A Hybrid Model based on Multi-Dimensional Features for Insider Threat Detection

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

***Abstract*—Insider threats have shown their power by hugely affecting national security, financial stability, and the privacy of many thousands of people. A number of techniques have been proposed to detect insider threats either by comparing the behaviors among different individuals or by comparing the the behaviors across different time periods of the same individual. However, due to the fact that the behaviors of insider threats are always complex and diverse, both of them always fail to identify certain kinds of inside threats. This paper focuses on proposing a hybrid model to detect insider threats based on multi-dimensional features. First, based on the isolation Forest algorithm, an Across-Domain Anomaly Detection(ADAD) model is proposed to dentify anomalous behaviors that deviates from the behaviors of their peers by using multi-source information. Second, we propose an Across-Time Anomaly Detection(ATAD) model to measure the degree of unusual changes of a user's behavior by implementing an improved Markov model. Finally, our proposed hybrid model present a fusion method to integrate the evidences from the above two models. With the data lasting 17 months, we evaluate our proposed models comprehensively. The results demonstrate the robust performance of the ADAD and ATAD models and the hybrid model is demonstrated to outperform the two separate models**

***Index Terms*—anomaly detection, insider threat detection, information fusion, machine learning**

I. Introduction

Insider threats are threats with malicious intent directed towards organizations by people internal to the organization. These include physical sabotage activities, theft of confidential data and business secrets, and fraud. Financial loss and reputation damage caused by this "known unknow" cybersecurity threat far outweighs that caused by external attacks. One of the most recent articles from CSO magazine [3] compared the cost between external and internal attacks and noted that while it takes about 50 days to fix a data breach caused by an internal attack, it only takes 2 to 5 days in the case of external attacks. Moreover, attacks by malicious insiders are also the costliest to fix ($145,000), followed by denial of service ($127,000) and Web-based attacks ($96,000), indicating the severity of this problem. Researchers have proposed different models aimed at preventing or detecting the presence of attacks.

Existing literature focuses on two types of the detection models: data driven detection [4] [5] and behavior driven detection [2] [6]. The first model aims to find a normal portrait in all users data in order to detect insider threat that deviates from this normal portrait by comparing with their peers. Figure 1 shows an example of insider threats which can be detected by this kind of model [10]. After getting off work, the general behavior is that everyone will go home to have a rest. Few of users begin to secretly deal with the company's confidential documents when others are sunk in sleep at midnight and are more likely to be a insider threat. However, merely considering this scenario, this kind of models fails to detect a malicious insider tries to behave like a normal user to cover up his evil [10]. The second model regards the abnormal change in the behavior as the basis for inside threat detection by comparing behaviors of themselves in different time periods. Figure 1 also show an example to depict a malicious insiders detection by finding the user's unusual changes based on the behavior-driven method [8]. Over a long period in the past, the regular behavior of the user is to click on the browser after booting, and view the email then reply to a series of messages. But one day, he act abnormally: he connected the mobile device after booting, and had a series of operations on the file-copy to steal the company's confidential documents. These behaviors don't match his usual style of doing things. It is worth to note that although the malicious insider tries to behave like a normal user, the threat can also be detected based on the deviation from his regular behavior routines. However, the model unable to recognise situations where a user systematically attacks an organisation for long-term period. [2]. Thus, Insider threat surveys [9] suggest this problem cannot be considered only as a data or behavior driven problem. Therefore,it is necessary to model the problem of insider threat detection as data and behavior driven problems, based on which a effective model is to proposed to combine the two factors.

Thus, this paper proposes a hybrid model based on multi-dimensional features that combines a data driven model with a

Identify applicable funding agency here. If none, delete this.

behavior driven model to detect insider threat in a more robust and accurate manner. The multidimensional features extracted from data collected from the enterprise network is formatted and fed separately into the two separate models, and each model generates an abnormal score to represent the degree of users' unusual behaviors. Then, the abnormal scores of two models are fused for each user as the final abnormal scores of users. A user will be detected a insider threater if the anomaly score exceeds the threshold. After a wide range experiment, it is verified that the hybrid model can detect insider threats in a more robust and accurate manner. Overall, the contributions of this paper can be summarized as follows:

1) We apply isolate Forest to detect behavioral inconsistencies in the behavior of most users. In this model, we extract and combine temporal features from multi-sources of information comprehensively for the construction of the user's portrait.
2) We propose an improved model based on Markov to identify users' unusual change. The proposed model considers all historical behavior of users as the contextual information. As compared with the existing predictor of Markov model [10], our proposed approach show higher efficiency.
3) We propose a hybrid model for accuracy detection of insider threats by fusion methods which is able to determine malicious insiders that not only act inconsistent with peers but also have unusual change compared with their historic behavior and achieve the accuracy of 95%.

The remainder of this article is structured as follows. Section 2 presents the context and related work about the research. Then, section 3 we detail our approach and discuss algorithms used and how we train our model to detect insider threats. Next, in Section 4 we detail our implementation and analyze the reault. Finally we conclude in Section 5, also presenting limitation of our work and outlining avenues for future research.

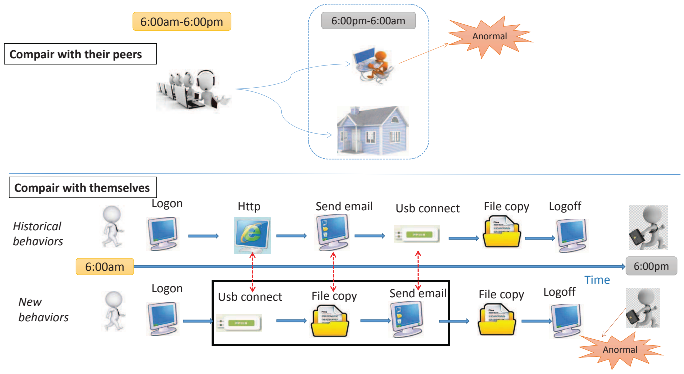


Fig. 1: Example of insider threat.

II. Related work

The topic of insider threat has attracted much attention recently in different fields. Researchers have proposed different models aimed at preventing or detecting the presence of attacks ([11] [12]). To elicit the state of art, the work presented here is focused on the approaches of detecting insider threat based on data-driven and behavior-driven.

Methods based on data analytics include the work done by Mathew et al. [13] on account of user access patterns, the work of Eberle et al. [14] on using social graphs to detecting the abnormal. More recently, Eldardiry et al. [15] have also proposed a system for insider threat detection based on feature extraction from user activities. Michael Goldsmith use some methods based on fusion of multi-source information and user behavior [16]. There have also been various approaches based on specific insider threat scenarios to detect abnormal [7] [8]. However, they did not factor in the change of user behavior over time. We note that while a common activity not be suspicious, a rare change of the order common activity can be.

There are several papers based on behavioral models [10] [11] [12]. Hoda et al. [13]detect peer groups of users and modeling user behavior with respect to these peer groups, and subsequently detect insider activity by identifying users who deviate from their peers with respect to the user behavior models. Tabish Rashid [**?**] takes the change of user behavior over time to detect the anomaly and achieved some results. However, these models based on behavior-driven just work out based on unusual change of user behavior but will miss recognising situations where a user systematically attacks an organisation over an extended time-framework.

Considering these two factors, we propose a hybrid model based on multi-dimensional features for accuracy detection of anomalies. We fusion two method based on data and behavior. The first one, we compare behaviours across different individuals, while in the latter, we compare behaviours acorss different times for the same individual. We abtain the accuracy of 95% through the proposed model, which is of great significance in industry and scientific research.

III. Overview of proposed approach

We propose a hybrid model based on multi-dimensional features for insider threat detection which is illustrated in Figure2. First, we count and analyze the characteristics of each domain(e.g., logons) from multi-source information as parameters into Across-Domain Anomaly Detection(ADAD) to get anomaly scores. Second, we extract the features of time-series behaviors of users from multi-source information, then we identify changes in activities of a user compared to that users past activities through Across-Time Anomaly Detection(ATAD) to get anomaly scores. Finally, we combine the anomaly scores from the two components to detect insider threat. Next, we will introduce the experimental data.

A. The Data Set

Due to the lack of availability of proper insider threat datasets, we have utilized the insider threat dataset published by CERT Carnegie Mellon University for this research [**?**]. The dataset R4.2.tar.bz has been used for this analysis. According

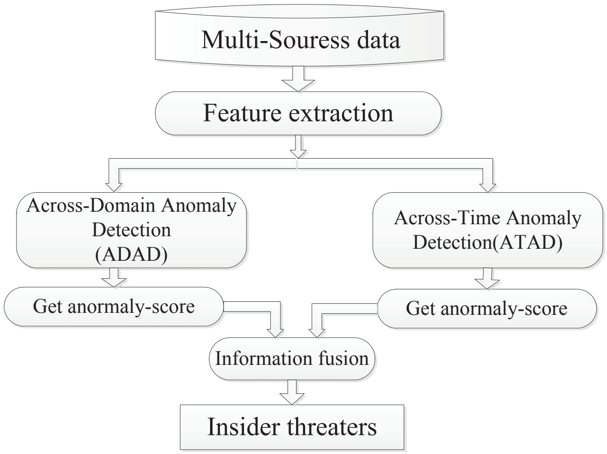


Fig. 2: Anomaly Detection Framework.

to the dataset owners, this is a dense needle dataset with a fair amount of red team scenarios. This dataset consists of six broad types of data records (HTTP, logon, device, file, email and psychometric) of 1000 employees over a 17 months period. All HTTP records contain user, PC, URL and web page content with time stamps. Logon.csv consists of user logon/logoff activities with the corresponding PC with timestamps. Logon activity corresponds to either a user login event or a screen unlock event, while the Logoff event corresponds to user logoff event. The third data file device.csv is a collection of data records of removable media usage. It indicates insert/remove actions with the relevant user, PC, and timestamp. Details of file copies are stored in file.csv file with date, user, PC, filename, and content. We should note that the CERT Dataset contains the ground truth for each user (when they are acting maliciously or not), which allows us to monitor the success or failure of our experiment.

IV. KEY METHODOLOGIES

In this section, we describe our across-domain anomaly detection(ADAD) model in detail at first. We next describe our proposed improved MM based on across-time to detect insider threats. Finally, we introduce the fusion method combining ADAD and ATAD.

A. Approach 1:Across-Domain Anomaly Detection(ADAD)

This framework will utilize multidimensional inputs, such as user interactions with hardware assets, logon records and operation on file, to identify anormal users who behave differently from their peers. The figure 2 reports the structure of the approach.

1) Feature Extraction: **Individual Logon-Logoff Behaviors.** This parameters can be used in identifying users abnormal logon/logoff activities, as most disgruntled insiders tend to commit malicious activities after hours [**?**]. Identifying users baseline behavior on system/device access is an essential part of malicious insider threat detection problem. For normal users and abnormal users, two parameters (the average of their maximum and mode) logon and logoff values have been calculated for every hour shown as Figure4.

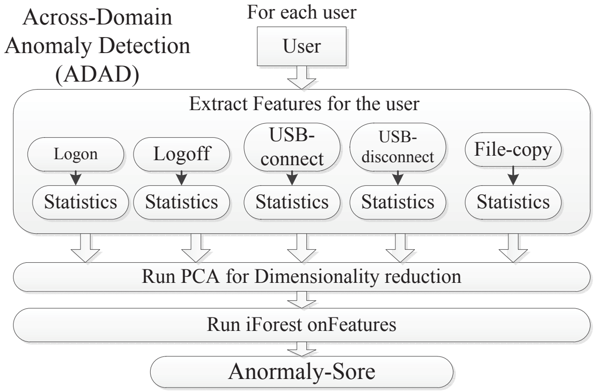


Fig. 3: An overview of the Across-Domain Anomaly Detection(ADAD).

TABLE I: Selected parameter set.

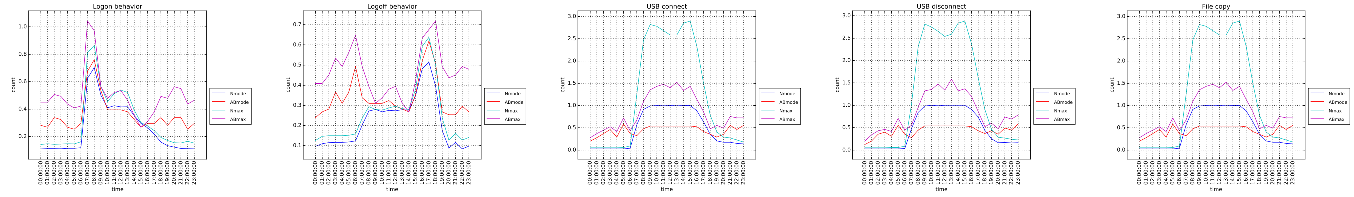
Module	Parameter (00:00-06:00)(06:00-12:00) (12:00-18:00)(18:00-24:00
Logon events	Maximum/Mode Logon counts
Logoff events	Maximum/Mode Logoff counts
Removable Media	Maximum/Mode Connect counts Maximum/Mode Disconnect counts
File copy events	Maximum/Mode Filecopy counts

Removable media usage. Removable media is among the most popular method used in theft of Intellectual Property (IP) in extracting confidential information from organizations [**?**]. Tracking the use of removable media can be an excellent information source for identifying suspicious events by trusted insiders . Baseline behavior of removable media usage is captured by the average of their maximum and mode time of Insert and Remove activities as in the logon/logoff event analysis. The average of the number of files copied per hour by normal and abnormal is also used in this analysis. Figure5 shows it.

From above, we found that there is a a big difference in behavior between normal user and abnormal user at different times, so we decide to merge the times of behavior every 6 hours as the parameter to input to our ADAD model. Figure 5 is an illustration of the parameters.

After the experiment we found that the features we extracted had noise effects (which in detail in the experimental part), in order to achieve higher accuracy, we use PCA [**?**] for denoising. All feature columns are normalized before the PCA decomposition is performed. By default, we consider a decomposition of the features to a 2-D space.

2) Anomaly Detection: Due to the complex nature of insider threat problem, it is extremely hard to pinpoint a user as a malicious insider. Therefore, the first step should be the identification of possible malicious insiders who are maximally deviating from peers as well as their normal behavior. Therefore, as the second stage of our analysis, we will focus



(a) Logon behavior (b) Logoff behavior (c) USB connect (d) USB disconnect (e) File copy

Fig. 4: Removable media usage behavior

on implementing an anomaly detection algorithm based on the the properties identified at the previous stage of this analysis. The anomaly detection algorithm adopted in this analysis is the Isolation forest algorithm, which stands out in effectively separating anomalous events from the rest of the instances [**?**].

B. Approach 2: Across-Time Anomaly Detection(ATAD)

Markov Model (MM) [**?**] is an extremely powerful tool to model temporal sequence information. It has been widely used in temporal pattern recognition problems (e.g., speech recognition, bioinformatics, gesture recognition) due to its high detection rate [**?**]. MM has also been used in the general area of intrusion detection by some notable works ([**?**]). Since we model users behaviors as a temporal sequence of observable query anomaly scores in our work in this section, and proposed an improved MM model to detect insider threats. Figure 6 shows a graphical overview of the processing approach and pipeline we use, and the remainder of this section is dedicated to describing each stage in it.

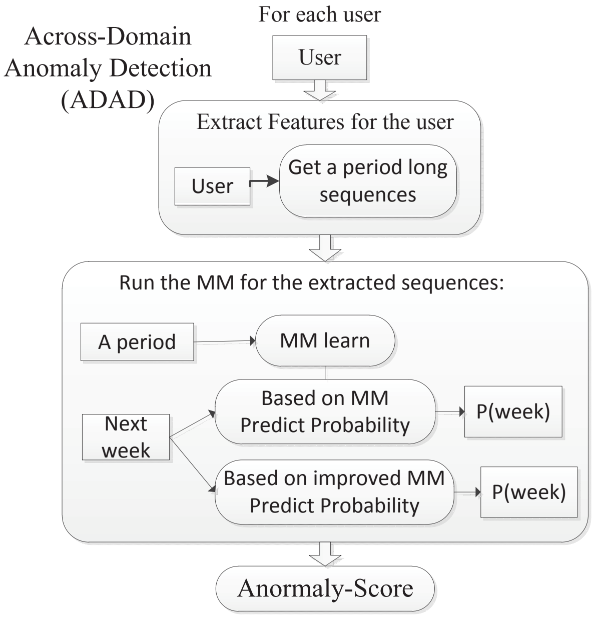

Across-Domain
Anomaly Detection
(ADAD)
For each user
User
Extract Features for the user
User
Get a period long sequences
Run the MM for the extracted sequences:
A period
MM learn
Next week
Based on MM Predict Probability
P(week)
Based on improved MM Predict Probability
P(week)
Anormaly-Score

Fig. 5: An overview of the Across-Time Anomaly Detection(ATAD).

1) Feature Extraction: Users have defferent behaviors on computers every day. When a user finish a behavior and bingen the next behavior, a new record is created. The historical behaviors may be a sequence of observations $B = (b_1, b_2..., b_n)$,

We partition all behaviors by each user and then sort them based on the respective timestamps. Finally we get a list of actions and the time at which they undertook them for each user. The behavior of each user may be a sequence of observations $B = (b_1, b_2..., b_n)$,

We assign a symbol to each behavior based on the set of features we are defined in ADAD.

We defined the temporal behavior in the recent past by opening up an observation window of size N on the continuous steam of audit events to view the last N audit events from the current time t: E t-(N-1)=t-N+1, , Et, where E stands for event.

2) Anomaly Detection: For insider threat detection, we propose an improved Markvo method(IM) to model behavior change. First,we need to build a period time norm profile of behavior by modeling behavior change for each user during a period of time. For every user, we computer the likelihood of transitions between one behavior to next behavior to get a transition probability matrix of behaviors. Second, we compare the behavior in the recent past to the period-time norm profile to get the anomaly score for each user and if the anomaly score exceeds a threshold we detect the user as a insider threater. Next, the remainder of this section is dedicated to describing the model.

Improved Markov model (IM). First, the behavior transition probability matrix can be learned from the observations of the user behavior in the past. We extract the behavior history $B = (b_1, b_2, ..., b_n)$ and the current context $C = (b_{12}, b_{13}, ..., b_{ij})$, b_{ij} denotes that behaver j happend behind behavior i. To calculate the probability of one behavior(b_i) transitions to the next behavior(b_j) and the initial probability distribution(q_i) of behavior i, we use a simple estimator P_{ij} from the historical behaviors B based on the assumption of Markov models [**?**]:

$$P_{b_{ij}} = N(b_{ij}, C)/N(b_i, B) \tag{1}$$

$$q_{b_i} = N(b_i, B)/N(b, B) \tag{2}$$

Where *N($s\prime$, s)* represents the number of times the substring $s\prime$ appears in the string s. We calculate the probability of all behavior to get the behavior transition probability matrix for the next step.

Second, we get a new behavior sequences of the user in the next week (b_{11},b_{12},......b_{DT}), where b_{DT} denotes the behavior

T happens on the day D. The anomaly score R_{trend} for the user is calculated by estimating the users transition likelihood over time. The anomaly probability score is computed depending on the behavior transition probability matrix caculated in the first step as

$$R_{trend} = 1/D \sum_{d=1}^{D} q_{b_1} \prod_{t=2}^{T} P_{b_{d(t-1)} b_{d} t} \tag{3}$$

Third, after abtaining the anomaly score, we detect the user as a insider threater if his anomaly score exceeds a threshold.

C. *Information fusion*

Our goal is to combine suspicion/anomaly scores that have been generated from each of the aforementioned methods to detect anomalies. Therefore, we developed a technique based on weight fusion to combine acorss-domain and across-time of evidence from multiple domains. In a broad sense, each source of information provides a suspicion score and the goal is to combine these scores by wight(W) in order to identify anomalies with greater accuracy. The combined suspicion score is computed as

$$R_{combined} = w_1 * R_{ADAD} + w_2 * R_{ATAD} \tag{4}$$

$$r = \begin{cases} w_2/w_1 & \text{if } w_1 > 0 \\ 1 & \text{if } w_1 \equiv 0 \end{cases} \tag{5}$$

V. Experiment

This section is dedicated to a comprehensive discussion of results obtained through our analysis. In the next sections, we will introduce the performance metrics for evaluating the three detection models separately. We then provide a comparative assessment of our proposed models with detecting insider threat methods.

A. *Across-Domain Anomaly Detection(ADAD)*

In this section, we introduce the performance metrics for evaluating the detection model. We then provide a comparative assessment of our proposed models with different parameters inputing.

*1) Evaluation Metrics and Baselines:*To quantitatively evaluate the models with different parameters inputing, we consider the metrics detection *Accuracy, Precision* and *Recall*. *Accuracy* represents the proportion of the number of users who are detected correctly in all users by the model. *Precision* represents the proportion of the number of insider threats are detected correctly in the insider threats who are detected by the model and *Recall* represents the proportion of the number of insider threats are detected in all insider threats by the model.

2) Comparative Evaluation: Table II shows the results of running our model ADAD with different parameters. We found that removable media domains is the most detectable (with the highest detective accuracy), while Logon and file domains are harder to detect. It appears that users show great variation in their logon and file behavior, but are more uniform in device usage . Except that, the property of the mode is more effective than maximum to detect, which indicates that the mode represents the difference between normal and abnormal. It is obvious that some features interfere with the detection results. To improve the accuracy, we use *PCA* (Principal Component Analysis) to decompose of the features to a 2-D space, and get a higher predictive accuracy.

TABLE II: The experiment result of ADAD.

iForest Input		Accuracy	Precision	Recall
Logon	MAX	89%	32%	47%
	MODE	87%	22%	34%
Logoff	MAX	88%	23%	34%
	MODE	85%	14%	21%
Connet	MAX	78%	65%	27%
	MODE	79%	70%	28%
Disconnect	MAX	80%	73%	30%
	MODE	79%	69%	28%
Filecopy	MAX	80%	60%	28%
	MODE	77%	44%	20%
All properties	MAX	82%	68%	34%
	MODE	79%	52%	31%
All properties	dimension=2	87%	79%	35%

B. *Across-Time Anomaly Detection(ATAD)*

In this section, we introduce the performance metrics for evaluating the detection model. We then provide a comparative assessment of our proposed models with existing MM detection methods.

1) Evaluation Metrics and Baselines: To quantitatively evaluate the models, we consider the metrics detection Receiver Operating Characteristic curves (or *ROC* curves) curve. These curves are used to plot the true-positive (correct) rate against the false-positive rate for the different possible points in a diagnostic test. In order to produce the ROC curves, we use a process of threshold varying. After obtaining the anomaly score for each user from our model, we set a value for the threshold and then check each user one-by-one and classify them as anomalous if *score¿Threshold*. We then increase the threshold and repeat until no users are classified as anomalous.

Comparison Methods. The improved Markov (IM)is respectively compared with existing detection Markov:

a) Markov: when caculating the anomaly score, this model regard the behavior sequences of the week as whole and treat the evaluation result of the sequences as the final anomaly score for the week.

b)improved Markov (IM): this model evaluate the sequence of every day in the week and get the anomaly score for each day, and then get this week's average anomaly score as the last anomaly score for this week.

2) Comparative Evaluation: **Comparison between IM and MM.** We apply our Across-Time Anomaly Detection(ATAD) detection method as follows. It's obvious that the IM model is more effective than MM. For MM, if a user has once unusual change during this week, which will has a important influence on the final result, maybe the change just is working overtime. This will has effect on the detection. However, it can be avoided by IM. Because it will reduce the impact of accidental unusual change on scoring the user's behavior by using the week's average anomaly score to evaluate a user's behavior.

In other words, IM is more likely to detect abnormal changes in behavioral trends over a period of time, which detect insider threat more accurately.

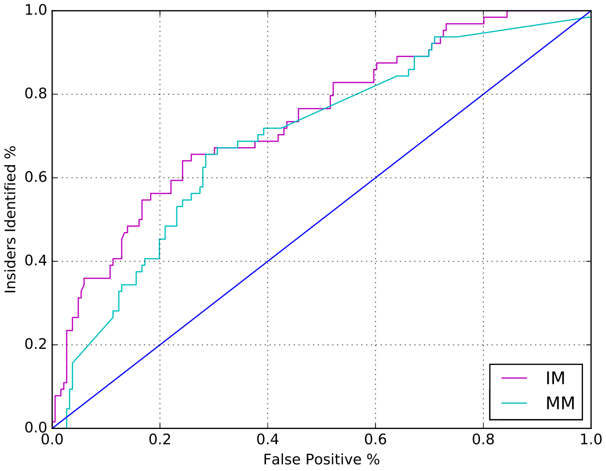


Fig. 6: ROC Curve showing the differences between MM algorithm and improved MM algorithm.

C. Information fusion

In this section, we use the same performance metrics of ADAD for evaluating the detection model. we then provide a comparative assessment of our fusion method with above two models.

We get the final anomaly score based on weight fusion as shown in Table 2. Parameter *r* represents the proportion of the weight of the two methods score. The scores are normalized before the weight fusion(combine). When r = 0, it is the result of individual ADAD method, when r = 1, it is the result of individual ATAD method. We can see that when r is 3/7, the accuracy and recall rate is the highest, where the combination of weights is the best. Indeed, it is remarkable that a suitable combination of the individual ATAD and ADAD scores in an appropriate fashion leads to significant improvement in performance relative to any of the individual ATAD or ADAD sources.

Figure 8 is an indication of how the anomaly scores are distributed when r is 3/7 in this analysis. The graph indicated few points above the Red color horizontal line which is equivalent to an anomaly score of 1.3 which is determine by *ROC* curves. Users belong to those points can be considered as anomalous users. To get a better understanding and visualization of results based on our approach, we mark the positive sample with red and negative sample with blue. It reached 95% of the accuracy.

VI. Conclusion And Future work

In this paper we have examined the problem of insider threat detection. Our main contribution is that a novel fusion approach for robust detection of anomalies is discussed. The two main components of our framework - ADAD and ATAD - improve anomaly detection prediction accuracy by combining information from multiple domains and time-instances.

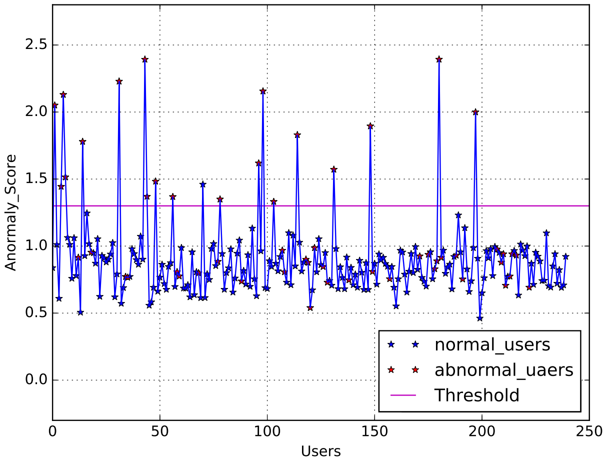


Fig. 7: Anomaly score distribution.

Future work. We use a combined method based on weight fusion to integrate the ADAD and the ATAD, which has a high precision but a pessimistic recall. To solve this problem, we could apply on some complex and effective fusion scheme to combine their information in order to improve the recall of anomaly detection(e.g., [**?**]). In addition, we could take user role into account in generates user Normal portrait that can describe the full extent of activities that users perform within the organization based on role to improve the recall of anomaly detection(e.g., [**?**]).

Acknowledgment

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

References

[1] Gamachchi, Anagi, L. Sun, and S. Boztas. "A Graph Based Framework for Malicious Insider Threat Detection." Hawaii International Conference on System Sciences 2017.

[2] Rashid, Tabish, I. Agrafiotis, and J. R. C. Nurse. "A New Take on Detecting Insider Threats: Exploring the Use of Hidden Markov Models." International Workshop 2016:47-56.

[3] By the numbers: Cyber attack costs compared, 2016, accessed on 31/05/2016. [Online]. Available: http://www.csoonline.com/article/3074826/security/bythe-numbers-cyber-attack-costs-compared.html

[4] D. Cappelli, A. Moore, and R. Trzeciak, The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud). Addison-Wesley Professional, 2012

[5] Young, William T, et al. Use of Domain Knowledge to Detect Insider Threats in Computer Activities. 2013.

[6] Eldardiry H, Sricharan K, Liu J, et al. Multi-source fusion for anomaly detection: using across-domain and across-time peer-group consistency checks[J]. Computing & Informatics, 2014, 31(3):575-606.

[7] Gamachchi, Anagi, L. Sun, and S. Boztas. "A Graph Based Framework for Malicious Insider Threat Detection." Hawaii International Conference on System Sciences 2017.

[8] Sherali Zeadally, et al. "Detecting Insider Threats: Solutions and Trends." Information Security Journal A Global Perspective 21.4(2012):183-192.

TABLE III: The experiment result of information fusion.

r	0(ADAD)	1/9	2/8	3/7	4/6	5/5	6/4	7/3	8/2	9/1	1(ATAD)
Precision	79.17%	90%	94.44%	**95%**	74.73%	90%	85%	82.35%	66.67%	61.9%	60%
Recall	35.19	35.18%	31.48%	**35.19%**	33.33%	33.33%	31.48%	25.92%	25.92%	24.07%	27.78%

[9] D. Cappelli, A. Moore, and R. Trzeciak, The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud). Addison-Wesley Professional, 2012

[10] W. L. Winston, Operations Research: Applications and Algorithms. Belmont, CA: Duxbury Press, 1994.

[11] P. A. Legg et al., Towards a conceptual model and reasoning structure for insider threat detection, J. Wireless Mobile Netw., Ubiquitous Comput., Dependable Appl., vol. 4, no. 4, pp. 2037, Dec. 2013.

[12] M. Bishop et al., Insider threat detection by process analysis, in Proc. IEEE SPW, 2014, pp. 251264.

[13] Sunu Mathew, Michalis Petropoulos, Hung Q Ngo, and Shambhu Upadhyaya. A data-centric approach to insider attack detection in database systems. In Recent Advances in Intrusion Detection, pages 382401. Springer, 2010.

[14] William Eberle, Jeffrey Graves, and Lawrence Holder. Insider threat detection using a graph-based approach. Journal of Applied Security Research, 6(1):3281, 2010.

[15] H. Eldardiry et al., Multi-domain information fusion for insider threat detection, in Proc. IEEE SPW, May 2013, pp. 4551.

[16] Automated Insider Threat Detection System Using User and Role-Based Profile Assessment Philip A. Legg, Oliver Buckley, Michael Goldsmith, and Sadie Creese