[1] **Marshet Tamirat**
[2] **Anteneh Girma (PhD)**
[3] **Tilahun Melak (PhD)**

Current Detection Methods for Insider Threats and Social Engineering Attacks: Enhancements and Analysis Using Deep Learning

Journal of Electrical Systems

Abstract: - Despite advancements in technology, insider threats and social engineering attacks continue to pose significant challenges. Current threat detection methods often fail to effectively identifies insider threats, leaving organizations vulnerable. This systematic review thoroughly examines and evaluates existing detection methods for insider threats and social engineering attacks, performs comparative gap analyses, assesses detection effectiveness, identifies inherent challenges, and proposes conceptual system architecture. A primary challenge is distinguishing between normal and malicious insider activities, which exceed the capabilities of current network intrusion detection systems. Although machine learning and deep learning-based intrusion detection systems have been developed continuously, issues such as false positive and false negative rates persist due to the human elements involved in insider threats and social engineering attacks. The review focuses on identifying current network and host-based detection methods, analyzing existing gaps, and proposing a detection framework that integrates user behavior analysis with network and host-based detection and deep learning techniques to enhance detection accuracy and cost-effectiveness. Incorporating user cybersecurity behavior into existing intrusion detection systems and making detection unified (comprehensive) will result a high-performance threat detection system specifically for malicious insiders and social engineering attacks.

Keywords: Insider threats, Social engineering attacks, Detection methods, Deep learning, Cybersecurity.

I. INTRODUCTION

Cybersecurity (CS) remains a critical concern in today's digital landscape, necessitating robust defenses against evolving threats [1] and [2]. As organizations increasingly rely on interconnected systems, the risk of cyber-attacks targeting sensitive data and network integrity escalates [3]. Effective CS strategies are essential to mitigate these risks and ensure operational continuity [4]. CS remains a critical concern in today's digital landscape, necessitating robust defenses against evolving threats [1]. As organizations increasingly rely on interconnected systems, the risk of cyber-attacks targeting sensitive data and network integrity escalates [4]. Effective CS strategies are essential to mitigate these risks and ensure operational continuity [1]. The CS faces increasing challenges from insider threats and SEAs [5] and [3]. Insider threats occur when they utilized their authorized access for malicious activities, whether intentional or unknowingly, leading to critical financial losses and reputational harm [6]. These risks are arise by attackers, through exploiting SEAs via insiders [7]. Despite advancements in ML techniques to bolster security, there remains a pressing need to enhance the detection performance of current IDSs [8, 9]. Organizations deploy diverse security measures encompassing technical, physical, and administrative controls [10, 5]. However, each countermeasure presents own limitations, underscoring the absence of a universal and comprehensive solution the cost of detection also expensive [10]. Traditional cybersecurity measures often fall short against increasingly sophisticated threats like insider threat and attacks [11].

Machine learning (ML) and deep learning (DL) techniques have emerged as promising solutions by significantly boosting detection capabilities through advanced pattern recognition and anomaly detection methods [12], [13] and [14]. However, despite these advancements, the effective application of ML in CS faces persistent challenges stemming from dataset complexities such as imbalanced classes and feature redundancy [15]. Resolving these issues is pivotal for the development of dependable intrusion detection systems (IDS) [13] and [12]. Consequently, this study aims to explore how insider network and host level CS behavior with DL based detection approaches leverage the existing methods, giving a especial focus to insider threat and SEAs.

This study aims to conduct a comprehensive review of methods for detecting insider threats and SEAs. By examining and comparing ML and DL approaches, the study seeks to propose an advanced IDS framework based on DL that integrates insider behavior into network-based IDSs. This proposed solution aims to enhance threat

[1] Department of Computer Science and Engineering (CSE), School of Electrical Engineering and Computing (SoEEC), Adama Science and Technology University (ASTU), Adama, Oromia, Ethiopia. Email: Marshet.Tamirat@astu.edu.et
[2] Associate Professor of Computer Science and Cyber Security at the University of the District of Columbia. Email: An-teneh.girma@udc.edu
[3] Department of Software Engineering, College of Engineering, Addis Ababa Science and Technology University (AASTU), Addis Ababa, Ethiopia, P.O.Box 16417. Email: the_melak@yahoo.com

detection performance. Our objectives include reviewing current methods for detecting insider threats and SEAs, providing an overview of existing insider threat detection methods, and identifying ML and DL approaches for threat detection and attack mitigation. We aim to explore the advantages, challenges, limitations, and gaps in current detection methodologies. The subsequent sections

of this paper are structured as follows: Section II presents an overview of insiders in cybersecurity, highlighting their roles and impact on organization. Section III delves into insider threats and attacks, detailing the various threat and attack vectors related with malicious insiders. Section IV explores the role of insiders in social engineering and phishing attacks, emphasizing their engagement and the trick attackers employed. Section V discusses the detection and prevention techniques for insider threats and attacks, providing an overview of different methods and approaches, including IDS, signature and anomaly-based detection, digital and behavioral warning signs, and network and host-based intrusion detection. Section VI focuses on the datasets and algorithms used in various studies, including the selection and application of datasets and methods. Section VII outlines the methodological approaches employed in this research, detailing the inclusion and exclusion criteria, preliminary collection of papers, and the article screening process. Section VIII identifies the challenges and gaps in the detection of insider threats, providing an overview and comparative analysis of existing solutions, particularly DL-based detection methods. Section IX discusses the application of machine and DL techniques in detecting insider threats, including data preprocessing and performance evaluation metrics. Section X proposes new approaches to improving the detection of insider threats, summarizing the potential benefits and future directions. Section XI concludes the paper, summarizing key findings. References compile the comprehensive sources cited throughout the paper.

II. Literature Review

A Overview

Cybersecurity (CS), the practice of safeguarding computers, systems, networks, and programmes from cyber-attacks, plays a crucial role in today's digital environment [16]. With the increasing reliance on networks and the Internet, organizations face numerous security threats that challenge the safe continuation of their operations [6]. Cyberspace, encompassing inter- connected networks of information technology infrastructures, has become a vital domain, but it also exposes vulnerabilities to various cyber threats [17]. Intentional or unintentional insider threats have been attributed to a substantial number of data breaches, resulting in significant financial losses and damages [18]. While external threats remain a concern, insider threats have emerged as a formidable challenge in CS due to the direct access insiders have to organizational networks [19].

Both insider threats and attacks can result in various negative consequences, including significant financial losses that cost companies an average of $ 16 million per incident, exceeding those caused by external attacks [10]. Motivations that drive insider threats can vary [7], however, financial gain, revenge, or personal dissatisfaction are mostly often causes [20]. Individuals with malicious intent may exploit their knowledge and expertise in internal systems, processes, and vulnerabilities to carry out their harmful activities [21], [22] and [23]. Internal vulnerabilities can also create opportunities for external attacks [24] and [8]. According to [14], the insider threat distribution revealed that negligence accounted for 56%, intellectual theft for 18%, and malicious intent for 26%. Figure 1 shows the distribution of these types of threat and attack incidents.

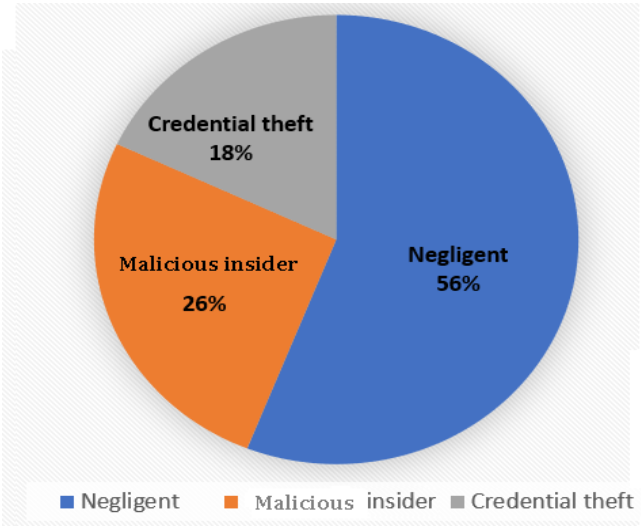

Figure 1: Distribution of Insider Threat Types

Insider threats can facilitate external attacks by either intentionally or unintentionally enabling unauthorized bodies to get access to sensitive information weaken security measures [18]. Malware or malicious software, presents a substantial risk caused by insiders aiming to harm computers and networks operations [25] and [26]. Malware can intrude systems through Malicious software, also known as malware, has the ability to infiltrate computer systems in many ways, including attachments or infected software downloads [25], to safeguard against malware, it is crucial for individuals and organizations to regularly update software [10]. Ransomware refers to malicious software that encrypts files or restricts system access, with attackers demanding payment in cryptocurrencies to maintain transaction anonymity. This results in the victim's files becoming inaccessible unless a ransom is paid [27]. Zero-day attacks are the exploitation of software vulnerabilities that are either unrecognized by vendor that have not yet been addressed with a patch [28]. Malicious actors take advantage of these vulnerabilities before a solution is developed, presenting challenges for organizations in their efforts to defend it. Zero-day attacks are also often directed at specific targets and can result in significant damage [19]. The continuous progress of technology has given rise to powerful techniques like machine and DL [12] and [19]. These advanced technologies have significantly bolstered organizations' security measures by detecting anomalies through analysis from massive amounts of data [29] and [30]. DL takes threat detection a step further by excelling in processing unstructured data, identifying complex patterns, and making accurate predictions [28] and [31]. In the realm of IDSs, supervised ML also had proven its capabilities specifically for structured datasets [32]. Researchers have proposed that combining different supervised ML, can further enhance IDS performance [15]. For instance, Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), greatly increased the accuracy of threat detection and decrease false alarms [12]. DNNs, which consist of at least three hidden layers, have promising algorithms for developing efficient threat detection systems [33]. DL accept large amounts of data, has the potential to improve current threat and attack detection by utilizing complex algorithms that have the ability to capture sophisticated patterns [24]. Auto DNNs are a promising technology as they possess the capability to improve true positive rates and decrease false positives and negatives alarms [18]. Currently, DL algorithm frequently employed in network based IDS, for instance, IF AE utilized in fog network, it classifies inbound traffic as either normal packets or malicious attacks [34, 30].

Insiders are often targeted in SEAs, making them a critical weak point in CS defenses [35]. SEAs, such as phishing, exploit human psychology to deceive individuals into revealing sensitive information or performing malicious actions [36]. Neglecting insider threats and SEAs can lead to increased CS risks and costs for organizations [37, 5]. Phishing attacks, characterized by deceptive electronic communication, aim to extract sensitive information or induce individuals to perform malicious actions [38]. These attacks often target insiders, leveraging social engineering techniques to deceive them into divulging confidential information [39]. Mitigating the risk of phishing attacks requires recognizing common tactics and implementing robust CS measures [38]. Many scholars have tried to develop ML and DL-based ID models, however, the performance still needs improvement in order to detect insider threats accurately. Due to the SEAs, the existing insider threat detection methods have limitations. Though accessing real and current datasets is now one of the challenges in DL model training, here, the UNSWB15 and NSL KDD online intrusion detection public datasets play significant roles. Insider threats can create vulnerabilities that social engineering attackers exploit, often using phishing as a tool [40]. Detecting insider threats is an essential element of CS as it involves identifying the risks posed by individuals who have authorized access to an organization's systems and data, commonly known as insiders [37]. To detect and prevent SEAs, organizations can implement security awareness training, multi-factor authentication, and continuous monitoring of user behavior [41], [42] and [43]. Detection and prevention techniques for SEAs can also be effective in detecting and preventing phishing attacks, as they share similar characteristics and objectives [44]. By understanding the interconnected nature of these threats, organizations can develop comprehensive strategies.

Preventive measures, including access controls, such as data loss prevention (DLP) systems and intrusion detection and prevention systems (IDPS), can also be used to detect and prevent insider threats. However, these techniques are often ineffective against sophisticated insider threats, social engineering and phishing attacks. Although existing detection techniques have made progress, there remains a pressing need for more comprehensive and unified solutions to address insider threats and SEAs. Developing new, efficient methods that integrate network, host, and insider behavior analysis is essential, and one promising approach to consider is the implementation of deep neural networks (DNNs) [45]. Figure 2 displays the classification of insider threats.

B Detection and Prevention

Threat detection and prevention represent critical challenges that demand sophisticated IDS methods and strategies [18]. Despite diverse IDS in CS measures, the persistent threat posed by insiders and SEAs presents

formidable obstacles to detection and prevention efforts [46]. This section seeks to explore the theoretical classification of IDSs, examining IDS detection challenges and gaps. The network-based IDS (NIDS) monitors network traffic for suspicious activities and anomalies, analyzing packets to detect potential threats. Conversely, host-based detection systems (HIDS) focus on individual host systems, scrutinizing log files and system calls for signs of malicious activities [47], signature-based IDS, anomaly-based IDS, and hybrid IDS [48]. Signature-based IDS rely on predefined patterns of known threats, making them effective against known attacks but vulnerable to zero-day exploits [48]. Anomaly-based IDS establishes a baseline of normal behavior and raises alerts for deviations, offering adaptability to emerging threats but being susceptible to false positives [48]. Hybrid IDS combine multiple detection techniques to enhance overall threat detection capabilities [48]. Currently, IDS faces challenges such as the non-technical psychological tricks of attackers now making it difficult for the IDS to generate, high FP rates, and difficulties in detecting sophisticated insider threats and SEAs [49]. Overcoming these challenges requires ongoing research and innovation to develop more robust and adaptive IDS solutions capable of defending against evolving cyber threats in today's dynamic digital landscape [48].

Signature-based detection, or rule based detection, relies on predefined patterns or signatures of known threats to identify malicious activities. It involves comparing network traffic or system behavior against a database of signatures. When a match is found, the system raises an alert or takes appropriate action. For instance, antivirus software utilizes signature-based detection to identify and block known malware strains [38].

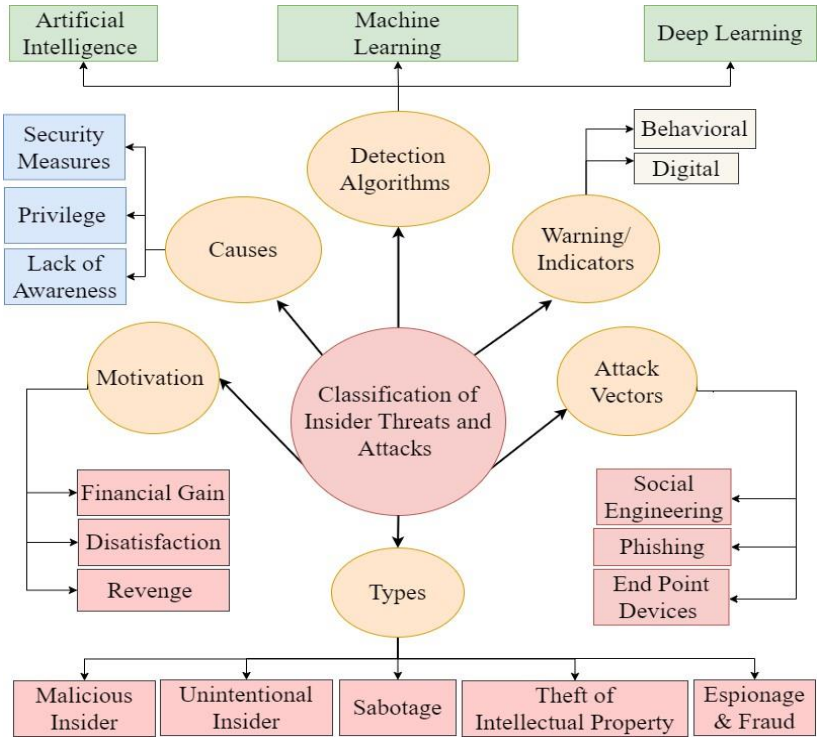

Figure 2: Taxonomy of Insider Threat and Detection Approaches

However, signature-based detection challenged by zero-day attacks, which are previously unknown threats [50]. While signature-based IDS excel in identifying known threats, they are susceptible to zero-day exploits, contrasting with anomaly-based IDS that adapt to emerging threats but are prone to FPs [48]. Hybrid IDSs, leverage overall detection. however it confront challenges such as limited high FP rates and struggles in detecting insider threats and SEAs [48] and [49]. Addressing these challenges necessitates continuous research and innovation to develop more resilient and adaptive IDS solutions capable of effectively countering evolving cyber threats in today's dynamic digital landscape [48]. Figure 3 illustrates the prevention and detection approaches of insider threats. Anomaly-based detection focuses on identifying deviations from normal patterns of behavior. It establishes a baseline of expected behavior and flags any activities that significantly differ from it. This approach utilizes machine learning (ML) algorithms to analyze network traffic, system logs, or user behavior and detect anomalies that may indicate potential threats [51]. Anomaly-based detection is effective in detecting zero-day attacks and other previously unknown threats [29]. ML and DL techniques have emerged as powerful tools in threat detection. ML involves training computers to learn algorithms and perform tasks. In the context of threat detection, supervised ML techniques have proven to be effective, and combining different supervised ML techniques can further enhance ID

performance [15]. For example, the fusion of CNN and RNN has shown improvements in network traffic classification accuracy, reducing false positives and improving overall detection rates [31]. DL, on the other hand, utilizes DNNs with multiple hidden layers to develop efficient TDSs, and have demonstrated better accuracy compared to conventional supervised and unsupervised ML methods.

The ability of DL to handle massive amounts of data and extract complex patterns and correlations makes it a promising approach [33, 24]. Autoencoders (AEs) have demonstrated significant efficacy in both network and host-based intrusion detection systems (IDS). Notably, the research conducted by [13] explores the application of AEs, particularly Variational AEs (VAEs) and Deep AEs, for enhancing anomaly detection systems. These methodologies have also shown promising results in detecting anomalies within fog networks. Furthermore, the integration of AEs with the Isolation Forest algorithm (AE-IF) has been developed as an effective approach for classifying inbound traffic as either legitimate or malicious [18] and [30]. The combination of these advanced machine and DL techniques holds great potential in enhancing detection capabilities and improving the overall security of systems and networks [34]. Figure 3 have presented IDS categorizes and various detection methods employed in network security [27], [52], [53], [54], [55] and [36].

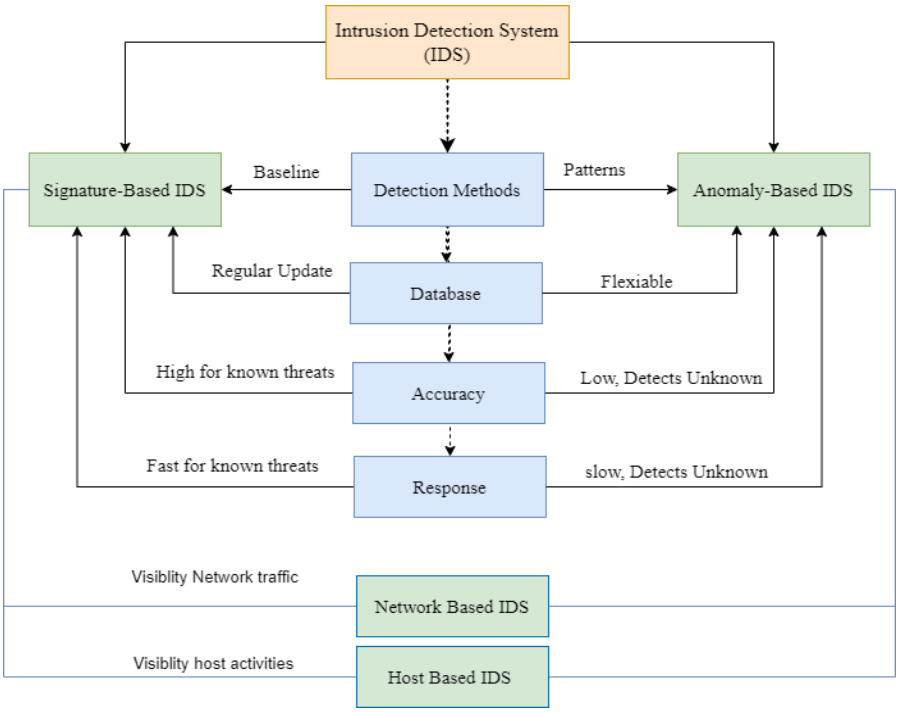

Figure 3: Intrusion Detection Methods

Digital warning signs serve as crucial indicators of potential signs, for instant accessing sensitive data not associated with the organization's function, downloading or accessing a substantial amount of data, accessing data outside of their unique job roles, network crawling and searches for sensitive data, using unauthorized storage devices such as USB drives, and emailing sensitive data outside the organization can be used to detect insider threats [13]. By paying attention to these signals, organizations can stay alert and distinguishing between legitimate and malicious user activities. Signs of behavioral changes, such as difficulties in adhering to security protocols, high stress levels, dissatisfaction at work, negative interactions with colleagues, frequent presence in the office during non-working hours, breaches of company rules, and discussions about resignation or new opportunities, can serve as indicators of potential insider threats [32]. By being vigilant to these signs, organizations can remain proactive and take necessary measures to avert potential risks. Methods for detecting and preventing SEAs can also be applied to phishing attacks, given their shared traits and goals [44]. Recognizing the intertwined nature of these threats allows organizations to devise comprehensive strategies to reduce their risks [10]. Preventive steps, including access controls like Data Loss Prevention (DLP) systems and Intrusion Detection and Prevention Systems (IDPS), can be employed to identify and prevent insider threats [56]. Despite the advancements in existing detection techniques, there is an urgent solutions to tackle insider threats and SEAs [31]. One potential approach worth exploring is the use of DNNs [13] and [45]. Many of the current approaches rely on traditional signature-based detection methods, which excel in identifying known threats but struggle to detect newly emerging threats such as SEAs [43] and [57]. Additionally, these methods often concentrate solely on network-based or host-based detection [12], [15] and [58],

resulting in fragmented detection approaches and these disjointed detection methods inefficient in detecting and making blind spots. Review reveals that existing approaches are prone to high rates of false positives and false negatives [7], leading to inaccurate detection and potentially compromising the overall efficacy of the detection system. Figure 4 demonstrate the digital and behavioral warning indicators.

The network allows interaction, collaboration, and sharing of expensive resources within the organizations, However, insider threats significantly harm the network infrastructures of organizations [10]. To effectively address the network security issue, it is crucial to improve and update the network-based detection systems as a preventive measure. As insiders are users of organizations who have access to resources [18], they get access to different disallowed resources as a result major Cs problems for organizations, companies, and enterprises will arise. Digital warning signs such as accessing sensitive data not associated with the organization's function, downloading or accessing a substantial amount of data, accessing data outside of their unique job roles, network crawling and searches for sensitive data, using unauthorized storage devices such as USB drives and emailing sensitive data outside the organization can be used to detect insider threats [59]. By paying attention to these signals, organizations can stay alert and distinguishing between legitimate and malicious user activities. Researchers and Practitioners are now enhancing NIDS using various DL algorithms. However, improving detection performance still needs extra research [18]. Host-based intrusion detection plays a crucial role in CS by safeguarding computer systems against unauthorized access and malicious attacks. Recent research articles have highlighted the use of ML techniques to enhance the effectiveness and efficiency of IDSs. In the case of host-based IDS, individual hosts are equipped with host-based (HIDS) to monitor their activities and detect any signs of potential attacks or malware [15]. Studies have demonstrated that employing ML techniques in host-based intrusion detection offers a promising approach to strengthening computer systems against insider threats. However, further research is needed to address existing challenges in this field, including improving detection capabilities and ensuring resilience against emerging threats and attacks.

The DARPA dataset, UCI ADFA dataset, CAOD dataset, Windows Logon Activity dataset, UNSW-NB15 dataset, CICIDS2017 (Canadian Institution of Cyber) [13] dataset, and CERT dataset are widely used in the field of CS for various purposes [60] and [33]. The UCI ADFA dataset is a network IDS dataset collected from an Australian Defense Force Academy (ADFA) campus network [61]. The CAOD dataset, also known as the Cyber Attack On-Demand dataset, is a collection of real-world cyber attack traces [61]. The Windows Logon Activity dataset contains logon activity data from Windows systems, used for user behavior analysis in anomaly detection [9]. UNSW-NB15 is a network IDS dataset, collected from a real-world environment. The CICIDS2017 dataset is a NID dataset that contains a variety of network traffic features [61]. The CERT dataset, provided by the Computer Emergency Response Team, is used by scholars for insider threat detection [32].

ML techniques have become pivotal in enhancing the detection of insider threats by identifying anomalous behaviors that may indicate malicious activity within organizations [62]. It offers robust solutions for securing critical cyber-physical systems by detecting vulnerabilities and responding to attacks in real time [63]. DT and SVM classification techniques have proven effective in detecting various types of cyber intrusions [40]. Furthermore, ML approaches utilized to analyze user behavior patterns to identify phishing attacks, thereby improving the detection of malicious emails [44]. These ML applications underscore the adaptability of MLs in CS, offering automated threat detection. By learning from past incidents, ML systems can continuously evolve and provide enhanced security measures. For example, the fusion of CNNs and RNNs, significantly improve the accuracy of network traffic classification, leading to a reduction in false positives and an overall enhancement in the detection rate [32] and [33]. The unsupervised DL algorithms frequently employed by researchers for classification, and prediction and shown higher outcomes such as accuracy, precision, and recall when compared to other supervised methods [13], [62] and [30]. Ability to accept large amounts of data and automatic features extraction capability of the unsupervised DL, provide extra potential to improve current threat and attack detection use cases [24]. AE DNNs are a promising technology as they possess the capability to improve true positive rates and decrease false positives and negatives alarms [62], [30] and [18]. Intrusion detection solution in networks environments currently utilizes DL algorithms, for anomaly detection scenarios [34] and [30], a simple DNNs consists of at least three hidden layers [33], discussed about IDS trained using supervised ML techniques. On the other side, there also authors proposed a ML-based approach and show this approach efficiency in the detection of intrusion in IDS. They suggested that combining different supervised ML algorithms would improve the accuracy of intrusion detection. And [15] proposed an efficient network intrusion detection and classification system using ML techniques. They combined the CNNs and Recurrent Neural Network RNN to classify network traffic more accurately. Their model reduced FPs and improved the detection rate, making it more effective than the traditional IDSs.

A number of studies suggest that DL based methodologies can increase the accuracy of insider threat detection compared to traditional supervised and unsupervised learning methods. Various algorithms such as CNN, RNN, and AEs are used for insider threat detection studies by [31] and [56] have addressed UBM based on activity logs and anomaly detection algorithms respectively [18]. Proposed a unsupervised ensemble technique as an effective approach to detect anomalous user activities in ITD scenarios. [54], proposed a 1D CNN model for network IDS. In recent studies, various ML approaches, such as ANN, SVM, PCA, and DT, have been employed for feature selection and model training in CS applications [13] and [32]. A growing body of research has focused on integrating DL techniques due to their capacity to handle large datasets and complex feature sets, akin to big data, which significantly enhances threat classification and anomaly detection capabilities [24]. Auto DNNs have also emerged as powerful tools for anomaly detection [18]. The incorporation of DL has significantly improved traditional IDS, enabling them to learn and identify more intricate patterns of malicious behavior [34].

Data pre-processing for an insider threat and SEA detection model requires the collection of datasets. This process may involve feature selection, normalization of numerical values, handling any missing values, and reducing less relevant variables using statistical functions [36]. Equation 1 is commonly used for normalization using a min-max scaler, and

Figure 4: Detection and Prevention of Insider threat and SEAs

To select relevant features, many scholars utilized formula given in equation 1.

$$r(x,y) = \frac{\sum_{i=1}^{n}(xi-\dot{x})(yi-\dot{y})}{\sqrt{\sum_{i=1}^{n}(xi-\dot{x})^2}\sqrt{\sum_{i=1}^{n}(yi-\dot{y})^2}} \quad (1)$$

Performance evaluation of machine and deep learning for threat detection model trainings' often relies on metrics such as accuracy, precision, recall, and F1-score. Table 1 illustrates evaluation metrics commonly used in these articles [12], [13] and [15].

Table 1: Confusion matrix for threat detection

	Predicted: Normal	**Predicted: Malicious**
Actual: Normal	True Positive (TP): Model accurately predicted normal user activity as normal.	False Positive (FP): Model flagged a normal user activity as malicious
Actual: Malicious	False Negative (FN): Model incorrectly labeled a malicious activity as normal.	True Negative (TN): Model correctly identified a malicious activity as malicious

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{F1} - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

III. METHODOLOGY

The methodology adopted for this comprehensive systematic review was centered on the detection of insider threats and SEAs within the CS domain. To ensure the relevance and quality of the selected papers, there are specific inclusions and exclusions criteria were established. The screening process was facilitated using the Rayyan web application. Initially, a preliminary collection of pertinent papers was conducted by searching for keywords related to cybersecurity (CS), insider threats, IDS datasets, network intrusion detection, host intrusion detection, SEA, current trends in insider threats, malware detection, insider threat detection and prevention, performance evaluation in insider threat detection, machine or deep learning-based threat detection, anomaly-based detection, and signature-based threat detection. The initial search, conducted through Google Scholar, targeted articles published from 2018 onwards, resulting in the retrieval of approximately 323 papers from various scholarly sources. Figure 5 shows the overall study methodology. To refine the selection, the Rayyan web application was employed for paper filtering and screening. This methodological approach ensured a comprehensive and systematic review of the literature while maintaining rigor and relevance in the selection of papers for analysis.

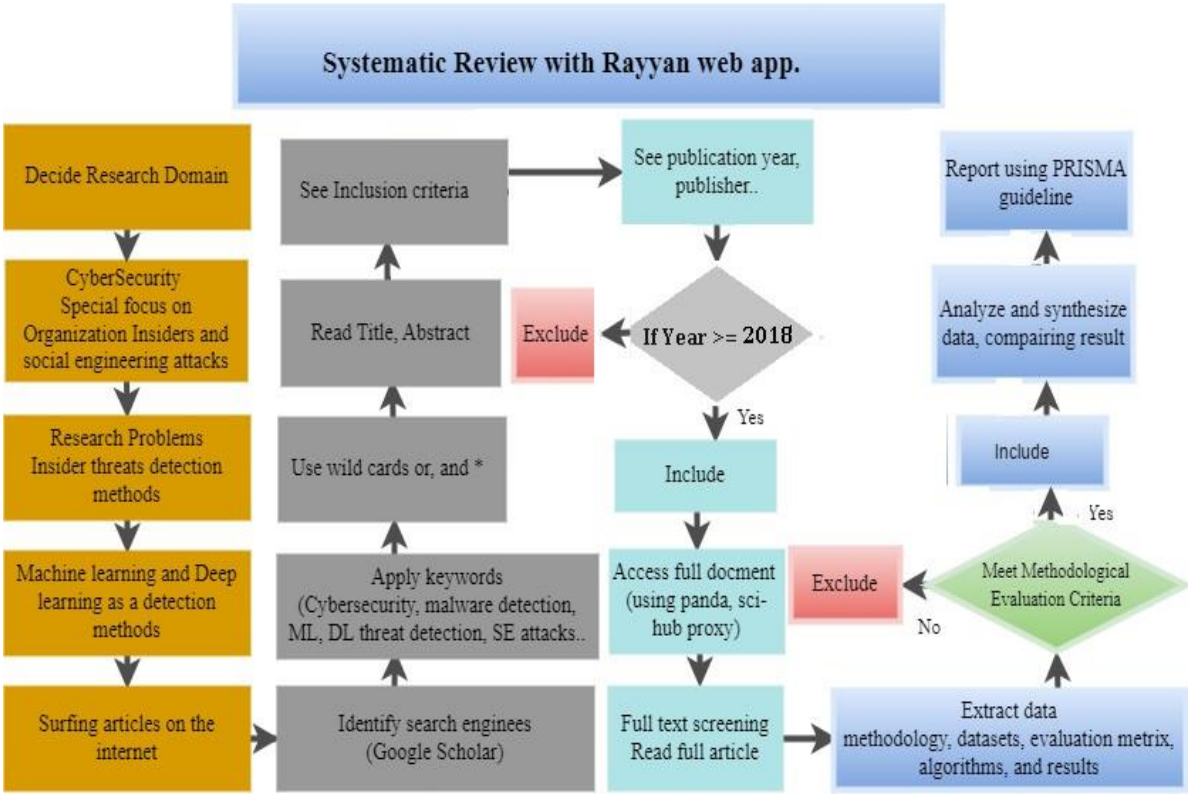

Figure 5: Study Methodology

A Inclusion and Exclusion Criteria

The inclusion criteria encompassed papers that focused on the detection of insider threats and SEAs within the CS domain. Additionally, papers that explored the use of ML or DL techniques for threat detection were included. Furthermore, studies addressing signature-based threat detection and anomaly-based detection were considered. The exclusion criteria encompassed papers that did not specifically address insider threats or SEAs, as well as papers that were published before 2018. Moreover, papers that focused on other aspects of CS or did not provide sufficient information were also excluded. Out of 323 total research articles, 76 papers have included.

B Preliminary Collection of Papers

The preliminary collection of relevant papers was conducted by searching for keywords related to cybersecurity, insider threats, intrusion detection system, datasets, network detection, host intrusion detection, SEAs, current trends in insider threats, malware, performance evaluation in ITD, machine or DL-based threat detection, anomaly-based detection, and signature-based threat detection. This initial search yielded approximately 323 papers from various sources, including the Google search engine.

C Article Screening Process

The screening process involved the use of the Rayyan web app to assess the relevance of the collected papers. The predefined inclusion and exclusion criteria were applied to filter out the intended papers that met the specific research objectives. The screening process in the Rayyan web app involved a thorough assessment of the title, abstract, year of publication and peer reviewed status to determine its relevance to the research topic. Papers that passed this initial screening were subjected to full text review, where their content and methodology were thoroughly evaluated to ensure their suitability for inclusion in the comprehensive systematic review. Figure 6 presents the articles screening summary.

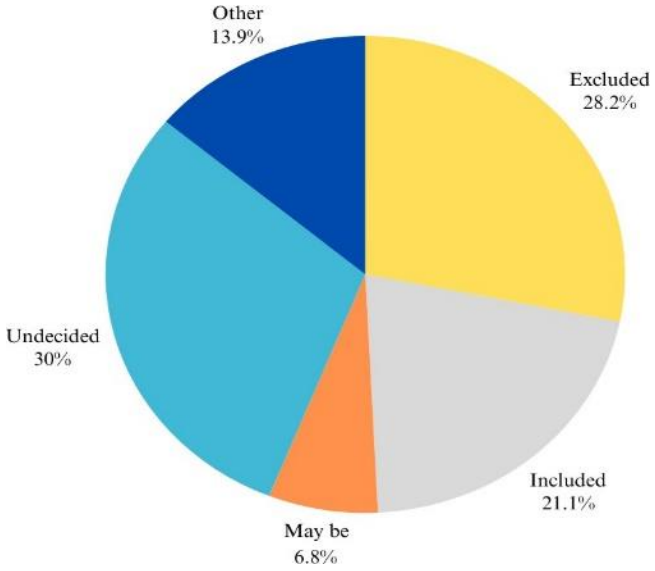

Figure 6: Article Screening Summary

IV. Challenges And Gaps In Detection

Threat detection faces several challenges and gaps that need to be addressed. One challenge is the complexity and diversity of attacks, which require advanced detection techniques and tools [10]. Another challenge is the ability to detect insider threats, as insiders have privileged access and can exploit their knowledge and privileges to carry out malicious activities [31]. Additionally, SAE manipulate human behavior, making them difficult to detect using traditional CS measures [11]. Finally, the rapid evolution of attacks and the emergence of new attack vectors pose a constant challenge to threat detection [13]. Threat and attacks are significant challenges in CS that needs effective tools for detection and mitigation. While organizations have implemented various CS measures, the complexity of insiders and SEAs make difficult the detection and prevent tasks. The aim of this section is to address the theoretical classification on insiders, detecting challenges and gaps analysis. Additionally, the paper explores the application of machine and DL methods in threat detection and the use of AE deep neural network (DNNs) architecture. Insiders are the cause of CS risks, especially phishing attacks. Due to the advanced technologies used by the attackers and the ineffective practice of giving insiders awareness training, the attackers are still difficult to identify, and

shockingly, the insiders themselves initiate the attacks. Additionally, insiders frequently conduct insider threats during work hours, making detection challenging [22]. A study suggested by [18] focused on activity logs on IDSs. However, the model was inefficient in detecting unknown attacks [48]. Detecting insider threats in separate approaches makes detection inefficient [27]. As a result enterprises and organizations able to have integrated solutions.

Numerous scholars have been developing techniques for detecting insider threats, however, these threats often emerging exponentially [64], [65] and [13]. Lack of proper training, limitation in CS expertise, lack of accountability, employee awareness, and security polices insider threats become sever stages [49] and [30]. The NSL-KDD dataset contains a diverse range of network traffic data, including normal and malicious traffic patterns. This dataset offers researchers a comprehensive set of scenarios to test and refine their intrusion detection algorithms [13]. Scholars employ the CIC-TruthSeeker 2023 dataset to analyze social media post texts [66] with ML algorithms, enhancing the robustness of their text analyses and developing more effective methods. Other commonly utilized datasets like DARPA, UCI ADFA, CAOD, Windows Logon Activity, and CICIDS2017 provide valuable resources for various ML and DL-based intrusion detection techniques, user behavior analysis, and insider threat detection [13] and [33]. These datasets collectively contribute to advancing the field of CS research, empowering researchers to address emerging challenges and safeguard digital environments.

A Comparative analysis of Deep Learning based Detection

In recent years, Intelligence, ML and DL have been applied to insider threat detection. Articles such as [31, 67], focus on user behavior modeling and anomaly detection algorithms while [56] explore the classification of ML techniques, datasets, open challenges, and recommendations related to insider threat detection. SEA is a type of cyber attack in which the attacker uses social interaction [36], it is one of the most significant risks at present since it uses non-technical methods and focuses on individuals [35] and [68] as a means to gain sensitive information or even access restricted services[58]. SE is taking advantage of human behavior and natural tendency. According to researches, many organizations allow employees to work from home and connect via various Internet services without providing adequate cyber protection [46]. This massive proportion underwent rapid changes in many activities, bringing gigantic problem in security and privacy [69]. As a result threat incidents dramatically increased more than ever before [70]. Now attackers use complex tools and get into the target organization, then they exploit their threat that causes financial loss, identity theft, and sabotage. Among the many different CS threats that occur on the internet, Distributed denial of service (DDoS) attacks is the most frequent [5]. The number of DDoS attacks around the world is continuously rising [57]. Even if, CS products are continually invented and deployed, attackers disrupt physical infrastructure by infiltrating the digital systems that control physical processes, damaging specialized equipment, and disrupting vital services [71]. These risks are becoming more complicated [11]

To mitigate the risk of attacks different CS solutions have been proposed and different methods were invented [72]. Since insiders are human beings [46], technical solutions are not sufficient methods of protection and detection, so in-depth investigation on user behavior is crucial. The advances in ML tools designed to detect insider malicious activities often fail due to a number of detection challenges, such as dataset for evaluation issues[8]. According to research done by [73], the COVID-19 epidemic has created ideal circumstances for SEAs to increase dramatically. Most organizations allow employees to work from home and connect via various Internet services without providing adequate cyber protection [46]. This massive proportion underwent rapid changes in many activities, bringing gigantic problem in CS [69]. As a result, threat incidents dramatically increased more than ever before [70]. Another study done by a CS firm showed that 78% of the firm IT organization leaders have to think cyber threat risks become higher because insiders get access permission from remote locations including employees' homes [46]. Now attackers use complex methods and get into the target organization, then they exploit their threat, that causes financial loss, identity theft, and sabotage. Among the many different CS threats that occur on the internet, Distributed denial of service (DDoS) attacks are the most frequent [5]. The number of DDoS attacks around the world is continuously rising [57]. Even if CS countermeasures are continually advanced and deployed, attackers disrupt physical infrastructure by infiltrating the digital systems that control physical processes, damaging specialized equipment, and disrupting vital services without a physical attack. These dangers are still complicated [71]. Insider, the human being gives less priority in the process of strengthening CS countermeasures [18].

Similarly, [5] proposes a CS culture framework to identify insider threats by promoting employee awareness and training. Meanwhile, [22] suggests implementing an employee awareness model to enhance knowledge of SEAs in Saudi Arabia's public sector. On the other hand, A review by [74] explores social engineering defense mechanisms and information security policies. The authors suggest that solving the root cause of cyber-attacks,

requires policymakers and promoting organization wide CS culture [5]. Similarly, [75, 8] points out that personal characteristics, such as trustfulness, can make employees more susceptible to SEAs. In general, insider threats and SEAs are significant challenges in organization security that require security policies, security culture, training, awareness, and the development of effective IDSs. The article presented by [13], brings a hybrid IDS framework that combining rule-based detection with one-dimensional CNNs. In addition, [21, 7, 43, 33] are recent works that bring solutions for intrusion detection by combining the classical ML algorithms DL. Table 2 compares different papers on threat detection, discussing their approaches, datasets utilized, advantages, limitations, and proposed improvements. The first research, conducted by [31], employs a variety of machine learning and deep learning techniques like artificial neural networks, support vector machines, principal component analysis, and decision trees to identify insider threats. The combination of these methods in the study yields superior outcomes, and the authors recommend further experimentation on diverse datasets for comprehensive findings. Another study, by [13], achieves high accuracy in intrusion detection but lacks real-time testing, resulting in slower outcomes. The authors also propose more thorough evaluations for response scenarios to data breaches. In conclusion, these works underscore the significance of integrating different ML and DL method for effective threat detection, underscoring the necessity for continual enhancement and assessment.

Table 2: Comparison of Threat Detection Articles

Title	Methods	Dataset	Advantages	Limitations, Proposed Solutions
Insider Threat Detection Based on User Behavior Modeling and Anomaly Detection Algorithms [31]	UBM, ANN, SVM, PCA, DT	UNSWB15	Combines multiple methods for finding anomalies, detects malicious activities	Further experiments needed on different datasets with varying numbers of users; Detailed measures and comparison techniques could be included
ID using CNN feature extraction with EPCA for dimensionality reduction [76]	CNN, EPCA	Not specified	High accuracy in intrusion detection, categorizes multiple at- tacks	Lacks real-time testing leading to slower results; Further evaluation needed for data breach response scenarios
Employee Awareness Model to Enhance Awareness of Social Engineering [22]	Interviews, surveys	N/A	Highlights existing security awareness practices, conceptual framework for awareness improvement	Limited access to data due to privacy regulations
Social Engineering Attacks Prevention: A Systematic Literature Review [39]	Systematic literature review	N/A	Comprehensive review of social engineering attack prevention, identifies countermeasures	May not cover the latest SEA tactics; Could explore emerging SEA techniques
A novel study of preventing cybersecurity threats [77]	Systematic re- view	Not specified	Theoretical frame- work for prevention methods	Focus on prevention methods, less on specific attacks; More details needed on specific prevention techniques
Insider Threat Detection Using Machine Learning Approach [32]	RF & SVM	CERT (Computer Emergency Response Team)	Provide real time detection	May produce malicious flags for normal user behavior & Required optimal performance and detection time

Table 3 presents a comparative analysis of various detection methods used for threat detection. DAE method, by [13], shows an accuracy of 0.90, precision of 0.95, recall of 0.95, and an F1-score of 0.92, with the limitation that optimization is required. The VAE method, also from [13], performs slightly better with an accuracy of 0.92, precision of 0.96, recall of 0.96, and an F1-score of 0.94, but needs further evaluation on other datasets. The combination of LR with DAE, according to [62], achieves an accuracy of 0.90, precision of 0.92, recall of 0.92, and an F1-score of 0.95, though its performance is highly dependent on parameter tuning. Combining AE with Isolation Forest (IF), by [30], shows the highest performance with an accuracy of 0.95, precision of 0.95, recall of 0.97, and an F1-score of 0.96, but it has limited scalability for large-scale deployments.

Table 3: Performance Comparison of Existing Detection Methods

Method	Accuracy	Precision	Recall	F1-Score	Limitation
DAE [13]	0.90	0.95	0.95	0.92	Optimization required
VAE [13]	0.92	0.96	0.96	0.94	Need further evaluation on other datasets
LR + DAE [62]	0.90	0.92	0.92	0.95	Parameter tuning critical for optimal performance
AE + IF [30]	0.95	0.95	0.97	0.96	Requires significant computational resources; Limited scalability for large-scale deployments

V. Proposed Approaches

We propose an integrated framework for detecting insider threats and social engineering attacks (SEAs) that merges multiple datasets representing insider behavior, network traffic flows, and host-based activities. This framework will utilize advanced data preprocessing techniques including data cleaning, exploratory data analysis (EDA), normalization, feature engineering, and dimensionality reduction to optimize the dataset for model training and evaluation. A key component of our methodology will be the conduct of thorough EDA to identify critical data distributions, correlations, and anomalies. This exploratory step is essential for informing feature selection and ensuring that the most relevant input variables are utilized during model training. To address class imbalances, particularly evident in datasets like UNSW-NB15, we will implement sophisticated sampling techniques, including SMOTE (Synthetic Minority Over-Sampling Technique), to enhance model sensitivity to minority classes while maintaining overall dataset integrity. <light>Our proposed detection solution will leverage multimodal deep learning (DL) architectures.</light> Specifically, we will employ a stacking DL approach that incorporates ensemble methods to refine model performance across diverse datasets. By utilizing ensemble learning techniques and hyperparameter optimization, we aim to fine-tune models to achieve heightened accuracy, sensitivity, and specificity. The deployment of our proposed conceptual system architecture aligns with the goal of creating a holistic detection framework that integrates user behavior analysis into existing network and host-based intrusion detection systems (IDS). This unified model will not only facilitate real-time detection but will also adapt to the evolving landscape of insider threats and SEAs. Figure 7 presents a visual representation of the proposed architecture, which integrates various data streams and detection methods into a cohesive system.

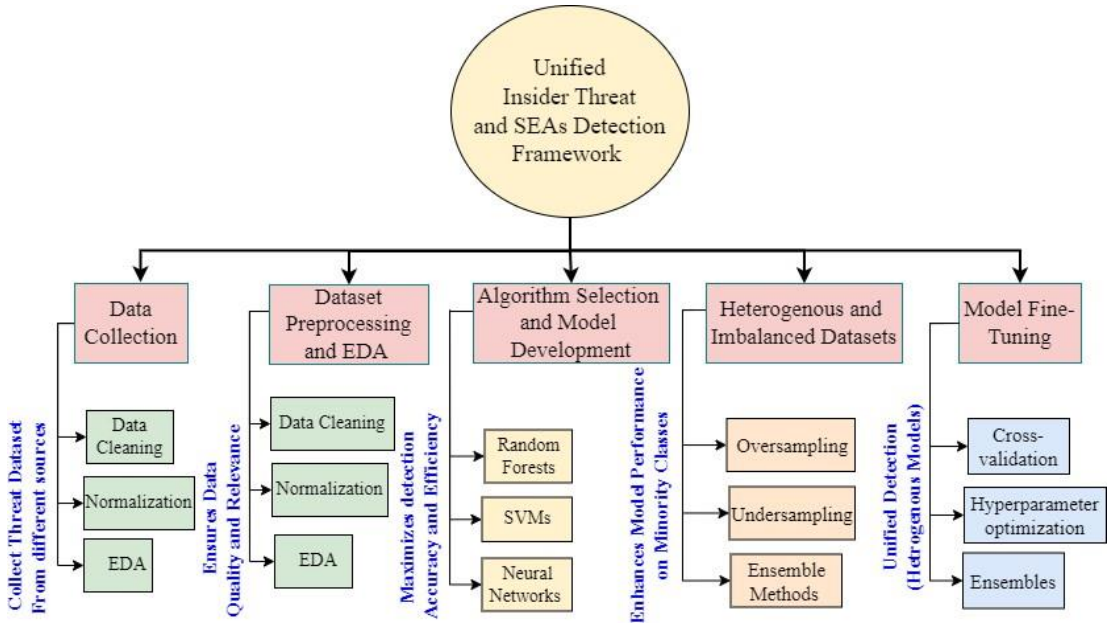

Figure 7: Proposed Conceptual System Architecture

VI. Conclusion

This comprehensive systematic review highlights the current challenges posed by insider threats and SEAs to organizational CS. In addition to there are several datasets available for evaluating detection solutions. However, it is necessity to enhance detection methods, as current network and host based threat detection approaches and

signature and anomaly based detection methods are insufficient in detecting newly emerging attacks such as insider and SEAs. This article also discusses current threat detection methods, focusing on insiders, SEAs, and phishing. It also underscores the limitations of existing detection systems and algorithms in countering these threats. Researchers have made valuable contributions to enhancing existing IDSs but securing organizations from insider threats remains a formidable challenge. Therefore, there is a pressing need to enhance detection systems. Here we proposed a DL based solution for better detection of insider threats and SEAs. Our research directions include exploring the integration of network traffic datasets with host-based detection datasets to create a unified and robust detection systems. Integrating insider CS behavior into the existing fragmented network and host based IDS system using DL methods is crucial for a comprehensive CS countermeasure. Ultimately, the integration of these disjointed detection approaches is essential to develop a holistic or unified and effective detection against insider and SEAs.

A Contribution

In the current digital world, CS threats, particularly insider threats and SEAs, present significant challenges to organizational integrity. Despite advancements in threat detection approaches, existing methods often struggle to accurately identify and mitigate these evolving threats. The primary contribution of this article is to address these newly emerging threats through a systematic review of current detection methods employed for insider threats and SEAs, analyzing the limitations of traditional IDS and their inability to adapt to heterogeneous complex threats and attack vectors, thereby providing a substantive understanding of existing gaps. Moreover, a central innovation of this work is the integration of diverse and heterogeneous user behaviors with ML and DL techniques into the detection framework. This innovative integration creates a resilient approach that effectively adapts to the ongoing variations in user behavior, network traffic patterns, host activities, and individual characteristics. Furthermore, this study introduces a holistic conceptual framework that harmonizes diverse user behavior analytics with both network and host-based detection techniques.

REFERENCES

[1] M. Singh, B. Mehtre, S. Sangeetha, and V. Govindaraju, "User behaviour based insider threat detection using a hybrid learning approach," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 4, pp. 4573–4593, 2023.

[2] W. S. Admass, Y. Y. Munaye, and A. Diro, "Cyber security: State of the art, challenges and future directions," *Cyber Security and Applications*, p. 100031, 2023.

[3] S. Kemp, "Digital 2021: Global overview report," *DataReportal. Recuperado de https://datareportal.com/reports/digital-2021-global-overview-report*, vol. 0, no. 1, 2021.

[4] M. Al-Fawa'reh, M. Al-Fayoumi, S. Nashwan, and S. Fraihat, "Cyber threat intelligence using pca-dnn model to detect abnormal network behavior," *Egyptian Informatics Journal*, vol. 23, no. 2, pp. 173–185, 2022.

[5] A. Georgiadou, S. Mouzakitis, and D. Askounis, "Detecting insider threat via a cyber-security culture framework," *Journal of Computer Information Systems*, pp. 1–11, 2021.

[6] S. Perumal, M. Tabassum, G. Narayana Samy, S. Ponnan, A. K. Ramamoorthy, and K. Sasikala, "Cybercrime issues in smart cities networks and prevention using ethical hacking," in *Data-Driven Mining, Learning and Analytics for Secured Smart Cities*, pp. 333–358, Springer, 2021.

[7] C. Liu, Z. Gu, and J. Wang, "A hybrid intrusion detection system based on scalable k-means+ random forest and deep learning," *IEEE Access*, vol. 9, pp. 75729–75740, 2021.

[8] A. G. Akpan, J. O. Ugah, and V. N. Ezeano, "Leveraging on cyber security for digital economy: Analysis of emerging cyber security threats and attacks,"

[9] N. Ahmed, A. b. Ngadi, J. M. Sharif, S. Hussain, M. Uddin, M. S. Rathore, J. Iqbal, M. Abdelhaq, R. Alsaqour, S. S. Ullah, *et al.*, "Network threat detection using machine/deep learning in sdn-based platforms: A comprehensive analysis of state-of-the-art solutions, discussion, challenges, and future research direction," *Sensors*, vol. 22, no. 20, p. 7896, 2022.

[10] J. Payne, "Annual data exposure report 2023." https://www.code42.com/resources/reports/ 2023-data-exposure, 2023. Code42.

[11] W. Syafitri, Z. Shukur, U. A. Mokhtar, R. Sulaiman, and M. A. Ibrahim, "Social engineering attacks prevention: A systematic literature review," *IEEE Access*, 2022.

[12] Z. Wang, Y. Liu, D. He, and S. Chan, "Intrusion detection methods based on integrated deep learning model," *Computers & Security*, vol. 103, p. 102177, 2021.

[13] E. Pantelidis, G. Bendiab, S. Shiaeles, and N. Kolokotronis, "Insider detection using deep autoencoder and variational autoencoder neural networks," *arXiv preprint arXiv:2109.02568*, 2021.

[14] E. System, "Insider threat statistics: Facts and figures." https://www.ekransystem.com/en/blog/ insider-threat-statistics-facts-and-figures, 2023. Accessed: 2024-06-14.

[15] Z. Ahmad, A. Shahid Khan, C. Wai Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4150, 2021.

[16] M. Lehto and J. Limnéll, "Strategic leadership in cyber security, case finland," *Information Security Journal: A Global Perspective*, vol. 30, no. 3, pp. 139–148, 2021.
[17] V. K. Fedorov, E. Balenko, Y. I. Starodubtsev, and P. Zakalkin, "Cyberspace: Key properties and traits," in *Journal of Physics: Conference Series*, vol. 2096, p. 012039, IOP Publishing, 2021.
[18] D. C. Le and N. Zincir-Heywood, "Anomaly detection for insider threats using unsupervised ensembles," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1152–1164, 2021.
[19] M. Verkerken, L. Dhooge, T. Wauters, B. Volckaert, and F. De Turck, "Unsupervised machine learning techniques for network intrusion detection on modern data," in *2020 4th Cyber Security in Networking Conference (CSNet)*, pp. 1–8, IEEE, 2020.
[20] H. Owen, J. Zarrin, and S. M. Pour, "A survey on botnets, issues, threats, methods, detection and prevention," *Journal of Cybersecurity and Privacy*, vol. 2, no. 1, pp. 74–88, 2022.
[21] A. Alshammari and A. Aldribi, "Apply machine learning techniques to detect malicious network traffic in cloud computing," *Journal of Big Data*, vol. 8, no. 1, pp. 1–24, 2021.
[22] M. F. Alghenaim, N. A. A. Bakar, R. C. M. Yusoff, N. H. Hassan, and H. Sallehudin, "Employee awareness model to enhance awareness of social engineering threats in the saudi public sector," in *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, pp. 1–6, IEEE, 2021.
[23] Verizon, "DBIR 2023 Data Breach Investigations Report." https://example.com/dbir2023, 2023. Accessed on June 1, 2023.
[24] Y. Li and Q. Liu, "A comprehensive review study of cyber-attacks and cyber security; emerging trends and recent developments," *Energy Reports*, vol. 7, pp. 8176–8186, 2021.
[25] F. Antonucci and M. M. Chowdhury, "Botnets as the modern attack vector," in *2022 IEEE World AI IoT Congress (AIIoT)*, pp. 585–590, IEEE, 2022.
[26] A. Alraizza and A. Algarni, "Ransomware detection using machine learning: A survey," *Big Data and Cognitive Computing*, vol. 7, no. 3, p. 143, 2023.
[27] N. Gupta, V. Jindal, and P. Bedi, "Cse-ids: Using cost-sensitive deep learning and ensemble algorithms to handle class imbalance in network-based intrusion detection systems," *Computers & Security*, vol. 112, p. 102499, 2022.
[28] Y. Guo, "A review of machine learning-based zero-day attack detection: Challenges and future directions," *Computer Communications*, vol. 198, pp. 175–185, 2023.
[29] L. Wang and R. Jones, "Big data analytics in cyber security: network traffic and attacks," *Journal of Computer Information Systems*, vol. 61, no. 5, pp. 410–417, 2021.
[30] K. Sadaf and J. Sultana, "Intrusion detection based on autoencoder and isolation forest in fog computing," *IEEE Access*, vol. 8, pp. 167059–167068, 2020.
[31] J. Kim, M. Park, H. Kim, S. Cho, and P. Kang, "Insider threat detection based on user behavior modeling and anomaly detection algorithms," *Applied Sciences*, vol. 9, no. 19, p. 4018, 2019.
[32] V. Kumar, D. Sinha, A. K. Das, S. C. Pandey, and R. T. Goswami, "An integrated rule based intrusion detection system: analysis on unsw-nb15 data set and the real time online dataset," *Cluster Computing*, vol. 23, pp. 1397–1418, 2020.
[33] M. Radhi Hadi and A. Saher Mohammed, "A novel approach to network intrusion detection system using deep learning for sdn: Futuristic approach," *arXiv e-prints*, pp. arXiv–2208, 2022.
[34] A. R. Khan, M. Kashif, R. H. Jhaveri, R. Raut, T. Saba, and S. A. Bahaj, "Deep learning for intrusion detection and security of internet of things (iot): current analysis, challenges, and possible solutions," *Security and Communication Networks*, vol. 2022, 2022.
[35] A. H. Washo, "An interdisciplinary view of social engineering: A call to action for research," *Computers in Human Behavior Reports*, vol. 4, p. 100126, 2021.
[36] E. Jaw and X. Wang, "Feature selection and ensemble-based intrusion detection system: an efficient and comprehensive approach," *Symmetry*, vol. 13, no. 10, p. 1764, 2021.
[37] A. Al-Harrasi, A. K. Shaikh, and A. Al-Badi, "Towards protecting organisations data by preventing data theft by malicious insiders," *International Journal of Organizational Analysis*, 2021.
[38] P. Bayl-Smith, R. Taib, K. Yu, and M. Wiggins, "Response to a phishing attack: persuasion and protection motivation in an organizational context," *Information & Computer Security*, vol. 30, no. 1, pp. 63–78, 2022.
[39] W. Syafitri, Z. Shukur, U. A. Mokhtar, R. Sulaiman, and M. A. Ibrahim, "Social engineering attacks prevention: A systematic literature review," *IEEE Access*, vol. 10, pp. 39325–39343, 2022.
[40] H. Alqahtani, I. H. Sarker, A. Kalim, S. M. Minhaz Hossain, S. Ikhlaq, and S. Hossain, "Cyber intrusion detection using machine learning classification techniques," in *Computing Science, Communication and Security: First International Conference, COMS2 2020, Gujarat, India, March 26–27, 2020, Revised Selected Papers 1*, pp. 121–131, Springer, 2020.
[41] S. Ali, S. U. Rehman, A. Imran, G. Adeem, Z. Iqbal, and K.-I. Kim, "Comparative evaluation of ai-based techniques for zero-day attacks detection," *Electronics*, vol. 11, no. 23, p. 3934, 2022.
[42] K. Kioskli, T. Fotis, S. Nifakos, and H. Mouratidis, "The importance of conceptualising the human-centric approach in maintaining and promoting cybersecurity-hygiene in healthcare 4.0," *Applied Sciences*, vol. 13, no. 6, p. 3410, 2023.
[43] Z. Wang, H. Zhu, and L. Sun, "Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods," *IEEE Access*, vol. 9, pp. 11895–11910, 2021.
[44] R. Eckhardt and S. Bagui, "A user-centric focus for detecting phishing emails," in *AI, Machine Learning and Deep Learning*, pp. 313–333, CRC Press.

[45] L. Liu, O. De Vel, Q.-L. Han, J. Zhang, and Y. Xiang, "Detecting and preventing cyber insider threats: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1397–1417, 2018.
[46] J. Groenendaal and I. Helsloot, "Cyber resilience during the covid-19 pandemic crisis: A case study," *Journal of Contingencies and Crisis Management*, vol. 29, no. 4, pp. 439–444, 2021.
[47] M. Essam, "Intrusion detection system with ml and dl." https://www.kaggle.com/code/essammohamed4320/ intrusion-detection-system-with-ml-dl, 2023. Accessed: 2024-05-27.
[48] R. A. Alsowail and T. Al-Shehari, "A multi-tiered framework for insider threat prevention," *Electronics*, vol. 10, no. 9, p. 1005, 2021.
[49] V. Zimmermann and K. Renaud, "Moving from a human-as-problem to a human-as-solution cybersecurity mindset," *International Journal of Human-Computer Studies*, vol. 131, pp. 169–187, 2019.
[50] Y. Aun, M. Gan, N. Wahab, and G. H. Guan, "Social engineering attack classifications on social media using deep learning," *Comput. Mater. Contin*, vol. 74, pp. 4917–4931, 2023.
[51] T. Boros, A. Cotaie, A. Stan, K. Vikramjeet, V. Malik, and J. Davidson, "Machine learning and feature engineering for detecting living off the land attacks.," in *IoTBDS*, pp. 133–140, 2022.
[52] M. A. Khan, "Hcrnnids: hybrid convolutional recurrent neural network-based network intrusion detection system," *Processes*, vol. 9, no. 5, p. 834, 2021.
[53] I. Martins, J. S. Resende, P. R. Sousa, S. Silva, L. Antunes, and J. Gama, "Host-based ids: A review and open issues of an anomaly detection system in iot," *Future Generation Computer Systems*, 2022.
[54] M. K. Hooshmand and M. D. Huchaiah, "Network intrusion detection with 1d convolutional neural networks," *Digital Technologies Research and Applications*, vol. 1, no. 2, pp. 25–34, 2022.
[55] M. Al-Qatf, Y. Lasheng, M. Al-Habib, and K. Al-Sabahi, "Deep learning approach combining sparse autoencoder with svm for network intrusion detection," *Ieee Access*, vol. 6, pp. 52843–52856, 2018.
[56] M. N. Al-Mhiqani, R. Ahmad, Z. Zainal Abidin, W. Yassin, A. Hassan, K. H. Abdulkareem, N. S. Ali, and Z. Yunos, "A review of insider threat detection: Classification, machine learning techniques, datasets, open challenges, and recommendations," *Applied Sciences*, vol. 10, no. 15, p. 5208, 2020.
[57] J. Singh and N. Jyoti, "A comprehensive review: Detection and mitigation solutions of ddos attacks in cps," *Security and Resilience of Cyber Physical Systems*, p. 61, 2022.
[58] I. Ahmad, Q. E. Ul Haq, M. Imran, M. O. Alassafi, and R. A. AlGhamdi, "An efficient network intrusion detection and classification system," *Mathematics*, vol. 10, no. 3, p. 530, 2022.
[59] S. G. Bhol, J. Mohanty, and P. K. Pattnaik, "Taxonomy of cyber security metrics to measure strength of cyber security," *Materials Today: Proceedings*, 2021.
[60] A. S. Dina and D. Manivannan, "Intrusion detection based on machine learning techniques in computer networks," *Internet of Things*, vol. 16, p. 100462, 2021.
[61] M. Ghurab, G. Gaphari, F. Alshami, R. Alshamy, and S. Othman, "A detailed analysis of benchmark datasets for network intrusion detection system," *Asian Journal of Research in Computer Science*, vol. 7, no. 4, pp. 14–33, 2021.
[62] Z. Wang, Y. Ren, H. Zhu, and L. Sun, "Threat detection for general social engineering attack using machine learning techniques," *arXiv preprint arXiv:2203.07933*, 2022.
[63] A. Raza, S. Memon, M. A. Nizamani, and M. H. Shah, "Machine learning-based security solutions for critical cyber- physical systems," in *2022 10th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1–6, IEEE, 2022.
[64] F. Yuan, Y. Cao, Y. Shang, Y. Liu, J. Tan, and B. Fang, "Insider threat detection with deep neural network," in *International Conference on Computational Science*, pp. 43–54, Springer, 2018.
[65] G. N. Samy, N. Maarop, B. Shanmugam, M. Radhakrishnan, S. Perumal, and F. A. Rahim, "Multidimensional insider threat detection model for organization," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 20, pp. 4770–4785, 2021.
[66] A. C. Mazari, N. Boudoukhani, and A. Djeffal, "Bert-based ensemble learning for multi-aspect hate speech detection," *Cluster Computing*, pp. 1–15, 2023.
[67] K. Fotiadou, T.-H. Velivassaki, A. Voulkidis, D. Skias, S. Tsekeridou, and T. Zahariadis, "Network traffic anomaly detection via deep learning," *Information*, vol. 12, no. 5, p. 215, 2021.
[68] K. Hughes-Lartey, M. Li, F. E. Botchey, and Z. Qin, "Human factor, a critical weak point in the information security of an organization's internet of things," *Heliyon*, vol. 7, no. 3, p. e06522, 2021.
[69] T. Pósa and J. Grossklags, "Work experience as a factor in cyber-security risk awareness: A survey study with university students," *Journal of Cybersecurity and Privacy*, vol. 2, no. 3, pp. 490–515, 2022.
[70] K. Khando, S. Gao, S. M. Islam, and A. Salman, "Enhancing employees information security awareness in private and public organisations: A systematic literature review," *Computers & Security*, vol. 106, p. 102267, 2021.
[71] M. Lehto, "Cyber-attacks against critical infrastructure," in *Cyber Security*, pp. 3–42, Springer, 2022.
[72] M. Imran, M. H. Durad, F. A. Khan, and A. Derhab, "Toward an optimal solution against denial of service attacks in software defined networks," *Future Generation Computer Systems*, vol. 92, pp. 444–453, 2019.
[73] S. Venkatesha, K. R. Reddy, and B. Chandavarkar, "Social engineering attacks during the covid-19 pandemic," *SN computer science*, vol. 2, pp. 1–9, 2021.

[74] D. Alharthi and A. Regan, "A literature survey and analysis on social engineering defense mechanisms and infosec policies," *International Journal of Network Security & Its Applications (IJNSA) Vol*, vol. 13, 2021.

[75] M. K. Alotaibi, *The Influence of Personal Characteristics and Other Factors on the Susceptibility of Public Sector Employees to Cyber-Social Engineering Through LinkedIn: A Mixed-Methods Sequential Explanatory Study*. PhD thesis, Trinity College Dublin, 2021.

[76] A. Kayyidavazhiyil and M. Silic, "Intrusion detection using deep (cnn) convolutional neural network feature extraction with (epca) enhanced principal component analysis for dimensionality reduction," *Global journal of Business and Integral Security*, 2022.

[77] M. I. Alghamdie, "A novel study of preventing the cyber security threats," *Materials Today: Proceedings*, 2021.