Insider Threat Detection: Exploring User Event Behavior Analytics and Machine Learning in Security Reviews

Ruba Altuwaijiri[1], Hanan AlShaher [2]
Department of Computer Sciences, College of Computer and Information Sciences, Majmaah University, Majmaah, 11952, Saudi Arabia
Emails: 441204474@s.mu.edu.sa; h.alshaher@mu.edu.sa

Correspomf Author: Ruba Altuwaijir , 441204474@s.mu.edu.sa

Abstract

With the exponential increase in technology use, insider threats are also growing in scale and importance, becoming one of the biggest challenges for government and corporate information security. Recent research shows that insider threats are more costly than external threats, making it critical for organizations to protect their information security. Effective insider threat detection requires the use of the latest models and technologies. Although a large number of insider threats have been discovered, the field is still limited by many issues, such as data imbalance, false positives, and a lack of accurate data, which require further research. This survey investigates the existing approaches and technologies for insider threat detection. It finds and summarizes relevant studies from different databases, followed by a detailed comparison. It also examines the types of data used and the machine learning models employed to detect these threats. It discusses the challenges researchers face in detecting insider threats and future trends in the field.

Keywords: User event behavior analytics; machine learning; detect insider threats.

1. Introduction

As information technology continually advances, wired and wireless Internet networks have become primarily responsible for information sharing. Information is among the most valuable concepts, but expanding information coupled with rapid technological developments has increased information security threats. The two main categories of known cyber threats are insider threats and external threats. Existing strategies and procedures used to counteract external threats to businesses include denial-of-service (DoS), phishing, and hacking. Insider threats are distinct from external threats in that they have a direct and legitimate authority to access data within a company. Moreover, insider threats can be challenging to identify and categorize because insider attackers possess access to or knowledge about information and may purposefully exploit it for their benefit [1].

The most recent technical report from the Computer Emergency Response Team Coordination Centre (CERT/CC) defines an insider threat as a malicious insider who knowingly takes advantage of their privileged access to an organizational network, system, and data to compromise the confidentiality, availability, or integrity of this information and the ICT infrastructures of the organization [2]. Thus, the primary factors that contribute to the rise of insider threats include increased use of end-user devices, higher volumes of data leaving the protected perimeter of the organization during "normal" use cases, increased use of data-leaping applications (such as web emails), and the migration of sensitive data to the cloud [3].

Insider threats can manifest themselves in various ways, including theft, sabotage, violence, espionage, and cybercrime. Espionage refers specifically to the activities of spying on people to obtain sensitive information for financial, political, or military gain. In contrast, violence involves actions that are intended to incite hostility or abuse. Theft includes the unlawful stealing of funds or intellectual property, and sabotage entails intentional acts

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

to damage an organization's infrastructure through both physical and virtual means. Finally, cybercrime involves using technology, gadgets, or the internet for theft, espionage, assault, or sabotage. Although non-malicious IT infrastructure exposure can also result in unintentional threats, purposeful threats are defined as malicious activities that use technology to perform an attack strategy, interrupt ordinary corporate operations, or obtain confidential information [4].

According to a global report for 2022 [5], insider threat events have increased by 44% over the last two years, and the average cost of such incidents has increased by more than a third over the same period to $15.38 million. Recent surveys indicate that 68% of firms consider insider threats to be becoming more frequent, and 70% of enterprises reported having experienced at least one insider attack in the preceding year [4]. The bitglass report for 2020 [6], also revealed that 61% of the companies included in the study had suffered from inside threats in the previous 12 months. Furthermore, Kaspersky discovered that employees were responsible for 22% of data leaks, compared to 23% caused by cyberattacks [7]. In general, insider threats are more harmful to companies and military organizations than threats from external systems because knowledgeable employees may exploit gaps in system implementation and operational procedures to easily engage in behavior that compromises system security, such as confidentiality breaches, information theft, commercial fraud, and system destruction [8].

Companies frequently invest in security defenses to fortify their network against malicious external attacks; however, these do not implement safeguards against possible attacks by malicious insiders [9]. Therefore, offering efficient techniques for spotting insider threats has become necessary. Insider threat detection is best achieved by data visualization and User Event Behavior Analytics (UEBA) [10]. UEBA, a recent advancement in information security, provides a workable way to find anomalies in user behavior by employing a range of techniques, such as statistical analysis and machine learning (ML) [11].

One important aspect of a network is its users. Daily, users typically complete a significant volume of work and activity [12]. As a result, the network experiences consistent usage patterns for different tasks. Consequently, routine chores and process activities might highlight a discernible pattern to map unique types of user behavior. If identified, irregular and aberrant activity can be intuitively discriminated from true user behavior. As such, ML techniques are the best option for identifying insider threats based on user event behavior analysis [13].

The purpose of this survey is to gain a better understanding of potential insider threats and user event behavior analyses. It aims to comprehend the factors contributing to the development and increase of these threats and what expert researchers have done to help mitigate and eliminate them. The main contributions of the paper are as follows:

- It will demonstrate that, by monitoring user behavior, we can identify most insider threats that go undetected by firewalls, intrusion detection systems, and security software.
- It provides a thorough literature analysis of the most recent methods for using machine learning (ML) to identify insider threats.
- Insider threats are responsible for many security incidents in both the public and private sectors. Therefore, the frequency and severity of insider threats emphasize the significance of adopting machine learning technologies to take proactive steps.
- There is a need to comprehend complex insider threats, recognize behavioral patterns, and employ advanced techniques for identifying and reducing insider threats.
- The variety of insider threats and their complex appearances, such as data theft and fraud, emphasize the need to know the many types of insider threats and their motivations.
- References the latest research (until 2023) in machine learning to detect insider threats.
- The survey presents current researchers' challenges in this area and provides suggestions for future research that can reduce or solve these threats.

The structure of this paper is as follows: In Section 2, you will find a detailed analysis of the literature review relevant to this research. Section 3 explains the methods used in this work. Section 4 focuses on insider threat detection and type. Sections 5 and 6 discuss the challenges researchers face in this field and present suggestions for future work. Finally, Section 7 concludes this paper. Figure 1 shows the sections of the paper in detail.

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

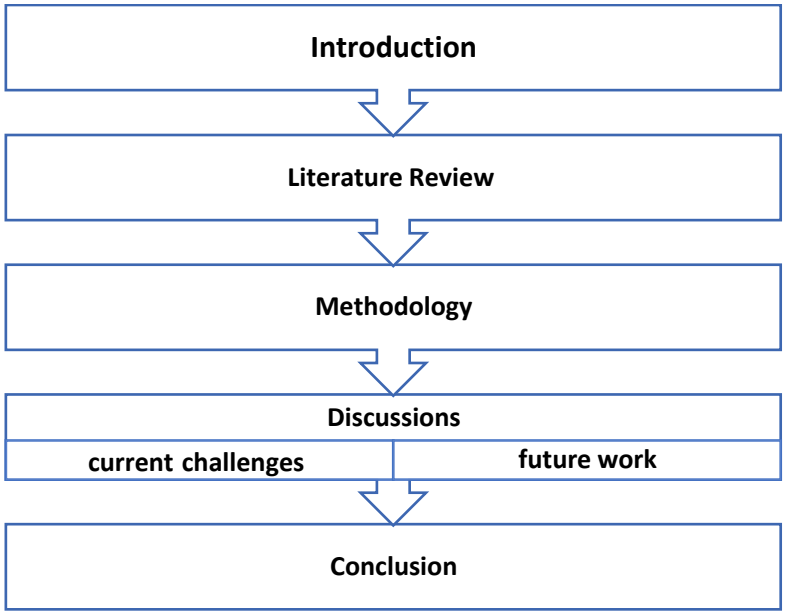

Figure 1: Paper outline.

2. Literature Review

Identifying threats is a broad topic requiring various models and techniques to analyze and visualize user behavior. Insider threats and risks posed by company employees are of particular concern. Therefore, it is necessary to analyze different techniques used to identify insider threats and the role previous research has played in alleviating those concerns to identify current research gaps and potential topics for future research.

Combining ML and user behavior analytics, Alshehri et al. [14] suggested a cutting-edge method for identifying cyberattacks. User actions in a network were represented by a sequence of events termed the window design, which enabled the model to determine the regularity of activity to characterize how the user approaches the network. The performance of the proposed RNN-LSTM model was evaluated relative to other models. Experimental results indicated that RNN-LSTM 1 worked most effectively, with an Area under Curve (AUC of 0.97) attained compared to others.

Mehmood et al. [15] developed a methodical approach to recognize diverse anomalous events that could signify irregularities and security concerns linked to privilege escalation, proposing an ML-driven system for identifying and categorizing insider threats. Greater prediction performance and improved machine learning outcomes were made possible by ensemble learning, which combines numerous models. To classify insider assaults, applied ML methods using a customized dataset made out of many CERT dataset files. Analyses were conducted on this dataset using four ML algorithms: Random Forest (RF), AdaBoost, XGBoost, and LightGBM, among which the LightGBM algorithm exhibited the best accuracy (97%); the other accuracy figures are 86% for RF, 88% for AdaBoost, and 88.27% for XGBoost.

Han et al. [16] applied ML to identify malicious behavior based on host process data. To categorize data as malicious or legitimate, several ML algorithms were used, including KNN, NB, RF, AE, and MemAE. Furthermore, the performance of the model classification was enhanced with the addition of dimension reduction techniques, including principal components analysis (PCA) and uniform manifold approximation and projection (UMAP), as well as the adaptive synthetic (ADASYN) and Synthetic Minority Over-sampling Technique (SMOTE) sampling approaches. Several metrics were used to assess model performance, including accuracy, recall, F1-Score, and area under the ROC curve (AUROC). The findings of this work showed that the preprocessed dataset (SMOTE) enhanced model performance and that all models outperformed the original dataset in terms of accuracy, recall, F1-Score, and AUROC. The preprocessed dataset (SMOTE) and MemAE model provided the best performance, yielding an F1 score of 1.00 and an AUROC value of 0.9826.

Peccatiello et al. [17] presented a methodology for insider threat detection that integrates several data science approaches, including data stream analysis, periodic retraining, and the use of supervised and semi-supervised ML techniques. The Elliptic Envelop, Local Outlier Factor, and Isolation Forest algorithms were utilized. Finally, the

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

ISOF algorithm provided the best results based on the F1-Score, precision, and recall measures, with a positive class (malign) recall of 0.78 and a negative class (benign) recall of 0.80.
Adun and Amadin [18] created, simulated, and evaluated the hybrid supervised machine learning model for the prediction of insider threats (HSMLM-IT), which uses the adaptive neuro-fuzzy inference system (ANFSI) for predictive learning and a support vector machine (SVM) for label classification. The ANFIS training blocks yielded an ANFIS accuracy of 91% and an ANFIS error of 9%, whereas the SVM blocks enabled a classification accuracy of 92% with a precision of 93%.

Bin Sarhan and Altwaijry [19] used a deep feature synthesis algorithm to generate behavioral features from historical data. Utilizing the 69,738 features generated for every user, PCA was applied to reduce the dimensionality of the data. Advanced ML algorithms were then applied to detect insider threats, with the anomaly detection and classification models achieving 91% accuracy. This was performed using the CERT insider threat dataset. The results obtained indicate that although the SMOTE balancing technique improves recall and precision at the expense of accuracy, it also reduces the impact of imbalanced datasets. Among all ML models, feature extraction methods and SVM models provided excellent results, with the classification model achieving 100% accuracy.

Haq et al. [20] suggested combining two deep-learning hybrid LSTM models with the Google Word2vec LSTM and Global Vectors for Word Representation (GLoVe) LSTM for insider threat detection. Using a real dataset, the models were assessed, indicating that ML-based models outperformed deep learning-based models. The models were compared against the most recent ML models, including AdaBoost, XGBoost, RF, KNN, and LR (Logistics Regression). XGBoost, an ML-based model, demonstrated an accuracy of 92%, whereas word2vecLSTM and GLoVeLSTM, two deep learning-based models, demonstrated accuracy values of 73.4% and 74.00%, respectively. The limitations of that study include the small amount of real data that is available, privacy and ethical concerns, and the large amount of data that requires accurate computing.

Alshehari and Alsowail [21] proposed a model for detecting insider data leakage events. Several ML algorithms – LR, DT, RF, NB, KNN, and KSVM – were trained on the CERT dataset to detect insider data leaks on unseen data. Subsequently, several experiments were carried out to evaluate the performance of the proposed model and determine the most suitable evaluation metrics for the optimal ML model. The results obtained showed that the proposed model could detect insider data leakage incidents with an AUC-ROC value of 0.99.

Nasir et al. [22] provided an insider attack detection framework based on deep learning. The primary purpose of establishing this method is to apply it to user technical data within an organization with modest processing and memory needs. To this end, the designed system was basic and customizable to the bare minimum domain knowledge needs. Various insider threat scenarios were applied to detect insider threats, and the 'LSTM-Autoencoder' was used. Subsequently, the model was trained and evaluated on CMU CERT V4.2. In addition, the performance of the suggested algorithm was compared to that of other known technologies, such as LSTM- CNN, RF, LSTM- RNN, One Class SVM, Markov Chain Model, Multi-State LSTM-CNN, Gated Recurrent Unit, and Skip-gram. Comparisons between these models indicated that the novel approach yielded a relatively high precision (97%), accuracy (90.60%), and F1 score (94%).

Zhang et al. [23] proposed a supervised insider threat detection method utilizing self-supervised learning and ensemble learning to identify malicious sessions. They employed TF-IDF feature extraction and over-bootstrap sampling to improve detection efficacy. Obtaining best-case AUCs of 99.2% and 95.3%, the experimental results demonstrate that the suggested method can successfully identify malicious sessions in the CERT4.2 and CERT6.2 datasets. The study also demonstrates how the TF-IDF feature extraction method, self-supervised learning, and ensemble learning approaches can increase insider threat detection's efficacy.

Gayathri et al. [24] used deep learning methods to solve the insider threat detection problem. Their method performs multiclass classification by combining generative models with supervised learning. To improve on the few data samples used, they performed data resampling analysis on the CERT insider threat dataset using generative adversarial networks (GAN). GANs were selected by applying three different resampling techniques to four different classification methods. The GAN method was nominated for its promising results relative to other resampling techniques.

According to Janjua et al. [25], the most significant cybersecurity challenge is preventing hostile insiders from acting maliciously within an organizational system. Classifying emails from the Wolf of SUTD (TWOS) dataset using a variety of ML techniques, they submitted this dataset to the following supervised learning techniques: AdaBoost, Naïve Bayes (NB), Logistic Regression (LR), KNN, LR, and SVM. Finally, after an experimental

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

approach, AdaBoost exhibited the optimal classification accuracy rate (98%) for benign and harmful emails. The original dataset was used to train the model, but with limited data; as such, a larger dataset could enhance the performance of this model.

Zou et al. [26] developed a data processing tool to process identified user activity and produce information-use events. Additionally, a data adjustment (DA) strategy was developed to change the weights of the majority and minority samples. Subsequently, the XGBoost model and DA strategy were combined to create an efficient ensemble approach for identifying anomalous behavior. Using the CERT dataset, this strategy was tested against an insider threat. Their results showed that the operation had an accuracy rate of 99.51% and an average recall rate of 98.16%, i.e., insider threats were successfully identified using the suggested strategy.

Tamanna [27] proposed an insider threat detection model based on user behavior and activity data, combining CNN and LSTM. Their proposed method was tested using the CMU CERT version 4.2 public dataset. Their experimental results indicated that the proposed model could successfully detect insider threats with a ROC of 0.914, i.e., CNN-LSMT is a practical method for identifying insider threats.

Le et al. [28] also suggested that insider threats are among the most dangerous and challenging-to-detect forms of assault on an enterprise since insiders have access to a company's networked systems and are familiar with its setup and security procedures. Insider malware detection faces distinct difficulties, including highly unbalanced data, a lack of ground truth, and behavioral shifts and drifts. Their study used ML to identify harmful behavior, particularly malicious insider threats, by analyzing data at multiple levels of detail under realistic conditions. In most cases, RF outperformed other ML techniques, obtaining good detection performance and an F1- score with low false-positive rates. The approach produced a false-positive rate of 0.78% and an accuracy of 85%.

Kim et al. [29] provided a framework for identifying insider threats at every stage of the user behavior modeling process based on anomaly detection and user behavior modeling techniques. The CERT database was the foundation for creating three datasets: email content, email communications, and user daily actions datasets. To find inside information, the authors employed classification algorithms. Their results indicated that the suggested framework could identify malicious insider activities with a good degree of accuracy. The dataset (CERT), which was used to develop the system, has been carefully curated to include a range of threat scenarios. Nevertheless, the dataset is artificially generated and simulated, which poses specific limitations to the study.

Singh et al. [30] studied user behavior profiles to monitor and analyze behavior and identify potential insider threats. An irregularity in external strings brought into behavior patterns can be determined using a hybrid ML approach that employs anomaly detection based on convolutional neural networks (CNN) and multistate long-term memory (MSLSTM). An uncomplicated single-state LSTM was found to be inferior to a multistate LSTM. Furthermore, with AUCs of 0.9047 on test data and 0.9042 on training data, the recommended LSTM multistate approach effectively detected insider threats.

To identify malicious insiders, Le and Zincir-Heywood [31] proposed a user-centered machine learning method with a high accuracy rate, specifically, a supervised learning ML model using popular techniques, including RF, LR, and ANN. Although their proposed system could identify malicious insiders with minimal training, the authors recommended employing more sophisticated data preprocessing techniques and feature analysis to improve the accuracy of the system.

Jiang et al. [32] proposed a model for analyzing user behavior by classifying user actions in great detail, identifying their qualities, and then detecting potential insider attacks. Using RF, SVM, and XGBoost approaches, they showed that the XGBoost algorithm, with an F-measure score of 99.96%, was superior to both SVM and RF algorithms for identifying insider threats. This study was based on experimental results applied to user behavior datasets.

Table 1: Comparison between various ML and DL techniques.

Sr.	Paper Ref.	Algorithms	Technique			Dataset	Results	Balancing
			ML	DL	Hybrid			
1	[14]	RNN-LSTM, SVM	√	√	√	CERT r4.2	RNN-LSTM 1 (AUC = 0.97) SVM (AUC = 0.91)	
2	[15]	RF, AdaBoost,	√	×	×	CERT	LightGBM (acc = 97%)	

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

Sr.	Paper Ref.	Algorithms	Technique			Dataset	Results	Balancing
			ML	DL	Hybrid			
		XGBoost, LightGBM						
3	[16]	KNN, NB, RF, AE ,MemAE	√	×	×	Host	MemAE (AUROC = 0.9826)	SMOTE
4	[17]	Elliptic Envelop (EV), Local Outlier Factor (LOF), and Isolation Forest (ISOF), RF	√	×	×	CERT r4.2	ISOF (Benign R = 0.80, Malign R = 0.78)	
5	[18]	SVM, ANFSI	√	×	√	CERT Insider threat from (Kaggle)	SVM (acc = 92 %) ANFSI (acc = 91 %)	
6	[19]	OCSVM, iForest, NN, SVM, AdaBoost, RF	√	×	×	CERT r4.2	SVM (acc =100%)	SMOTE
7	[20]	Word2vecLSTM, GLoVeLSTM, RF, AdaBoost, LR, XGBoost, KNN	√	×	√	Enron	XGBoost (acc = 92 %)	
8	[21]	LR, DT, RF, KNN, KSVM	√	×	×	CERT r4.2	KSVM (AUC-ROC = 0.99)	SMOTE
9	[22]	LSTM-Autoencoder LSTM-CNN, Random Forest, LSTM- RNN, One Class SVM, Markov Chain Model, Multi-State LSTM-CNN	×	√	×	CERT r4.2	LSTM-Autoencoder (acc = 90%)	
10	[23]	TF-IDF +Over Booting+ Self Supervised	√	×	×	CERT4.2 and CERT6.2	AUCs = 99.2% and 95.3%	
11	[24]	XGBoost, RF, MLP, 1DCNN	√	√	×	CERT r4.2	P = 83% R = 76%	GAN
12	[25]	SVM, NB, AdaBoost, LR, KNN, LR	√	×	×	TWOS	AdaBoost (AUC = 0.983, Acc = 98.3%)	
13	[26]	RF, RF+DA, GBT, GBT+ DT, XGBoost, XGBoost+ DA	√	×	×	CERT	XGBoost+ DA (AUC = 96.87%)	
14	[27]	CNN-LSMT	×	√	√	CERT r4.2	ROC = 0.914	
15	[28]	LR, NN, RF, XG	√	×	×	CERT r5.2	RF (acc = 85%)	SMOTE
16	[30]	Multi-State LSTM -CNN	×	√	√	CERT r6.2	AUC = 90.47%	
17	[32]	XGBoost, SVM, RF	√	×	×	CERT r6.2	XGBoost (F-measure = 99.96%)	SMOTE

3. Methodology

Identifying The Survey Methodology Section explains the research process used to examine existing studies on insider threat detection. The section covers a detailed and up-to-date collection of literature and a systematic

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

approach to selecting and analyzing previous studies. We will also discuss how current research was chosen based on particular inclusion and exclusion criteria. To achieve these objectives, a comprehensive database was used to conduct the survey. Relevant papers were found in online digital libraries through various digital repositories such as ScienceDirect, IEEE Xplore, and Springer. The scope of our survey is based on the following inclusion and exclusion criteria:

1. We focused on studies that included user event behavior analysis, deep learning, and machine learning models to detect insider threats.
2. We also emphasized that English is the selected language for reviewing the articles.
3. We focused on papers published between 2018 and 2023.
4. The emphasis was on studies whose central theme was the insider threat issue rather than articles in which the discussion of insider threats was tangential.
5. We replaced the old articles with the new ones, except when new ones had less information.
6. To find relevant articles in various digital repositories, we used the following keywords: "insider threats" "user event behavior analysis and insider threat" and "machine learning and deep learning".

4. Insider Threat Detection

Insider threats are among the most common and significant security threats facing various businesses, organizations, and governmental entities. These threats entail malicious actions conducted by employees with permission inside the organization. Insiders pose significant organizational security risks because they have access to its networked systems and are conversant with its protocols [28]. The Centre for the Protection of National Infrastructure (CPNI) [33], defines an insider as “a person who exploits, or has the intention to exploit, their legitimate access to an organization's assets for unauthorized purposes”. Such threats can compromise the availability, confidentiality, and integrity of privacy standards required of any secure defense system [34]. The IBM 2023 Cost of a Data Breach Report [35], states that the most costly data breaches were committed by malicious insiders, with an average cost of USD 4.90 million; this is 9.5% more than the cost of the average data breach (USD 4.45 million).
Saxena et al. [36], proposed the division of insiders into three different types, as shown in Figure 2 These insider types are defined as follows:

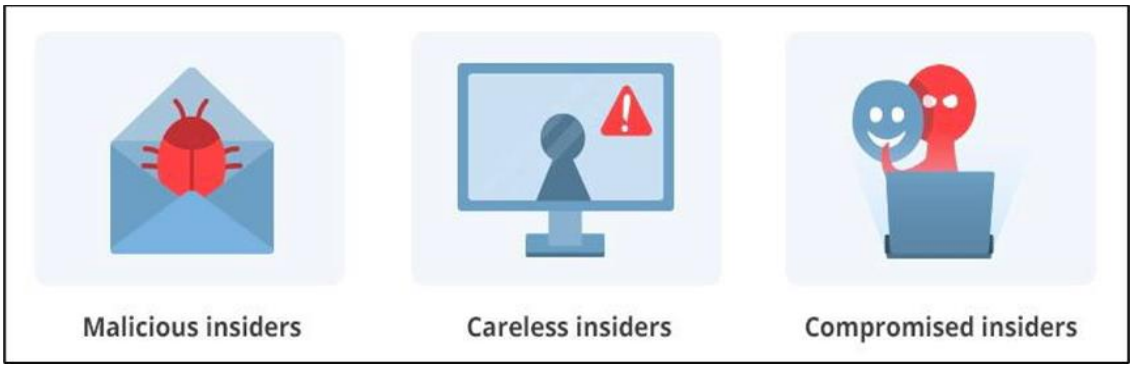

Figure 2: Types of Insiders [37].

- Malicious insider: A person who willfully misuses their access rights to hurt the organization for their benefit, for example, a dissatisfied worker who intentionally sells confidential information.
- Compromised insider: A person who has been the object of a malicious attack, the compromising of which allowed third parties access to personal data that allowed them to access insider systems. Examples include a person subjected to spear phishing, bribery, and social engineering targeting staff.
- Careless insider: Individuals who do not pay attention to or are ignorant of their company's security policies; as a result, they make mistakes that expose data and harm the company. For example, staff members may write down their credentials and fail to properly save them, resulting in their eventual acquisition by a third party.

The process by which external attackers breach secured networks has been the subject of numerous studies. Safeguarding a network from behind an organization's secured walls has received increased attention recently. In this regard, it is crucial to identify the factors that make an insider threat possible. Three elements are involved in this consideration, namely capability, opportunity, and motive, as shown in Figure 3 These insider threat elements involved are defined as follows: [38] :

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

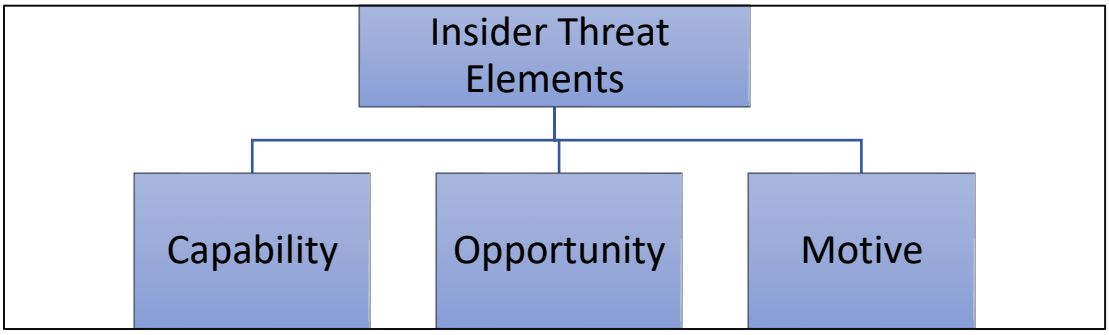

Figure 3: Insider Threat Elements.

1. Capability: When security measures are adequate, an insider threat must advance technologically. Technical knowledge enables sophisticated insider threats that carry out complex maneuvers and hide their traces from auditing systems. Furthermore, skilled attackers can change their behavior to avoid identification by detection schemes.
2. Opportunity: A crime cannot be committed without opportunity. A system administrator or other insider role within the organization could present an opportunity, as could the misuse of security mechanisms during their development, installation, or enforcement.
3. Motive: Given the frequency of criminal activity by insider threats, it is useful to consider the conditions that may encourage someone to do so. A person may become an insider threat due to discontent, outside influences, or financial hardships. In one study, four different categories of causes were identified: personality features, emotional states at the time, mental diseases, and a propensity for hostile behavior [39].

The CERT [40] defined three categories of criminal activity related to malicious insider threats: intellectual property (IP) theft, fraud, and IT sabotage, which are defined as follows:

- Insider fraud: Approximately 61% of businesses consider insider fraud to be extremely pervasive within their organization, making it one of the most prevalent types of attack [41]. When an insider leads a scam, their goal is typically to make money. This accentuates how firms must devise mitigating strategies to safeguard assets because financial gain plays a role in such attacks [39].
- Insider threat sabotage: Usually carried out by insiders with highly skilled and technical positions, this type of threat aims to compromise information systems by employing malware, such as advanced persistent threats (APTs) and privilege escalation tactics. These attacks are motivated by the desire to damage individual or organizational data as a result of stress, disappointment, or dissatisfaction [42].
- IP Theft: Intellectual property theft includes stealing important information, such as customer lists or source code. Technical tactics, including phishing emails and network transfers, are performed by those with authorized access [39].

5. Future research directions and challenges

This review of previous studies has presented the range of models and techniques used by researchers to detect insider threats. Many of these studies have focused on the complexities of dealing with insider threats, primarily because they differ from other types of cyberattacks. In most cases, cyberattacks are not initiated by malicious outsiders or software. Rather, they are conducted by insiders who have been granted access to organizational systems, data, or assets and misuse or abuse that access for malicious purposes. Threats have the distinct advantage of bypassing security controls and going undetected, which can cause significant damage to organizations, including data breaches, intellectual property theft, and vandalism. As a result, insider threats are very common and can have catastrophic effects on many organizations. Previous studies considering such threats have yielded insightful information, yet there are still several unresolved challenges and gaps:

1. Genuine and extensive public datasets of insider threats are rare. Because these incidents are so sensitive, organizations frequently find it challenging to provide researchers access to full information about their threat incident cases. Although the CERT dataset has made an effort to present information that is as realistic as possible, it is difficult to account for the inherent differences between synthetic data and real-world scenarios.
2. The quantity of false alarms increases as insider information increases because there is a lack of context, which requires considerable time, human, and material resources.
3. Conventional techniques for identifying insider threats focus on rules-based strategies developed by subject matter experts; however, their rigidity and lack of resilience limit their application.

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

4. Datasets are commonly imbalanced because there are significantly fewer malignant cases than normal cases.
5. Most companies emit large amounts of data, such as logs, network traffic, and user activity. Analyzing this data in near real-time to detect insider threats will require modern algorithms that can handle both the amount and speed of information.
6. Insider threat detection techniques, such as those based on UEBA and machine learning, may have difficulties achieving the necessary balance between the ability to detect potential threats and the prevention of false positives and negatives. False positives can lead to unnecessary investigations and outcomes, which can increase the workload of security professionals. On the other hand, false negatives can result in undetected insider threats. Therefore, striking the right balance between detection accuracy and false alarm reduction is a significant challenge for UEBA and machine learning research.
7. Insider threats are dynamic and change over time. UBEA techniques and machine learning frequently fail to adapt to new and developing threat patterns. Traditional UBEA and machine learning algorithms may be unable to detect improved insider threat techniques or novel attack vectors. Future research should focus on developing methods for detecting new insider threats and regularly updating UBEA models to stay up with the changing threat landscape.
8. False positives provide a substantial obstacle when dealing with unexpected changes in employee behavior. Alert fatigue caused by a high amount of false alerts can desensitize security personnel, resulting in genuine dangers being noticed or ignored. Previous research has shown that false positives have a negative impact on the performance of insider threat detection systems, emphasizing the significance of improving detection algorithms to reduce false alarms while retaining high detection rates [43].
9. One of the biggest technical challenges is distinguishing between normal user behavior and malicious actions without invading people's privacy. Traditional ways of spotting unusual activity, like using rules or statistics, often give many false alarms because of their inability to differentiate between normal and malicious activities accurately.

More effort is needed to detect insider threats and make them less harmful. Here are some recommendations for future work on insider threats:

1. Future efforts should focus on more refined artificial intelligence (AI) and machine learning algorithms (big data, deep learning, federated learning, and hybrid models) to spot and sort out the dangers of insider threats.
2. Threats present in textual data, such as e-mail and messaging between coworkers, may be detected by natural language processing (NLP) using advanced neural language programming techniques, which help understand their contents' meaning.
3. Future work may include increasing security awareness and annual training to minimize and eradicate insider threats.
4. Future research should use innovative technology solutions such as Isolation Forests or One-Class SVMs [17],[22] to detect insider threats without compromising employee privacy. These models can understand user behavior patterns, detect malevolent intent deviations, and reduce false positives through continuous learning and adaptation [43].
5. Future research should focus on developing more complex behavioral analytics methods to better understand and predict insider threat behavior. Examining user activity patterns, access logs, network traffic, and other relevant data sources can help spot anomalies and potential indicators of malicious activities.
6. Advanced detection methods for insider threats are needed, requiring the integration of multiple data sources, such as user behavior, network logs, access data, content analysis, and video monitoring. This comprehensive view aids in early detection and mitigates false positives.
7. Future efforts should focus on Endpoint Detection and Response (EDR) systems, which can detect and neutralize insider threats at the endpoint level. AI-powered EDR systems can continuously monitor endpoint activities, identifying patterns or unexpected actions that signal malware existence, enabling faster identification and reaction in distributed and heterogeneous enterprise contexts [44].

6. Conclusion

In conclusion, given the ever-evolving threat landscape, insider threats are a significant concern for security analysts. Detecting insider threats is essential to cybersecurity, particularly for companies that want to safeguard their intellectual property and sensitive data from malicious attacks. Insider threats can arise from employees, contractors, or anyone accessing the company's data and systems. Insider threats are detected and addressed via technology, laws, and behavioral analysis. Preventing and identifying harmful behavior is more crucial than ever. The fundamental purpose of surveillance is to utilize machine learning to classify user behavior as normal or malicious based on their activity. Although the boundaries and descriptions of insider threats remain unclear, much research is needed. Therefore, understanding and gaining insights into detecting insider threats is an important

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

research direction. The survey paper provides a comprehensive overview of insider threats by examining and classifying the available literature in digital repositories. It sheds light on the role of machine learning-based methods in detecting these malicious threats. Also, it highlights the challenges and vulnerabilities in insider threats, making it a valuable reference for researchers.

Funding: "This research received no external funding."

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] E. E. Schultz, "A framework for understanding and predicting insider attacks," Comput. Secur., vol. 21, no. 6, pp. 526–531, 2002.
[2] D. L. Costa, M. J. Albrethsen, M. L. Collins, S. J. Perl, G. J. Silowash, and D. L. Spooner, "An insider threat indicator ontology," SEI Pittsburgh PA USA Rep CMUSEI-007, 2016, Accessed: Aug. 05, 2023. [Online]. Available: https://apps.dtic.mil/sti/citations/tr/AD1044939
[3] Cybersecurity Insiders, "2020 Insider Threat Report," Technical report, Gurucul. Accessed: Nov. 01, 2023. [Online]. Available: https://www.cybersecurity-insiders.com/wp-content/uploads/2019/11/2020-Insider-Threat-Report-Gurucul.pdf
[4] Cybersecurity and Infrastructure Security Agency (CISA)., "Insider Threat Mitigation Guide," p. 133, Accessed: Oct. 28, 2023. [Online]. Available: https://www.cisa.gov/sites/default/files/2022-11/Insider%20Threat%20Mitigation%20Guide_Final_508.pdf
[5] P. Institute, "2022 Cost of Insider Threats Global Report", Accessed: Oct. 28, 2023. [Online]. Available: https://protectera.com.au/wp-content/uploads/2022/03/The-Cost-of-Insider-Threats-2022-Global-Report.pdf
[6] Bitglass, "2020 Insider Threat Report." Accessed: Oct. 28, 2023. [Online]. Available: https://pages.bitglass.com/rs/418-ZAL-815/images/CDFY20Q3Bitglass2020InsiderThreatReport.pdf
[7] Kaspersky, "Kaspersky 2022 IT Security Economics Survey." Accessed: Oct. 29, 2023. [Online]. Available: https://go.kaspersky.com/rs/802-IJN-240/images/IT%20Security%20Economics%202022_report.pdf
[8] L. Liu, O. De Vel, Q.-L. Han, J. Zhang, and Y. Xiang, "Detecting and preventing cyber insider threats: A survey," IEEE Commun. Surv. Tutor., vol. 20, no. 2, pp. 1397–1417, 2018.
[9] M. Omar, "Insider threats: Detecting and controlling malicious insiders," in New Threats and Countermeasures in Digital Crime and Cyber Terrorism, IGI Global, 2015, pp. 162–172.
[10] P. A. Legg, "Visualizing the insider threat: challenges and tools for identifying malicious user activity," in 2015 IEEE Symposium on Visualization for Cyber Security (VizSec), IEEE, 2015, pp. 1–7.
[11] S. Babu, "Detecting anomalies in Users-An UEBA approach," in Proceedings of the International Conference on Industrial Engineering and Operations Management, 2020, pp. 863–876.
[12] N. Khan, J. Abdullah, and A. S. Khan, "Defending malicious script attacks using machine learning classifiers," Wirel. Commun. Mob. Comput., 2017.
[13] N. A. Khan, M. Y. Alzaharani, and H. A. Kar, "Hybrid feature classification approach for malicious JavaScript attack detection using deep learning," Int. J. Comput. Sci. Inf. Secur., vol. 18, no. 5, 2020.
[14] A. Alshehri, N. Khan, A. Alowayr, and M. Y. Alghamdi, "Cyberattack Detection Framework Using Machine Learning and User Behavior Analytics.," Comput. Syst. Sci. Eng., vol. 44, no. 2, 2023.
[15] M. Mehmood, R. Amin, M. M. A. Muslam, J. Xie, and H. Aldabbas, "Privilege Escalation Attack Detection and Mitigation in Cloud using Machine Learning," IEEE Access, 2023.
[16] R. Han, K. Kim, B. Choi, and Y. Jeong, "A Study on Detection of Malicious Behavior Based on Host Process Data Using Machine Learning," Appl. Sci., vol. 13, no. 7, p. 4097, 2023.
[17] R. B. Peccatiello, J. J. C. Gondim, and L. P. F. Garcia, "Applying One-Class Algorithms for Data Stream-Based Insider Threat Detection," IEEE Access, 2023.
[18] I. J. ADUN and F. AMADIN, "A Hybrid Supervised Machine Learning Model for the Prediction of Insider Threats," J. Sci. Technol. Res., vol. 5, no. 3, 2023.
[19] B. Bin Sarhan and N. Altwaijry, "Insider Threat Detection Using Machine Learning Approach," Appl. Sci., vol. 13, no. 1, p. 259, 2022.
[20] M. A. Haq, M. A. R. Khan, and M. Alshehri, "Insider threat detection based on NLP word embedding and machine learning," Intell Autom Soft Comput, vol. 33, pp. 619–635, 2022.
[21] T. Al-Shehari and R. A. Alsowail, "An insider data leakage detection using one-hot encoding, synthetic minority oversampling and machine learning techniques," Entropy, vol. 23, no. 10, p. 1258, 2021.
[22] R. Nasir, M. Afzal, R. Latif, and W. Iqbal, "Behavioral based insider threat detection using deep learning," IEEE Access, vol. 9, pp. 143266–143274, 2021.

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024

[23] C. Zhang, S. Wang, D. Zhan, T. Yu, T. Wang, and M. Yin, "Detecting Insider Threat from Behavioral Logs Based on Ensemble and Self-Supervised Learning," Secur. Commun. Netw., pp. 1–11, 2021.
[24] R. G. Gayathri, A. Sajjanhar, Y. Xiang, and X. Ma, "Multi-class classification based anomaly detection of insider activities," arXiv, 2021.
[25] F. Janjua, A. Masood, H. Abbas, and I. Rashid, "Handling insider threat through supervised machine learning techniques," Procedia Comput. Sci., vol. 177, pp. 64–71, 2020.
[26] S. Zou, H. Sun, G. Xu, and R. Quan, "Ensemble strategy for insider threat detection from user activity logs," Comput. Mater. Contin., 2020.
[27] T. Tamanna, "Detection of Insider Threats Based on Deep Learning Using LSTM–CNN Model," PhD Thesis, Dublin, National College of Ireland, 2020.
[28] D. C. Le, N. Zincir-Heywood, and M. I. Heywood, "Analyzing data granularity levels for insider threat detection using machine learning," IEEE Trans. Netw. Serv. Manag., vol. 17, no. 1, pp. 30–44, 2020.
[29] J. Kim, M. Park, H. Kim, S. Cho, and P. Kang, "Insider threat detection based on user behavior modeling and anomaly detection algorithms," Appl. Sci., vol. 9, no. 19, p. 4018, 2019.
[30] M. Singh, B. M. Mehtre, and S. Sangeetha, "User behavior profiling using ensemble approach for insider threat detection," in 2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA), IEEE, 2019, pp. 1–8.
[31] D. C. Le and A. N. Zincir-Heywood, "Machine learning based insider threat modelling and detection," in 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), IEEE, 2019, pp. 1–6.
[32] W. Jiang, Y. Tian, W. Liu, and W. Liu, "An Insider Threat Detection Method Based on User Behavior Analysis," in Intelligent Information Processing IX, vol. 538, Z. Shi, E. Mercier-Laurent, and J. Li, Eds., in IFIP Advances in Information and Communication Technology, vol. 538. , Cham: Springer International Publishing, 2018, pp. 421–429.
[33] Centre for the Protection of National Infrastructure (CPNI), "Insider Data Collection Study," London, UK, 2013. Accessed: Nov. 06, 2023.
[34] S. Yang and Y. Wang, "Insider threat analysis of case based system dynamics," Adv Comput Int J ACIJ, vol. 2, pp. 1–17, 2011.
[35] IBM, "Cost of a Data Breach Report 2023," 2023, Accessed: Nov. 17, 2023. [Online]. Available: https://www.ibm.com/downloads/cas/E3G5JMBP
[36] N. Saxena, E. Hayes, E. Bertino, P. Ojo, K.-K. R. Choo, and P. Burnap, "Impact and key challenges of insider threats on organizations and critical businesses," Electronics, vol. 9, no. 9, p. 1460, 2020.
[37] Ekran System, "What Is an Insider Threat? Definition, Types, and Countermeasures," Ekran System. Accessed: Nov. 06, 2023. [Online]. Available: https://www.ekransystem.com/en/blog/insider-threat-definition
[38] M. A. Bridgeman, "A Survey of Methods for Detecting Intentional Insider Threats Against Digital Systems," PhD Thesis, Monterey, CA; Naval Postgraduate School, 2021.
[39] I. A. Gheyas and A. E. Abdallah, "Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis," Big Data Anal., vol. 1, no. 1, p. 6, Dec. 2016.
[40] D. M. Cappelli, A. P. Moore, and R. F. Trzeciak, The CERT guide to insider threats: how to prevent, detect, and respond to information technology crimes (Theft, Sabotage, Fraud). Addison-Wesley, 2012.
[41] P. Institute, "2013 Cost of Data Breach Study: Global Analysis." Accessed: Nov. 17, 2023. [Online]. Available: https://www.ponemon.org/local/upload/file/2013%20Report%20GLOBAL%20CODB%20FINAL%205-2.pdf
[42] A. P. Moore, D. M. Cappelli, T. C. Caron, E. D. Shaw, D. Spooner, and R. F. Trzeciak, "A Preliminary Model of Insider Theft of Intellectual Property," Carnegie Mellon University's Software Engineering Institute: Pittsburgh, PA, USA, 2011. Accessed: Nov. 06, 2023. [Online]. Available: https://insights.sei.cmu.edu/documents/2213/2011_004_001_15362.pdf
[43] M. N. Al-Mhiqani *et al.*, "A review of insider threat detection: Classification, machine learning techniques, datasets, open challenges, and recommendations," *Appl. Sci.*, vol. 10, no. 15, p. 5208, 2020, Accessed: Apr. 27, 2024. [Online]. Available: https://www.mdpi.com/2076-3417/10/15/5208
[44] N. Vemuri, N. Thaneeru, and V. M. Tatikonda, "Adaptive generative AI for dynamic cybersecurity threat detection in enterprises," 2024.

Doi : https://doi.org/10.54216/JCIM.130213
Received: January 07, 2024 Revised: Mrach 14, 2024 Accepted: May 03, 2024