Security Informatics
a SpringerOpen Journal

RESEARCH Open Access

Specializing network analysis to detect anomalous insider actions

You Chen[1*], Steve Nyemba[1], Wen Zhang[2] and Bradley Malin[1,2]

* Correspondence: you.chen@vanderbilt.edu
[1]Department of Biomedical Informatics, School of Medicine, Vanderbilt University, Nashville, TN, 37203, USA
Full list of author information is available at the end of the article

Abstract

Collaborative information systems (CIS) enable users to coordinate efficiently over shared tasks in complex distributed environments. For flexibility, they provide users with broad access privileges, which, as a side-effect, leave such systems vulnerable to various attacks. Some of the more damaging malicious activities stem from internal misuse, where users are authorized to access system resources. A promising class of insider threat detection models for CIS focuses on mining access patterns from audit logs, however, current models are limited in that they assume organizations have significant resources to generate label cases for training classifiers or assume the user has committed a large number of actions that deviate from "normal" behavior. In lieu of the previous assumptions, we introduce an approach that detects when specific actions of an insider deviate from expectation in the context of collaborative behavior. Specifically, in this paper, we introduce a specialized network anomaly detection model, or SNAD, to detect such events. This approach assesses the extent to which a user influences the similarity of the group of users that access a particular record in the CIS. From a theoretical perspective, we show that the proposed model is appropriate for detecting insider actions in dynamic collaborative systems. From an empirical perspective, we perform an extensive evaluation of SNAD with the access logs of two distinct environments: the patient record access logs a large electronic health record system (6,015 users, 130,457 patients and 1,327,500 accesses) and the editing logs of Wikipedia (2,394,385 revisors, 55,200 articles and 6,482,780 revisions). We compare our model with several competing methods and demonstrate SNAD is significantly more effective: on average it achieves 20-30% greater area under an ROC curve.

Keywords: Insider threat, anomaly detection, collaborative information system, specialized network, electronic health record, access logs

1 Introduction

The popularity of collaborative information systems (CIS) has exploded over the past decade. CIS are now considered to be critical infrastructure in a wide-range of application domains. They are utilized in popular Web 2.0 applications, such as wikis, dynamic bookmarking, social networking, and groupware [1]. CIS have also transitioned beyond public systems and have become central to environments that handle personal or strategic knowledge [2], such as healthcare operations [3,4] and intelligence-related activities [5,6].

The escalated use of CIS is due, in part, to their ability to offer several major benefits in comparison to their predecessors. First, there is evidence that they can increase the

efficiency of completing tasks posed to the members of an organization [7]. These gains may be realized through various routes, such as coordination of workflows [8], team-based communication and editing [9], and management of the life-cycle of knowledge and documents in heterogenous environments [10]. Second, they can improve the *quality* of the work produced by the users of the system [11]. The basis of such improvements is also varied and may arise by leveraging the wisdom of crowds [12], performing innovative brainstorming sessions [13], and provisioning information in a just-in-time manner from experts [14]. In general, these benefits are realized because CIS facilitate flexible participation and coordination between disparate users over common tasks. Yet, the flexible nature of a CIS and the enhanced service capabilities they deliver leave them vulnerable to a range of novel information security threats. These vulnerabilities arise because the environments in which CIS are deployed are inherently complex.

They often consist of a large number of users, permissions or functions, and *ad hoc* relationships between users and data elements - all of which fluctuate over time. The consequence of such complexity is that CIS are subject to misuse and abuse, which can ultimately corrupt or expose information [15]. This is particularly a concern in CIS that manage sensitive data, such as electronic health records (EHRs), where misuse of the system can lead to the exploitation of private medical details [16].

The insider threat has long been recognized as a challenging problem in information systems security [17,18] and CIS are not devoid of protections. Over the years, access control frameworks (e.g., [19-21]), anomaly detection strategies (e.g., [22-25], and expert-informed suspicious behavior classifiers (e.g., [26,27]) have been developed to prevent and detect illicit activities in CIS. However, the approaches developed to date do not point out the specific anomalous activities of users. This is because the current set of approaches tend to assume that either training information is available to build robust classifiers or that the user has committed a large number of actions that deviate from "normal" behavior. Given these limitations, the main contributions of this paper, which is an extension of [28], are as follow:

1. A local network approach to anomalous access detection

We recognize that detailed and well-defined semantics about the users and data in a CIS are not always available. As such, our primary goal is to develop an analytic framework to detect anomalous accesses by leveraging the inherent collaborative nature of the users in a CIS. We hypothesize that if a user is a threat, the similarity of the network of users will be higher when the user is suppressed from the network. Based on this hypothesis, we build a model for each subject (e.g., patient's medical record) in the form of a subnetwork of users (e.g., the set of authenticated healthcare workers); i.e., the *specialized* network. Our model then assesses if the similarity of the network with and without the user are significantly different. In contrast to [28], we provide justification for certain modeling decisions and the similarity measures employed.

2. An empirical evaluation of the capabilities of our method

Beyond developing the model, we perform an empirical investigation with the access logs of two distinct CIS. The first consists of the access logs from a large restricted-access EHR system. The second consists of the editing logs from a publicly-accessible

online wiki. In the context of these real domains, we simulate intruding behavior in several manners that are indicative of various known illicit actions. Our results illustrate that when a specialized network is intruded upon, its similarity often sufficiently decreases to detect the intrusion. Additionally, and perhaps more importantly, we demonstrate that relatively simple data mining techniques are more effective than complex network decomposition methods for this specific detection problem. Beyond the material in [28], this paper provides a more comprehensive analysis of the competing anomaly detection techniques and variants of the specialized network analysis approach.

The remainder of the paper is organized as follows. In Section 2, we review related research in methodologies for preventing and detecting insider threats. Next, in Section 3, we introduce our model, which is called *specialized network anomaly detection*, or SNAD. Then, in Section 4, we report on our empirical study and highlight specific findings. Finally, in Section 5 we discuss the merits, as well as limitations, of our approach and in Section 6 we conclude the paper with a summary of the work.

2 Related Work

Prior research into insider threats in collaborative environments can be roughly partitioned into two types: prevention and detection.

2.1 Prevention of Insider Threats

The prevention of insider threats requires a combination of policy and technical approaches. From a policy perspective, it is recommended that organizations' perform various duties, such as assess the trustworthiness of potential employees (or group members), clearly document procedures, and provide periodic security awareness training. The focus of our paper is on technical approaches, but we refer the reader to [29] for an excellent summary of policies and insider threat management.

From a technical perspective, organizations are encouraged to enforce separation of duties and limit privileges in a need-to-know manner. In this regard, formal access control frameworks (e.g., [19,20]), which are designed to prevent illicit accesses from authenticated users by appropriately defining (and restricting) permissions, have been proposed. To enable efficient and effective management of users and their privileges, users are often assigned to roles; i.e., role-based access control (RBAC). Beyond the basic RBAC model, certain access control frameworks address team [23] and context-enhanced scenarios [22,24,25,30]. For instance, [25,31] demonstrated that RBAC models can include logical contextual rules for expressing the relationship between a user and a subject. This situation-based access control model, or SitBAC, enables formal representation of access scenarios to subjects as an ontology of entities involved in data access, in the form of i) the subject, ii) the user, iii) the task, iv) the legal authorization and v) their relationships. Though more specific than RBAC, contextual access control models do not readily capture the dynamic relationships among users, the hallmark of CIS.

Moreover, if an organization is to apply RBAC-like models, they must define roles. These may be derived through either a top-down or bottom-up approach. In the top-down sense, roles can be engineered through scenario-driven interviews with members of an organization to address the expected needs of an organization [32,33]. Given the

scale and complexity of modern organizations, it has been suggested that roles may be derived through an alternative bottom-up manner, via data mining techniques collectively referred to as role mining [34,35]. The mined roles are subsequently adjudicated by expert security engineers and administrators. For instance, [34] proposed a method to discover RBAC roles by extracting patterns from an organization's database of users' permission lists. These types of role discovery techniques are scalable, but, to the best of our knowledge, they have yet to be applied to dynamic systems. It is not clear that they are sufficiently flexible to model stable role relationships.

2.2 Detection of Insider Threats

In a CIS, users tend to function as dynamic teams [36], which makes it difficult to differentiate between normal and abnormal accesses based solely on roles or permissions. As a result, technologies offered by industry tend to focus less on insider threats than on external vulnerabilities [17,37]. To overcome this deficiency, insider threat *detection* methods have been proposed to supplement access control frameworks. Insider threat detection can be categorized into two types: supervised and unsupervised learning strategies.

2.2.1 Supervised Methods

Supervised methods assume that there are examples of insider threats available for building a classification tool. These examples are often provided through documented evidence or derived from expert knowledge. Of particular note to our investigations, there have been several recent approaches proposed for detecting insider threat detection in the context of EHRs. [27], for instance, described an approach that assumes most accesses to EHRs occur for a valid clinical or operational reason. Based on this belief, the method filters out insider actions when any of three general types of "explanations" are observed: i) direct, ii) group, and iii) consultation. The remaining actions are subsequently considered interesting for investigation by a human.

As an alternative, [26] recently proposed a machine learning approach to suspicious access detection based on interviews with the privacy officials of several healthcare organizations. In this approach, access events are defined over a collection of twenty-six features about the subjects (i.e., patients) and users (i.e., care providers). Privacy officials then label a set of suspicious and non-suspicious access events using an iterative refinement process. Based on these events, a classification model is trained (logistic regression and support vector machines were specifically applied in the study) to predict which events in a test set of EHR accesses were suspicious. [26] demonstrated that machine learning approaches were more effective detectors than the strict rule-based approaches the privacy officials tend to use in their current investigations. However, this model assumes that experts are aware of the features associated with insider threats and that they are readily available.

2.2.2 Unsupervised Methods

Unsupervised learning methods model inherent patterns in a system, so that, in the context of insider threat detection, anomalies can be pinpointed. In collaborative systems, certain data structures and theories have shown promise. These include techniques rooted in behavioral modeling and community detection [38-41]. SNAD falls into this category and here we review several related techniques.

In the set of methods most related to SNAD [39,40,42], the system is initially represented as the adjacency of a bipartite graph with subjects as rows and users as columns. The cells of the matrix are filled with values representative of the affinity of a user to the subject in question. The premise of all of these methods is that typical users are likely to form and function as neighborhoods, such that the likelihood a strangely behaving user will be characterized by the neighborhoods is low.

The main difference between these methods is how such neighborhoods are discovered and applied for anomaly detection. In [40], for instance, a method to detect anomalies in networked systems is leveraged. Specifically, for each vertex of interest (e.g., user), the method performs a random walk over the graph to derive proximity to other vertices (e.g., other users), which are integrated into a neighborhood. The method identifies anomalies and vertices that are sufficiently distant from the neighborhood.

Instead of a random walk, [42] proposes a spectral method to detect anomalous instances. In this method, a spectral decomposition is applied to the adjacency matrix to derive communities in the form of eigenvectors, which are weighted by the strength of their corresponding eigenvalues. Each column in the original adjacency matrix is projected onto the new communities of users. Columns that are distant from the communities (weighted by the eigenvalues) are considered to be anomalies.

A social networking perspective is offered in [39] to explicitly focus on collaborative systems. In this method, the adjacency matrix is subsequently transformed into a social network composed of vertices from only one of the classes; i.e., the users. The edges of the network are weighted based on the degree to which users access common vertices of the other class; i.e., the patients. Then, like [42], the method applies a decomposition of the graph to derive communities. Anomalies are detected by comparing each user to its nearest neighbors in the network with respect to the communities.

SNAD differs from the current set of methods in several significant ways. First, none of these methods distinguish between the actions of a user. Instead, they assume all actions committed by an anomalous user are suspicious. Such techniques are relevant when a user's account has been compromised or the user is performing a significant number of actions outside of their normal routine, but their application in our scenario would lead to a gross inflation in the number of false alarms. Second, all of the methods use a *global* perspective of the system, which assumes that distant relations influence local behavior. While this may be true when considering aggregate behavior, we hypothesize that the global view is more error-prone than a more localized view when attempting to detect subtle illicit actions, such as the access (or amendment) of a single subject in the CIS.

Though all of these methods were designed for anomalous user detection, the methods in [40] and [39] cannot be directly applied to a local network. Consider, the random walk model in [40] was designed to walk over the entire network. As a result, a user would receive the same anomaly score regardless of the local network in which they are being investigated. The method in [39] suffers from a similar problem since it does not distinguish between users and their actions. In contrast, the method in [42] can be adapted for anomalous access detection because the it can infer communities from a local network based on a space which is different from that of the global

network. As a result, we compare our method to a variation of the spectral method of [42] as described in the following section.

3 Intruding Access Detection Model

This section introduces our model, which we call *specialized network anomaly detection* (SNAD). The approach is dubbed "specialized" because it focuses on a local view of the information system (as opposed to a global view) that is conditioned on specific subjects. We begin with a high-level overview of SNAD and then delve into the details of the particular methods it incorporates.

SNAD functions under the premise that normal and abnormal accesses will have sufficiently different influence on the similarity of the users in an access network. As depicted in Figure 1, SNAD can be represented as two general components: 1) Similarity Measurement (SNAD-SM), which feeds into 2) Anomaly Evaluation (SNAD-AE).

The SNAD-SM component extracts networks of users from the access logs of a CIS. More specifically, this component constructs a local access network for each subject. It then calculates the similarity of the users' access patterns in the network. Rather than focus on the individual features of the users or the subjects, SNAD aims for a more general representation to model the social behavior in the system by constructing and measuring the similarity of users' access networks.

The SNAD-AE component evaluates each access by comparing the similarity of an access network to its subnetwork. More specifically, SNAD-AE measures the similarity of the users that access a particular subject. This network is then compared to the similarity of a subnetwork that suppresses one of the network's users. If the similarity between the network and subnetwork are significantly different, then SNAD claims the suppressed user's access was an anomaly.

For reference, Table 1 summarizes the variables and notation used throughout the paper.

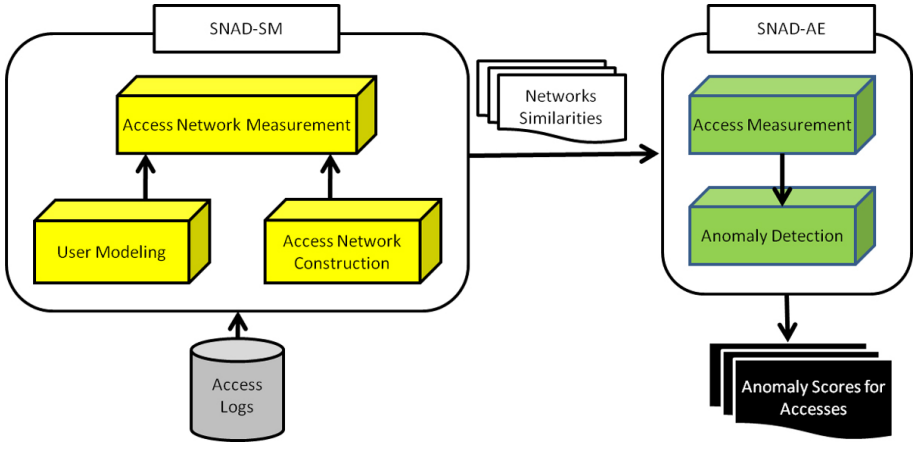


Figure 1 The specialized network anomaly detection (SNAD) framework. This figure depicts the architecture of SNAD. It consists of two components: 1) Similarity Measurement (SNAD-SM) and 2) Anomaly Evaluation (SNAD-AE). The SNAD-SM component constructs local access networks and measures the similarity of the users in such networks. The SNAD-AE component assesses the extent to which suppression of a user from a local access network changes its similarity.

Table 1 A summary of the variables applied in the SNAD framework.

Variable	Description
$S = \{s_1, s_2, \ldots, s_m\}$	The set of subjects in the CIS.
$U = \{u_1, u_2, \ldots, u_n\}$	The set of users in the CIS.
$u_j \rightarrow s_i$	An access of user u_j to subject s_i.
U_{s_i}	The set of users that access subject s_i.
Net_{s_i}	A complete graph of U_{s_i}.
SU	A binary matrix of subjects and users, the size of which is $m \times n$. If u_i accesses s_j, $SU(j, i) = 1$, else $SU(j, i) = 0$.
$\mathbf{U_i}$	A column vector of access history of u_i on all subjects. $\mathbf{U_i} = SU[:, i]$.
SU_IDF	A matrix with the same size as *SU*. Each cell value of *SU_IDF* corresponds to its inverse document frequency (IDF) transformation.
$\mathbf{B} = [1, 1, \ldots, 1]$	A vector of 1's of length *m*.
$\mathbf{IDF\_U_i}$	A column vector of access history of u_i on all subjects. $\mathbf{IDF\_U_i} = SU\_IDF[:, i]$.
PC'	A matrix created from *SU* or *SU_IDF*, the size of which is $l \times n$, where *l* is the number of selected principal components.
λ_k	The k^{th} eigenvalue
λ_{total}	The sum of the *l* eigenvalues.
$\lambda PC'$	A matrix created from *PC'*, where $\lambda PC'[k, :] = (\lambda_k/\lambda_{total}) \times PC'[k, :]$.
$\mathbf{C_i}$	A column vector of u_i on the selected *l* principal components. $\mathbf{C_i} = \lambda PC'[:, i]$.

3.1 SNAD Similarity Measurement

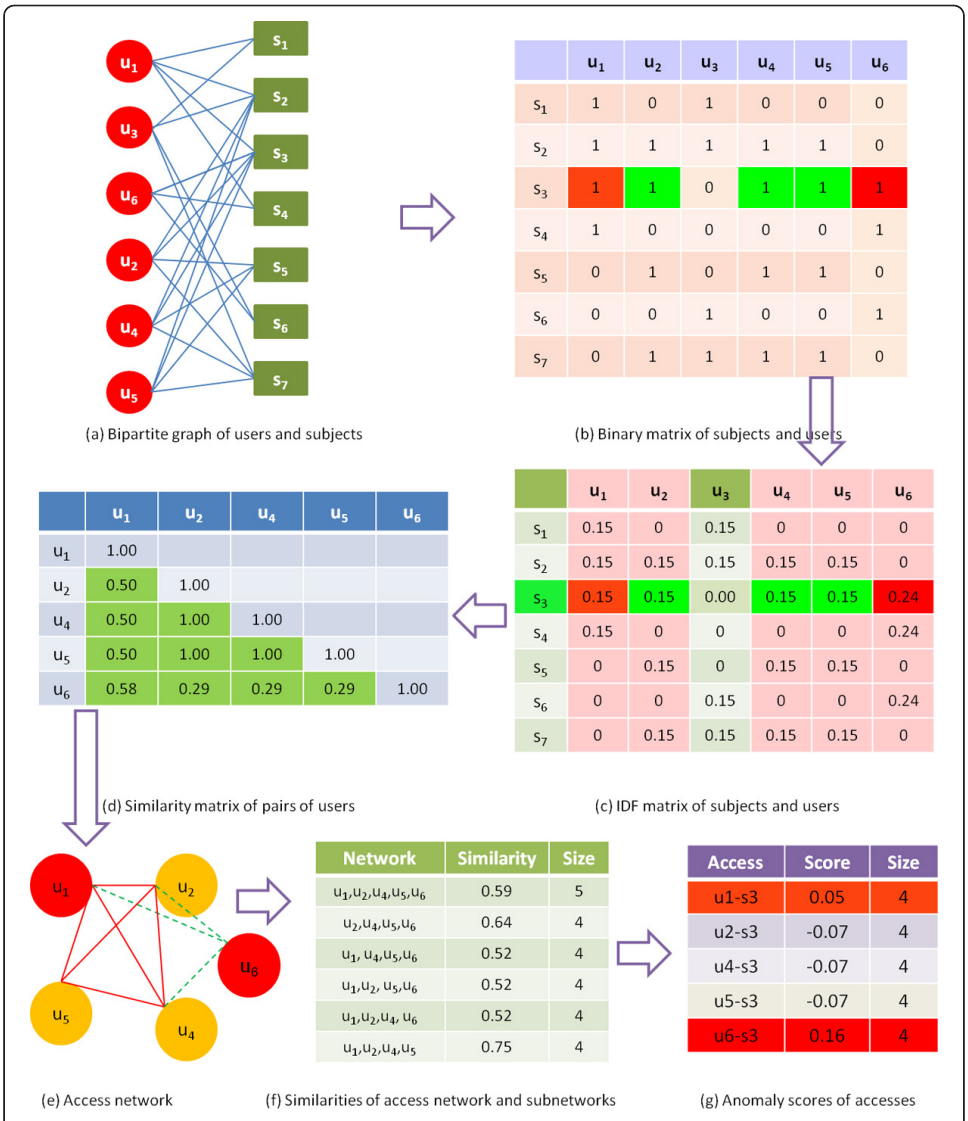
3.1.1 Access Network Construction

The SNAD-SM component transforms the CIS access logs into networks. The transformation begins by constructing a bipartite graph of the users and subjects that interact during a particular time period. Figure 2(a) depicts an example with six users and seven subjects modeled as vertices. Note, an edge represents a user accessed the subject's record.

Based on the graph, we define a local access network as follows. Let $S = \{s_1,..., s_m\}$ and $U = \{u_1, \ldots, u_n\}$ be the set of subjects and users, respectively. We define U_{s_i} as the set of users that accessed s_i in a certain time period (e.g., one day) and define Net_{s_i} as a complete graph of U_{s_i}, where the weight between a user pair is their similarity (defined below). For simplicity, we use cardinality $|\cdot|$ to represent the number of elements in a set. Figure 2(a) depicts an example where $U_{s_3} = \{u_1, u_2, u_4, u_5, u_6\}$ and the corresponding complete graph is Net_{s_3}, which is depicted in Figure 2(e).

3.1.2 User Modeling

Initially, we represent the subject-user bipartite graph as a binary matrix *SU*, as depicted in Figure 2(b). $SU(i, j) = 1$, if user u_j accesses subject s_i, and 0 otherwise. For reference, we represent u_i as the column vector of subject accesses, denoted $\mathbf{U_i}$. We use a binary matrix because in certain CIS, multiple components of the system record an access simultaneously. As a result, the number of user accesses to a particular subject may be artificially inflated. Moreover, the degree to which inflation exists is non-uniform across the components of the system. For instance, in an EHR, a user may access different components of a patient's medical record, such as laboratory reports, progress notes, or consult requests. These components may be accessed at different rates by different users and not all patients may have each type of note in their record.



	u_1	u_2	u_3	u_4	u_5	u_6
s_1	1	0	1	0	0	0
s_2	1	1	1	1	1	0
s_3	1	1	0	1	1	1
s_4	1	0	0	0	0	1
s_5	0	1	0	1	1	0
s_6	0	0	1	0	0	1
s_7	0	1	1	1	1	0

	u_1	u_2	u_3	u_4	u_5	u_6
s_1	0.15	0	0.15	0	0	0
s_2	0.15	0.15	0.15	0.15	0.15	0
s_3	0.15	0.15	0.00	0.15	0.15	0.24
s_4	0.15	0	0	0	0	0.24
s_5	0	0.15	0	0.15	0.15	0
s_6	0	0	0.15	0	0	0.24
s_7	0	0.15	0.15	0.15	0.15	0

	u_1	u_2	u_4	u_5	u_6
u_1	1.00				
u_2	0.50	1.00			
u_4	0.50	1.00	1.00		
u_5	0.50	1.00	1.00	1.00	
u_6	0.58	0.29	0.29	0.29	1.00

Network	Similarity	Size
u_1,u_2,u_4,u_5,u_6	0.59	5
u_2,u_4,u_5,u_6	0.64	4
u_1,u_4,u_5,u_6	0.52	4
u_1,u_2,u_5,u_6	0.52	4
u_1,u_2,u_4,u_6	0.52	4
u_1,u_2,u_4,u_5	0.75	4

Access	Score	Size
u1-s3	0.05	4
u2-s3	-0.07	4
u4-s3	-0.07	4
u5-s3	-0.07	4
u6-s3	0.16	4

Figure 2 An illustrative example for the SNAD model. This figure illustrates the workflow of the SNAD model. First, a bipartite graph of user-subject accesses is constructed (part a), which is subsequently represented as a binary matrix (part b). This matrix is then transformed using "inverse document frequency" (part c), from which the similarities between pairs of users are computed (part d) and from which a local access network associated with a subject s_3 is built(part e). The similarities of the local access network and its five subnetworks are calculated (part f) and finally, the scores for each access associated with subject s_3 are assessed.

As a consequence, models based on raw values could be biased toward users that work with the system in a more heavy manner or by subjects that have more information recorded in their records.

Prior research in social network analysis (e.g., [43]) suggests it is important to represent the affinity that a user has toward a particular subject when assessing the similarity of a group. There are various aspects of a user's relationship to subjects that could be leveraged for measuring similarity. To mitigate bias and develop a generic approach, we focus our attention on the number of subjects a user accessed. Using this feature, we employ the inverse document frequency (IDF) model, popularized by information retrieval systems and shown to be effective for weighting the affinity of individuals to

subjects in friendship networks [43]. IDF captures the affinity of a user to a subject relative to all subjects in the system. As such, the IDF transformation is defined as:

$$IDF(u_i) = log\frac{|S|}{1 + \mathbf{B} \cdot \mathbf{U_i}} \tag{1}$$

where $\mathbf{B} = [1, 1, \ldots, 1]$ is a vector of length m. Figure 2(c) provides an example of this transformation. After IDF transformation, the binary matrix SU is converted into IDF matrix SU_IDF. We use column vector $\mathbf{IDF\_U_i} = SU\_IDF\,[:, i]$ to represent u_i.

Relationships, or similarity, between pairs of users can be mined from their access vectors. The cosine similarity [44] is a measure that has been particularly successfully applied in various domains to measure the similarity of objects in a vector form. Following this reasoning, we compute the similarity of users u_i, u_j via the cosine of their IDF-transformed vectors:

$$Sim\,(u_i, u_j) = \frac{\mathbf{IDF\_U_i} \cdot \mathbf{IDF\_U_j}}{||\mathbf{IDF\_U_i}|| \times ||\mathbf{IDF\_U_j}||} \tag{2}$$

Figure 2(d) provides an example of user pair similarities.

3.1.3 Access Network Measurement

We hypothesize that if an insider wanders into a local network, then the network's similarity will decrease. To investigate this hypothesis we need to develop an appropriate similarity measure for an access network.

Different subjects have distinct local access networks. In order to compare the similarities of these across local networks, we define the similarity of an access network as the average similarity of all user pairs:

$$SIM(Net_{s_k}) = \frac{\forall u_i \neq u_j \in U_{s_k} \forall u_j \sum Sim\,(u_i, u_j)}{\frac{|U_{s_k}| \times (|U_{s_k}| - 1)}{2}} \tag{3}$$

where $|U_{s_k}|$ is the number of users in Net_{s_k}. When this value is high, the users are *close* to each other, such that they have a strong collaborative relationship with respect to subject s_k.

3.1.4 Access Measurement and Anomaly Detection

SNAD-SM provides a measure of similarity for an access network; however, to leverage such measures for anomaly detection, we need a formal approach to determine when a particular access is anomalous in the access network. In this regard, SNAD-AE evaluates each user's access in a network by calculating how the similarity of the network changes after the suppression of the user. SNAD-AE assumes that intruding accesses will lower the similarity of a network at a greater rate than a typical access.

We evaluate each access of a local network through similarity changes of the access network and its subnetworks as follows:

$$Score\,(u_j \rightarrow s_i) = SIM(Network_{s_{ij}}) - SIM(Network_{s_i}) \tag{4}$$

where $Network_{s_{ij}}$ is the network without user u_j. The larger the value for $Score(u_j \rightarrow s_i)$, the greater the likelihood that access $u_j \rightarrow s_i$ is an anomalous access. Notice, this approach assumes that scores are centered. We empirically demonstrate that this is the case for our datasets in Appendix A.

As an example, Net_{s_3} in Figure 2(e) consists of five users who accessed s_3. The process of access score calculation for every access associated with s_3 is depicted in Figure 3. In Net_{s_3}, the expectation is that if $u_j \rightarrow s_3$ is an intrusion, $Score(u_j \rightarrow s_3)$ will be larger than the subnetwork sans a typical user. Similarities of access networks and their subnetworks are depicted in Figure 2(f).

Figure 2(g) reports the scores (Equation 4) for all accesses involved with subject s_3. These scores were calculated from access network Net_{s_3}. The larger the score, the greater the probability the access is an intrusion. For the five accesses associated with network Net_{s_3}, u_1 and u_6 have scores which are larger than u_2, u_4 and u_5; 0.05 and 0.16, respectively. If we rank the scores and claim the highest as an anomaly, $u_6 \rightarrow s_3$ will be implicated by SNAD. Turning back to the *SU* matrix, it can be seen that u_2, u_4, and u_5 access common subjects, whereas u_6 only has s_3 in common. Except for s_2, s_3 and s_6, u_1 has no common subjects with u_2, u_4 and u_5.

3.2 Spectral Anomaly Detection Model

Though SNAD may appear to be a simplistic model, we find it is more appropriate for access-level insider threat detection in CIS than more sophisticated competitors. As evidence, we will compare SNAD to a well-regarded competitor, spectral anomaly detection [42], which we summarize here.

This model first decomposes the binary matrix *SU* (or IDF matrix *IDF_SU*) into its principle components. The results of the decomposition is a principal component matrix *PC*, where each row represents a principal component, and each column is a vector of real values that represent the affinity of a user to each principal component. Each vector is associated with an eigenvalue, which is proportional to the amount of variance each vector captures in the system.

The spectral anomaly detection method then computes how user relationships to subjects are related to the most significant principal components. To accomplish this computation, the model selects the l principal components with the highest scoring eigenvalues λ_k, which leads to a new principal component matrix *PC'*. The decomposed matrix *PC'* is then transformed into matrix $\lambda PC'$, where $\lambda PC'[k, :] = (\lambda_k/\lambda_{total}) \times PC'[k, :]$. We use a column vector $\mathbf{C_i}$ to model u_i on the selected l principal

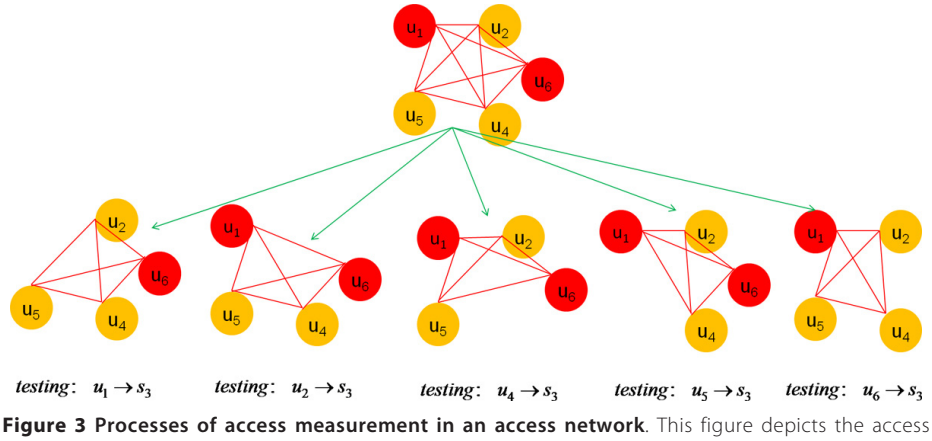


Figure 3 Processes of access measurement in an access network. This figure depicts the access measurement for the network depicted in Figure 2(e). The top of the figure corresponds to the local access network associated with subject s_3, whereas the bottom of the figure illustrates its five subnetworks.

components, where $\mathbf{C_i} = \lambda PC'[:, i]$. The model then computes the similarity between all pairs of users, where similarity is defined as:

$$Sim'(u_i, u_j) = \frac{\mathbf{C_i} \cdot \mathbf{C_j}}{||\mathbf{C_i}|| \times ||\mathbf{C_j}||} \tag{5}$$

Then, the average similarity of an access network is defined as:

$$SIM'(Net_{s_k}) = \frac{\forall u_i \neq u_j \in U_{s_k} \forall u_j \sum Sim'(u_i, u_j)}{\frac{|U_{s_k}| \times (|U_{s_k}| - 1)}{2}} \tag{6}$$

Finally, and similar to SNAD, the spectral model computes an access score for each user by measuring the change of distance in the access network after suppressing the user. As an example, access $u_6 \rightarrow u_3$ is scored by the spectral model as $Score(u_6 \rightarrow s_3) = SIM'(Net_{s_{3,6}}) - SIM'(Net_{s_3})$.

We apply the spectral anomaly detection model on both the pre- and post-IDF transformed *SU* matrices and refer to these models as Spectral-Binary and Spectral-IDF, respectively.

4 Experiments and Results

4.1 Datasets

For evaluation, we utilize datasets from CIS in two distinct domains: healthcare and online wikis. The first dataset corresponds to the real access logs of the Vanderbilt University Medical Center (VUMC) EHR system. This system has been in application for over a decade and is well-ingrained in healthcare operations. The logs document when an authenticated VUMC employee accessed a patient's record. We refer to this set of access transactions as the EHR dataset. This dataset contains 1,327,500 accesses, 6,015 users, and 130,457 patients and were collected over 30 weeks during the year 2006. Further details and analysis of this dataset can be found in [36].

The second dataset corresponds to the wiki-talk logs from Wikipedia [45]. In this dataset, each registered user has a "talk" page that they and other users can edit in order to communicate and discuss updates to various articles on Wikipedia. The logs cover the years 2002 to 2008. It contains 6,482,780 revisions, 2,394,385 users and 55,200 articles. For this study, we analyze the revisions documented over 50 weeks during the year 2007. We refer this set of transactions as the Wiki dataset.

In the EHR dataset, we refer to patient records as subjects, and user views of the records as accesses. Similarly, in the Wiki dataset, we refer to articles as subjects and user revisions as accesses. Summary statistics regarding the two datasets are provided in Table 2.

4.2 Detection Models

In this work, we evaluate four anomaly detection models. The first two are variants of SNAD, while the second two are variants of the spectral anomaly detection model.

Table 2 Summary statistics of EHR and Wiki datasets.

Dataset	Weeks	Users/week	Subjects/week	Accesses/week
EHR	30	2,281	13,148	44,250
Wiki	50	3,952	240	28,186

Specifically, we consider if the IDF transformation influences anomaly detection performance. As such, we evaluate both models using the raw binary matrix (*SU*) and the IDF-transormed matrix (*IDF_SU*). We refer to these models as SNAD-Binary, SNAD-IDF, Spectral-Binary, and Spectral-IDF.

4.3 Experimental Design

The datasets do not document which (if any) accesses were intrusions. As such, to conduct a controlled evaluation, we injected simulated actions into the logs (i.e., changed 0's to 1's in the *SU* access matrix).

For this study, we use three scenarios to assess the intrusion detection rate under various settings:

1. **Accesses Per User:** We select a user at random, inject between 1 to 100 new subject accesses, and execute the detection model. This process is repeated 15 times per week.
2. **User Per Access Load:** We investigate how the number of intruding users influences the detection rate. We select a set of users to inject three intruding accesses into. We perform this analysis over the range of 2 to 20 intruding users.
3. **Diverse Setting:** We emulate a more realistic environment by allowing for a variety of simultaneous intruding users and actions. Specifically, we inject a set of random subject accesses, between 1 and 100, into a random set of users, between 1 and 20.

Each of these scenarios is simulated on a per week basis.

We measure the performance of the models using the receiver operating characteristic (ROC) curve [46]. This is a characterization of the true positive rate versus the false positive rate for a binary classifier as its discrimination threshold is varied. The area under the ROC curve (AUC) reflects the relationship between sensitivity and specificity for a given test. Sensitivity (also called recall rate in some fields) measures the proportion of actual positives which are correctly identified as such (e.g., the percentage of simulated accesses which are correctly identified as anomaly). Specificity measures the proportion of negatives which are correctly identified (e.g., the percentage of real access which are correctly identified as normality). A higher AUC indicates better performance. In the first two simulation settings, we report on the average AUC per simulation configuration.

4.4 Results and Analysis

4.4.1 Complexity and runtime Analysis

We analyzed the complexity of SNAD from a theoretical and an empirical runtime analysis.

From a theoretical perspective, the complexity of SNAD is defined as follows. Let m, and n be the number of subjects and users in the CIS. For every subject, SNAD constructs a local network, which is accomplished in $O(m)$ time. For each local network, SNAD computes the similarity for all pairs of users, which consumes $O(logm \times (logn)^2)$ time. It is anticipated that the number of subjects accessed by a user is significantly

smaller than m and thus the time to compute the similarity of a pair of users is $O(logm)$. Since there are $(logn)^2$ pairs, the complexity of SNAD is $O(m \times logm \times (logn)^2)$.

The average runtime of the four models for EHR and Wiki datasets are depicted in Table 3. The runtime is the average seconds per scored access. And it is calculated over all experiments and settings in this paper. We would like to highlight three points based on the table. First, all models applied to the Wiki dataset have smaller runtime than the EHR dataset. This is because, for the same number of accesses, Wiki has a smaller number of local networks than EHR. In other words, models have to be executed over more local networks in EHR than in Wiki. Second, the SNAD models have smaller runtime than the spectral models. This is because spectral analysis on the local network is more time consuming than similarity calculation of the local network. Third, the SNAD and spectral models have nearly the same runtime when executed over their binary or IDF matrices. This is because the time required for transforming the binary matrix to the IDF matrix is negligible.

4.4.2 Similarity of Real Access Networks

Figure 4 depicts the distributions of access network similarity in the EHR and Wiki datasets for an arbitrary week (similar results were observed with other weeks). Notably, these environments capture different social phenomena. In the EHR dataset, for instance, the majority of access networks are small in size. And, as shown in the upper plot of Figure 4, the similarity approaches zero as the network size grows. This suggests that when a user is suppressed from a large network in the healthcare setting, the average similarity has little change. The main driving factor of this phenomenon is that large access networks in the EHR system tend to be varied in the user composition (i.e., complex dynamic teams of care providers).

In contrast, the lower plot of Figure 4 indicates that Wiki users in large access networks are relatively similar. This implies that when an intruder joins a large sized access network in the wiki world, the average similarity of the new access network could greatly decrease.

These observations suggest that SNAD may not be appropriate for a large sized access network whose average similarity is small. In this case, the removal of the intruder from an access network is expected to have little influence on the average similarity.

4.4.3 Accesses Per User

In the first experiment, we investigate how the number of intrusions committed by a single user influences detection.

Figures 5 and 6 depict the average AUC and one standard deviation of the detection models as a function of the number of simulated intrusions. It can be seen that SNAD-IDF has equal or larger AUC than both spectral models and SNAD-Binary. We note there are only two points at which the spectral models and SNAD-IDF were equivalent (3 intruding accesses in the EHR dataset and 5 intruding accesses in the Wiki dataset). Additionally, unlike the spectral models, the AUC of both SNAD variants tends to increase with the number of accesses. When the insider has only one

Table 3 Average runtime (second per access) of the four models for the datasets.

Dataset	SNAD-IDF	SNAD-Binary	Spectral-IDF	Spectral-Binary
EHR	1.04	**1.02**	1.16	1.14
Wiki	0.44	**0.43**	0.50	0.49

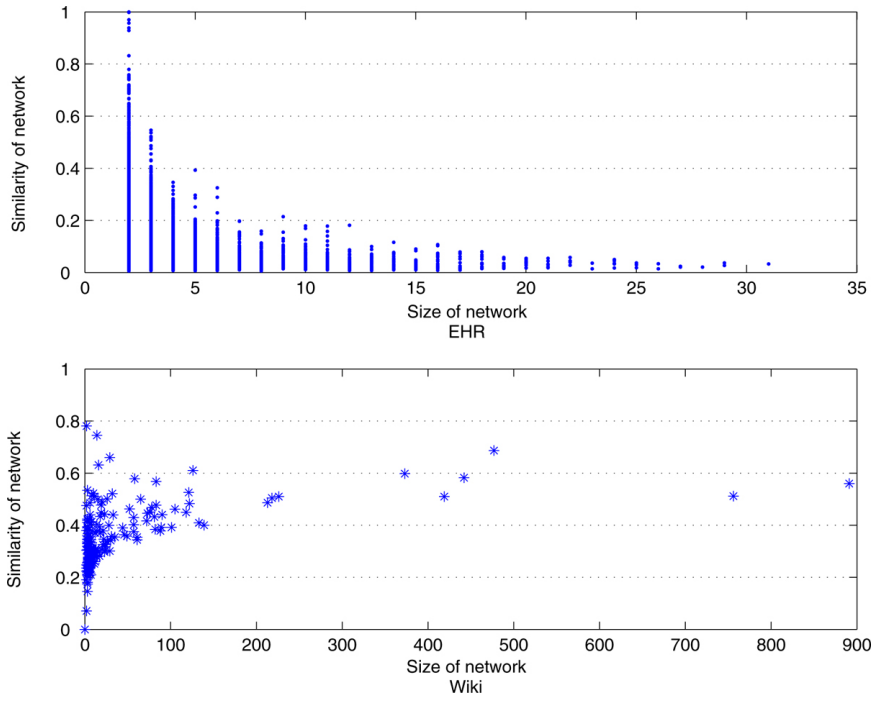


Figure 4 The distribution of access network similarity as a function of access network size. This figure shows the distributions of similarity of access networks for an arbitrary week of the study. In both plots, the x-axis corresponds to the number of in users in (or size of) the networks. The y-axis corresponds to the similarity of the access networks. The upper and lower plots are derived from the EHR and Wiki datasets, respectively.

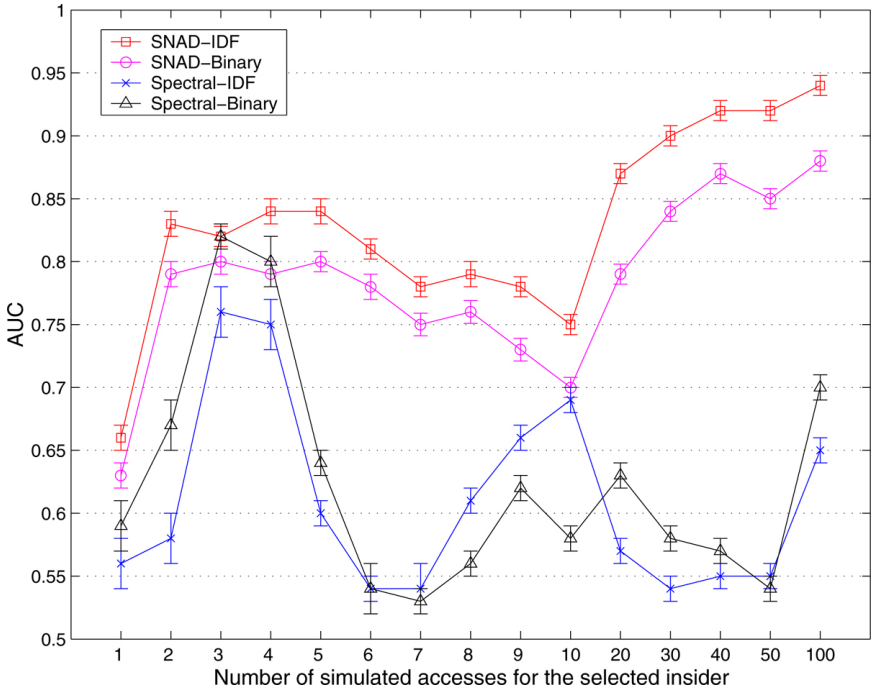


Figure 5 Average AUC of the detection models on one simulated intrusion in the EHR dataset. This figure summarizes the AUC (+/- one standard deviation) of the four detection models on varying quantities of simulated intrusions, where one user is intruding per simulation, in the EHR dataset. The x-axis corresponds to the number of simulated accesses for the selected insider, which ranges from 1 to 100. The y-axis is the average AUC and standard deviation of the models.

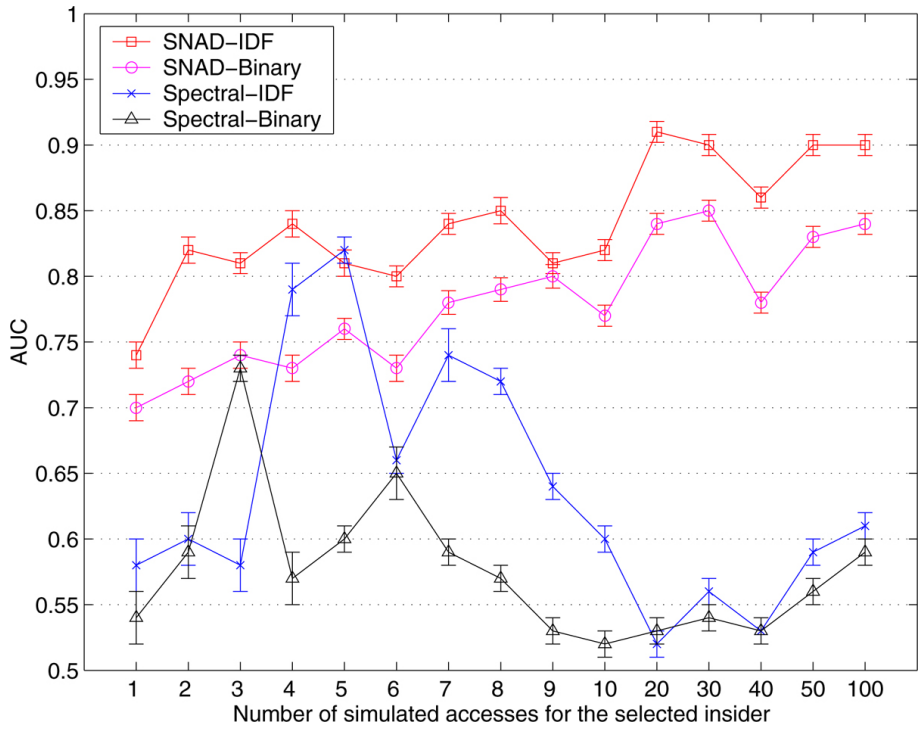


Figure 6 Average AUC of the detection models on one simulated intrusion in the Wiki dataset. This figure summarizes the AUC (+/- one standard deviation) of the four detection models on varying quantities of simulated intrusions, where one user is intruding per simulation, in the Wiki dataset. The x-axis corresponds to the number of simulated accesses for the selected insider, which ranges from 1 to 100. The y-axis is the average AUC and standard deviation of the models.

simulated access, SNAD-IDF's average AUC is nearly 0.65 compared to 0.59 of its nearest spectral decomposition competitor, Spectral-Binary in EHR dataset. When the number of simulated accesses is 30, SNAD-IDF's AUC reaches 0.9 in both datasets.

Also, we can see that SNAD-IDF is more effective than SNAD-Binary. This indicates that the affinity of a user to a subject relative to all subjects is an important factor in detecting strange insider actions. However, for spectral models, neither Spectral-Binary nor Spectral-IDF is a clear winner in terms of performance with respect to both datasets.

4.4.4 Users Per Access Load

In this experiment, we investigate how the number of intruding insiders influences detection. We fix the number of simulated accesses to 3. The results are depicted in Figures 7 and 8, which demonstrate the AUC for all models increase with the number of accesses for the EHR dataset, but only the AUC of the SNAD variants increase in the Wiki dataset. Nonetheless, both SNAD variants greatly outperform the spectral models at all evaluation points. We suspect this is because the insiders with simulated accesses substantially amend the local access network, but have little influence on the global network, which the spectral approach depends upon. The implication is that indirect relations, which are critical to the discovery of intruding user behavior *in general* [39,42], may be less important than the direct relations in the detection of *specific intruding accesses*. However, we recognize that a more detailed investigation, perhaps with more datasets, is necessary before such a conjecture can be confirmed.

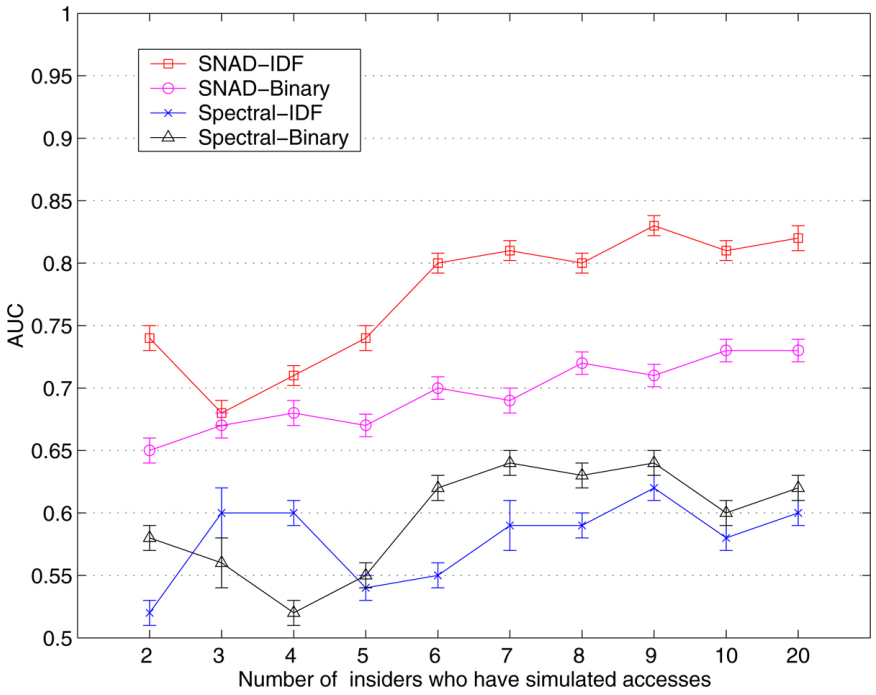


Figure 7 Average AUC of the detection models for a varying number of simulated insiders in the EHR dataset. This figure summarizes the AUC (+/- one standard deviation) of the four detection models on varying quantities of simulated intrusions, where a varying number of users are intruding per simulation, in the EHR dataset. Each simulated intruder issues three accesses. The x-axis corresponds to the number of intruding users, which ranges from 2 to 20. The y-axis is the average AUC and standard deviation of the models.

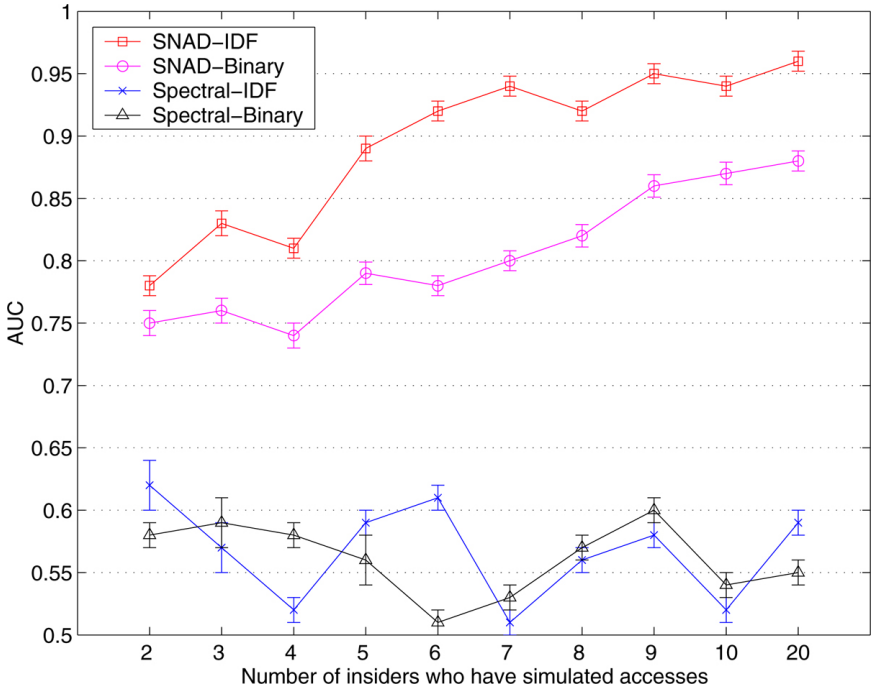


Figure 8 Average AUC of the detection models for a varying number of simulated insiders in the Wiki dataset. This figure summarizes the AUC (+/- one standard deviation) of the four detection models on varying quantities of simulated intrusions, where a varying number of users are intruding per simulation, in the Wiki dataset. Each simulated intruder issues three accesses. The x-axis corresponds to the number of intruding users, which ranges from 2 to 20. The y-axis is the average AUC and standard deviation of the models.

Figure 7 and 8 also demonstrate that the AUC increases with the number of intruding insiders. Here it can be observed that SNAD-IDF exhibits an AUC that is 20-30% higher, on average, than the spectral models, and 5-10% higher than SNAD-Binary.

4.4.5 Diverse Insider Setting

In the third experiment, we injected a random number of accesses into a random set of user vectors. Figures 9 and 10 provide a comparison of the ROC curves of the detection models for both datasets. It is apparent that SNAD has better performance than both of the spectral models and SNAD-Binary at every operating point.

Table 4 summarizes the average AUC scores (+/- one standard deviation) of the anomaly detection models in this setting. The results indicate that SNAD-IDF achieves the highest AUC, 0.83 and 0.91 in the EHR and Wiki datasets, respectively. This translates into AUC scores that are 10% - 20% higher, on average, than the spectral models.

5 Discussion

SNAD is an unsupervised learning model, based on social network analysis, for detecting anomalous accesses in CIS. SNAD uses a "local" view of a social network in that, for each subject, it considers only the direct relationships of the users that access the subject. This is a simpler approach than methods based on a "global" perspective, such as neighborhood formation, community detection, and spectral decomposition models. Though the latter models may invoke more heavy computational machinery, our empirical investigations with real world data from healthcare and an online wiki domains suggest that they are less effective for detecting specific behavioral deviations.

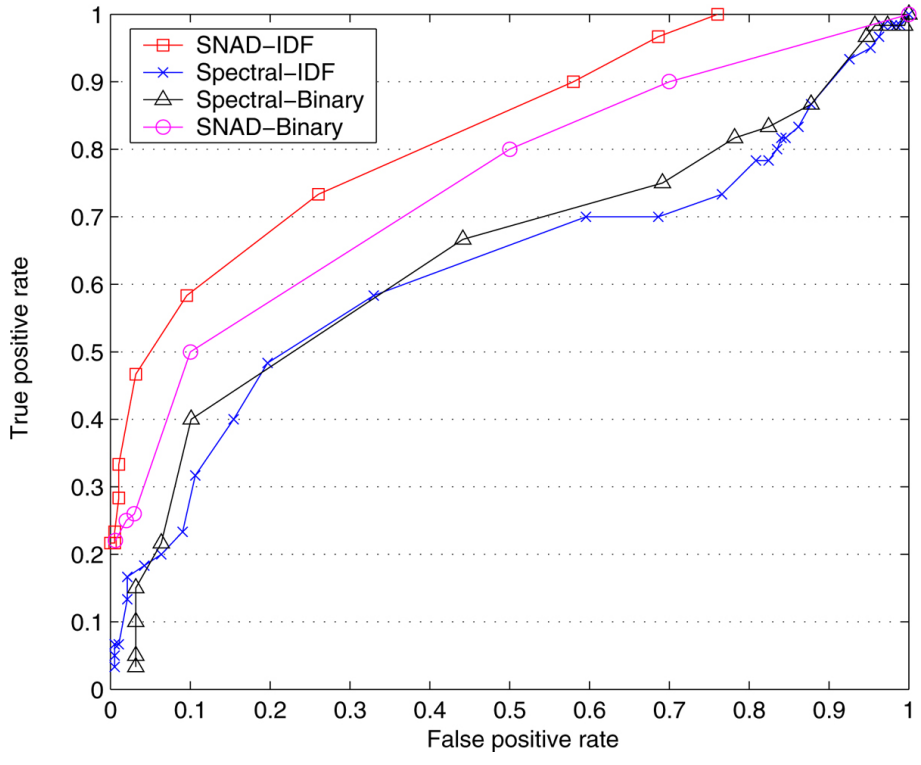


Figure 9 ROC curves for the detection models in a diverse setting for the EHR dataset. This figure summarizes the ROC scores for the four detection models, as applied to the EHR dataset, when the number of intruders and the quantity of intruding accesses are randomly generated.

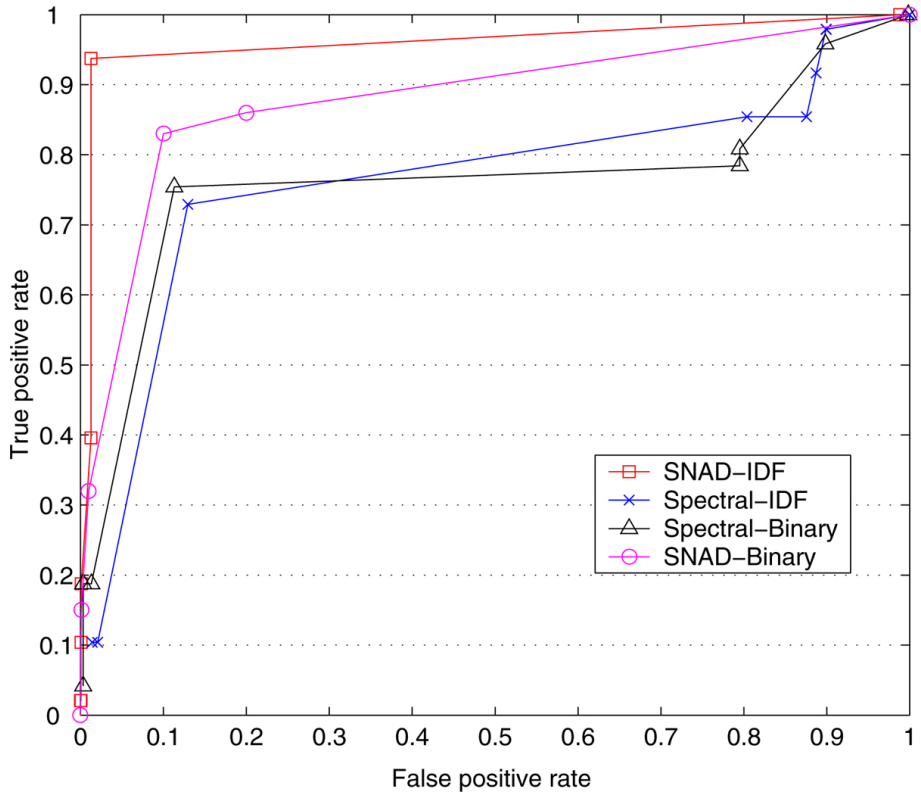


Figure 10 ROC curves for the detection models in a diverse setting for the Wiki dataset. This figure summarizes the ROC scores for the four detection models, as applied to the Wiki dataset, when the number of intruders and the quantity of intruding accesses are randomly generated.

Our results suggest the local view is significantly more adept at determining when a user has wandered into a network of users associating with a particular subject.

Despite the development of a more robust CIS access anomaly detection model, there are several limitations of the study that we wish to point out to serve as a guidebook for future research on this topic.

First, the results of the experiments outline the scope and context in which SNAD is applicable. SNAD appears to be suited to environments where the access networks exhibit high similarity. Additionally, the performance of SNAD seems to increase as a function of the number of illicit insiders and the quantity of suspicious accesses they execute. However, SNAD may not be suited for large networks in environments where the users are collaborating in physical settings, such as healthcare. Our results suggest that in these environments, the similarity of the users in local networks is relatively low. We suspect that this is due to the ad hoc team-based nature of dynamic organizations that function in the real world. For instance, in a hospital setting, there often over 100 job titles or specialities and large teams are often constructed based on who

Table 4 AUC scores (+/- one standard deviation) of the detection models on the datasets.

Dataset	SNAD-IDF	SNAD-Binary	Spectral-IDF	Spectral-Binary
EHR	**0.83±0.03**	0.79±0.03	0.69±0.06	0.74±0.05
Wiki	**0.91±0.02**	0.85±0.02	0.76±0.04	0.64±0.04

from which speciality is available to work. In contrast, in the online Wikipedia setting, it appears that large teams are not limited by speciality, but who is interested in collaborating over a subject. We believe further investigation is necessary before any concrete conclusions can be drawn about such behavior. In particular, users may have different roles, such as administrators and gateways, which may influence their relationships to other users, and thus some of accesses associated with these users will be falsely detected as anomalous by SNAD. In this paper, we do not consider the impact of creating this type of noise to our SNAD models, but believe it is a fruitful direction for future research.

Additionally, SNAD accounts for the relationship between users and subjects, but neglects the semantics of the relation. For instance, SNAD does not model the intention of a user while executing an action, which serves as the foundation of the recent model proposed in [27]. In a CIS, the system is often mission-oriented, such that the semantics of the users and subjects are informative. Consider, in an EHR system, patients are assigned diagnoses and procedures, while users are affiliated with various departments and assigned certain roles within a healthcare organization. Rather than treat each user and patient equally, we believe that detection sensitivity could be improved by integrating such information into the network modeling process.

Finally, SNAD was evaluated on only one type of attack; i.e., when a user issues an intruding access randomly. Yet, in real systems, there may be many types of attacks [27], some which are more complex and require different simulation methods.

6 Conclusion

In this paper we proposed a *specialized network anomaly detection model*, or SNAD, to discover anomalous actions in collaborative information systems (CIS). SNAD differs from previous insider threat detection techniques in that it is engineered to assess specific event-related actions as opposed to global patterns. The foundation of SNAD is an efficient unsupervised learning method, such that it can be deployed in real systems. We evaluated our technique against several competitors, based on spectral decomposition, with real EHR access and Wiki "talk" logs. The empirical results demonstrate that SNAD exhibits better performance than its competitors in almost every assessed scenario. Nonetheless, there are limitations of SNAD, such as a requirement for highly similar behavior of all users accessing a specific subject, which may be difficult to realize in large networks in the physical world. We believe that extending the model to account for the semantics of the users and subjects will improve its effectiveness in such cases and anticipate extending SNAD in this direction.

Endnotes

A Validity of Similarity Measure

The anomaly detection measure proposed for SNAD assumes that the "similarity" scores are approximately distributed around a well-centered mean. To verify if this is the case, we modeled the distribution of access scores for an arbitrary week of the EHR and the Wiki datasets. The distributions of scores are depicted in Figure 11 and Figure 12. From the figures, we observed that the distribution of access scores in EHR and Wiki datasets is close to a Laplace distribution, which is centered.

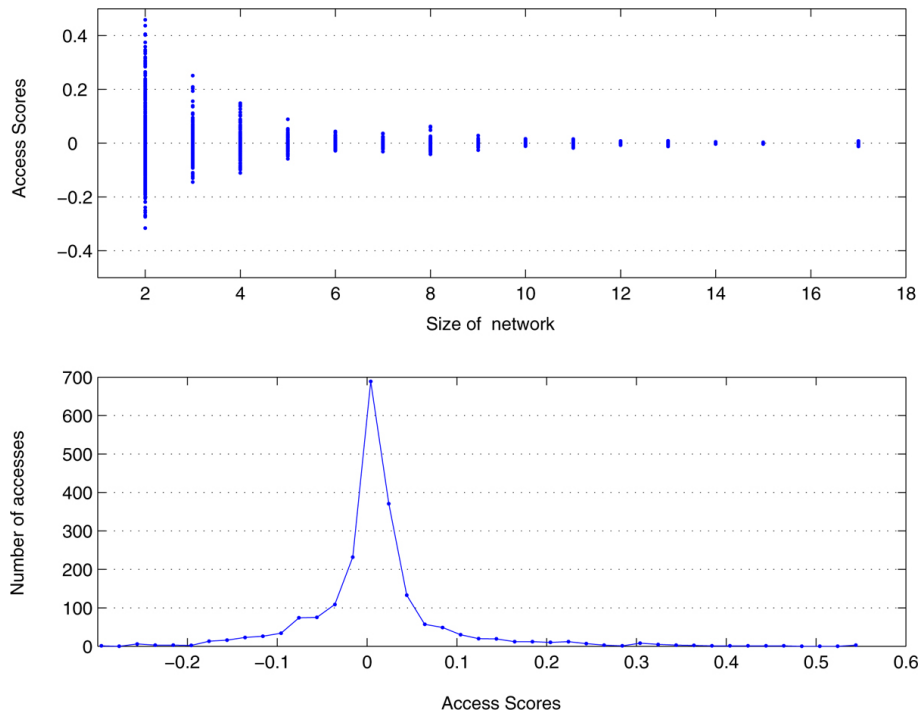


Figure 11 Density function of access score in an arbitrary week of the EHR dataset. In this figure, the upper plot corresponds to the access scores on access networks of various size in the EHR dataset. The lower plot corresponds to the distribution of the access scores.

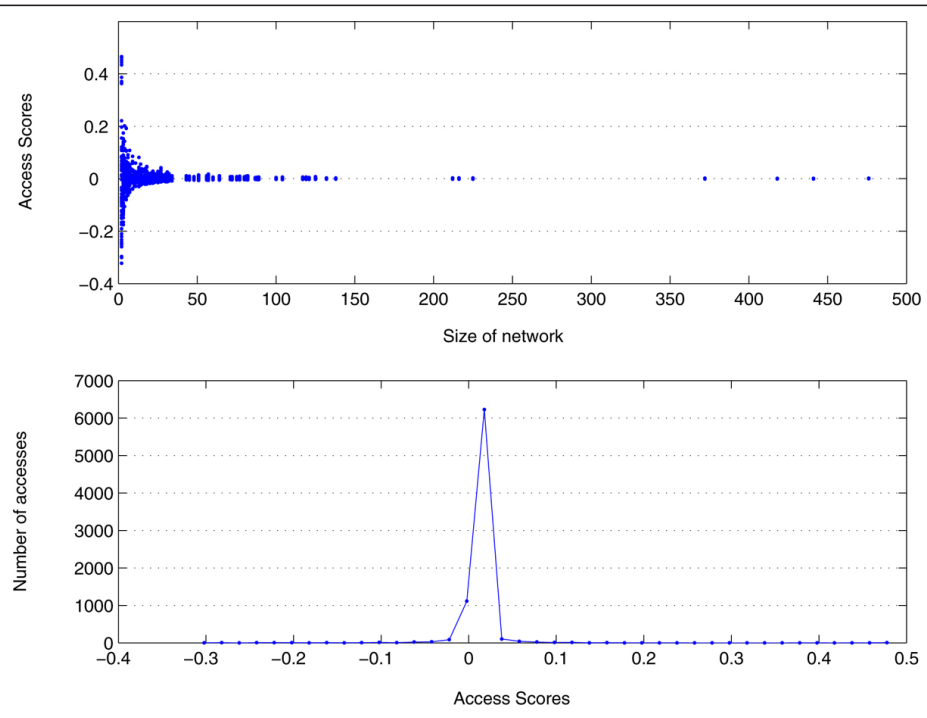


Figure 12 Density function of access score in an arbitrary week of the Wiki dataset. In this figure, the upper plot corresponds to the access scores on access networks of various size in the Wiki dataset. The lower plot corresponds to the distribution of the access scores.

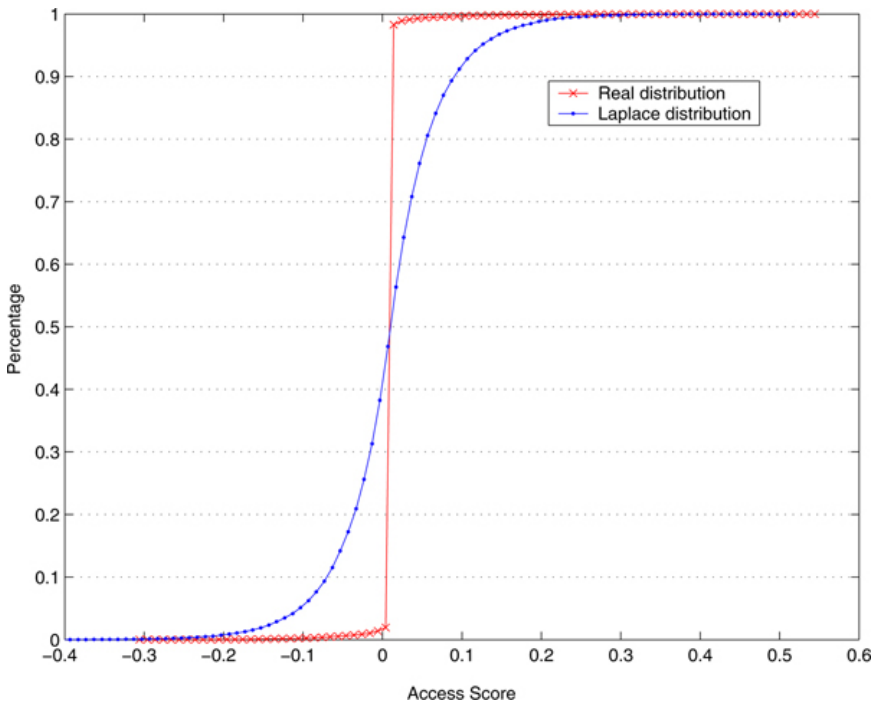


Figure 13 Cumulative distribution function of access scores in an arbitrary week of the EHR. This figure reports the frequency of access scores, as a cumulative distribution function for one week of the EHR dataset. In this figure, the x-axis corresponds to the access scores. The y-axis corresponds the fraction of the scores with value less than a specific score X_i.

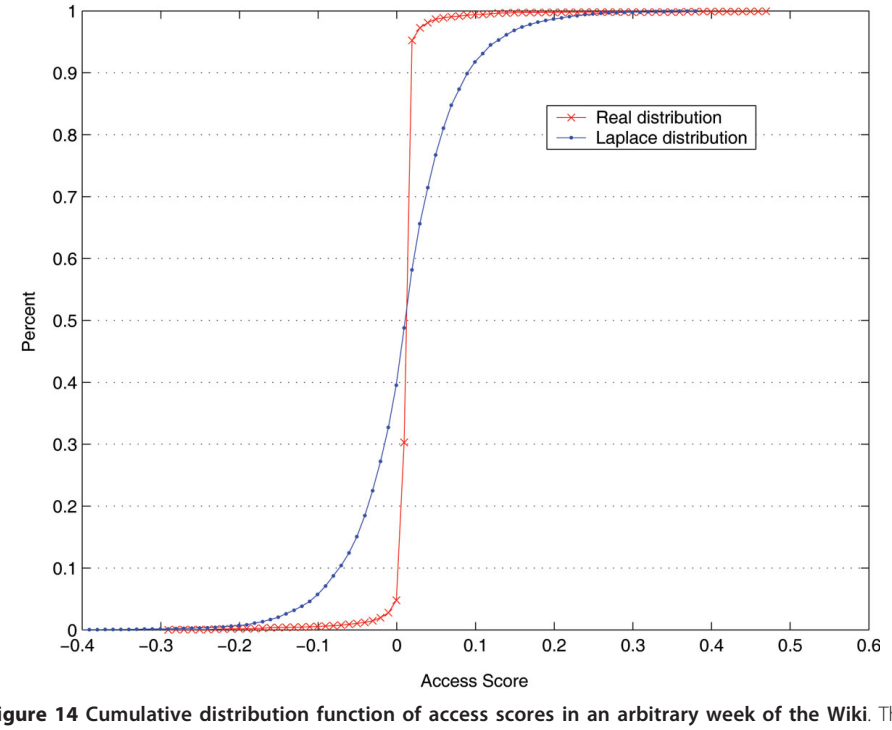


Figure 14 Cumulative distribution function of access scores in an arbitrary week of the Wiki. This figure reports the frequency of access scores, as a cumulative distribution function for one week of the Wiki dataset. In this figure, the x-axis corresponds to the access scores. The y-axis corresponds the fraction of the scores with value less than a specific score X_i.

In order to assess the relationship between the distributions of access scores and a Laplace, we created a random dataset from an ideal Laplace distribution. Our simulation has two parameters: location a and scale b. The location a is set to 0, and the scale b is set to 0.05. In the simulation, we generated the cumulative distribution functions of the ideal Laplace and access scores, which are in Figure 13 and Figure 14.

We calculated the correlation between the distributions via a Pearson coefficient, ρ. The ρ value is 0.866 in the EHR dataset, and 1 in the Wiki dataset, which suggests that the real access score distributions are highly correlated with Laplace. The relation between the real and Laplace distributions can be observed in Figure 13 and Figure 14.

List of Abbreviations

AUC: Area Under the ROC Curve; CADS: Community-based Anomaly Detection System; CIS: Collaborative Information Systems; EHRs: Electronic Health Records; IDF: Inverse Document Frequency; RBAC: Role-Based Access Control; ROC: Receiver Operating Characteristic; SitBAC: Situation-Based Access Control; SNAD: Specialized Network Anomaly Detection; SNAD-SM: SNAD-Similarity Measurement; SNAD-AE: SNAD-Anomaly Evaluation; VUMC: Vanderbilt University Medical Center.

Acknowledgements

The authors thank D. Giuse for supplying the access logs, studied in this paper. The authors also thank E. Boczko, J. Denny, C. Gunter, D. Liebovitz, as well as the members of the Vanderbilt Health Information Privacy Laboratory for discussions. The authors also extend their gratitude for the excellent comments and suggestions provided by the anonymous reviewers of [28] and earlier versions of this paper. This research was sponsored, in part, by grants CCF-0424422 and CNS-0964063 from the National Science Foundation and R01LM010207 from the National Institutes of Health.

Author details

[1]Department of Biomedical Informatics, School of Medicine, Vanderbilt University, Nashville, TN, 37203, USA
[2]Department of Electrical Engineering and Computer Science, School of Engineering, Vanderbilt University, Nashville, TN, 37203, USA

Authors' contributions

YC performed the method design, analysis of the models, design, evaluation and interpretation of the experiments, and writing of the manuscript. SN performed extraction and formatting of the data, assisting in model design, and writing of the manuscript. WZ assisted in the design of the models and writing of the manuscript. BM performed the method design, design and interpretation of the experiments, and writing of the manuscript. All authors read and approved the final manuscript.

Authors' Information

YC received the PhD degree in computer science from the Chinese Academy of Science. He is currently a postdoctoral research fellow at the Department of Biomedical Informatics at Vanderbilt University. His main research interests are related to data and application security and privacy. His research focuses on the construction and evaluation of data privacy models for personal information that is collected, stored, and shared in large complex systems.

SN received the MS degree in Software Engineering from Southern Adventist Univeristy. He is currently a software engineer at the Department of Biomedical Informatics at Vanderbilt University. He has extensive experience in health care transactional systems and customer facing applications systems design and implementation. He has contributed to various open source projects and is the primary author of open source projects: Jx framework and stored procedure generator for PostgreSQL. In 2008, he earned the Business Objects Award for excellence from Emdeon Inc.

WZ is currently a PhD student in computer science at Vanderbilt university. He obtained his BS and MS from the University of Science and Technology of China. His main research interests are data mining and information security. He is particularly interested in methodologies to detect and prevent privilege abuse using automated learning methods in the context of large biomedical information systems.

BM received the PhD degree in computer science from Carnegie Mellon University. He is currently an Associate Professor of Biomedical Informatics and an Associate Professor of Computer Science at Vanderbilt University. His current research interests include privacy in health and genetic databases, auditing and surveillance in electronic medical record systems, and the design and execution of model-based clinical information systems. Among various honors, Dr. Malin received the Presidential Early Career Award for Scientists and Engineers (PECASE).

Competing interests

The authors declare that they have no competing interests.

Received: 16 September 2011 Accepted: 27 February 2012 Published: 27 February 2012

References

1. Gruber T: **Collective knowledge systems: where the social web meets the semantic web.** *Journal of Web Semantics* 2007, **6**:4-13.
2. Chen H, Zeng D, Atabakhsh H, Wyzga W, Schroeder J: **COPLINK: managing law enforcement data and knowledge.** *Communications of the ACM* 2003, **46**:28-34.
3. Reddy M, Spence P: **Collaborative information seeking: A field study of a multidisciplinary patient care team.** *Information Processing and Management* 2008, **44**:242-255.
4. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, Taylor R: **Can electronic medical record systems transform health care?** *Health Affairs* 2005, **24**:1103-1107.
5. Bier E, Card S, Bodnar J: **Principles and tools for collaborative entity-based intelligence analysis.** *IEEE Transactions on Visualization and Computer Graphics* 2010, **16**:178-191.
6. Westphal CR: **Data Mining for Intelligence, Fraud & Criminal Detection: Advanced Analytics & Information Sharing Technologies.** *CRC Press* 2008.
7. Eldenburg L, Soderstrom N, Willis V, Wu A: **Behavioral changes following the collaborative development of an accounting information system.** *Accounting, Organizations and Society* 2010, **35**:222-237.
8. Sheth A: **From contemporary workflow process automation to adaptive and dynamic work activity coordination and collaboration.** *Proceedings of the 8th International Workshop on Database and Expert Systems Applications* 1997, 24-27.
9. Spinellis D, Louridas P: **The collaborative organization of knowledge.** *Communications of the ACM* 2008, **51**:68-73.
10. Horvath G, Bolinger J, Ramanathan J, Ramnath R: **Document-centric collaborative spaces for increased traceability in knowledge-intensive processes.** *Proceedings of the International Symposium on Collaborative Technologies and Systems* 2009, 400-407.
11. Huang CC, Li TY, Wang HC, Chang CY: **A collaborative support tool for creativity learning: idea storming cube.** *Proceedings of International conference on Advanced Learning Technologies and Technology-enhanced Learning* 2007, 31-35.
12. Leimeister J, Huber M, Bretschneider U, Krcmar H: **Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition.** *Journal of Management Information Systems* 2009, **26**:197-224.
13. Bao P, Gerber E, Gergle D, Hoffman D: **Momentum: getting and staying on topic during a brainstorm.** *Proceedings of the 28th International Conference on Human Factors in Computing Systems* 2010, 1233-1236.
14. Reddy M, Jansen B: **A model for understanding collaborative information behavior in context: a study of two healthcare teams.** *Information Processing and Management* 2008, **44**:256-273.
15. Doss G, Tejay G: **Developing insider attack detection model: a grounded approach.** *Proceedings of IEEE International Conference on Intelligence and Security Informatics* 2009, 107-112.
16. Menachemi N, Brooks R: **Reviewing the benefits and costs of electronic health records and associated patient safety technologies.** *Journal of Medical Systems* 2008, **30**:159-168.
17. Stolfo S, Bellovin S, Hershkop S, Keromytis A, Sinclair S, Smith SW: *Insider attack and cyber security: beyond the hacker* Springer; 2008.
18. Schultz E: **A framework for understanding and predicting insider attacks.** *Computers and Security* 2002, **21**:526-531.
19. Thomas RK, Sandhu SR: **Task-based authorization controls (TBAC): a family of models for active and enterprise-oriented autorization management.** *Proceedings of the 11th IFIP International Conference on Database Security XI: Status and Prospects* 1997, 166-181.
20. Seo YW, Sycara K: **Cost-sensitive access control for illegitimate confidential access by insiders.** *Proceedings of IEEE International Conference on Intelligence and Security Informatics* 2006, 117-128.
21. Gunter CA, Liebovitz DM, Malin B: **Experience-based access management: a life-cycle framework for identity and access management systems.** *IEEE Security and Privacy Magazine* 2011, **9**:48-55.
22. Byun JW, Li N: **Purpose based access control for privacy protection in relational database systems.** *International Journal on Very Large Data Bases* 2008, **17**:603-619.
23. Georgiadis C, Mavridis I, Pangalos G, Thomas R: **Flexible team-based access control using contexts.** *Proceedings of ACM Symposium on Access Control Models and Technologies* 2001, 21-27.
24. Kulkarni D, Tripathi A: **Context-aware role-based access control in pervasive computing systems.** *Proceedings of ACM Symposium on Access Control Models and Technologies* 2008, 113-122.
25. Beimel D, Peleg M: **The context and the SitBAC models for privacy preservation - an experimental comparison of model understanding and synthesis.** *IEEE Transactions on Knowledge and Data Engineering* 2010, **10**:1475-1488.
26. Boxwala AA, Kim J, Grillo JM, Machado LO: **Using statistical and machine learning to help institutions detect suspicious access to electronic health records.** *Journal of the American Medical Informatics Association* 2011, **18**:498-505.
27. Fabbri D, LeFevre K: **Explanation-based auditing.** *Proceedings of the VLDB Endowment* 2011, **5(1)**:1-12.
28. Chen Y, Nyemba S, Zhang W, Malin B: **Leveraging social networks to detect anomalous insider actions in collaborative environments.** *Proceedings of the IEEE International Conference on Surveillance and Security Informatics* 2011, 119-124.
29. Cappelli D, Moore A, Shimeall T, Trzeciak R: **Common sense guide to prevention and detection of insider threats.** *Carnegie Mellon University CyLab* 2009.
30. Beimel D, Peleg M: **Comparing the context and the SitBAC models for privacy preservation in terms of model understanding and synthesis.** *Proceedings of the American Medical Informatics Association Annual Symposium* 2008, 874.
31. Peleg M, Beimel D, Dori D, Denekamp Y: **Situation-based access control: privacy management via modeling of patient data access scenarios.** *Journal of Biomedical Informatics* 2008, **41**:1028-1040.
32. Coyne E: **Role engineering.** *Proceedings of the first ACM Workshop on Role-based Access Control* 1996, 4.
33. Neumann G, Strembeck M: **A scenario-driven role engineering process for functional RBAC roles.** *Proceedings of the 7th ACM Symposium on Access Control Models and Technologies* 2002, 33-42.
34. Kuhlmann M, Shohat D, Schimpf G: **Role mining - revealing business roles for security administration using data mining technology.** *Proceedings of the 8th ACM Symposium on Access Control Models and Technologies* 2003, 179-186.

35. Vaidya J, Atluri V, Guo Q: **The role mining problem: finding a minimal descriptive set of roles.** *Proceedings of the 12th ACM Symposium on Access Control Models and Technologies* 2007, 175-184.
36. Malin B, Nyemba S, Paulett J: **Learning relational policies from electronic health record access logs.** *Journal of Bimedical Informatics* 2011, **44**:333-342.
37. Probst CW, Hansen RR, Nielson F: **Where can an insider attack?** *Proceedings of USENIX Conference on File and Storage Technologies* 2006, 127-142, Edited by 4691 L.
38. Noble CC, Cook DJ: **Graph-based anomaly detection.** *Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining* 2003, 631-636.
39. Chen Y, Malin B: **Detection of anomalous insiders in collaborative environments via relational analysis of access logs.** *Proceedings of ACM Conference on Data and Application Security and Privacy* 2010, 63-74.
40. Sun J, Qu H, Chakrabarti D, Faloutsos C: **Neighborhood formation and anomaly detection in bipartite graph.** *Proceedings of IEEE International Conference on Data Mining* 2005, 418-425.
41. Eberle W, Holder LB: **Applying graph-based anomaly detection approaches to the discovery of insider threats.** *Proceedings of IEEE International Conference on Intelligence and Security Informatics* 2009, 206-208.
42. Shyu ML, Chen SC, Sarinnapakorn K, Chang L: **A novel anomaly detection scheme based on principal component classifier.** *Proceedings of International Conference on Data Mining* 2003, 172-179.
43. Adamic LA, Adar E: **Friends and neighbors on the Web.** *Social Networks* 2003, **25**:211-230.
44. Sarwar B, Karypis G, Konstan J, Riedl J: **Item-based collaborative filtering recommendation algorithms.** *Proceedings of ACM World Wide Web Conference* 2001, 285-295.
45. **Wiki Talk Dataset.** [http://snap.stanford.edu/].
46. Bradley A: **The use of the area under the ROC curve in the evaluation of machine learning algorithms.** *Pattern Recognition* 1997, **30**:1145-1159.

doi:10.1186/2190-8532-1-5
Cite this article as: Chen *et al.*: Specializing network analysis to detect anomalous insider actions. *Security Informatics* 2012 **1**:5.