operations, or cause financial and reputational damage. Traditional cybersecurity approaches such as firewalls, intrusion detection systems, and access control mechanisms are primarily designed to defend against external attackers. However, these methods often struggle to detect insider threats because insiders already possess legitimate access privileges and knowledge of organizational systems. As a result, detecting abnormal user behavior becomes a key factor in identifying potential insider attacks.

Recent advancements in artificial intelligence and machine learning have provided new opportunities for enhancing cybersecurity systems. Neural networks, in particular, are capable

Received: 13-02-2026 | Accepted: 21-03-2026 | Published: 27-03-2026

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

of learning complex patterns from large volumes of data and identifying subtle anomalies in user behavior. By analyzing user activity logs, network traffic, and system interactions, neural network models can effectively distinguish between normal and suspicious activities. This research proposes a Neural Network-Based Insider Threat Detection System that leverages machine learning techniques to monitor user behavior and detect potential insider attacks in real time. The system collects and analyzes behavioral data, extracts relevant features, and applies neural network models to classify activities as normal or malicious. By integrating AI-driven analysis into cybersecurity frameworks, the proposed system aims to improve threat detection accuracy, reduce response time, and enhance overall organizational security.

2.LITERATURE REVIEW

Insider threat detection (ITD) in IoT-enabled infrastructures has garnered significant research attention due to the increasing frequency and sophistication of malicious activities executed by authorized personnel. Traditional security mechanisms often fail to identify such threats, as the perpetrators operate within the bounds of legitimate access. Over the years, researchers have proposed numerous techniques employing machine learning, deep learning, and graph-based models to improve ITD performance. This literature survey presents a comprehensive analysis of existing techniques, identifies their limitations, and highlights the motivations for our proposed Graph Convolutional Network (GCN)-based ITD framework.

Traditional Approaches for Insider Threat Detection

Early ITD models primarily relied on rule-based, statistical, and heuristic methods to identify malicious activities. These methods were often static, manually defined, and susceptible to evolving attack strategies.

- **Rule-Based Models**

  Rule-based systems, such as intrusion detection systems (IDS), define patterns based on predefined rules. While these models are efficient for well-known threats, they fail to adapt to novel insider tactics. For instance, threshold-based anomaly detection algorithms perform poorly when user behaviors subtly change over time.

- **Statistical and Heuristic Model**

  Statistical models use historical data to identify anomalies by analyzing deviations from normal patterns. However, they often yield high false positive rates in dynamic environments like IoT infrastructure. For example, stochastic models have been applied to identify suspicious login attempts but struggled to differentiate between legitimate access anomalies and actual malicious activity.

- **Behavioral Analysis Techniques**

  Behavioral profiling involves monitoring and modeling user activities to detect deviations. While these techniques effectively identify certain malicious patterns, they are susceptible to noise in high-dimensional data and often misclassify benign outlier behaviors.

Limitations:

1. Lack of adaptability to evolving insider behaviors.
2. High false positive rates.
3. Inability to handle large-scale IoT-generated datasets.

Machine Learning-Based Insider Threat Detection

Machine learning (ML) models introduced a data-driven perspective to ITD by leveraging historical user activity records to build predictive models.

Supervised Learning Approaches

Supervised learning models require labeled datasets to distinguish between benign and malicious activities.

- **Logistic Regression and Decision Trees:**

  Basic models like logistic regression and decision trees have been employed for ITD due to their simplicity and interpretability. However, these models struggle with imbalanced datasets and high-dimensional features, as observed in CMU CERT logs.

- **Random Forest and Gradient Boosting Machines:**

  Ensemble methods, such as Random Forest and XGBoost, have improved classification performance by aggregating multiple decision trees. Despite their effectiveness in capturing complex patterns, these models fail to account for temporal and relational aspects inherent in insider activity.

Limitations:

1. Poor performance on imbalanced datasets.
2. Inability to capture sequential and relational user behaviors.
3. Limited generalization to dynamic IoT environments.

Unsupervised Learning Approaches

Unsupervised learning techniques identify anomalies without requiring labeled data.

- **Clustering Techniques**

  Methods like k-means clustering and DBSCAN have been applied to group users based on behavioral similarity. For

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

example, anomaly detection models cluster users into normal and suspicious groups. However, these techniques are sensitive to noise and may misinterpret legitimate anomalies as threats.

- **Autoencoders for Anomaly Detection**

  Autoencoder-based models reconstruct normal user activity patterns and flag deviations as potential insider activities. While effective, autoencoders struggle to capture relational interactions, which are crucial in collaborative environments.

Limitations:

1. High susceptibility to noise and outliers.
2. Inability to capture graph-structured dependencies.
3. Poor interpretability of anomaly scores.

Semi-Supervised Learning Approaches

Semi-supervised models utilize both labeled and unlabeled data, making them suitable for insider threat scenarios where malicious activities are infrequent.

- **Hybrid Models**

  Combining supervised and unsupervised techniques, hybrid models improve detection accuracy. For instance, semi-supervised anomaly detection frameworks have been applied to ITD but often encounter issues related to feature selection and model interpretability.

Limitations:

1. Limited scalability with increasing data volumes.
2. Inconsistent performance across diverse insider threat scenarios.
3. Complexity in balancing labeled and unlabeled samples.

Deep Learning-Based Insider Threat Detection

Deep learning (DL) models have significantly advanced ITD by automatically learning feature representations from complex, high-dimensional data.

Recurrent Neural Networks (RNNs)

RNNs are commonly used to model sequential user activities.

- **Long Short-Term Memory (LSTM) Networks:**

  LSTM-based ITD models capture temporal patterns in activity logs, enabling the detection of time-dependent insider behaviors. For instance, an LSTM autoencoder was used to analyze login sequences and detect abnormal patterns in CMU CERT datasets. However, LSTM models are prone to overfitting, especially with imbalanced datasets.

- **Bidirectional LSTM (BiLSTM):**

  BiLSTMs process sequences from both past and future contexts, enhancing the model's understanding of activity patterns. However, these models require significant computational resources and struggle with relational insights.

Limitations:

1. Inefficient handling of imbalanced datasets.
2. High computational overhead.
3. Inability to capture structural relationships between users and activities.

Convolutional Neural Networks (CNNs)

CNNs have been adapted for ITD by treating activity logs as temporal or spatial patterns.

- **1D CNNs for Log Analysis**

  1D CNNs process activity sequences by learning spatial dependencies. Although effective in capturing local activity patterns, CNNs fail to grasp the global relational structure within IoT infrastructures.

Limitations:

1. Poor performance in capturing long-term dependencies.
2. Limited adaptability to dynamic behavioral changes.
3. High sensitivity to noise.

Autoencoder-Based Models

Autoencoders reconstruct activity patterns and identify deviations.

- **Variational Autoencoders (VAEs):**

  VAEs have been employed to learn probabilistic activity distributions and detect anomalies. While effective for noise reduction, VAEs struggle with imbalanced datasets and require significant tuning.

Limitations:

1. Inconsistent performance across different scenarios.
2. Vulnerability to adversarial attacks.
3. Complexity in selecting optimal hyperparameters.

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

Graph-Based Models for Insider Threat Detection

Graph-based models have gained attention for their ability to capture relational and structural patterns inherent in ITD scenarios.

- **Graph Convolutional Networks (GCNs):**

  GCNs extend traditional convolutional operations to graph-structured data, enabling the detection of anomalous interactions between users, devices, and resources. Recent studies have employed GCNs to analyze collaborative patterns and identify suspicious activities in network environments.

- **Dual Domain Graph Convolutional Network (DD-GCN):**

  A study proposed a DD-GCN to compute user behavior similarity using weighted feature metrics. While this approach improved detection performance, it failed to address data imbalance effectively.

- **Hypergraph Convolutional Networks (HGCNs):**

  HGCNs model complex multi-entity relationships by representing activity logs as hypergraphs. However, these models encounter difficulties when applied to large-scale datasets with dynamic behavioral patterns.

Limitations:

1. Class imbalance remains a persistent challenge.
2. High computational cost for real-time applications.
3. Limited generalization across diverse datasets.

Insights for Our Work:

The limitations observed in existing graph-based models highlight the need for an optimized GCN architecture with mechanisms to address data imbalance, computational complexity, and evolving user behaviors.

Data Imbalance Handling Techniques

Class imbalance is a critical issue in ITD, as malicious activities are relatively rare compared to normal actions. Various techniques have been proposed to address this challenge:

1. **Oversampling Methods:**
   - **SMOTE (Synthetic Minority Oversampling Technique):** SMOTE generates synthetic minority samples by interpolating between existing data points. However, it can lead to overfitting, especially in high-dimensional datasets.
   - **ADASYN (Adaptive Synthetic Sampling):** ADASYN focuses on samples near the decision boundary but may introduce noise when applied to complex datasets.
2. **GAN-Based Data Augmentation:**

   Generative Adversarial Networks (GANs) have been used to synthesize adversarial samples. Variants like Conditional GANs (CGANs) and Wasserstein GANs (WGANs) demonstrate improved performance but often suffer from mode collapse and training instability.

Limitations:

1. Synthetic samples may not accurately represent real activity patterns.
2. GANs are sensitive to hyperparameter configurations.
3. Oversampling methods can induce overfitting in downstream models.

Hyperparameter Optimization Techniques

Hyperparameter optimization significantly influences model performance, particularly for deep learning models in ITD.

- **Grid Search:**

  A brute-force method that exhaustively searches hyperparameter combinations but is computationally expensive.

- **Random Search:**

  Randomly samples hyperparameter values, providing better efficiency than grid search but lacking systematic exploration.

- **Bayesian Optimization:**

  Bayesian optimization uses probabilistic models to identify promising hyperparameter settings. The acquisition functions, such as Expected Improvement (EI) and Probability of Improvement (PI), guide the search process. Bayesian optimization has shown promise in tuning deep learning models, particularly with high-dimensional, non-convex objective functions.

Limitations:

1. High computational costs in large-scale applications.
2. Performance degrades with noisy objective functions.

Gaps in Existing Literature

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

A critical analysis of the surveyed techniques reveals several persistent challenges:

1. **Insufficient Graph-Based Modeling:**

   Few studies have effectively leveraged GCNs to model relational dependencies in insider activities.

2. **Ineffective Data Imbalance Handling:**

   Existing oversampling techniques either induce overfitting or fail to generalize across different datasets.

3. **Lack of Dynamic Adaptation:**

   Most ITD models do not adapt to evolving user behaviors, leading to performance degradation over time.

4. **Suboptimal Hyperparameter Optimization:**

   Conventional models often rely on manual or simplistic optimization methods, neglecting advanced techniques like Bayesian optimization.

Motivation for the Proposed Model

Based on the insights derived from this literature survey, we propose an ITD framework that addresses the aforementioned gaps by:

1. Utilizing a **GCN-based model** to capture relational patterns in insider activity logs.
2. Applying **IPCA** for efficient feature extraction and dimensionality reduction.
3. Introducing **EBiGAN** with an additional discriminator to generate high-quality adversarial samples.
4. Leveraging the **PI acquisition function of Bayesian optimization** for optimized hyperparameter tuning.
5. Evaluating the model's performance across diverse insider threat scenarios using the **CMU CERT dataset (r6.1 & r6.2)**.

This integrative approach is designed to enhance detection performance, minimize false alarms, and provide a scalable solution for ITD in dynamic IoT environments.

3.SYSTEM ANALYSIS

Insider threat detection (ITD) has been an area of extensive research due to the growing concern over malicious activities conducted by authorized individuals. Existing ITD models predominantly rely on machine learning (ML), deep learning (DL), and statistical techniques to identify anomalies in user behavior. While these methods have demonstrated success in identifying certain patterns of malicious activity, they often fail to address key challenges, such as data imbalance, high dimensionality, and the dynamic nature of insider behavior. This section reviews the existing system's methodologies, their limitations, and the necessity for an advanced, graph-based solution.

Existing Approaches for ITD

The traditional and contemporary approaches to insider threat detection can be broadly classified into the following categories:

Statistical and Rule-Based Methods

Early ITD models employed rule-based and statistical techniques to monitor and detect anomalies in user activity.

- **Signature-Based Detection:** Signature-based models use predefined rules and patterns to identify insider activities. For example, if a user accesses confidential files during non-working hours, the system flags the behavior. **Limitation:** These models are ineffective against zero-day insider attacks and rely heavily on manual rule updates.

- **Anomaly Detection with Statistical Methods:** Methods like mean, standard deviation, and correlation analysis have been applied to detect deviations from normal patterns. **Limitation:** These techniques yield high false positive rates in dynamic IoT environments due to fluctuations in legitimate activities.

Machine Learning-Based ITD

Machine learning models introduced data-driven insights into ITD by training algorithms on historical user activity data.

- **Supervised Learning:** Techniques like logistic regression, decision trees, and support vector machines (SVMs) classify user activities as benign or malicious. **Limitation:** These models require large, labeled datasets, which are often unavailable in real-world ITD scenarios.

- **Unsupervised Learning:** Methods like K-means clustering and DBSCAN identify anomalous patterns without labeled data. **Limitation:** High sensitivity to noise and poor interpretability hamper their real-world applicability.

- **Semi-Supervised Learning:** These models combine labeled and unlabeled data to enhance detection performance. **Limitation:** Performance depends heavily on the representativeness of the labeled subset.

Deep Learning-Based ITD

Deep learning models, particularly Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), have gained popularity in ITD due to their ability to learn complex, temporal patterns.

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

- **RNN and LSTM Models:** RNNs and LSTMs track sequential patterns in user activity logs. **Limitation:** These models suffer from vanishing gradients and struggle with long-term dependencies.

- **Autoencoders and GANs:** Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) have been employed to generate synthetic minority samples. **Limitation:** GANs often encounter mode collapse and generate samples that poorly represent malicious activity distributions.

Graph-Based Models

Graph-based models analyze the relationships between entities, such as users, devices, and files, within the network.

- **Graph Convolutional Networks (GCNs):** GCNs capture relational dependencies in graph-structured data by applying convolutional operations on nodes and edges. **Limitation:** Existing GCN implementations lack robust mechanisms to handle data imbalance and optimize hyperparameters effectively.

Limitations of the Existing System

The primary limitations of the existing ITD models include:

1. **Inability to Handle Data Imbalance:** Insider activity datasets often contain far fewer malicious instances compared to benign activities, causing models to exhibit bias toward the majority class.

2. **High Dimensionality and Noise:** IoT infrastructures generate vast amounts of high-dimensional data, complicating feature selection and extraction.

3. **Static and Rule-Based Approaches:** Signature-based methods require continuous manual updates to detect new insider tactics.

4. **Lack of Relational Insight:** Most models overlook the relational dependencies between users, devices, and activities, which are crucial for identifying collaborative attacks.

5. **Computational Inefficiency:** Deep learning models frequently face computational bottlenecks when handling large-scale, real-time IoT data.

6. **Suboptimal Hyperparameter Tuning:** Manual or grid search-based hyperparameter optimization methods are time-consuming and often yield suboptimal configurations.

7. **High False Alarm Rates:** False positives reduce the credibility of ITD systems, overwhelming analysts with benign activity alerts.

PROPOSED SYSTEM

To address the limitations of existing ITD models, this research proposes an advanced detection framework that leverages a **Graph Convolutional Network (GCN)** alongside data augmentation and Bayesian optimization techniques. The proposed model introduces innovations in feature extraction, data balancing, and hyperparameter tuning to improve detection accuracy and reduce false alarms.

Architectural Overview

The proposed ITD model consists of three core components:

1. **Feature Extraction and Clustering:** Utilize Improved Principal Component Analysis (IPCA) and outlier-resistant K-means clustering.

2. **Data Augmentation:** Generate high-quality adversarial samples using Enhanced Bidirectional GAN (EBiGAN).

3. **Graph Convolutional Network (GCN) with Bayesian Optimization:** Implement a GCN model optimized with the Probability of Improvement (PI) acquisition function of Bayesian optimization.

Key Components and Methodology

Feature Extraction and Clustering

The model begins with the extraction of user activity features from IoT-generated log records. These records are preprocessed to convert categorical variables into numerical representations using ordinal encoding. Next, **Improved Principal Component Analysis (IPCA)** is applied to reduce the dimensionality of the dataset while preserving its underlying structure.

Process:

- Calculate covariance matrices.

- Compute eigenvalues and eigenvectors.

- Retain principal components based on the Kaiser criterion (eigenvalue > 1).

The extracted features are then grouped using **outlier-resistant K-means clustering**, which incorporates a Huber loss function to minimize noise and identify scenario-based user activity patterns.

Advantage:

- Reduces computational complexity.

- Mitigates noise impact during clustering.

- Captures distinct behavioral patterns across different IoT scenarios.

Data Augmentation with EBiGAN

Received: 13-02-2026 | Accepted: 21-03-2026 | Published: 27-03-2026

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

The extracted samples reveal a significant class imbalance, with malicious activities being severely underrepresented. To address this, the model employs an **Enhanced Bidirectional Generative Adversarial Network (EBiGAN)** for data augmentation.

EBiGAN Architecture:

- **Generator:** Produces adversarial samples resembling malicious activities.
- **Encoder:** Encodes real data into the latent space.
- **Discriminator:** Distinguishes real from generated samples.
- **Additional Discriminator:** Evaluates the quality of encoded samples.

Innovative Aspect:

- The additional discriminator improves sample quality by ensuring the generated adversarial samples maintain structural similarity to real malicious activities.
- Unlike SMOTE or traditional GANs, EBiGAN reduces mode collapse and enhances the model's generalization capacity.

GCN-Based Insider Threat Detection

The core of the detection framework is a **Graph Convolutional Network (GCN)** implemented using PyTorch Geometric.

GCN Workflow:

1. Construct a graph representation where nodes represent users or devices, and edges represent interactions.
2. Apply graph convolution operations to aggregate node features based on their local neighborhoods.
3. Pass the aggregated representations through multiple GCN layers to capture higher-order dependencies.

Advantages of GCN:

- Captures relational patterns in insider activity data.
- Effectively models collaborative behaviors and anomalous interactions.
- Scales efficiently with increasing dataset size.

Bayesian Optimization with PI Acquisition

The model's performance depends heavily on the selection of optimal hyperparameters, including learning rates, dropout rates, and layer configurations. Manual tuning is inefficient and prone to suboptimal results. To address this, the proposed model employs **Bayesian optimization** with the **Probability of Improvement (PI)** acquisition function.

Optimization Process:

- Construct a probabilistic model of the objective function.
- Use the PI acquisition function to balance exploration (searching new hyperparameter values) and exploitation (refining known good values).
- Select hyperparameters that maximize validation accuracy and minimize loss.

Benefits:

- Automates hyperparameter tuning.
- Ensures efficient convergence.
- Reduces computational overhead compared to grid or random search methods.

Model Workflow

The proposed model follows these steps:

1. **Data Preprocessing:**
   - Convert raw activity logs into numerical representations.
   - Normalize data and handle missing values.
2. **Feature Extraction and Clustering:**
   - Apply IPCA for dimensionality reduction.
   - Use outlier-resistant K-means to cluster activity patterns into predefined scenarios.
3. **Data Augmentation:**
   - Generate adversarial samples using EBiGAN to balance the dataset.
4. **Graph Construction and GCN Training:**
   - Represent user activities as graph structures.
   - Train the GCN model to learn node embeddings and classify activities as benign or malicious.
5. **Model Optimization:**
   - Apply Bayesian optimization with PI acquisition for hyperparameter tuning.
6. **Evaluation and Detection:**
   - Evaluate model performance using accuracy, precision, detection rate, false

Received: 13-02-2026 | Accepted: 21-03-2026 | Published: 27-03-2026

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

positive rate, and false negative rate metrics.

- Identify and classify insider threats in real-time operational scenarios.

Advantages of the Proposed Model

The proposed ITD framework offers several improvements over existing models:

1. **Improved Relational Analysis:** GCN architecture captures complex relationships and collaborative behaviors in activity logs.
2. **Better Data Balance:** EBiGAN generates high-quality adversarial samples to rectify class imbalances without inducing overfitting.
3. **Robust Feature Selection:** IPCA ensures that only the most informative features are retained, reducing dimensionality and computational costs.
4. **Dynamic Hyperparameter Optimization:** Bayesian optimization with the PI acquisition function dynamically adjusts hyperparameters, improving model generalization.
5. **Reduced False Alarm Rate:** The integration of outlier-resistant clustering techniques and balanced training data minimizes false positives.

Comparative Analysis with Existing Models

The proposed system addresses the core challenges identified in the literature:

Challenge	Existing Models	Proposed Model
Data Imbalance	SMOTE/ADASYN (Overfitting issues)	EBiGAN with additional discriminator for realistic samples
Relational Behavior Modeling	ML models lack relational insights	GCN captures interaction patterns between entities
Feature Extraction	Manual and error-prone	IPCA automates feature selection based on eigenvalues
Hyperparameter Optimization	Grid/Random search	Bayesian optimization with PI acquisition function
Model Generalization	Overfitting and poor adaptation	Balanced training with EBiGAN and optimized GCN layers

Conclusion

The proposed GCN-based ITD model introduces a comprehensive solution to the limitations observed in existing ITD frameworks. By integrating advanced feature extraction, data augmentation, and Bayesian optimization techniques, the model enhances detection accuracy, reduces false positives, and provides a scalable, adaptable solution for securing IoT-enabled institutional environments against malicious insiders.

4.SYSTEM ARCHITECTURE

Neural Network-Based Insider Threat Detection for Enhanced Cybersecurity

The **system architecture** for the Neural Network-Based Insider Threat Detection system is designed to monitor user activities, analyze behavior patterns, and detect malicious insider actions within an organization. The architecture integrates **data collection, preprocessing, neural network analysis, and alert mechanisms** to provide enhanced cybersecurity protection.

Data Collection Layer

This layer gathers security-related data from different sources within the organization.

Sources of Data:

- System login and logout records
- Network traffic logs
- File access and modification logs
- Email activity records
- User behavior and activity patterns

These logs are collected from **servers, workstations, and network devices** and sent to the processing module.

Data Preprocessing Layer

The collected raw data is processed before being fed into the neural network model.

Key preprocessing steps:

- Data cleaning and removal of irrelevant information
- Feature extraction from log data
- Normalization of values
- Conversion of activity logs into structured datasets

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

This step ensures that the neural network receives **relevant and high-quality input data**.

Feature Extraction and Selection Module

Important behavioral features are extracted from the processed data.

Examples of features:

- Frequency of login attempts
- File access frequency
- Unusual access times
- Data transfer volumes
- Privilege escalation attempts

Feature selection improves the **efficiency and accuracy of the neural network model**.

Neural Network Detection Module

This is the core component of the system.

- A **Neural Network model (such as Artificial Neural Network or Deep Neural Network)** is trained using historical user activity data.
- The model learns patterns of **normal user behavior and malicious insider behavior**.
- When new data is analyzed, the system classifies activities as **normal or suspicious**.

Threat Detection and Analysis Module

The output of the neural network is analyzed to determine potential insider threats.

- Identifies **anomalous or suspicious activities**
- Classifies the threat level (low, medium, high)
- Generates detection reports for security administrators

Alert and Response Module

If suspicious activity is detected, the system takes appropriate actions.

Possible responses:

- Generate security alerts
- Notify administrators
- Temporarily block suspicious user accounts
- Log incidents for further investigation

Monitoring and Visualization Dashboard

A dashboard provides real-time monitoring and analysis.

Functions:

- Display detected threats
- Show user activity statistics
- Provide security reports and logs
- Assist administrators in making security decisions

CONCLUSION

In the IoT infrastructure, insider threat detection is challenging because authorized employees access sensitive data within institutions. Cybersecurity researchers have developed effective ITD models using neural networks to protect IoT enabled institutions. However, the existing ITD models have complications when working with a benchmark insider dataset, such as data imbalance problems and confrontation, to maintain the model's generalizability and interpretability. The issues discussed above are addressed by the proposed detection model to detect insider threats in the institutional log records of CMU CERT. The proposed detection model comprises three modules: IPCA for extracting the important user functionality-based samples and outlier estimators of k-means clustering for grouping the scenario-based user functionality samples to achieve dimensionality reduction and EBiGAN data augmentation to avoid a skewed class distribution in the dataset and ensure the encoded quality of the generated adversarial samples. The DNN-PI identifies the insiders of the IoT infrastructure and improves the overall performance of the model by achieving a high detection rate and minimal false alarm rate. The proposed model works well for the standard insider threat data of IoT enabled institutions. However, the model's limitations remain: it is difficult to adapt to different environments, it requires infeasible resources, and it lacks interpretability in real-time environment. The model ensures enhanced security of infrastructure although it faces complexity. In the future, the proposed model can be employed to generate samples and detect suspicious activities in other real-time scenarios.

REFERENCES

1. Khan, A. Y., Latif, R., Latif, S., Tahir, S., Batool, G., & Saba, T. (2020). *Malicious Insider Attack Detection in IoTs Using Data Analytics*. *IEEE Access*, 8, 11743–11753. DOI: 10.1109/ACCESS.2019.2959047
2. Marbut, A. R., & Harms, P. D. (2023). *Fiends and Fools: A Narrative Review and Neo-Socioanalytic Perspective on Personality and Insider Threats*. *Journal of Business and Psychology*, 39, 679–696. DOI: 10.1007/s10869-023-09885-9
3. Storchak, Y. (2024). *Insider Threat Statistics for 2024: Facts, Reports & Costs*. Available Online
4. Burhan, M., Rehman, R., Khan, B., & Kim, B.-S. (2018). *IoT Elements, Layered Architectures and Security Issues: A Comprehensive Survey*. *Sensors*, 18(9), 2796. DOI: 10.3390/s18092796
5. Pahlevanzadeh, B., Koleini, S., & Fadilah, S. I. (2021). *Security in IoT: Threats and Vulnerabilities,*

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

Layered Architecture, Encryption Mechanisms, Challenges and Solutions. *Communications in Computer and Information Science*, Springer, 267–283.

6. Peccatiello, R. B., Gondim, J. J. C., & Garcia, L. P. F. (2023). *Applying One-Class Algorithms for Data Stream-Based Insider Threat Detection*. *IEEE Access*, 11, 70560–70573. DOI: 10.1109/ACCESS.2023.3293825

7. Kim, A., Oh, J., Ryu, J., & Lee, K. (2020). *A Review of Insider Threat Detection Approaches with IoT Perspective*. *IEEE Access*, 8, 78847–78867. DOI: 10.1109/ACCESS.2020.2990195

8. Al-Shehari, T., & Alsowail, R. A. (2021). *An Insider Data Leakage Detection Using One-Hot Encoding, Synthetic Minority Oversampling and Machine Learning Techniques*. *Entropy*, 23(10), 1258. DOI: 10.3390/e23101258

9. Al-Mhiqani, M. N., Ahmed, R., Zainal, Z., & Isnin, S. N. (2021). *An Integrated Imbalanced Learning and Deep Neural Network Model for Insider Threat Detection*. *International Journal of Advanced Computer Science and Applications*, 12(1), 573–577. DOI: 10.14569/IJACSA.2021.0120166

10. Yuan, F., Shang, Y., Liu, Y., Cao, Y., & Tan, J. (2020). *Data Augmentation for Insider Threat Detection with GAN*. *IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*, 632–638. DOI: 10.1109/ICTAI50040.2020.00104

11. Zhu, D., Huang, X., Li, N., Sun, H., Liu, M., & Liu, J. (2022). *RAP-Net: A Resource Access Pattern Network for Insider Threat Detection*. *International Joint Conference on Neural Networks (IJCNN)*, 1–8. DOI: 10.1109/IJCNN55064.2022.9892637

12. Rao, T. K., Darapaneni, N., Paduri, A. R., Kumar, A., & Ps, G. (2023). *Insider Threat Detection: Using Classification Models*. *15th International Conference on Contemporary Computing (ICCC)*, 307–312. DOI: 10.1109/ICCC54662.2023.1000297

13. Li, X., Li, X., Jia, J., Li, L., Yuan, J., Gao, Y., & Yu, S. (2023). *A High Accuracy and Adaptive Anomaly Detection Model with Dual-Domain Graph Convolutional Network for Insider Threat Detection*. *IEEE Transactions on Information Forensics and Security*, 18, 1638–1652. DOI: 10.1109/TIFS.2023.3245413

14. Sepehrzadeh, H. (2023). *A Method for Insider Threat Assessment by Modeling the Internal Employee Interactions*. *International Journal of Information Security*, 22(5), 1385–1393. DOI: 10.1007/s10207-023-00670-1

15. Alshehri, A., Khan, N., Alowayr, A., & Alghamdi, M. Y. (2023). *Cyberattack Detection Framework Using Machine Learning and User Behavior Analytics*. *Computer Systems Science and Engineering*, 44(2), 1679–1689. DOI: 10.32604/csse.2023.026526