Inside the Mind of the Insider: Towards Insider Threat Detection Using Psychophysiological Signals*

Yessir Hashem, Hassan Takabi†, Mohammad GhasemiGol, and Ram Dantu
Department of Computer Science and Engineering
University of North Texas, Denton, TX, USA
YassirHashem@my.unt.edu, {Takabi, Mohammad.ghasemigol, Ram.Dantu}@unt.edu

Abstract

Insider threat is a great challenge for most organizations in today's digital world. It has received substantial research attention as a significant source of information security threat that could cause more financial losses and damages than any other threats. However, designing an effective monitoring and detection framework is a very challenging task. In this paper, we examine the use of human bio-signals to detect the malicious activities and show that its applicability for insider threats detection. We employ a combination of the electroencephalography (EEG) and the electrocardiogram (ECG) signals to provide a framework for insider threat monitoring and detection. We empirically tested the framework with ten subjects and used several activities scenarios. We found that our framework able to achieve up to 90% detection accuracy of the malicious activities when using the electroencephalography (EEG) signals alone. We then examined the effectiveness of adding the electrocardiogram (ECG) signals to our framework and results show that by adding the ECG the accuracy of detecting the malicious activity increases by about 5%. Thus, our framework shows that human brain and heart signals can reveal valuable knowledge about the malicious behaviors and could be an effective solution for detecting insider threats.

Keywords: Insider Threat, Brain Computer Interface, Electrocardiogram, Electroencephalography, Physiological Indicators

1 Introduction

Insider threat has been a source of critical threats for government and industry organizations. The term insider represents authorized and trusted current or former entities in an organization who have certain privileges and intimate knowledge of internal organizational system structure. Insiders often have information and capabilities not known to external attackers and as a consequence, can cause serious harm [2]. A lot of research has been undertaken in assessment of insider threat attacks in the last decades, however, business organizations, institutions, and government agencies have been suffering from such attacks.

Nowadays, the threat of insider continues to be a major concern in both the public and private sectors [9]. Many surveys demonstrate the severity of insider threat. For example, the Cybercrime report by PwC states that the most serious threat cases were committed by insiders [29]. In 2015, a Federal cybersecurity survey of 200 federal IT managers shows that 76% of the participants are concerned about leaks

Journal of Internet Services and Information Security (JISIS), volume: 6, number: 1 (February 2016), pp. 20-36

*This paper is an extended version of the work originally presented at the 7th ACM CCS International Workshop on Managing Insider Security Threats (MIST'15), Denver, Colorado, USA, October 2015 [11]

†Corresponding author: Computer Science and Engineering Department, University of North Texas, 3940 N.Elm St, Denton, TX 76207, Tel: +1-940-565-2385, Web: `http://www.cse.unt.edu/~takabi`

from insider threats [28].
Since it is almost unattainable to stop an insider threat at the gate, the early detection is the best solution. There have been several detection and monitoring approaches that aim to detect insiders by monitoring the insider's behaviors on using network and organization resources [12]. However, solutions to insider threat are primarily concerned with users for which some psychosocial models have been proposed [10]. As the interest increases in psychological aspects of cyber security, scientists are more concerned with discovering new models that represent these psychological behaviors. However, these models rely on humans to identify the signs and record the behaviors for detecting insider threats. In the meantime, skilled insiders are capable of feigning normal behavior and bypassing the security system. On the other hand, psychophysiological metrics such as electroencephalography (EEG), electrocardiogram (ECG), electromyography (EMG), etc. show a number of advantages over behavioral assessments alone. They are generated involuntarily, meaning that it's too difficult to be changed or mimicked, and they are also continuously available and can be measured automatically.
In this paper, we extend our previous work [11], by analyzing the effectiveness of adding another bio-signal in building our framework to monitor and detect the malicious insider threats. More specifically, we focus on using the electrocardiogram (ECG) signals that arise from the user's heart activities in combination with the electroencephalogram (EEG) signals that arise from the user's brain activities and use them in insider threat detection. The proposed system analyzes the user's bio-signals, extracts features, and classifies them in order to detect the malicious activities. Our experiments involve human subjects to collect the signal samples and evaluate the proposed monitoring framework. The proposed insider threat monitoring framework uses the following procedure:

- Data Acquisition: In order to investigate and analyze the bio-signals, first step is to record the signals. We use a non-invasive Brain Computer Interface (BCI) device (Emotiv EPOC headset [13]) that consists of fourteen active electrodes recording the signals from different parts of the brain and transmitting them via Bluetooth to the computer. We also use the OpenBCI 32-bit Board [15] to record the electrocardiogram (ECG) signals with 256 sampling rates.

- Signals pre-processing and features extraction: In order to extract useful features, we first remove potential noise and artifacts that could be included in the signals during the recording, especially those arising from head movements, eye movements and blinks. Then, we implement feature extraction algorithms to analyze the recorded signals.

- Signals Classification: Next, we apply classification algorithms to the feature vectors we extract to distinguish between regular and malicious activities.

- Evaluation: We evaluate our framework by examining the accuracy of the classifier using the recorded bio-signals samples for different real world scenarios.

The rest of the paper is organized as follows. In Section 2, we present background information of the biological signals. In Section 3, we present related work. In section 4, we describe our threat model. Section 5 describes the framework design and methodology including the signal pre-processing and feature extraction components. Section 6 describes the experiment setup, result, and discusses the limitations. Finally, section 7 discusses the future work, and concludes the paper.

2 Background

Human biological signals such as electroencephalography (EEG), electrocardiogram (ECG), and electromyography (EMG) signals are electrical signals generated by some biological activity inside the hu-

man body and can be continually measured and monitored. The bio-signals vary in frequency and amplitude are measured using amplifiers recording the difference between two electrodes attached to the skin. Electroencephalography (EEG) is a way of monitoring the brain's electrical activities using multiple electrodes placed on the scalp that measure the voltage alteration between the neurons of the brain [22]. These measurements are commonly used in medical and research areas. For example, it has been used to diagnose mental diseases, sleep disorders, and encephalopathy [24]. Using these EEG signals is becoming increasingly popular, especially with the arrival of the low-cost brain computer interfaces (BCI) devices where the signals are analyzed and translated to commands using machine learning techniques [3]. This opens the door widely for new research and proposing new applications and has been used in video games where players use their brain to control the games [8].

These EEG signals capture many important aspects of the human brain activities, and by analyzing the frequency or time domains we could extract valuable features reflecting the conducts of unusual thoughts or behavior changes for the users and provide a number of advantages over behavioral assessments alone as mentioned in section 2. These advantages make it a suitable tool for detecting the malicious activities and identifying insider threats.

The electrocardiogram (ECG) signals have been used in the medical domain for diagnosing many cardiac diseases [19]. However, recently many algorithms have been developed for the exertion of analyzing and classifying the signals and these signals are being used for non-medical purposes. The electrocardiogram (ECG) signals have a specific pattern in a normal life situation. This pattern could be changed due to many reasons such as heart diseases, the body's physical and mental activities, and the emotion or the stress levels.

Recognizing the pattern changes combined with other aspects such as the beat detection which determines the heart rate offer a suitable tool in detecting changes in user behavior and could lead to identifying the insider threats. Recently, ECG signals have been used to identify many aspects of users including emotion and fear with the help of the new generation of the inexpensive consumer grade ECG devices available on the market [27].

There are a number of relatively low cost and low risk EEG devices and several companies such as Emotive [13] and NeuroSky [14] offer variety of inexpensive consumer grade EEG devices. More recently, OpenBCI offers a customizable and fully open brain-computer interface platform, an affordable analog-to-digital converter that can be used to sample EEG, ECG, EMG, and more [15]. Emotiv EPOC is one of the most widely studied of the inexpensive off-the-shelf EEG systems. It is a compact, wireless headset that requires comparatively little effort to set up. It allows increased flexibility and mobility over traditional EEG. Thus, providing an inexpensive tool that developers can use to measure EEG. With the benefit of being noninvasive to the wearer, it is a tool that is practical for our purpose. In this study, we use Emotive EPOC and OpenBCI to perform our experiments and collect data.

3 Related Work

Insider threat problems have been increasing in recent years. This turns the attention to the research community that has been striving to provide solutions and approaches capable of mitigating these kinds of security threats. Many research studies have been investigating and analyzing these type of threats, and raising many research questions such as: what is the insider threat, what is the scope for involving human and psychological factors, and what are the proper approaches to detect and effectively protect the information system. Most of these studies and proposed solutions are based on technological and behavioral theories and intended to detect the attack before impacting the system and posing irrecoverable damages [18][25][26][30].

Recently, several new techniques propose numerous models, and frameworks to deceive insiders and

detect them before achieving their goal [5]. Most of these techniques focus on the anomalous activities of the insider on the network, document accesses, queries, and introduce decoys in order to entrap adversarial insiders onto the network. For example, *Park et al*. propose a software-based decoy system for deceiving insiders [25]. *Kaghazgaran et al*. present a model that shows how to plant and integrate honey permissions into role-based access control model [18]. *Salem et al*. use machine learning, one-class SVM, and one-class Hellinger distance-based to detect malicious intent in information gathering commands [26]. *Thompson et al*. use a Hidden Markov Model to propose a content-based framework to recognize insider anomalies in accessing documents and queries [30].

Although the insider threat has been subject of extensive study, there are very few studies that focus on the biological behaviors of the insider for designing insider threat monitoring and detection systems. However, some studies have investigated the use of heart pulses, voice change, and skin conductivity. For example *Lee et al*. introduce a real-time data leakage prevention system based on biometrics signals recognition technique. Their model detect the unusual changes of biometrics signals such as pulse, electrocardiogram, and skin conductivity when insider tries to leak internal information [21]. *Almehmadi et al*. design an insider threat monitoring system called Physiological Signals Monitoring (PSM) that is able to detect the insider threat [1]. Their system works based on the PSM measurement and the deviation rate of electrocardiogram (ECG) amplitude, Galvanic Skin Response (GSR), and skin temperature that occurs before the threat is executed [1].

On the other hand, there are several studies using the biological signal for emotion identification. For example, *Chanel et al*. propose an approach to classify emotions into three main areas of the valence-arousal space by using physiological signals from both the peripheral nervous system (PNS) and the central nervous system (CNS) [6]. *Kim et al*. use the combination of music and story as stimuli to develop a novel emotion recognition system based on the electrocardiogram (ECG), skin temperature variation, and electrodermal activity [20]. *Wang et al*. analyze the characteristics of EEG features for emotion classification to address the trajectory of emotion changes with manifold learning [31].

There are a few studies that focus on the insider threat issue using human bio-signals. The most recent study done by a company named Veritas proposed an insider threat monitoring system used to monitor both U.S. Army troops and contractors to detect any signs of disloyalty [16]. The system called Hand-Shake implemented in special helmet that can pick up the electromagnetic signals (EEG) and functional near-infrared imaging (fNIRs), which detect the blood flow changes in the brain. The company claim that they reach 80-90% accuracy in detecting the potential insider threat [16].

In summary, most of the previous approaches for insider detection utilize the insider behavior, some of them are using some of the physiological signals to measure the stress and the emotional change for the insiders based on the heart signals, body temperature, and skin conduct. Most of the experiments have also used pictures or audio to stimulate the subjects and raise their stress and emotional level. However, stress or emotion change is not the only characteristic involving insider threats. They could not represent a true scenario for such attack. People could be stressing for different factors such as personal or job-related issues, sickness, or fatigue. In our study we propose a real time insider threat monitoring framework based on the users' biological signal patterns and use real world insider threat scenarios to validate the results.

Our framework depends completely on the bio-signals generated involuntarily by the users and these signals are very difficult to mimic or change. Contrary to the other physiological signals monitoring approaches, we use the electroencephalography (EEG) signals as the main indictor. Also, our approach differs from the Hand Shake system which requires special helmet to measure the functional near-infrared imaging (fNIRs), an expensive equipment and uncomfortable to use, while our approach use inexpensive consumer-grade devices with increased flexibility and mobility to record the EEG and ECG signals.

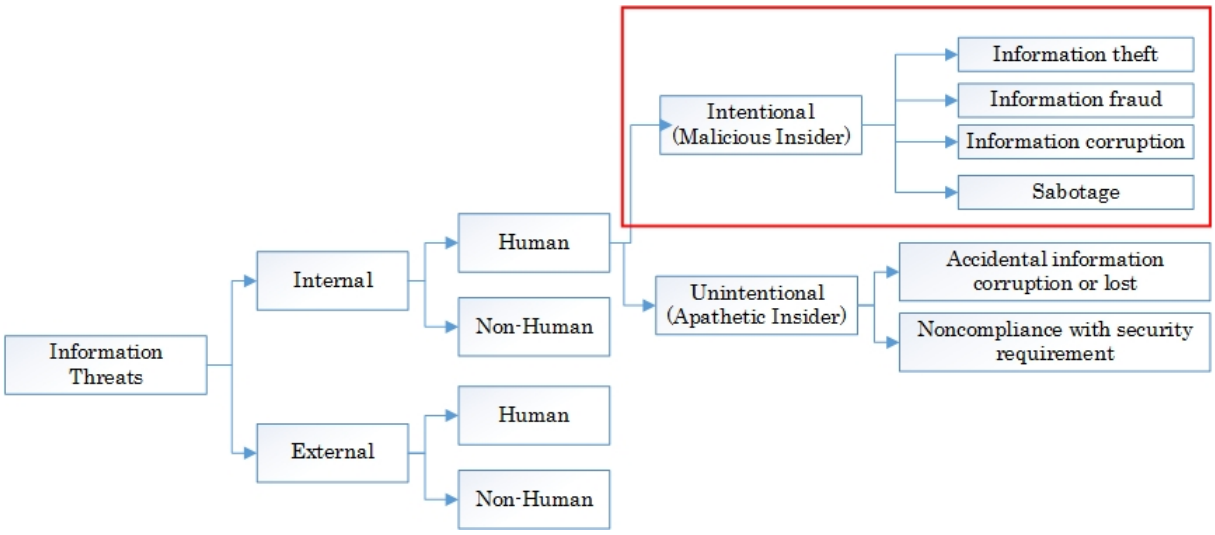

Figure 1: Information system threat taxonomy

4 Insider Threat Taxonomy

In this section, we explain what exactly insider threat means and what kind of insider we target in this work. Forming our threat model is a very essential in explaining our insider threat monitoring and detection framework.

Information threats in general could rely on two sources: human and non-human, such as a hardware failure, power outage, etc. The human source is the most difficult to control and could be represented as external or internal factors. The external factors represent the outside source for the malicious activities such as the adversaries who try to access the system and launch their attacks. On the other hand, the internal factor of threat from the human source is the focus of our research and is one of the major information security research problems. Insiders could be current or former employees or others who have access to the system combined with the good knowledge of the internal organizational processes that may allow them to exploit weaknesses [32].

Insider abuse is categorized into two categories: malicious and apathetic. Malicious threats are when the employee intentionally abuses the system with the aim of manipulating the information or causing damage to the organization. This malicious activity could be in different forms such as fraud, theft, or corruption of information. Apathetic or careless employee could also represent another form of insider threat and could abuse the information system unintentionally by accidentally entering wrong information that can threaten data integrity or sharing important information with unauthorized parties. This also includes the dereliction in doing the job responsibilities such as sending sensitive data without encryption, forgetting to make backup, or negligence in following the security requirements and policies such as using sophisticated passwords, maintaining the account security tools, and keeping computer locked when leaving the desk. Figure 1 adapted from *Loch et al*. show the information system threat taxonomy and how each threats are categorized [23].

In this paper, we focus on the malicious insider threat where the employee or other insider tries to intentionally abuse the system. The adversary in our threat model has access to the system and good knowledge of the internal organizational processes and structures. Our study will provide a strong real-time monitoring framework that is able to detect malicious activities from insiders based on the changes in the bio-signal patterns.

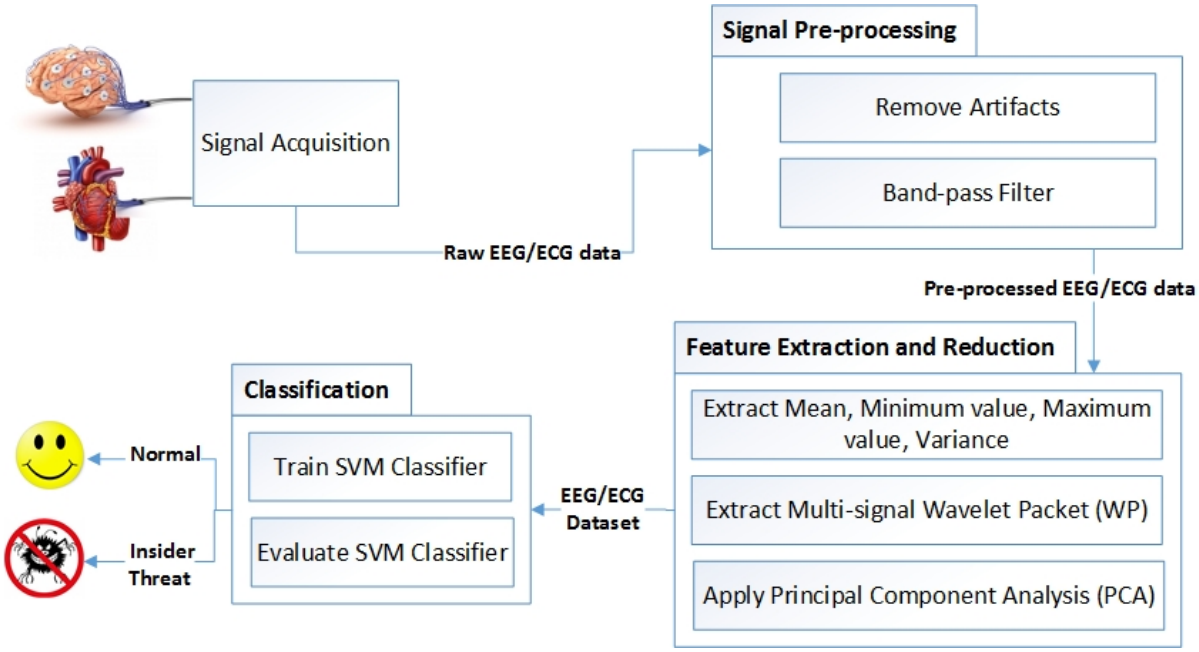

Figure 2: The proposed insider threat monitoring and detection framework

5 The Proposed Insider Threat Monitoring and Detection Framework

Our proposed framework design contains four major units as shown in Figure 2. The signals acquisition unit include sensors that records users' electroencephalography (EEG) and the electrocardiogram (ECG) signals in real-time. The signals are sent to the pre-processing unit for filtering and removing any noise or artifacts and to prepare the signals for further processing. This unit also samples the recorded signals into specific time frames that represent the time window being monitored in the users' brain and heart signals.

Next, the signals are fed to the feature extraction and reduction unit. The unit applies various algorithms to each signals sample to extract features that will be used for detecting malicious activities. In our experiments, we use two types of activity tasks: the regular and the malicious activities. The regular activities record the subject's EEG and ECG signals while the subject performs regular office job activities. The malicious activities record the same signals while subjects perform malicious tasks that represent an insider threat. Based on that, we label each signal with the users' activities. Finally, the labeled features vectors are then sent to the classification unit to help distinguish between the regular activity and the malicious activity. Our trained classifier is able to detect insider threats on the system before an incident occurs and sends an alarm to the monitoring system. The following section describe each component in detail.

5.1 Signals Acquisition

EEG devices can provide knowledge about the user-state during experiments. The EEG data provides signals that are continuously available and could be logged without the user's conscious awareness. This creates an objective measure of the user's brain state.

As mentioned before, there are a number of relatively low cost and low risk EEG and ECG devices that could be used to acquire brainwave and heatwave signals. In this work, we use a consumer-grade BCI device developed by Emotiv [13]. Emotiv EPOC is one of the most widely used of the inexpensive off-

the-shelf EEG systems. It is a compact, wireless headset that allows increased flexibility and mobility over traditional EEG. Thus, providing an inexpensive tool that developers can use to measure EEG. It can be worn like a headset and is equipped with 16 sensors out of which 14 are used to record the EEG signals from different parts of the brain. It has been widely used with different BCI applications and many studies have been using it in their experiments to record subjects' brain signals. Figure 3 shows one of our experiment subjects wearing the Emotiv headset and performing the experiment. For the heart signals, we used the OpenBCI 32-bit board to record the electrocardiogram (ECG) signals with 256 sampling rates [15]. OpenBCI offers a customizable and fully open brain-computer interface platform, an affordable analog-to-digital converter that can be used to sample EEG, ECG and EMG.

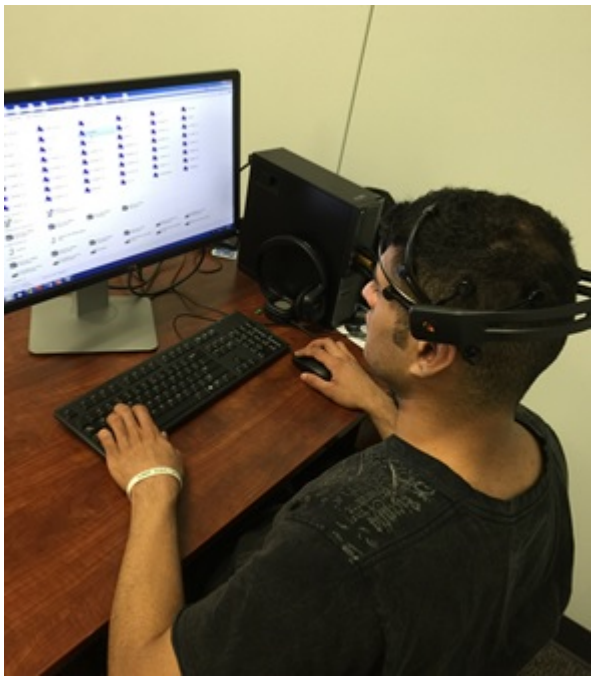

Figure 3: The subject wearing the consumer-grade BCI device capturing the EEG signals

5.2 Signal Pre-processing

The EEG signals have very low frequencies between 0.1 Hz to 50 Hz and very low electrical activity power measured by microvolt. When we record the brainwave signals, typically there are some noises due to unintentional movements that can cause unwanted data. In order to detect the EEG signals, we need to filter the recorded signals from the high frequencies and remove the artifacts. First, we filter the recorded signals from the Electromyography (EMG) noise by removing the higher frequencies using band-pass filter and selecting our frequency range (0.1-30 Hz). We then remove the electro-oculogram (EOG) artifacts that could be included in the signals during the recording. The same techniques are used with the ECG signals with the exception of the use of the band-pass filter to select the (10-30Hz) frequency band in order to extract pure and clear ECG signals.

5.3 Features Extraction and Reduction

The feature extraction component is the most important part of the framework. This is where we extract useful information from the EEG and ECG signals. We use time domain and frequency domain to extract

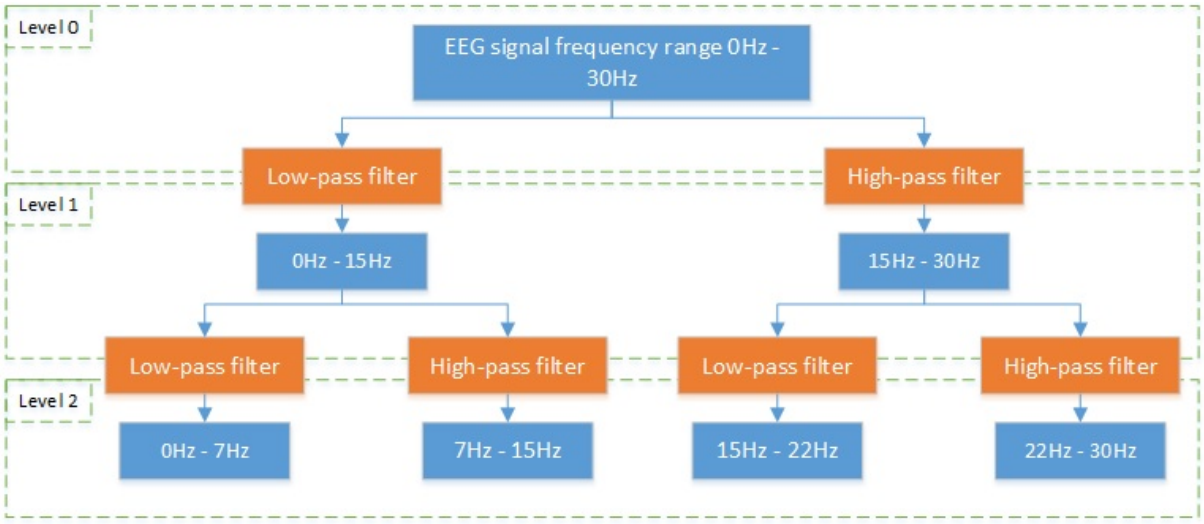

Figure 4: Wavelet Packet Decomposition graph for the recorded EEG signals on three levels decomposition

the features by applying linear and nonlinear measures. For the EEG signals we use wavelet packet decomposition (WPD) method to decompose the EEG frequencies into different frequency subsets and extract features from each subset. It is done when the wavelet packet breaks down the signal to a number of sub-bands using the wavelet function. This could be presented as a sub-band tree where each level of the tree is computed by passing the previous approximation coefficients over high and low pass filters. In our case, we decompose our filtered EEG signals with the frequency of 30 Hz to smaller sub-bands and get the energy for each band. To do this, we use three levels of decomposition that give us six different bands as show in Figure 4. For the ECG signals, we use the same wavelet packet decomposition but with six levels of decomposition.

In the time domain, we use one second window size for the five second sampling time frame. The features are extracted from all the recording electrodes (channels) for the same time frame and added to one feature vector representing that specific time frame. We also extract other features such as Microvolts (μV) mean value, maximum μV, minimum μV, number of peaks and the distance between the high and low μV, from each five-second time frame. Since the number of extracted features is high and not all of them are relevant in the analysis process, we choose only the features that have the most impact on the final results. We apply principal component analysis (PCA) [17] to reduce the number of features, decrease redundancy and consequently improve the system performance. Principal Components Analysis (PCA) is one of the best statistical tools that has been widely used for evaluating and pre-processing data, exposing strong patterns and revealing hidden structures.

5.4 Classification

In the final process in our framework, we fed our features vectors to our classification sachem for training then to distinguish between regular activities and malicious activities. We use the support vector machines (SVMs) as our classification algorithm, it is one of the most practical and powerful classifier used widely in machine learning [7]. The reason behind choosing this classification algorithm is because it provides the best performance on the training set among the other classification algorithms.

We create the training set by labeling each feature vector by its activity task based on the experiment as explained in more detail in section 6. In our experiments, we use two activity tasks: the regular activity

recording the EEG and ECG signals for the subject during doing regular office job activities and malicious activities recording the subject EEG and ECG signals during doing an insider threat attack. The labeled vectors will then be fed to the classifier.

6 Experiments and Evaluation

The main target of our experiments is to define an approach to distinguish between normal activities and malicious activities for the purpose of insider threat detection. To do this, we recorded EEG and ECG signals of a number of participants while performing three different tasks. Each experiment emulates a real-life scenario and we tried to choose scenarios that are very close to a normal work environment and as realistic as possible [11]. The experiments are conducted with the approval of Institutional Review Board (IRB) from the University of North Texas. In the followings, we describe the study and the experiments.

Participants: We had a total of 10 subjects of which 5 were male and the other five were female. All the subjects were between the age of 18 and 33 years old. Most subjects were students from the department of computer science and engineering at the University of North Texas.

Procedure: The experiments were done for each participant separately and at different times during the day. The participants were briefed on the objectives of the study and given a written informed consent explaining the experiment procedure and their right on participating to read and sign. Once the consent was obtained, the participants were seated in a comfortable chair. After the relevant areas of their face and mastoids were cleaned, the Emotiv EEG headset was positioned on the participant's head as shown in Figure 3. Then the OpenBCI sensor was positioned on the participant's right arm in order to record the ECG signals. The examiner verified impedances in connections between each electrode and the participant's scalp.

The EEG headset has 14 electrode channels with channel locations: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 as show in Figure 5. The headset is wirelessly connected to the computer, so subjects can move their head and body comfortably. Each subject met with two investigators in a quiet closed-room setting for three 10 minutes sessions. The experiment was divided into 10 minutes activity sessions for three sessions: one session as regular activity task and two sessions as malicious activity tasks.

Once the participants were seated in a comfortable chair in front of a computer that was used for the experiments, the investigators then proceeded to mount the emotive device on the participants' head. The participants were given a brief verbal explanation of the tasks. The subjects were given five minutes to relax and to feel comfortable with the recording device and the test environment.

In the first scenario, the participants performed regular office job activities such as browsing the internet, using computer applications or using the email account. This scenario observes the brain reaction to the regular daily activities done by most of the employees in any organization. In the second and third scenarios, the participants were asked to perform malicious activities by trying to access information they were not authorized for. We used two realistic scenarios when employee uses the remote access or the network to access unauthorized information. We recorded the ECG signal and the EEG signals generated by the brain in reaction to these malicious activities during each experiment. The EEG and ECG signals recorded during these experiments are used by our framework to extract the features and train the classifier. In the following, we describe each task in more detail.

Regular activity task: The tasks performed by the participants involved the participants doing the following activities for a period of 10 minutes:

- Use the internet browse to find answers for a set of questions provided by the investigator.

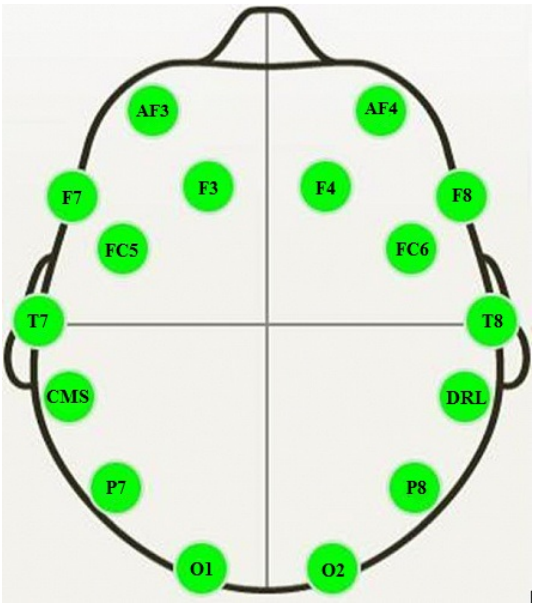

Figure 5: Sensors location map adapted from Emotiv website

- Type the answers in word processor file.
- Attach the file and send to the investigator by email.

These tasks cover three normal office job activities, namely browsing the internet, using computer applications such as Microsoft office, and sending an email.

Malicious activity task I: This task requires the participants to remotely connect to two computers in the same room and do the following:

- Login to the computers remotely and try to find a folder containing students grades in an excel files.
- Search the file to find the information for some specific students (list of students names are given to the subject during the explanation part).
- Copy the information to a new file in the experiment computer desktop.
- Close the files and exit without leaving any trace and making the account owner notice the access.

The participants were asked to perform these tasks without the computer account owners noticing and as an incentive, they received extra reward if they could complete these tasks without leaving any trace. To make the situation as realistic as possible, the account owners were physically present in the same room but away from their computer monitors. For both tasks, we recorded the participants' EEG and ECG signals, and the experiment computer desktop activities.

Malicious activity task II: In the second malicious activity task, the participants were told that a project's source code was located somewhere on the network and they were going to make a copy of the code without anyone noticing. They needed to perform the following activities for a period of 10 minutes:

- Search the lab network and find the project's code folder.

- Access the folder.
- Copy the files from the folder to their computer's hard drive. In order to avoid raising the suspicion of the network administrator, they were told that no file larger than 1 MB should be transferred. If the file was larger than 1 MB, the participants had to split the file into smaller files and transfer them separately.
- At the end of the task, the participant needed to close the files and the folder and exit without leaving any trace.

The participants were asked to perform the task without alarming the network administrator who was in the same room. The participants needed to be careful and copy a small amount of data for each transaction (less than 1MB). The participant were told that a successful transfer of the files without leaving any visible trace will get them an extra reward. During the tasks, we recorded the participants EEG and ECG signals, and the experiment computer desktop activities.

6.1 Experimental Results

In order to evaluate our framework's effectiveness, we used several well-known and widely used metrics. All of these evaluation metrics can be derived from four combinations of a classifier result: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). A TP occurs when a classifier correctly classifies a malicious action, whereas a FP occurs when a normal action is misclassified as being malicious. Likewise, a TN is generated whenever a normal action is correctly classified, while a FN occurs when a malicious action is not detected by the classifier. According to these variables, we define Accuracy, Precision, Recall, F-measure, and Error-rate as follows:

$$Accuracy = (TP+TN)/(TP+TN+FP+FN) \tag{1}$$

$$Precision = TP/(TP+FP) \tag{2}$$

$$Recall = TP/(TP+FN) \tag{3}$$

$$F-measure = 2 \times Recall \times Precision/(Precision+Recall) \tag{4}$$

$$Errore-rate = (FP+FN)/(TP+TN) \tag{5}$$

Our results section is divided into two parts, the first one presents the results of using the EEG data recorded during the experiments, and the second one presents the results of combining the EEG and ECG signals for the same subjects. The goal here is to show the capability of our framework in including multiple bio-signals and also to present how these signals are able to affect the results.

6.1.1 EEG signals results

After the experiments were performed and the signals were collected, we categorized the recorded signals into three groups: regular activity group, malicious I activity group, and malicious II activity group. Each group consists of 10 signal samples represented by each participant in the experiments. We then analyzed and pre-processed the signals by removing the high frequency noise and the electrooculogram

Class type	Classification Accuracy	precision	recall	F-measure	Error Rate
Normal Class	0.8445	0.8383	0.8231	0.8513	0.1555
Malicious Class	0.8390	0.8430	0.8557	0.8320	0.1610

Table 1: Results of EEG dataset containing the first malicious scenario

Class type	Classification Accuracy	precision	recall	F-measure	Error Rate
Normal Class	0.8995	0.8874	0.8929	0.8969	0.1005
Malicious Class	0.8960	0.9090	0.9045	0.8985	0.1040

Table 2: Results of EEG dataset containing the second malicious scenario

Class type	Classification Accuracy	precision	recall	F-measure	Error Rate
Normal Class	0.8501	0.8865	0.7734	0.8876	0.1499
Malicious Class	0.8555	0.7797	0.8877	0.7856	0.1445

Table 3: Results of EEG dataset containing both malicious scenarios

(EOG) artifacts.
Next, we applied our feature extraction algorithm for each five second timeframe and labeled the result features vector with its represented group type. Then, we applied the SVM classifier using k-fold cross validation (k=5) to evaluate the accuracy of the SVM to classify normal and malicious activity. Table 1 shows the results for the EEG dataset containing the first malicious scenario. The results show 84% accuracy in detect the malicious activities and the regular activities, the results also show around 84% precision and about 85% recall in detecting malicious activities. Figure 6. shows that the best results are achieved when sigma (kernel function parameter) is set to 5.

// Similar to the first test, we applied the SVM on EEG dataset containing the second malicious activity scenario. The results show 90% accuracy in detecting both regular and malicious activities for EEG dataset containing the second malicious activity scenario as shown in Table 2.
In the third test, we applied the SVM on both malicious scenarios to evaluate the detection accuracy for normal and malicious in general. The results show up to 86% accuracy for detecting both regular and malicious activities as show in Table 3. In general, our results show detection accuracy above 84% for all the scenarios which demonstrates that electroencephalography (EEG) can reveal valuable knowledge about user behaviors and could be a good solution for the continuous insider threat monitoring.

6.1.2 Combined EEG and ECG signals results

In this section, we show our framework capability of adding other bio-signals and show how the results will improve. To do this, we chose five subjects with the lowest detection accuracy from the EEG results. We then combined their EEG signals with their ECG signals recorded during the same experiment. We ran our classifier for each case and compared the results. Note that all the ECG signals have been pre-processed and filtered and passed the same process as the EEG signals. We also have used the band-pass filter and chosen the (10-30Hz) frequency band and applied our feature extraction algorithm for each five second time frame and labeled the result feature vector with its represented group type.
We applied the SVM to the EEG features dataset containing the regular and both malicious activities for five subjects using k-fold cross validation (k=5), and then we added the ECG features and ran the

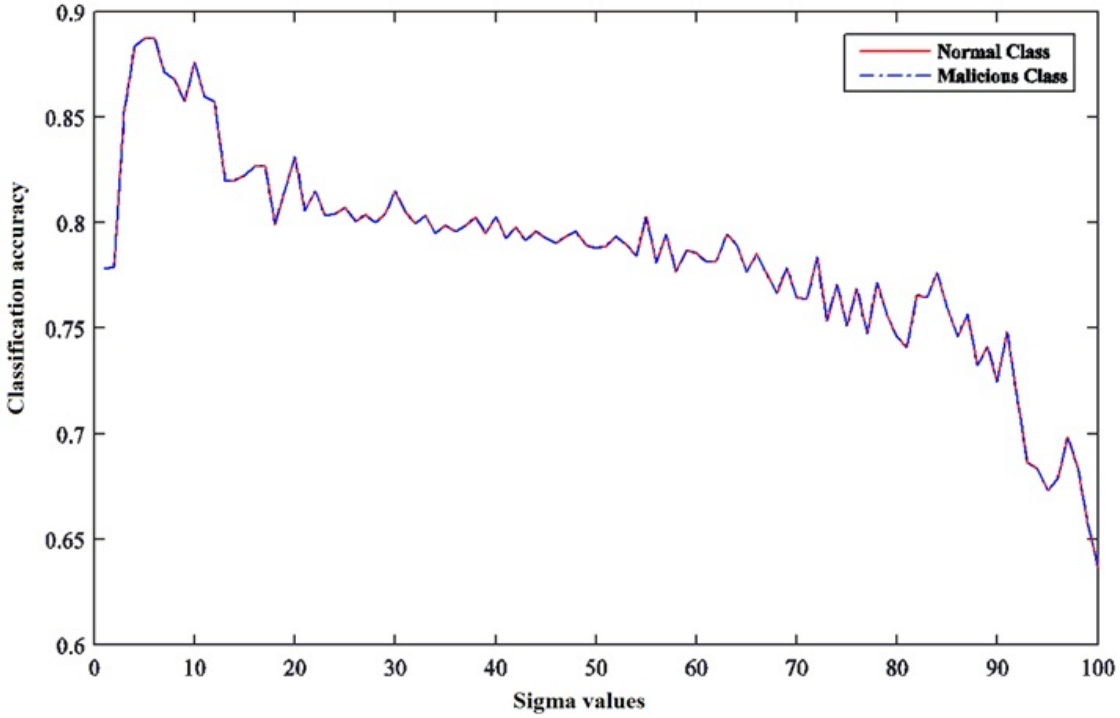

Figure 6: Best results tuning the RBF kernel parameter in EEG dataset are occurred for Sigma=5

Experiment	Class type	Accuracy	precision	recall	F-measure	Error Rate
EEG Signals	Normal Class	0.8191	0.8856	0.8228	0.8531	0.1809
	Malicious Class	0.8116	0.7066	0.8194	0.7588	0.1884
EEG + ECG Signals	Normal Class	0.8620	0.9266	0.8720	0.8984	0.1380
	Malicious Class	0.8691	0.7490	0.8483	0.7956	0.1309

Table 4: Results of combined EEG and ECG signals

classification process again. Table 4 shows the results for the two cases, the first one using the EEG signal features only and the second one using both EEG and ECG signals features for the same subjects and the same time frames. The results show up to 5% increase in the classification accuracy, precision and recall for both normal and malicious activities scenarios using both bio-signals. Figure 7 shows the receiver operating characteristic (ROC) curve plotting the true positive rate against the false positive rate and comparing the results for the two experiments. The graph shows that the combination of EEG and ECG signals provide the best performance for both malicious and normal activities.

6.2 Limitations

Despite the promising results, our experiments have some limitations. One limitation of our proposed framework is that it relies on users wearing the headset in order to continuously monitor users' EEG signals. This is inconvenient and it is not realistic to expect users to wear the device all the time. However, this could be used in military environments for example or for subset of users who have access to highly sensitive information. Additionally, with recent advances in wearable technologies and devices, the limitation will not exist in near future. For example, most smartwatches have sensor that can record ECG

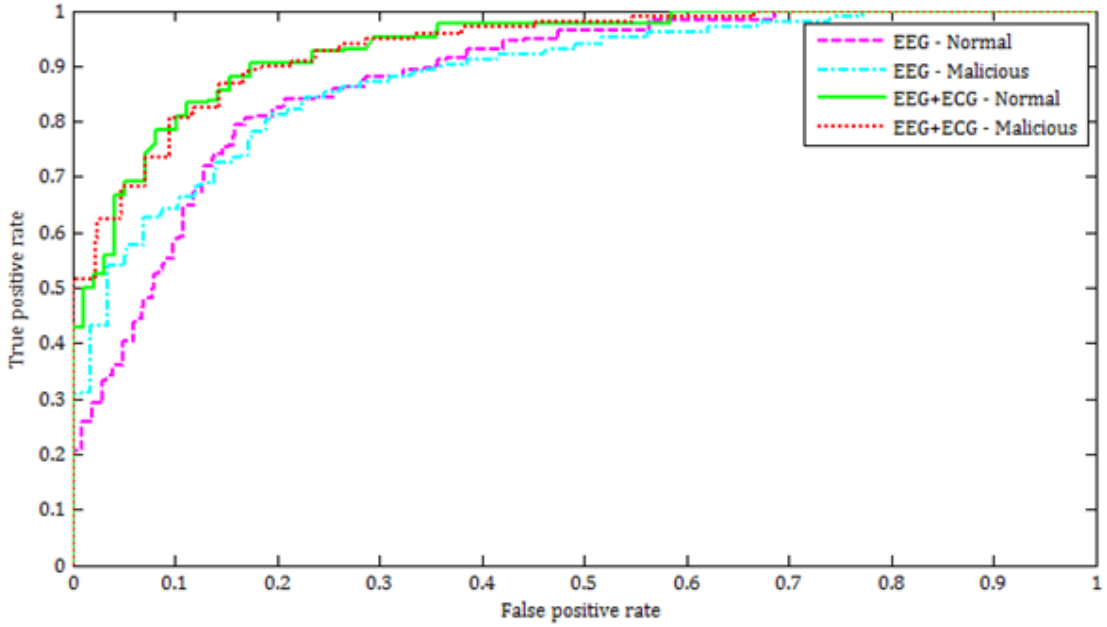

Figure 7: ROC comparison for EEG experiment and EEG+ECG experiment

signals and it is very likely that we will soon see wearable devices capable of recording EEG signals in the market.

Another concern with the proposed framework is potential privacy issues of monitoring bio-signals. Bio-signals like EEG and ECG carry a wealth of information about user's health and other personal information. Previous research has shown that EEG signals could be used to extract sensitive personal information such as religious beliefs, sexual preferences, etc [4]. Our proposed framework doesn't store the recorded signals and only uses them to detect potential malicious activities. Although this answers some of the concerns, accessing the raw bio-signals and processing them can still raise privacy concerns. In our future work, we will develop algorithms to process these signals in a privacy preserving manner.

7 Conclusion and Future Work

Insider threat is considered as one of the major concerns for cybersecurity. As it is almost impossible to stop an insider threat at the gate, a sufficient insider threat monitoring and detection framework is needed for early detection. It is recognized that solutions to insider threat are mainly user centric. In this paper, we present a real-time insider threat monitoring framework based on the EEG and ECG signals.

Our framework analyzes the users' biological signals, extracts features, and classifies them into normal or malicious activities. We evaluated our framework by conducting an experiment including ten subjects doing three different scenarios that indicate regular and malicious activities. Our framework shows up to 86% average accuracy in detecting the malicious insider using EEG signals. We also evaluated our framework's capability by adding another bio-signal (ECG). When we used the ECG signal features in combination with the EEG features, our results show about 5% increase in the detection accuracy. The results demonstrate that bio-signals can reveal valuable knowledge about the user's behavior and could be a very effective solution for detecting the insider threats.

For future work, we plan to define new features to detect the insider threats with higher accuracy. We also plan to integrate other biological aspects such as the eye movements into our framework.

References

[1] A. Almehmadi and K. El-Khatib. On the possibility of insider threat detection using physiological signal monitoring. In *Proc. of the ACM 7th International Conference on Security of Information and Networks (SIN'14), Glasgow, UK*, pages 223–230. ACM, September 2014.

[2] E. Bertino and G. Ghinita. Towards mechanisms for detection and prevention of data exfiltration by insiders: Keynote talk paper. In *Proc. of the 6th ACM Symposium on Information, Computer and Communications Security (ASIA CCS'11), Hong Kong, China*, pages 10–19. ACM, March 2011.

[3] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, S. Haufe, C. Maeder, L. Ramsey, I. Sturm, G. Curio, et al. The berlin brain–computer interface: non-medical uses of bci technology, December 2010. [Online; Accessed on Feabruary 20, 2016] `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3002462/`.

[4] T. Bonaci and H. J. Chizeck. Privacy by design in brain-computer interfaces. Technical Report UWEETR-2013-0001, Dept of EE, University of Washington, 2013.

[5] B. M. Bowen, M. B. Salem, A. D. Keromytis, and S. J. Stolfo. Monitoring technologies for mitigating insider threats. In C. W. Probst, J. Hunker, and D. G. M. Bishop, editors, *Insider Threats in Cyber Security*, volume 49 of *Advances in Information Security*, pages 197–217. Springer US, July 2010.

[6] G. Chanel, K. Ansari-Asl, and T. Pun. Valence-arousal evaluation using physiological signals in an emotion recall paradigm. In *Proc. of the 2007 IEEE International Conference on Systems, Man and Cybernetics (ISIC'07), Montreal, Canada*, pages 2662–2667. IEEE, October 2007.

[7] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, September 1995.

[8] A. L. cuyer, F. Lotte, R. B. Reilly, R. Leeb, M. Hirose, and M. Slater. Brain-computer interfaces, virtual reality, and videogames. *IEEE Computer*, 41(10):66–72, October 2008.

[9] J. Glasser and B. Lindauer. Bridging the gap: A pragmatic approach to generating insider threat data. In *Proc. of the 2013 IEEE Security and Privacy Workshops (SPW'13), San Francisco, CA, USA*, pages 98–104. IEEE, May 2013.

[10] F. L. Greitzer, L. J. Kangas, C. F. Noonan, A. C. Dalton, and R. E. Hohimer. Identifying at-risk employees: Modeling psychosocial precursors of potential insider threats. In *Proce. of the 45th IEE Hawaii International Conference System Science (HICSS'12), Hawaii, USA*, pages 2392–2401. IEEE, January 2012.

[11] Y. Hashem, H. Takabi, M. GhasemiGol, and R. Dantu. Towards insider threat detection using psychophysiological signals. In *Proc. of the 7th ACM CCS International Workshop on Managing Insider Security Threats (MIST'15), Denver, Colorado, US*, pages 71–74. ACM, October 2015.

[12] J. Hunker and C. W. Probst. Insiders and insider threats—an overview of definitions and mitigation techniques. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2(1):4–27, March 2011.

[13] E. Inc. Emotiv epoc / epoc+: Scientific contextual eeg. [Online; Accessed on Feabruary 20, 2016] `www.emotiv.com`.

[14] N. Inc. Eeg and ecg biosensor solutions. [Online; Accessed on Feabruary 20, 2016] `www.neurosky.com`.

[15] O. Inc. Openbci 32bit board kit. [Online; Accessed on Feabruary 20, 2016] `www.openbci.com`.

[16] V. S. Inc. Handshakes test and technologies. [Online; Accessed on Feabruary 20, 2016] `http://www.veritas.blueleveragemedia.com/products/handshake/`.

[17] I. Jolliffe. *Principal component analysis*. Springer-Verlag New York, 2002.

[18] P. Kaghazgaran and H. Takabi. Toward an insider threat detection framework using honey permissions. *Journal of Internet Services and Information Security (JISIS)*, 5(3), August 2015.

[19] S. Karpagachelvi, M. Arthanari, and M. Sivakumar. Ecg feature extraction techniques-a survey approach. *International Journal of Computer Science and Information Security*, 8(1):76–80, May 2010.

[20] K. H. Kim, S. Bang, and S. Kim. Emotion recognition system using short-term monitoring of physiological signals. *Medical and biological engineering and computing*, 42(3), May 2004.

[21] H. Lee, J. Jung, T. Kim, M. Park, J. Eom, and T. Chung. An application of data leakage prevention system based on biometrics signals recognition technology. In *Proc. of the 3rd International Conference on Network and Computing Technology (ICNCT'14), Jeju Island, South Korea*, pages 93–94, April 2014.

[22] G. A. Light, L. E. Williams, F. Minow, J. Sprock, A. Rissling, R. Sharp, N. R. Swerdlow, and D. L. Braff. Electroencephalography (eeg) and event-related potentials (erps) with human participants, July 2010. [Online; Accessed on Feabruary 20, 2016] http://www.ncbi.nlm.nih.gov/pubmed/20578033.

[23] K. D. Loch, H. H. Carr, and M. E. Warkentin. Threats to information systems: today's reality, yesterday's understanding. *Mis Quarterly*, 16(2), June 1992.

[24] N. L. of Medicine – National Institutes of Health. electroencephalogram (eeg), February 2015. [Online; Accessed on February 20, 2016] http://www.nlm.nih.gov/medlineplus/ency/article/003931.htm.

[25] Y. Park and S. J. Stolfo. Software decoys for insider threat. In *Proc. of the 7th ACM Symposium on Information Computer and Communications Security (ASIA CCS'12), Seoul, Republic of Korea*, pages 93–94. ACM, May 2012.

[26] M. B. Salem and S. Stolfo. Masquerade attack detection using a search-behavior modeling approach. Technical Report CUCS-027-09, Computer Science Department, Columbia University, 2009.

[27] J. Selvaraj, M. Murugappan, K. Wan, and S. Yaacob. Classification of emotional states from electrocardiogram signals: a non-linear approach based on hurst, May 2013. [Online; Accessed on Feabruary 20, 2016] http://www.ncbi.nlm.nih.gov/pubmed/23680041.

[28] SolarWinds. Solarwinds survey investigates insider threats to federal cybersecurity, January 2015. [Online; Accessed on February 20, 2016] http://www.solarwinds.com/company/newsroom/press_releases/threats_to_federal_cybersecurity.aspx.

[29] P. Taiwan. Cybercrime: Protecting against the growing threat - events and trends, March 2012. [Online; Accessed on February 20, 2016] http://www.pwc.tw/en_TW/tw/publications/events-and-trends/assets/e256.pdf.

[30] P. Thompson. Weak models for insider threat detection. *International Society for Optics and Photonics*, 5403(3), April 2004.

[31] X.-W. Wang, D. Nie, and B.-L. Lu. Emotional state classification from eeg data using machine learning approach. *Neurocomputing*, 129, April 2014.

[32] R. Willison and M. Warkentin. Beyond deterrence: An expanded view of employee computer abuse. *MIS Quarterly*, 37(1), March 2013.

Author Biography

Yassir Hashem is Ph.D. candidate in Computer Science and Engineering at the University of North Texas. He received a Master in Computer Science and Engineering from University of North Texas and BS in Computer Science (2005) from university of Mosul, and has 6 years of industrial experience in united stated worked with Foxconn Hon Hai Precision Industry Co. and Ingram micro mobility. He has done broad research on security topics such as insider threat, context aware computing, mobile privacy, and privacy and security of online social networks.

Hassan Takabi is an Assistant Professor of Computer Science and Engineering at the University of North Texas, Denton, Texas, USA. He is director and founder of the INformation Security and Privacy: Interdisciplinary Research and Education (INSPIRE) Lab and a member of the Center for Information and Computer Security (CICS), which is designated as National Center for Academic Excellence in Information Assurance Research (CAE-R) and Education (CAE-IAE). His research is focused on various aspects of cybersecurity and privacy including advanced access control models, insider threats, cloud computing security, mobile privacy, privacy and security of online social networks, and usable security and privacy. He is member of ACM and IEEE. Contact him at takabi@unt.edu.

Mohammad GhasemiGol is a PhD candidate in computer engineering at Ferdowsi University of Mashhad (FUM). He will join the Department of Computer Engineering at the University of Birjand in fall 2016. He received the B.S. degree in Computer Engineering from Payame Noor University (PNU), Birjand, Iran, in 2006. He also received the M.S. degree in Computer Engineering at FUM, Iran, in 2009. November 2014 to July 2015, he was with the Department of Computer Science and Engineering, University of North Texas, Denton, TX, USA as a visiting research scholar. His research interests include network security, intrusion detection and response systems, alert management, machine learning and data mining, and optimization problems.

Ram Dantu has 15 years of industrial experience in the networking industry, where he worked for Cisco, Nortel, Alcatel, and Fujitsu and was responsible for advanced technology products from concept to delivery. He is a full professor in the Department of Computer Science and Engineering, University of North Texas (UNT). During 2011, he was a visiting professor at Massachusetts Institute of Technology (MIT) in the School of Engineering. He is the founding director of the Network Security Laboratory (NSL) at UNT, the objective of which is to study the problems and issues related to next-generation networks. He is also the director of the Center for Information and Computer Security at UNT. He has received several NSF awards in collaboration (lead PI) with Columbia University, Purdue University, University of California at Davis, Texas A&M University and MIT. During the last 6 years he received 10 research awards from the National Science Foundation (NSF) for a total of $5M. He was selected as a member of Innovation Corps of NSF in 2011.