IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 10 September 2022, accepted 6 October 2022, date of publication 14 October 2022, date of current version 21 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3214819

RESEARCH ARTICLE

Threats From Unintentional Insiders: An Assessment of an Organization's Readiness Using Machine Learning

M. M. HAFIZUR RAHMAN[1], MOHAMMED ABDULAZIZ AL NAEEM[1], AND ADAMU ABUBAKAR[2]

[1]Department of Computer Networks & Communications, CCSIT, King Faisal University, Al Hassa 31982, Saudi Arabia
[2]Department of Computer science, KICT, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia

Corresponding author: M. M. Hafizur Rahman (mhrahman@kfu.edu.sa)

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, under Grant 1222.

ABSTRACT Today's organisations are facing a number of challenges, one of the most significant of which is ensuring the safety of their digital data. This is as a result of the fact that they are frequently faced with internal and external threats that can put the data they have been entrusted with in jeopardy of being compromised. As a result of this, this study investigates the dimension of threats associated to unintentional internal user of an organisation and utilises NARX to model and test a detection scheme associated to the menace. In addition, this study aims to provide a better understanding of the current state of the threat landscape. The data adopted for this research is primarily a ''user activity logs'' dataset from CERT (release version r4.2). From the data, the study conceptualized ''Access'', ''Motivation'', and ''Action'' to be the key dimensions influencing ''insider'', whereas ''Intent'', ''+Action'', ''Method'', and ''knowledge'' are the key dimension influencing ''threats''. Experimental analyses conducted by NARX within several numbers of partitions of the data point to a good detection capacity, with the greatest value of R^2 coming in at 0.97. This indicates that NARX was able to detect the crucial dimension that was formulated for by the research to be the detections parameter of an inadvertent insider threat when operating under the best partition. In light of these findings, organisations can use the proposed approach to assess their preparedness for Insider attacks.

INDEX TERMS Unintentional insiders, threats, network, attacks, data breaches.

I. INTRODUCTION

Getting into an organization's computer system is the first and most essential step in the rise of the insider threat [1]. Therefore, the capability to gain access to an organization's information technology resources is the first step in the establishment of an internal threat in an organization [2]. When an ''individual working in a company'' gains access to the system of an organisation, there is a possibility that a ''insider threat'' will emerge. This is the case regardless of whether the individual has the intention to cause harm or is acting inadvertently. The latter refers to potential dangers that could occur as a consequence of human error or the violation of a policy [3]. The ''intended insiders threat'' and the ''unintended insiders threat'' are both categories that can be used to describe the ''insiders threat''. Intended insiders are individuals who have the ability to carry out deliberately malicious actions directed at any organisation for a variety of reasons [4]. Insiders who are not who they claim to be can be classified as either traitors or masqueraders [5]. On the other hand, unintended insiders are people who accidentally launch attacks within an organisation due to inadvertent actions such as breaking security policy [6]. These types of attacks are more common than intended insider attacks [4] Employees can get unintentional access to sensitive information and then accidentally change or delete it, even if they have sufficient authorization and the ability to access the information in the first place [7]. An unintended policy violation, on the other

The associate editor coordinating the review of this manuscript and approving it for publication was Songwen Pei.

hand, does not constitute a security policy breach in the same way that a hostile policy violation does [8]. Threats from within the organisation make it more difficult to differentiate between actions that are authorised and those that are not. It may be difficult to differentiate between appropriate and harmful behaviour in the workplace if users have access to internal resources [9]. This indicates that insiders pose a risk because they have the potential to act maliciously and possess privileged access to either physical or virtual networks [10].

The problem statement associated to this study lies with the fact that everyone is aware that the goal of attacks carried out by inadvertent insiders is to compromise IT resources that have a low level of protection or are otherwise susceptible [11]. Access limitations are frequently sidestepped by insiders who do it on purpose. If they have administrative privileges, they will be able to conceal their own activities if they so want [12]. Employees and administrators are two types of people that could be considered unintentional insiders. They are able to cause more damage to the organization in proportion to the amount of power they possess [13]. Employees within a corporation typically pose a greater threat than those working outside the corporation due to the fact that employees have more authority to carry out their jobs and more in-depth knowledge of the corporation's internal workings [14]. The danger posed by internal threats is typically greater than that posed by exterior dangers. Employees are familiar with the company's systems and have the authority to use such systems in order to carry out the obligations that are assigned to them. Break-ins are typically riskier than other types of internal invasions. Every employee at a company ought to be aware of their own shortcomings and be able to turn those shortcomings to their own advantage [15].

Another research problem that this research is willing to solves lies with the fact that the costs associated with events that take place within the company are higher than those associated with events that take place outside of the company. It is the responsibility of employers to ensure that their staff members are aware of their own vulnerabilities as well as the means by which to exploit those flaws to their own advantage [16]. When compared to dealing with difficulties on the outside, dealing with problems within the organisation is more expensive. An insider has the ability to take advantage of the company's security flaws since they are aware of those flaws [17]. Those with access to insider information have quick and easy access to information since they are aware of where to look. The insider has a significant advantage over the outsider because of their understanding of the company's security protocols and any holes they may have [18]. Those who have access to confidential information have the ability to access a variety of information [19]. Insiders may target weak points in data protection or security, such as exposed information. Those with access on the inside are able to bypass the security measures [20]. Administrators that have the capacity to establish accounts and adjust settings are able to take on the persona of another user so that they can carry out their responsibilities incognito [21].

Considering the problems identified by this current research, and despite the greater possibility of unintentional attacks to occur at any time, the research community has not yet covered the large region required to detect this form of insider threat. Even though today's most advanced firewalls, intrusion detection systems (IDS), anti-virus, anti-spam, and email security software are all equipped with powerful tools to do so, there is no way to detect an insider threat [22]. This is the case despite the fact that there is no way to detect an insider threat. The most prevalent reasons for unintended insider threats are unauthorised network access, sabotage, data manipulation, illegal network access, and the usage of unapproved network ports [23]. Unintentional insider risks can also be caused through unlawful network access. Abuse of a network can have extremely negative effects on how an employee's network access is managed, which can have serious repercussions [24]. In the previous studies, there was a lack of individual behavior-based detection approaches. The unintended insider risk also includes a significant component that involves the unauthorised use of data obtained from an outside source [25]. Access to the company's networks is relatively easy to come by, but not everyone has direct access to the company's information technology resources [26]. In the prior studies, it was difficult to track down findings associated to those different resources. This is due to the fact that the negligent activities of company insiders can cause damage to businesses by exploiting vulnerabilities in their computer networks and data storage systems.

This research has established that it is a good idea to use machine learning to investigate threats from inadvertent insiders for an organization's readiness, since ''Access Management'' by insiders is a challenge for the modern organisational network. This problem can be solved by using machine learning. Those individuals who have a genuine requirement to make use of the system are able to do so. There is the possibility of putting together a structured dataset that can be put to use for gaining a deeper understanding of a variety of situations. It is wise to use the specifics of the major access control systems for future anomaly detection in order to protect the infrastructure of the organisation from both external and internal threats. This may be accomplished by using the information obtained from large access control systems. Utilizing this type of detection is one way to further secure one's safety. This study does not plan how experts would analyse the various levels of access that employees have based on their jobs and behaviours, which is an important oversight, but to ensure that it captured the dataset for complete access to organizational system and considering each member of the company that ought to be given a certain position within the system. Roles can be subdivided in addition to the connected activities of the user within organization by the current data available. That is why the objective of this current study is to examine threats from unintentional insiders for an organization's readiness using machine learning.

The remaining parts of the paper are organized as follows: after this current section, which provides an overview of the

research, section two presents the previous research studies on insider threats and the detection techniques involved in undertaking detection of an insider threat. Following the discussion of the research methodology in the third section, the experimental analysis and the findings of the study are presented in the fourth section. A comprehensive analysis of the findings and a discussion of the implications of the study are presented in section five, while the conclusion to the study is presented in section six.

II. RELATED WORK

Many studies have been conducted on ''Insider Threats''. Despite the fact that many other analyses classify it as an issue of security. Typical to that is the work involving ''Security Analysis'' [12], emerging threats in cybersecurity [13], ''Network threats'' [15], weak point in the information security of an organization [16], Cyber threats [17], Insider Threats to IT Security [20] and most important lie with Unintentional Insider [23]. Access is a significant component of insider threats, as proven by observations from the real world, assertions from academics, and discoveries from the information technology sector. The bulk of research that have been conducted on the topic of insider threats have come to the conclusion that the primary reason that insiders steal data from their organisations is because they were given authorized access to the material that was subsequently taken [23], [24], [25], [26]. Seventy-five percent of thefts committed with authority were carried out, as reported by CERT, by workers or other insiders [1]. Around seven out of ten of the overall workforce had active accounts throughout the time of the incident, which indicates that the majority of authorized personnel had accounts at that time. The vast majority of formerly employed employees did not qualify in any way as authorized insiders. This is true across the board. The content is accessible to eighty-eight percent of the insiders who have been authorized to view it.

There have been a number of studies conducted in the past that have been associated with researching the behaviour of insiders, specifically their activities, to identify whether or not they provide a risk to the information and technology systems of a firm. These studies have been done in the past. According to Mazzarolo et al. [23], Cyber Misconduct on a Social Media Site is a cybersecurity risk for every business. A key factor in this is what people share online, given that almost everything about oneself is shared online, including demographic, family, activities, and work-related information. The research contends that if an organization's rules, training, and technology fail to effectively fix issues together, there may be a risk in every corporation. The report asserts that vital firm information can often be jeopardised by an employee's behaviour. Due to the fact that social media has evolved into a reconnaissance tool for bad actors and that user accounts are now seen as a gold mine by cybercriminals. The study gathers and examines open-source data from LinkedIn, finds data leaks, and uses software as a service to identify personality types (SAAS).

The study recognises that a person's attitude toward releasing sensitive information is predicted by behavioural characteristics. An insider threat prediction technique related to personal misuse of information technology (IT) resources was put forth by Magklaras and Furnell [24]. The study found that little effort is put into tackling the issue of internal IT misuse, despite the well-documented and growing insider danger to information systems. Further, it was asserted that most misuse prevention strategies target abuse coming from outside sources, specifically the perceived threat posed by unauthorized users. Because of this, the study recommends a method for calculating the degree of harm that is most likely to come from a specific insider by implementing a threat evaluation system based on specific profiles of user behaviour.

Regarding insider risks to information security and the nature of loyalty and betrayal in the context of organizational, cultural, and shifting economic and social circumstances, Colwil [25] analyses some of the important themes. According to the study, insiders represent security threats because of their legitimate access to resources and information, familiarity with the organization, and knowledge of where significant assets are located. Insiders will understand how to make the biggest impact while leaving the fewest clues. However, it's possible that organisations didn't use efficient risk management procedures to deal with the pace and scope of change. According to Colwi [25], outsourcing may cause protection barriers and controls to disintegrate and a rise in the number of people who are treated as full-time employees. According to the study, focusing on insiders during security risk assessments and compliance regimes is essential for containing the insider threat. According to the research, using technology alone won't result in solutions.

The majority of the time, the detection of data theft in ''Infrastructure As A Service'' (IAAS) is carried out by employing machine learning in conjunction with the K Nearest Neighbor (KNN) algorithm [27]. Although Nikolai and Wang's [27] study focuses on categorization of issues related to insider threat both intentional and unintentional. This study is closely associated to their approach, however quite a lot of issues are not clarified by their studies. They came to the conclusion that cloud security is just as important as any other security based on the Cloud Security Alliance's lists of data theft and insider assaults. In order to train, monitor, and detect a pattern of inside attacker data theft on Infrastructure as a Service nodes, they suggested using the k-nearest neighbours anomaly detection technique. They were able to observe a good detection rate for anomalous login activity and data copies to external systems using their suggested methodology.

KNN has been used in a somewhat different approach to study the message patterns of networks that are in charge of data transmission in order to determine whether or not something is incorrect [28]. In order to find a common anomaly on an organisational system, a user database, file access patterns, and statistical analysis methods were

developed and applied. Both of these tools were used to do this. According to Punithavathani et al. [28], the insider threat is only slightly addressed by most information security policies, despite the fact that insiders offer the greatest threat to a company through a variety of harmful actions. In order to combat insider threats, the research suggested a surveillance mechanism that would involve both monitoring the organisational network for any incoming and outgoing packets as well as monitoring the individual packets captured to analyse their features in order to gather any data relating to suspicious activities. The research then mines several log files that are thought to include crucial traces of information for insider attacks.

Additionally, rule mining, feature analysis using the One-Class Support Vector Machine (OCSVM), and other cryptographic techniques including watermarking were used to detect an insider threat [29]. Furthermore, according to Hu et al. [29], that a sizable portion of employees steal data when switching jobs. In other words, insider attackers who are authorised to access an organization's best-kept secrets provide a significant risk to organisational security. The investigation was conducted because not enough was being done to stop data exfiltration activities. The study profiles authorised users' proper use of file repositories using statistical techniques. With file access histories, the proposed model's efficacy was evaluated.

According to Feng et al. [30], as more businesses use cloud file-sharing services, this opens up a new avenue for potential insider threats to corporate data and intellectual property. In order to identify abnormalities, they develop a two-stage machine learning technique. In order to identify outliers for insider threats, they coordinated the establishment of an outlier function. The proposed system has been implemented in a real corporate context, and some case studies have shown its usefulness. Data gathered from ''the upload'', ''the download'', and ''web-browsing'' were used to study the insider activities linked with their transaction. Machine learning algorithms were used in order to establish patterns that can help identify insider risks [30]. This was done in order to study the insider activities linked with their transaction. K-Means++ clustering and supervised machine learning in the form of SVM classifiers are both utilised in big data analytics in order to successfully identify potential insider threats. This is performed through the usage of large data. This is done in order to manage scalability as well as anticipate network traffic [31]. A strategy for doing research on the properties of memory access that makes use of a semi-supervised method of investigation. The runtime memory access patterns of programmes that are now running on distributed compute cluster slave nodes are the basis for the proposed method [32].

It was demonstrated that integrating rules and regulations into insider operational events is a valid method for determining whether or not the behaviour of employees corresponds to the rules and regulations [33]. This was the case in the majority of the cases that were examined. This was demonstrated by referring to many everyday occurrences. Users should encrypt their credentials before sharing them with Internet apps as a sign to indicate that any insider could not possibly be able to view the content of the data even if it was intercepted. Users should share their credentials with Internet apps as a sign to indicate that they have encrypting their credentials. This is a warning notice indicating that the data's content cannot be viewed by any person who is not an authorized user of the system. This is because, in general, a password is one of the main tools for securely guarding any kind of digital transmission and storage [34]. As a result, this is why it is so important to have a password.

Scanning network packets is one of the ways that has been employed in the detection of dangers that are posed by insiders [35]. The analysis of the users' network traffic makes use of some characteristics of an insider, such as the user's device capacity, the user's security level, and the state of the network connection, in order to discover a new set of contextual approaches that can be used to determine how trustworthy the given context of an access request is [36]. These contextual approaches can be used to determine how trustworthy the given context of an access request is. It has been demonstrated that one can acquire uncomfortable psychological inclinations of individual users by analyzing the content of electronic interactions. After that, these traits can be utilized to forecast and identify threats emanating from within the organization [37]. A honey pot sensors-based method has also been used within a local network to identify the behaviors of insiders in order to generate certain patterns [38]. This has been done in order to generate certain patterns.

The research of Mazzarolo et al. [23] suffers from a lack of quantitative analysis of the behavioural factors related to security issues. due to the diversity of behaviours. prior investigation. By defining the characteristics associated with assessing behavior-threat, this research addresses the shortcomings of Mazzarolo et al. [23] and focuses in particular on accidental insider behaviour using a CERT dataset of 1000 people and 700 occurrences of insider threat definition. Potentially, this could probe unspoken concerns about insider dangers. Mazzarolo et al. [23] paper analysis, data use, and methodological approach are the main ways it differs from the present study. In the current study, a well-known dataset was used, and time-based modelling (NARX), an objective approach that has been shown to be as accurate as feasible, was used to evaluate the data. Therefore, the current paper's contribution consists of offering an accurate model, namely one with high accuracy in determining the dimensions related with insider threats.

By defining the characteristics associated with analysing behavior-threat as a model and focusing on unintentional insider behaviour with a CERT dataset of 1000 people and 700 instances of an insider threat developed and tested, the current research addresses the shortcomings of Magklaras and Furnell [24]. This provides the opportunity to study underreported insider threat issues. This paper's main distinction from previous studies is related to the

methodological approach, data utilisation, and validation. The model and dataset used in this study are well-known, in contrast to the validations needed for Magklaras and Furnell's [24] suggested tool. Therefore, the current paper's contribution is to provide an accurate model, namely one with high accuracy in identifying the dimensions related to insider threats.

Through the development and testing of a constructed model, the current study addresses Colwil's [25] flaw. While Colwi [25] analyses the situations and offers some important remarks related to them, he does not do any empirical research. This paper's main distinction from previous research relates to the validation testing; Colwi [25] instead of testing or establishing a model, he instead conducts an analysis to evaluate the circumstance. Therefore, the current paper's contribution is to provide an accurate model, i.e., a model for identifying the dimensions related to insider threats.

The shortcomings of Nikolai and Wang's [27] research to analyse unintentional behavioural variables linked to insider danger is one of its flaws. In particular, it focuses on unintentional insider behaviour using CERT dataset to investigate the potential of the silent issues related to unintentional insider threats. This research addresses the shortcomings of Nikolai and Wang [27] by establishing the dimensions associated to analysing unintentional behavior-threat. This paper's main distinction from earlier work has to do with classification and time-series analysis. Nikolai and Wang [27] focus on categorization issues, but this paper expands the investigation to time-series utilising the NARX model. Therefore, the current paper's contribution consists of offering an accurate model, namely one with high accuracy in determining the dimensions related with insider threats.

The shortcomings of Punithavathani et al. [28] to test and validate the model related to the issue at hand is one of their weaknesses. By specifying the dimensions, creating a model, and testing the model with a benchmarking well-known dataset, the current research fixes the flaw of Punithavathani et al. [28]. Punithavathani et al. [28] work solely with generated data that was created for their research situation and has not been confirmed. In contrast to Punithavathani et al. [28] work, our study was unable to validate the model pertinent to the problem at hand. In addition, this research added dimensions to a model and tested it against a benchmark dataset, which Punithavathani et al. [28] do not do. Only generated data that was produced for Punithavathani et al. [28] research scenario and has not been verified is used in their work. The current paper's contribution is thus to provide an accurate model, namely one that is highly accurate in identifying the dimensions associated with insider threats.

While the study by Hu et al. [29] had certain flaws, one of them was that it was unable to be dynamic, meaning that it was not based on time series and would be rendered useless by any change to any of the statistical model's proposed variables. A time series model was therefore applied in our research to address Hu et al. [29] flaw. Using a benchmarking well-known dataset, a dynamic model linked to time series was tested. Because the study is not dynamic, or time series-based, where any change to any proposed statistical model variable would render the model unusable, it differs from earlier research in this area. We consequently addressed the problem raised by Hu et al. [29] by using a time series model. An experimental dynamic model linked to time series was tested using a benchmarking well-known dataset. The current study makes a contribution by offering a precise model, namely one that is quite accurate in calculating the dimensions related to insider threats.

The shortcomings of the Feng et al. [30] research to first test and validate models that are related with behavioural aspects associated is where its flaws lie. By creating the dimensions related to evaluating behavior-threat, this research addresses the flaw of Feng et al. [30] and focuses in particular on unintentional insider activity derived from the CERT dataset. The primary ways in which the present study differs from this paper are in its analysis, data use, and methodological approach. Time-based modelling (NARX), an objective method that has been demonstrated to be as accurate as practical, was employed in the current study to analyse the data using a well-known dataset. The current paper's contribution is thus to provide an accurate model, namely one that is highly accurate in identifying the dimensions associated with insider threats.

III. RESEARCH METHODOLOGY

The research methodology that underpins this current study, which focuses on the context of the insider threat, is comprised of the conceptualization of the insider threat and the machine learning experimental analysis of the detection of unintentional insider threats. Both of these elements are essential to the context of the insider threat (see Figure 1). One of the most significant challenges that must be surmounted in order to detect and prevent attacks carried out by insiders is the fact that very few studies on this subject have been designed to find a solution to the problem in a broad and comprehensive manner. This is one of the obstacles that must be overcome. The majority of the models that have been investigated for the purpose of this research center on the insider threat and how it relates to specific problems that arise within specific organizations. Recently, several models to detect and prevent insider threats have been presented, the majority of which have focused on technical issues. These models have been presented in recent times. On the other hand, very few of these models have addressed the social, cultural, and demographic factors that are at play.

The steps involved in the research methodology are illustrated in Figure 1. It encompasses the entirety of the insider threat that is present on the surface. The ''insider'', the ''threat'', and the ''insider threat detection'' are the subcomponents that make up the ''insider threat''. What

Figure 1 entails is a thorough process of data identification required for experimental analysis employing a model associated with detecting unintentional insider threat. The data adopted for this research is primarily a ''user activity logs'' dataset from CERT (release version r4.2).

In order to carry out evaluations, the dataset keeps a record of the user activity logs for a long period of time, and these logs come from a variety of different companies. This study operationally extracted and defined attributes based on ''Insider'', ''Threat'', and ''Insider threat'' in order to appropriately evaluate the number of malicious users associated to a specific set of attribute. This was necessary in order to account for the fact that insider threat events are typically uncommon in the real world. The ''Insider'' and ''Threat'' subparts make up the initial component linked to the dataset. As for the last section, it focuses on the detection of one-dimensional insider threats using machine learning. The ''Insider'' role requires (Access, Motivation, and Action), where Intention, +Action, Method, Trust, and Insider Knowledge are all components of the threat equation. At Last, an Insider. The term ''Detection'' refers to the process of identifying potential threats, whereas ''Behaviors'' and ''Performance Analysis'' are related concepts.

A. INSIDER LEVEL

At the insider level, the purpose of this study is to define the data that is associated with having ''access'' as a construct of an insider threat/attack. Three elements, ''Access'', ''Motivation'', and ''Action'', are outlined at the ''insider level''. When we reach this point, everyone who works for a company has access to the company's internal network and other forms of information technology up to a maximum access level that has been set by the organisation. That is to say, everyone has the right to use the available hardware, software, and networks for processing data and facilitating communication within the company. They become inspired because of the access, and everyone responds to a situation in a way that is consistent with their level of motivation.

B. ACCESS LEVEL OF INSIDER CONSTRUCT

At the access level, the access principles that this research outlines apply to all of an organization's ''Hardware'', ''Software'', and ''Data'', as well as its ''Network''. Organisation in which, in the event that an insider acquired access to them, he or she may be a potential to inflict harm, either purposefully or unintentionally threat. As a consequence of this, when determining whether or not an insider threat and attack was intentional, it is necessary to cover the relevant facts connected with the ability to access. The reason for this is that the study's initial stage is to isolate it. The datasets from CERT that were utilised for this research are connected to component of access that makes insider threats and attacks possible. Finaly this data that is generated under the ''access'' is linked with an arrow to motivation conceptually, it does not mean that

C. MOTIVATION LEVEL OF INSIDER CONSTRUCT

It is very likely that an attacker may have a motive to set a trap for an insider while the insider was ignorant that the trap was being set. In a similar vein, It is also quite feasible that an insider could have reason to violate a trust that was given to him. This research focuses on ''motivation'' in each of these scenarios, as a significant step that follows when an insider obtains access to a company's information and technological resources, this is also known as ''insider trading''. Because of this, therefore when doing research to accurately estimate and evaluate accidental insider threats and attacks, data associated to motive are extremely important. The CERT dataset that was used in this research had the motivational dimensions that were employed.

D. ACTION LEVEL OF INSIDER CONSTRUCT

Within the first phase of the methodological flow, the final step is referred to as ''action''. Taking into consideration the fact that access must first be granted to an insider, then the motivation must be formed to cause a negative consequence, and ultimately the attack must be carried out. The activity that is being referred to here is the degree to which all of the possibilities that lay within compromising the information technology resources of a company. This is the reason why it is distinct from the ''activity'' that is presented in the following stage, which involves the actual occurrences that lead up to the attack. According to the findings of this study, the information concerning the ''activities'' involved in carrying out an insider assault, and more precisely an unintentional insider attack, is essential from the very beginning. CERT dataset is included with it, which is the reason why this research chooses to make use of it.

E. THREATS LEVEL

The next step in the flow of the methodology for this study is to investigate the ''threats'' that are related with insiders. The critical issue at hand is having an understanding of the fundamental flaws in how information technology is vulnerable. Because of this, within the methodological flow of ''threats'', the ''intent'', ''added action'', on the perceived motivation, as well as the ''method'', ''level trust'' and ''Insider Knowledge'' are required to model an assessment of an organization for an organization's readiness of the unintentional insider attack.

F. INTENT AND (+ ACTION) LEVELS OF THREATS CONSTRUCT

This research defines ''intent'' at the sub level of ''Threats'' as a high-level indicator of an action that will be conducted to trigger a unitentional insider attack. This is the reason why the arrow from ''action'' on the previous stage is set to immediately connect to ''intent'' in the current state, and also connected to sub construct (+ action) ''added action'' is needed in order to comprehend the strategy that is being utilised. The difference between action and (+ action) is the

consequences, and the lesson learnt in how to avoid those consequences. To put it in another way, the earlier is geared toward carrying out the early tasks in order to ascertain what the end outcome will be. The influence of the result is then utilised in the performance of a subsequent action, which defines the later (+ action). However, depending on the degree of the (+ action), it will either show up as the action that is part of the ''insider'' or as a ''additional action''. The actions of an individual can, in the vast majority of cases, be utilised to deduce the meaning that they were going for. It is possible for the ''added action'' to appear either as an action within the ''insider'' or as a ''added action'', depending on the level of the ''+ action'', despite the fact that the ''added action'' is a decision zone in which both possible outcomes of the decision are favourable. In light of this, ''intent'' can be seen to be the most significant controlling variable of every threat that an organisation will come up against. This is due to the fact that it sheds light on the reasons behind the threat. The variables that were associated with Intent were taken from the CERT dataset and included in this study as a direct result of this. As a result of testing the model with them, we are able to obtain a comprehensive picture of how we are able to comprehend the unintended insider attack.

G. METHOD LEVEL OF THREATS CONSTRUCT

This research operationally defines a ''method'' as how an activity will be conducted to cause a unitentional insider attack at the ''method'' level. This is the reason why the arrow from ''+ action'' on the stage before this one is set to directly connect to ''method'' in the current state. Method is also connected to ''level of trust,'' which is required in order to understand how the method will be applied, so this is the reason why it is directly connected. The degree to which an individual can be trusted can, in most circumstances, be used to infer the course of action that may be adopted. In light of this, ''degree of trust'' can be understood to be the most essential controlling variable of any threat that an organisation would confront. This suggests that the ''degree of trust'' that an insider person was granted by the organisation is related to the action that the insider person conducts that puts the organisation in danger. However, this does not rule out the possibility that there is a direct connection between the two. As a direct result of this, the data linked with Intent were extracted from the CERT dataset and included in this study. Through testing the model with them, we can have an overall picture of how we can understand the accidental insider attack.

H. INSIDER KNOWLEDGE LEVEL OF THREATS CONSTRUCT

This study requires the data related with the insider day-to-day existence in an organisation associated to the entirety of the secreat that the organisation maintains in order to reach the level of ''insider knowledge''. Insider knowledge is a form of information that is regarded to be a source of information that is contradictory to the conventional or official instructions, and it is given in the form of recommendations from individuals with insider knowledge in a range of different organisations. Because of this, being able to perceive the risk will be dependent on one's acquaintance with the procedures that are involved in making use of information and communication technology. If a person was highly knowledgeable and enjoyed a high level of trust within an organisation, there is a significant possibility that they could have been complicit in an insider attack that went unnoticed, and the same is true for the other way around. This is because an insider attack is more likely to go undetected if the person committing the attack is trusted and has a high level of knowledge about the organisation. The collection of relevant data is therefore of the utmost importance.

I. INSIDER THREATS DETECTION

The last phase of the study is the experimental component, and it involves the assessment of threats from inadvertent insiders. This assessment will define an organization's preparation for the next part of the study. This section discusses ''detection'', which is accomplished by examining ''behaviours'' and determining the ''performance'' of the model that was proposed in the work.

J. DETECTION

At a detection level. For the purpose of detection analysis, the dataset obtained from the first two steps, which includes the dimensions ''Insider'' and ''Threats'', will be utilised. With the help of NARX, an experimental machine learning approach was modelled with the purpose of identifying unexpected insider threat. A time-series model that allows for the investigation and experimentation of the detection of inadvertent insider threats.

K. BEHAVIOURAL EXTRACTIONS

At the level of behavioural analysis, numerous different approaches to splitting the dataset are utilised. The research divided the dataset into several ratios of training, testing, and validation, and the detection analyses derived from the data were set out and tested on these multiple partitionings. In order to accomplish this, the data must be segmented into smaller pieces, each of which must be constructed independently from the others. This can boost the availability of detection by removing the potential for a single point of failure and allowing for the autonomous management of smaller parts of the dataset. Additionally, this can allow for an increase in the overall size of the dataset. This is accomplished by organising the dataset in such a way that individual elements of it can be processed in isolation from one another.

L. PERFORMANCE MEASUREMENT

At the level of performance, the NARX Machine learning algorithms are able to be contrasted and evaluated using a wide number of different performance metrics. When evaluating the effectiveness of a machine learning algorithm that has been applied to a detection issue, there is typically a

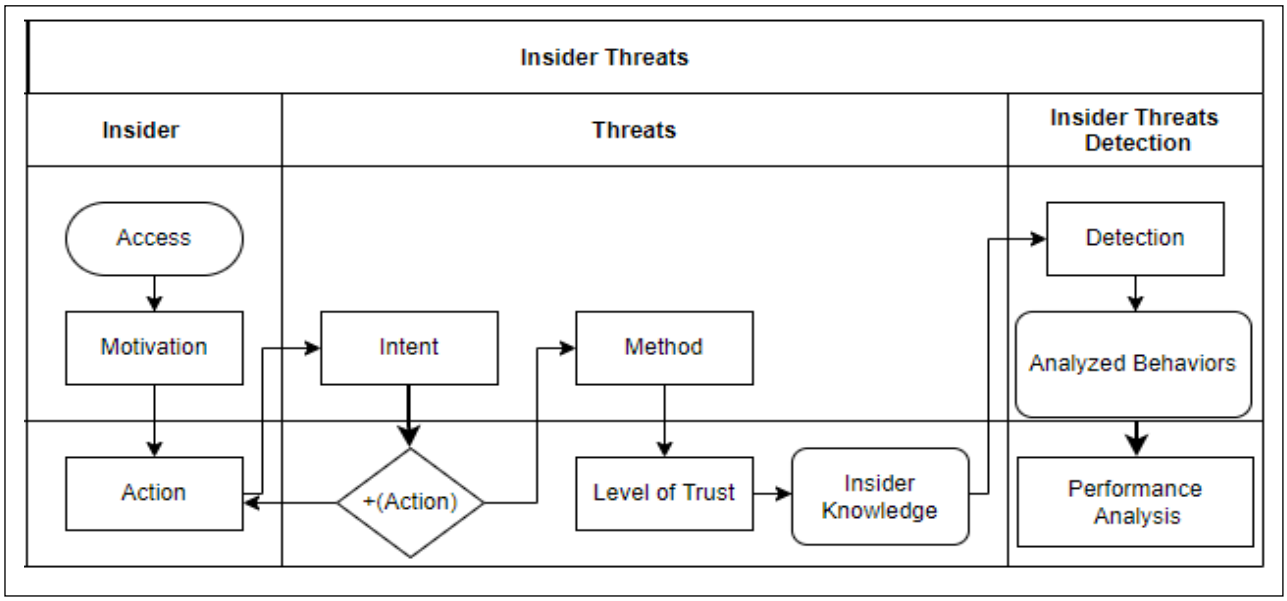

FIGURE 1. The research methodological process.

significant amount of data to process. The purpose of this research is to conduct an analysis of the performance of the model in order to gain an understanding of how various issues have impacted the NARX network model. Examining the outcomes of the performance analysis is the first step that needs to be taken before we can proceed with the calculation of the model error. The dependability of the findings provides support for the utility of this modelling strategy. On top of that, the detection performance of the NARX network model will be optimised in regions where it has bad performance in places that have poor performance.

M. CONCEPTUALIZATION

An experimental machine learning approach to detecting unintended insider risks is used in conjunction with a conceptualization of the insider threat to create the access to an information system is the first key criteria for Insider ''Insider threat'' as it occurs with an Intent or unintentional insider. Thus a neural network intended for the prediction and detection of unintentional insider threat made up of a large number of simple processing units that are highly interconnected with one another are utilized [39]. Because the data involve is a time series, a nonlinear autoregressive network with exogenous inputs (NARX) was applied, which is a recurrent dynamic kind of ANNs networks [40]. The processing units collaborate in order to generate massive amounts of parallel processing through the use of a computer architecture known as distributed computing. Through the use of the architecture and operation of neural networks, it is feasible to duplicate some of the functions that are performed by biological brains and neural systems [41]. The ability of neural networks to self-organize, learn in a way that is adaptive to their environment, and tolerate errors is among the most significant of their many benefits. Pattern recognition applications are beginning to make greater use of neural networks as a result of the excellent capabilities offered by these systems [42]. A variety of neural models, including as backpropagation networks, high-order neural networks, time delays, and recurrent neural networks, can be quite helpful [43].

Feed-forward networks are utilised in the vast majority of pattern recognition applications. In this context, the phenomenon where there is no response to the input being provided is referred to as ''feed-forward'' [39]. By providing feedback to the patterns that are being input, neural networks are able to learn from their mistakes in a manner that is analogous to how individuals gain knowledge from their errors. Reconstructing input patterns and ensuring they are free of errors using this form of feedback is one way that performance of neural networks could be improved. Building neural networks of this complexity requires a significant amount of time and labour. The term ''autoassociative neural networks'' has been given to refer to these types of networks [40]. They implement back-propagation strategies, which is rather self-explanatory given their name. Back-propagation algorithms frequently face the challenge of dealing with the appearance of local minima. In addition to these difficulties, neural networks also face issues regarding the rate at which they learn new information as well as the design that they employ [41], [42], [43]. In addition to that, they have issues with modularity and size. The potential advantages of neural networks are extremely extensive, despite the fact that they are beset by their own unique set of challenges and issues.

N. DATASET PRESENTATION

This research uses the CERT dataset to investigate the dangers posed by insiders [44]. The CERT insider threat datasets

are open to the public and can be used for the research, development, and testing of insider threat mitigation strategies. The dataset known as "CERT R4.2" mimics a firm with a total workforce of 1000 people, of whom there are 70 employees who are actively working to undermine the organization's security. as part of a study being conducted at Carnegie Mellon University (CMU). It was compiled from the experiences of 4000 users who, over the course of 18 months, were involved in more than 700 instances of an insider threat. The dataset contains a variety of user activities such as logging in and out, performing le operations, emailing, browsing the web, and using USB drives, among other things. It was produced by simulating the following three possible forms of attack: The first incident was caused by malicious insiders who used removable devices to inject malware into the system, which led to the sabotage of the system. The second type of data exfiltration was carried out by dishonest employees using the cloud or removable media. The final scenario involved an attack on data integrity that was carried out by masqueraders through the use of email attachments. This enables us to undertake experiments with a greater degree of flexibility in order to detect an inadvertent insider danger and provide a better knowledge of the behaviours exhibited by insiders.

O. DEVELOPMENT OF NARX MODEL

Recurrent and dynamic artificial neural networks (ANNs) are what make up NARX neural networks [45]. They got their names from the basic nervous systems that animals have, which served as the inspiration for their overall structure. When this is taken into consideration, it is possible that one or more synthetic neurons will be regarded as being connected to each node. Each node in a network contains a single artificial neuron that is responsible for processing one or more input signals, computing an output based on the sum of the input signals, and then passing on that output to the following node in the network. On the other hand, the activation function of the NARX network is nonlinear along its whole length [46].

This is true throughout its entirety. The information flow that ANNs are able to process can be used to differentiate one type of ANN from another. On the other hand, a NARX makes it possible to develop connections between neurons that are located in the same or previous layers. This is due to the fact that information goes in both directions within a NARX, whereas within a ANN it only travels in one direction. NARX is superior to other types of recurrent neural networks in terms of its speed of convergence as well as its lower minimum threshold for the number of neurons that need to be calibrated [47]. As a consequence of this, it is superior to other recurrent neural networks in terms of its ability to identify long-term dependencies. Because of the exogenous inputs that the NARX network receives, external time series can have an impact on the time series that are of interest. The NARX model equation is:

$$y(t+1)\left(y(t), y(t-1), \ldots, y\left(t-n_y\right), z(t+1), z(t), \times z(t-1), \ldots, z\left(t-n_y\right)\right) \quad (1)$$

where the NARX neural network model is built via a multilayer perceptron that possesses both a time delay as well as output feedback. This structure is used to create the model [48]. The output of the multilayer perceptron can be generated in this manner if one so chooses.

$$\hat{y}(t) = f\left[\left(y(t), y(t-1), \ldots, y\left(t-n_y\right), z(t+1), z(t), \times z(t-1), \ldots, z\left(t-n_y\right) + e(t)\right)\right] \quad (2)$$

where $y(t)$ and $\hat{y}(t)$ are the outputs that should be achieved, as well as those that are expected $z(t)$ is the variable that is being fed into the network.

n_t and n_y represent the lags that occur between the input and the output variables $e(t)$. The input layer of the neural network is where the parameters for the neural network's input are stored, while the hidden layer is where the processes that take place during the transition between the input and output layers are situated as

$$W_i(t) = f_1\left[\sum_{r=0}^{n_z} w_{ir} u(t-r) + \sum_{i=0}^{n_y} w_{ir} y(t-l) + \Lambda_i\right] \quad (3)$$

For each model, a unique collection of extra input values was taken into consideration [49]. The values between the input neuron and the output neuron were incorporated as inputs to the first model and the $z(t-r)$ hidden neuron connection weight, where w_{ir} is the connection and the bias of the $y(t-l)$ hidden neuron, and f $1(\cdot)$ is the activation function for the hidden layer and combining the hidden layer output, the prediction is:

$$\hat{W}_i(t) = f_1\left[\sum_{t=1}^{n_z} w_{ji} z(t) + \Lambda_j\right] \quad (4)$$

In the hidden layer, the activation function was a sigmoid, whereas the output layer had a linear activation function and only one neuron [49]. The hidden layer was constructed using this activation function. The training process optimised the weight and bias of the NARX model for the output layer, where the model weight represents the relationship between the hidden neuron and the anticipated output and w_{ji} is the predictability of the output dependent on the nature of the relationship that is tied to it [50].

Due to the fact that a variety of machine learning-related detection models already exist, including autoregressive integrated moving average (ARIMA), long short-term memory (LSTM), regression, artificial neural network (ANN), integrated farm system model (IFSM), and trend analysis, among others [45], [46]. The focus of the present study is on "Time series analysis". Data that can arise from user activity logs of an organisation, for example, can be used to identify a signature of an insider threat, and so permits forecasts utilising connected historical data with the fewest forecasting mistakes. In addition, Time series analysis is based on a sequence of data that are arranged in chronological order. Simply said, time series forecasting constructs a quantitative model based on past data to make predictions about the future.

Organizational data on user actions can be analysed as a time series for patterns of behaviour that may indicate

TABLE 1. Dataset partition used for the simulation.

Partition	Training		Testing		Validation		Hidden	Delay
1	80	3202000	10	400250	10	400250	10	2
2	80	3202000	15	600375	5	200125	10	2
3	75	3001875	15	600375	10	400250	10	2
4	75	3001875	10	400250	15	600375	10	2
5	70	2801750	20	800500	10	400250	10	2
6	70	2801750	10	400250	20	800500	10	2

the presence of threats. Unfortunately, ARIMA, a popular machine learning method based on Time series, operates by utilising lags of the differenced series and lags of the projected errors, both of which necessitate the data to be stationary, which is achieved through differencing [47]. ARIMA model is based on the principle of using data with precise values and without any measurement mistakes, which has the advantages of flexibility and better seasonal patterns, but it is less accurate than any other models in time series. Because of this shortcoming, we use NARX, a nonlinear generalisation of the Autoregressive Exogenous (ARX), a common tool that can function also in linear black-box system identification. Time-series modelling is just one example of the many uses for NARX models. They may also be used to represent large-scale nonlinear dynamic systems.

P. PERFORMANCE EVALUATION METRIC

The effectiveness of the NARX network was evaluated with the help of four different measures for assessment: the coefficient of determination R^2, which demonstrates how the amount of variation in one factor is connected to the amount of variation in another component that is associated in the model of the experimental data. The Mean Absolute Error, or MAE, is a statistic that illustrates the degree to which the actual values differ from those that were predicted [51]. The Root Mean Squared Error (RMSE), also known as the square root of the average squared error, is a quantitative method for evaluating how closely the values produced by a model match the anticipated values [52]. It is necessary to evaluate the accuracy of the prediction utilising the data that was used after the model was tested to see whether or not the data fit the model. Putting performance assessment measures into action, taking these kinds of measures is what's known as undertaking performance evaluation [53].

IV. PRESENTATION OF RESULTS AND DISCUSSION

It is necessary to finish the task of gathering all of the relevant details before beginning the analysis. This must be done before the analysis can begin. There were a few distinct experimental setups that were built, and all of them were put to use. As more and more evidence is collected, it has been demonstrated that the model has an accuracy of 99.12 percent when it comes to making accurate predictions. The precision and the recall are both at a level of 99.05 percent, while the false acceptance rate sits at 21.13 percent and the false rejection rate sits at 19.42 percent respectively.

There were a number of different simulations, and in each one, a two-stage approach to the NARX modelling was used. In each partition, the NARX network-based model was trained and evaluated for accuracy (seeTable1). There were a total of six partitions used, and the entries were divided into ''Training'', ''Testing'', and ''Validation''. The total number of entries was 4002500.

The major goal of finding an ideal level for the model in question in terms of how it evaluated the dataset that was placed into an experiment is the purpose of finding an optimal level for partitioning datasets [54]. It is absolutely necessary to divide a data set into a training set and a test set before analysing it. In specifically, the training set is put to use in the process of learning a model, and the test set is put to use thereafter in evaluating how well the model learnt from the training set actually performed. The division of the data into the two sets, on the other hand, and the effect it has on the performance of the model are both determined by the optimal proportion that should be used for each set. As a result, the standard procedure involves randomly dividing the data into around 70 percent for training and 30 percent for testing purposes. If not, 70 percent will go into training, while the remaining 30 percent would be used to evaluate and assess performance [55]. According to Vrigazova [56] utilising an alternative structure for the test set (such as 30/70 or 20/80, for example) may be able to further optimise the performance of the machine learning technique. The NARX-based model's training and testing phases were carried out on the datasets that were contained within the first partitions of each of the first partitions. The data were analyzed with the help of the model. It demonstrates outstanding performance across the board for every single model and partition in the entire lineup.

The second model produced the best results ($R2 = 0.9517$, $MAE = 0.9609$, $RMSE = 1.1811$, $RAE = 0.0414$); (see Table 2) despite the fact that the model used only 5% of the validation parameters in its input variables, the prediction was still extremely accurate, with only minor differences when compared to the other models. The R2 value for the second model was 0.9517, while the MAE value was 0.9609, the RMSE value was 1.1811,

Due to the fact that, in order to validate the result of the first setoff, the experiment essentially has the

TABLE 2. The results of the first round of analysis.

Model	R^2	MAE	RMSE	RAE
1	0.8102	0.9327	1.1232	0.0391
2	0.9517	0.9609	1.1811	0.0414
3	0.8739	0.9019	1.2762	0.0336
4	0.9137	0.9308	1.0632	0.0314
5	0.9217	0.9008	1.2122	0.0709
6	0.9125	0.9347	1.1301	0.0834

characteristics of the insider thread, which is basically considering the unintentional insider threat, the time series of the events happening matters the most, the more noticeable decrease in performance (see Table 3) is observed passing from the experimental scenario on the dataset. The experiment has the characteristics of an insider thread because its purpose is to verify the outcome of the initial setoff. Because of this autocorrelation, the predictive ability of the model has a tendency to exhibit some variations, despite the fact that the best evaluation metrics indicate that the results are still accurate at model 5 (R2 = 0.9177, MAE = 0.8748, RMSE = 1.2071, RAE = 0.617). Despite these variations, the results are still accurate.

TABLE 3. The adjusted result of the analysis.

Model	R^2	MAE	RMSE	RAE
1	0.8154	0.9067	1.1181	0.0299
2	0.9069	0.9049	1.176	0.0322
3	0.8791	0.8759	1.2711	0.0244
4	0.9189	0.9048	1.0581	0.0222
5	0.9269	0.8748	1.2071	0.0617
6	0.9177	0.9087	1.125	0.0742

During the training and testing stages of the NARX-based model, the datasets that were contained within the first partitions of each of the first partitions were utilised in order to improve the adjusted result of the analysis that was performed on the data. This was done by using the datasets that were contained inside of each of the partitions. Utilization of the model was necessary in order to decipher the information that was contained within the material. The complete line of products, including each individual model and division, is capable of achieving exceptional levels of success in terms of performance. There were very minimal discrepancies observed between the models' predictions (R^2 = 0.9321, MAE = 0.9488; RMSE = 1.202; RAE = 0.0525). This was the case despite the fact that an earlier model likewise worked well in its input variables (see Table 4).

As a result of the fact that the experiment has the characteristics of an insider thread, which is essentially the same as taking into consideration the unintentional insider threat, the time series of the events that are taking place matters the most, and the most noticeable decrease in performance (see Table 5) is seen when moving from the experimental scenario to the dataset. As a consequence of the fact that, in order to validate

TABLE 4. The result of the analysis on improving the values.

Model	R^2	MAE	RMSE	RAE
1	0.8206	0.8807	1.113	0.0207
2	0.9021	0.8089	1.1709	0.023
3	0.8843	0.8499	1.266	0.0152
4	0.9241	0.8788	1.053	0.013
5	0.9321	0.9488	1.202	0.0525
6	0.9229	0.8827	1.1199	0.065

the outcomes of the initial setoff, the experiment, in essence, possesses the characteristics of an insider thread. Because of this, an insider thread is required in order to validate the outcomes of the preliminary stages of an experiment. This autocorrelation has a tendency to change the prediction ability of the model, despite the fact that its best assessment metrics indicate that the findings are still accurate at model 2 (R2 = 0.97122, MAE 0.9586, RMSE = 1.1683, RAE = 0.0236). This is because the model has a propensity for exhibiting specific differences, which explains why this is the case. The findings, notwithstanding the contradictions, should still be considered reliable.

TABLE 5. The final adjusted model results.

Model	R2	MAE	RMSE	RAE
1	0.8918	0.9804	1.1104	0.0213
2	0.97122	0.9586	1.1683	0.0236
3	0.89342	0.9496	1.2634	0.0158
4	0.93322	0.9785	1.0504	0.0136
5	0.94122	0.9485	1.1994	0.0531
6	0.93202	0.9824	1.1173	0.0656

Table 6 shows that when going from the experimental scenario to the dataset, performance suffers the greatest. This is because the experiment has the characteristics of an insider thread, which is basically the same as taking into consideration an unintended insider threat. Taking into account an accidental insider threat is equivalent to considering the experiment to have insider thread characteristics. There's a good reason for this: The experiment shares many characteristics with an insider post. Since it was essential to double-check the preliminary results, this experiment was obviously carried out by someone with special expertise. As a result of this, it is necessary to verify the results of the preliminary competition. Therefore, an insider's perspective is required to demonstrate the accuracy of early experiment data. Although the model's best evaluation metrics (R2 = 0.94234, MAE = 0.97529, RMSE = 1.0478, RAE = 0.0142) has a propensity to affect the model's capacity to forecast. While the model's ability to forecast may be affected by this autocorrelation, this is nevertheless the case. As a result of this, the situation described above has occurred since the model has a tendency to emphasise certain distinctions. Although there are some contradictions, the findings should nevertheless be taken seriously.

TABLE 6. The overall performance of the results.

Model	R2	MAE	RMSE	RAE
1	0.963	0.97719	1.1078	0.0219
2	0.90034	0.95539	1.1657	0.0242
3	0.90254	0.94639	1.2608	0.0164
4	0.94234	0.97529	1.0478	0.0142
5	0.90034	0.94529	1.1968	0.0537
6	0.94114	0.97919	1.1147	0.0662

V. DISCUSSION OF THE FINDINGS

In this study, evaluations are carried out using a dataset that maintains a record of the user activity logs for an extended period of time. These logs come from a variety of different companies, and within each of these logs, a sub-construct that is associated with ''Insider Level'' and ''Threats Level'' is defined. This research conceptualised that there are sub constructs within these two components that include ''Access'', ''Motivation'', and ''Action'' within the insider behaviour, and that there are dangers that arise from insider activities that include ''Intent'', ''+Action'', ''Method'', and ''knowledge''. Therefore, after doing an analysis, it was determined that the NARX-based model's training and testing reveal a high detection rate (R2 0.97122) at a variety of partitioning, with the optimal partitioning being discovered to be at (80:15:5). This indicates that the dataset that keeps a record of the user activity logs for a long period of time should be partitioned at the ratio to which this research reveals if it is to be used for the analysis of detecting an unintentional insider attack and if it is to be used for a long period of time.

The significance of the findings of this study to the field of theory resides in the fact that it was demonstrated that the procedure of identifying a breach of trust in an individual working for an organisation can result in the extraction of a number of sub constructs. ''Access'', ''Motivation'', and ''Action'' have been able to be extracted from this present research as being associated with ''Insider'', whereas ''Intent'', ''+Action'', ''Method'', and ''knowledge'' have been able to be extracted as being associated with ''Threats''. As a result, these are linked to the behaviours of users, which can be collected and gathered over an extended period of time.

Another influence of this study related to theory resides in the fact that carrying out attacks due to the fact that an insider led to the question of who within an organisation is unable to be a victim. This is a direct result of the fact that this study was conducted. The answer to the question lies in the fact that the characteristics of unintentional insider threat datasets cannot really isolate or tell more on who in an organisation can or cannot be involved in an insider attack; however, it can have associated timestamps, which is the tendency to exhibit patterns that can either be linear or nonlinear, or even both of these at the same time; this is the answer to the question in this research regard. As a result, it is possible to trace this to the individuals involved. Because of this, the researchers in this study elected to use the conventional autoregressive model. Exogenous prediction has been given a nonlinear twist thanks to the deployment of the NARX Neural Network prediction. After the linear model had been applied to a portion of the dataset that contained an insider threat, this technique was put into use. Because of this, the high performance detection measure that was acquired during this research has demonstrated that the model is able to predict the problems.

The findings of the study regarding the effectiveness of the methods are derived using scaled conjugate gradients. That is to say, the NARX model that was utilised is superior at training and testing the flow of the time series regularisation processes after several iterations and changes to the total number of neurons in the hidden layer. The performance metrics provide proof that training, testing, and validation are all improving, and they do so by supplying this evidence. This evidence shows that training, testing, and validation are all growing better. The result of this study from various analysis indicate that ''Time-series modelling'' technique, utilising the ''NARX'' computer programme, was the following step to take. The predictive strategy that was utilised in this investigation attempts to locate the most suitable neural network architecture for the total dataset. In addition, the parameters of the NARX model are confirmed using a number of different simulations. The model that was used not only produces the best results, but it is also plainly evident time series analysis used is effective in its own right. This is because the model R2 values from the findings are all above the threshold that was used produces the best outcomes. It would appear, on the basis of the findings of this research, that predictive or detective capability for the detection of insider threats is crucial based functionality. Particularly useful in in terms of setting out how datasets can be associated to the unintentional threat.

Another effect of this study impact lies with association of the neural network with NARX. It is crucial to understand that while Neural Network makes training the weights of the network significantly less difficult than it would have been otherwise, NARX is also associated to that approach for which it has been revealed from the result of this study. The insider threats estimation was set out to indicate exactly when the incident took place. Each layer of the NARX neural network comprises a different number of neurons. A vast number of neurons are used to construct a neural network that has been evolved to its maximum potential degree. This study has helped with the architectural modelling, which has led to the detection of the accidental insider danger. In the Neural network, the operation would be able to provide an estimate of the unintentional threat posed by an insider towards error estimation due to hidden layer evaluation associated with the neurons. However, there is a different amount of neuron density in each layer of the NARX neural network than in the previous layer used in this research. Neural networks with eight input neurons, ten hidden neurons, and two delays can

achieve the highest possible network effect. This is because these networks have a total of twenty-two neurons. Thus, the finding of this study strongly supported the mean square error as well as the detective metric (R2). The training cycle within some partitions, causes the partitioning to be modified as necessary. The model was initially set up with randomly generated weights and biases. Each training cycle saw a new set of weights and biases generated for the vectors. Following this, the model was utilised to choose relevant detective traits. Hence, in terms of NARX ability to model unintentional insider be appropriate.

Comparisons between this study's findings and those of other studies are quite important. Specifically, our research allows the proposed NARX model adopted to keeps the datasets within their patterns. Due to the focus on detection in this study's performance analysis, the key differences between this investigation and Mazzarolo et al.'s [23] work lie in the two groups' respective methods of data analysis, data selection, and overall methodology. However, in contrast to Mazzarolo et al. [23], the model used in this study achieved a good detection performance. When contrasting the results of this study with those of Magklaras and Furnell [24], it is worth keeping in mind that the former study's authors offered a method that was thought to quantify the level of threat that is likely to originate from a specific insider.

In contrast to Magklaras and Furnell [24], this study had a detection analysis success rate of almost 97%. Colwi [25] assesses internal risks rather of evaluating the issue through testing or model building, contradicting the results of the current investigation. In contrast to the present state of affairs, Nikolai and Wang [27] employ k-nearest neighbours to perform anomaly detection and report a perfect detection rate without disclosing their datasets. This study improved detection in terms of R2 in the domains of classification and time-series analysis, where the largest changes may be found. Whereas Punithavathani et al. [28], who utilise KNN with 87% accuracy, just established dimensions in a model and then evaluated it, the current investigation actually established dimensions in a model and then evaluated it to reach a high estimate of components related with insider threats. Hu et al. [29] established a statistical model for detecting insider threats, but the current study builds on machine learning to a greater degree of detection. The analysis is where this work departs most significantly from Feng et al. [30]. While Feng et al. [30] measure the temporal patterns of the user behaviours of insider threat, they do not provide any quantitative value linked with the outcome, but the current work was able to accomplish a good detection.

VI. CONCLUSION

This study employed a detection model, developed with the aid of the NARX algorithm, to generate predictions about the future trajectory of the insider threat caused by accident. This research aims to better understand the causes of accidental insider threat transactions. To model and forecast outcomes, this study employs machine learning. The research concluded that inadvertent insider risks are more likely to be uncovered. To maximise the accuracy of the model that forecasts the risk caused by unintended insiders, we conducted a few trials and performed some experimental analysis. The finding of the analysis indicates that 0.97 detection rate of NARX to detect unintentional insider threat. The implication of the research dwells on having a positive outlook regarding accidental insider threats by organization. The need of accurately identifying an accidental insider threat is widely recognized throughout the academic community. As so, this demonstrates how a company can ensure its security against an internal threat that was not intended. The academic world has recognised its significance and acts accordingly. because it equips you with the resources you need to ascertain whether or not an employee has purposefully harmed the business. This research has the potential to affect both the level of awareness of accidental insider threats and the accompanying level of ambiguity. Threats from unwitting insiders posed several problems, all of which were linked to trust inside the organisation. In summary, the study provided comprehensive and empirical direction to all issues about the investigation of inadvertent insider risks.

REFERENCES

[1] A. Kim, J. Oh, J. Ryu, and K. Lee, "A review of insider threat detection approaches with IoT perspective," *IEEE Access*, vol. 8, pp. 78847–78867, 2020.
[2] R. A. Alsowail and T. Al-Shehari, "Empirical detection techniques of insider threat incidents," *IEEE Access*, vol. 8, pp. 78385–78402, 2020.
[3] M. N. Al-Mhiqani, R. Ahmad, Z. Z. Abidin, K. H. Abdulkareem, M. A. Mohammed, D. Gupta, and K. Shankar, "A new intelligent multi-layer framework for insider threat detection," *Comput. Electr. Eng.*, vol. 97, Jan. 2022, Art. no. 107597.
[4] M. N. Al-Mhiqani, R. Ahmed, Z. Zainal, and S. N. Isnin, "An integrated imbalanced learning and deep neural network model for insider threat detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 1, pp. 1–6, 2021.
[5] A. Wall and I. Agrafiotis, "A Bayesian approach to insider threat detection," *J. Wireless Mobile Netw. Ubiquitous Comput., Dependable Appl.*, vol. 12, no. 2, pp. 48–84, 2021.
[6] A. Erola, I. Agrafiotis, M. Goldsmith, and S. Creese, "Insider-threat detection: Lessons from deploying the CITD tool in three multinational organisations," *J. Inf. Secur. Appl.*, vol. 67, Jun. 2022, Art. no. 103167.
[7] A. D. Williams, S. N. Abbott, N. Shoman, and W. S. Charlton, "Results from invoking artificial neural networks to measure insider threat detection & mitigation," *Digit. Threats, Res. Pract.*, vol. 3, no. 1, 2021, pp. 1–20.
[8] 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.
[9] S. Yuan and X. Wu, "Deep learning for insider threat detection: Review, challenges and opportunities," *Comput. Secur.*, vol. 104, May 2021, Art. no. 102221.
[10] P. Chapman, "Defending against insider threats with network security's eighth layer," *Comput. Fraud Secur.*, vol. 2021, no. 3, pp. 8–13, Jan. 2021.
[11] S. Prabhu and N. Thompson, "A primer on insider threats in cybersecurity," *Inf. Secur. J., Global Perspective*, vol. 31, no. 5, pp. 602–611, 2021.
[12] M. P. Singh, S. Sural, J. Vaidya, and V. Atluri, "A role-based administrative model for administration of heterogeneous access control policies and its security analysis," *Inf. Syst. Frontiers*, pp. 1–18, Jul. 2021, doi: 10.1007/s10796-021-10167-z.
[13] J. Jang-Jaccard and S. Nepal, "A survey of emerging threats in cybersecurity," *J. Comput. Syst. Sci.*, vol. 80, no. 5, pp. 973–993, 2014.
[14] T. T. Y. Alabdullah, E. R. Ahmed, M. Almashhadani, S. K. Yousif, H. A. Almashhadani, R. Almashhadani, and E. Putri, "How significantly to emerging economies benefit from board attributes and risk management in enhancing firm profitability?" *J. Accounting Sci.*, vol. 5, no. 2, pp. 104–113, Jul. 2021.

[15] R. Dastres and M. Soori, "A review in recent development of network threats and security measures," *Int. J. Inf. Sci. Comput. Eng.*, pp. 162–173, Feb. 2021.

[16] 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, Mar. 2021, Art. no. e06522.

[17] W. Steingartner and D. Galinec, "Cyber threats and cyber deception in hybrid warfare," *Acta Polytech. Hungarica*, vol. 18, no. 3, pp. 25–45, 2021.

[18] B. Acciaio, A. M. G. Cox, and M. Huesmann, "Model-independent pricing with insider information: A Skorokhod embedding approach," *Adv. Appl. Probab.*, vol. 53, no. 1, pp. 30–56, Mar. 2021.

[19] X. Zhang, J. Lu, and D. Li, "Confidential information protection method of commercial information physical system based on edge computing," *Neural Comput. Appl.*, vol. 33, no. 3, pp. 897–907, 2021.

[20] I. Gaidarski and Z. Minchev, "Insider threats to IT security of critical infrastructures," in *Digital Transformation, Cyber Security and Resilience of Modern Societies*. Cham, Switzerland: Springer, 2021, pp. 381–394.

[21] Q. A. Al-Fatlawi, D. S. Al Farttoosi, and A. H. Almagtome, "Accounting information security and IT governance under COBIT 5 framework: A case study," *Webology*, vol. 18, pp. 294–310, Apr. 2021.

[22] J. Liang and Y. Kim, "Evolution of firewalls: Toward securer network using next generation firewall," in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2022, pp. 752–759.

[23] G. Mazzarolo, J. C. F. Casas, A. D. Jurcut, and N. A. Le-Khac, "Protect against unintentional insider threats: The risk of an employee's cyber misconduct on a social media site," in *Cybercrime in Context*. Cham, Switzerland: Springer, 2021, pp. 79–101.

[24] G. B. Magklaras and S. M. Furnell, "Insider threat prediction tool: Evaluating the probability of IT misuse," *Comput. Secur.*, vol. 21, no. 1, pp. 62–73, Jan. 2001.

[25] C. Colwill, "Human factors in information security: The insider threat—Who can you trust these days?" *Inf. Secur. Tech. Rep.*, vol. 14, no. 4, pp. 186–196, Nov. 2009.

[26] E. A. Lavrov, A. L. Zolkin, T. G. Aygumov, M. S. Chistyakov, and I. V. Akhmetov, "Analysis of information security issues in corporate computer networks," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 1047, no. 1, Feb. 2021, Art. no. 012117.

[27] J. Nikolai and Y. Wang, "A system for detecting malicious insider data theft in IaaS cloud environments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[28] D. S. Punithavathani, K. Sujatha, and J. M. Jain, "Surveillance of anomaly and misuse in critical networks to counter insider threats using computational intelligence," *Cluster Comput.*, vol. 18, no. 1, pp. 435–451, Mar. 2015.

[29] Y. Hu, C. Frank, J. Walden, E. Crawford, and D. Kasturiratna, "Profiling file repository access patterns for identifying data exfiltration activities," in *Proc. IEEE Symp. Comput. Intell. Cyber Secur. (CICS)*, Paris, France, Apr. 2011, pp. 122–128.

[30] W. Feng, W. Yan, S. Wu, and N. Liu, "Wavelet transform and unsupervised machine learning to detect insider threat on cloud file-sharing," in *Proc. IEEE Int. Conf. Intell. Secur. Informat. (ISI)*, Jul. 2017, pp. 155–157.

[31] M. Mayhew, M. Atighetchi, A. Adler, and R. Greenstadt, "Use of machine learning in big data analytics for insider threat detection," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2015, pp. 915–922.

[32] S. Aditham, N. Ranganathan, and S. Katkoori, "Memory access pattern based insider threat detection in big data systems," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2016, pp. 3625–3628.

[33] Z. Li and K. Liu, "An event based detection of internal threat to information system," in *Proc. Int. Conf. Harmony Search Algorithm*. Cham, Switzerland: Springer, 2019, pp. 44–53.

[34] S. Rajamanickam, S. Vollala, R. Amin, and N. Ramasubramanian, "Insider attack protection: Lightweight password-based authentication techniques using ECC," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1972–1983, Jun. 2020.

[35] D. Patil and B. Meshram, "Network packet analysis for detecting malicious insider," in *Proc. 3rd Int. Conf. Converg. Technol. (I2CT)*, Apr. 2018, pp. 1–8.

[36] A. Shaghaghi, S. S. Kanhere, M. A. Kaafar, E. Bertino, and S. Jha, "Gargoyle: A network-based insider attack resilient framework for organizations," in *Proc. IEEE 43rd Conf. Local Comput. Netw. (LCN)*, Oct. 2018, pp. 553–561.

[37] S. S. Tan, J. C. Na, and S. Duraisamy, "Unified psycholinguistic framework: An unobtrusive psychological analysis approach towards insider threat prevention and detection," *J. Inf. Sci. Theory Pract.*, vol. 7, no. 1, pp. 52–71, 2019.

[38] M. M. Yamin, B. Katt, K. Sattar, and M. B. Ahmad, "Implementation of insider threat detection system using honeypot based sensors and threat analytics," in *Future Inf. Commun. Conf.* Cham, Switzerland: Springer, 2019, pp. 801–829.

[39] S. Chakraborty, R. Krishna, Y. Ding, and B. Ray, "Deep learning based vulnerability detection: Are we there yet?" *IEEE Trans. Softw. Eng.*, vol. 48, no. 9, pp. 3280–3296, Sep. 2022.

[40] D. N. Fabio, S. I. Abba, B. Q. Pham, A. R. M. T. Islam, S. Talukdar, and G. Francesco, "Groundwater level forecasting in northern Bangladesh using nonlinear autoregressive exogenous (NARX) and extreme learning machine (ELM) neural networks," *Arabian J. Geosci.*, vol. 15, no. 7, pp. 1–20, Apr. 2022.

[41] L. V. Jospin, H. Laga, F. Boussaid, W. Buntine, and M. Bennamoun, "Hands-on Bayesian neural networks—A tutorial for deep learning users," *IEEE Comput. Intell. Mag.*, vol. 17, no. 2, pp. 29–48, May 2022.

[42] H. Patel and K. P. Upla, "A shallow network for hyperspectral image classification using an autoencoder with convolutional neural network," *Multimedia Tools Appl.*, vol. 81, no. 1, pp. 695–714, 2022.

[43] U. Thakkar and H. Chaoui, "Remaining useful life prediction of an aircraft turbofan engine using deep layer recurrent neural networks," *Actuators*, vol. 11, no. 3, p. 67, Feb. 2022.

[44] *Cert Dataset*. Accessed: Jan. 19, 2022. [Online]. Available: https://kilthub.cmu.edu/articles/dataset/Insider_Threat_Test_Dataset/12841247/1

[45] T. Lin, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in NARX recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 7, no. 6, pp. 1329–1338, Nov. 1996.

[46] J. Wang and Y. Chen, "Using NARX neural network to forecast droughts and floods over Yangtze river basin," *Natural Hazards*, vol. 110, no. 1, pp. 225–246, Jan. 2022.

[47] A. H. Dhafer, F. M. Nor, G. Alkawsi, A. Z. Al-Othmani, N. R. Shah, H. M. Alshanbari, K. F. B. Khairi, and Y. Baashar, "Empirical analysis for stock price prediction using NARX model with exogenous technical indicators," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–13, Mar. 2022.

[48] S. Amirkhani, A. Tootchi, and A. Chaibakhsh, "Fault detection and isolation of gas turbine using series–parallel NARX model," *ISA Trans.*, vol. 120, pp. 205–221, Jan. 2022.

[49] Z. Hussain and N. Zainul Azlan, "Estimation of the torques produced by human upper limb during eating activities using NARX-NN," *Appl. Artif. Intell.*, vol. 36, no. 1, pp. 1–20, Dec. 2022.

[50] Y. Yang, K.-R. Kim, R. Kou, Y. Li, J. Fu, L. Zhao, and H. Liu, "Prediction of effluent quality in a wastewater treatment plant by dynamic neural network modeling," *Process Saf. Environ. Protection*, vol. 158, pp. 515–524, Feb. 2022.

[51] M. Hadwan, B. M. Al-Maqaleh, F. N. Al-Badani, R. U. Khan, and M. A. Al-Hagery, "A hybrid neural network and Box–Jenkins models for time series forecasting," *Comput., Mater. Continua*, vol. 70, no. 3, pp. 4829–4845, 2022.

[52] S. Kouadri, C. B. Pande, B. Panneerselvam, K. N. Moharir, and A. Elbeltagi, "Prediction of irrigation groundwater quality parameters using ANN, LSTM, and MLR models," *Environ. Sci. Pollut. Res.*, vol. 29, no. 14, pp. 21067–21091, Mar. 2022.

[53] X. Liu, W. Liu, H. Huang, and L. Bo, "An improved confusion matrix for fusing multiple K-SVD classifiers," *Knowl. Inf. Syst.*, vol. 64, no. 3, pp. 703–722, 2022.

[54] K. Korjus, M. N. Hebart, and R. Vicente, "An efficient data partitioning to improve classification performance while keeping parameters interpretable," *PLoS ONE*, vol. 11, no. 8, Aug. 2016, Art. no. e0161788.

[55] H. Liu and M. Cocea, "Semi-random partitioning of data into training and test sets in granular computing context," *Granular Comput.*, vol. 2, no. 4, pp. 357–386, 2017.

[56] B. Vrigazova, "The proportion for splitting data into training and test set for the bootstrap in classification problems," *Bus. Syst. Res. J., Int. J. Soc. Adv. Innov. Res. Economy*, vol. 12, no. 1, pp. 228–242, May 2021.

[57] V. A. Memos, K. E. Psannis, and Z. Lv, "A secure network model against bot attacks in edge-enabled industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 24, no. 11, pp. 234–250, Nov. 2022.

[58] C. Stergiou, K. E. Psannis, B. B. Gupta, and Y. Ishibashi, "Security, privacy & efficiency of sustainable cloud computing for big data & IoT," *Sustain. Comput., Informat. Syst.*, vol. 19, pp. 174–184, Sep. 2018.

M. M. HAFIZUR RAHMAN received the B.Sc. degree in EEE from KUET, Khulna, Bangladesh, in 1996, and the M.Sc. and Ph.D. degrees in information science from the JAIST, in 2003 and 2006, respectively. He is currently working as an Assistant Professor with the Department of CN, CCSIT, KFU, Saudi Arabia. Prior to joining KFU, he was an Assistant Professor with Xiamen University, Malaysia & IIUM, Malaysia, and an Associate Professor with the Department of CSE, KUET. He was also a Visiting Researcher with the School of Information Science, JAIST, in 2008, and a JSPS Postdoctoral Research Fellow at the Graduate School of Information Science (GSIS), Tohoku University, in 2009, and the Japan & Center for Information Science, JAIST, Japan, from 2010 to 2011. His current research interests include hierarchical interconnection networks, optical switching networks, and software defined networks.

MOHAMMED ABDULAZIZ AL NAEEM received the B.Sc. degree in CIS from the College of Management Science and Planning, King Faisal University, in 2005, and the M.Sc. degree in networks and communications (specialization in information security) and the Ph.D. degree in networks and communications (specialization in wireless networks) from Monash University, Australia, in 2009 and 2015, respectively. He is currently the Chairperson for the Department of Computer Networks and Communications, King Faisal University. His research interests include wireless networks, network security, machine learning, artificial intelligence, and pattern recognition.

ADAMU ABUBAKAR received the B.Sc. degree (Hons.) in geography, the P.G.D. degree in computer science from Bayero University, Kano, Nigeria, and the M.Sc. and Ph.D. degrees in computer science from IIUM. He is currently an Associate Professor with the Department of Computer Science, International Islamic University Malaysia (IIUM). Prior to his completion of M.Sc. degree program, from 2005 to 2007, he has worked in various area of information technology. During his Ph.D. degree, he worked on many research projects. He received many awards research. He had published many papers in conferences, journals, and book chapters during his Ph.D. degree, postdoctoral studies, and at the current position of an Assistant Professor. He has been a Professional Member of Association for Computing Machinery (ACM) and Institute of Electrical and Electronics Engineers (IEEE), since 2014. He is working in the areas of computer network, 3D mobile map, information retrieval technologies, digital watermarking, steganography, artificial neural network, and wireless sensor network.