Enhancing Challenge-based Collaborative Intrusion Detection Against Insider Attacks using Spatial Correlation

Wenjuan Li*†, Weizhi Meng*†, Javier Parra-Arnau‡, and Kim-Kwang Raymond Choo§
*Institute of Artificial Intelligence and Blockchain, Guangzhou University, China
†Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark
‡Karlsruhe Institute of Technology, Germany
§Department of Information Systems and Cyber Security, University of Texas at San Antonio, USA
Email: {weme@dtu.dk, javier.parra@urv.cat, raymond.choo@fulbrightmail.org}

2021 IEEE Conference on Dependable and Secure Computing (DSC) | 978-1-7281-7534-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/DSC49826.2021.9346232

***Abstract*—With cyber-attacks becoming more complicated and the networks increasingly interconnected, there has been a move towards using collaborative intrusion detection networks (CIDNs) to identify cyber-threats more effectively. However, insider attacks may remain challenging to mitigate in CIDNs, as the intruders are able to control one or more internal nodes. Challenge-based trust mechanism is one promising solution to help safeguard CIDNs against common insider attacks, but not necessarily against advanced attacks such as passive message fingerprint attacks. In this work, we focus on challenge-based trust mechanism and advocate that considering additional level of trust can enhance the robustness of CIDNs. Specifically, we design an enhanced trust management scheme by checking spatial correlation among nodes' behavior, regarding forwarding delay, packet dropping and sending rate. Then, we evaluate our approach in a simulated environment, as well as a real-world environment in collaboration with an IT organization. Experimental results demonstrate that our approach can help enhance the robustness of challenge-based trust mechanism by detecting malicious nodes faster than similar approaches (i.e., reducing time consumption by two to three days).**

***Keywords*-Collaborative Intrusion Detection, Spatial Correlation, Advanced Insider Threat, Challenge-based Trust Management, Trust Computation.**

I. Introduction

Computer networks, including Internet of Things (IoT), are constantly targeted by a variety of cyber threats, such as malware. For example, cryptojacking attacks were identified as a commonly detected cyber threat, partly due to its low barrier of entry and minimal overhead [40]. Cyber criminals could exchange information in many hidden, underground hacker forums and purchase a wide range of malicious tools and shady services (also known as cybercrime-as-a-service in the literature) [24], i.e., launching (sophisticated) attacks on existing or emerging technologies and platforms.

To defeat cyber attacks, intrusion detection systems (IDSs) have been widely implemented to detect and mitigate a broad range of threats. Existing IDSs can be either network-based or host-based system [25, 26]. The former mainly monitors network events for any violations, while the latter detects abnormal events in the local system, e.g., system logs. An IDS can also be categorized based on the specific detection methods. The conventional signature-based IDS can identify a potential threat by performing a comparison between its signatures and the incoming traffic. The anomaly-based IDS determines or marks potentially malicious events by identifying a great deviation between the current condition and the normal status in the system or network. If a malicious situation is found, IDS can notify security administrators to investigate further.

Accessing to contemporary technologies and more computational resources can ease the launching of more complex and potentially impactful attacks, and complicate the challenge of IDS in distinguishing potentially malicious cyber activities from the legitimate network traffic. This reinforces the importance of collective cyber threat intelligence. Collaborative intrusion detection systems (CIDSs) or networks (CIDNs) are thus designed to acquire and share intelligence between a set of IDS nodes [42]. By having access to the collective cyber threat intelligence, such systems can have a global view of the cyber threat landscape and hence, can better respond to emerging cyber threats. However, not all IDS nodes are 'equal' in the terms of trustworthiness. There are a number of trust mechanisms proposed in the literature, including the challenge-based trust mechanism [8], where a special message of *challenge* is sent to help determine the reputation of other nodes within the same network. Then, a trust value can be derived by comparing the received answer with the expected feedback.

Motivations. Challenge-based trust mechanism can help identify and mitigate most common insider attacks such as newcomer attacks. However, such a mechanism is ineffective against advanced insider attacks, i.e., when a malicious node is capable of choosing a strategy to send untruthful feedback. For example, under the *passive message fingerprint attack (PMFA)*, several malicious nodes can collude by exchanging information to collectively figure out a normal request [16]. Then, these colluding malicious nodes can make a plan and send untruthful feedback to normal request only, in order to maintain their trust values.

Contributions. In this work, we advocate that additional level of trust can be considered to improve the robustness

of challenge-based CIDNs, particularly against complicated insider attacks. It is observed that existing computing nodes are also required to help relay packets, and cyber criminals would often misbehave in comparison to normal ones, for example by manipulating network packets to launch attacks. This may cause dropping and delaying of numerous packets. Thus, it is highly likely to detect potentially malicious node by investigating the correlation among neighboring nodes. Specifically, a node can monitor the behavior of other nodes and help determine abnormal nodes. Motivated by this observation, we propose an approach by combining challenge-based trust with nodes' behavioral trust. Our contributions can be summarized as below.

- We propose a hybrid trust management approach to improve the robustness of challenge-based CIDNs through combining challenge-based trust with behavioral trust of CIDN nodes. We adopt an algorithm to help identity suspicious nodes based on spatial correlation regarding traffic dropping and time delay. The main merit of such algorithm is that it does not require any prior knowledge about normal or malicious nodes.
- To derive the behavioral trust of nodes, we update the architecture of challenge-based CIDNs through developing a sub-component called behavioral trust under the trust management component, which can finally output an overall trust value for measuring the reputation of a node within the network.
- To investigate the performance, we test our approach under both simulated and real CIDN environments (i.e., co-working with an IT organization). The collected data and results demonstrate that our approach can enhance the robustness of challenge-based CIDNs against common, and even advanced insider attacks, by leveraging the spatial correlation in the neighborhood activities to measure the reputation of a node.

The structure of this paper is shown as follows. Section II introduces related research on distributed and collaborative intrusion detection. In Section III, we introduce the revised architecture of challenge-based CIDNs including major components and interactions. Section IV describes how we can calculate the behavioral trust based on spatial correlation among neighbor nodes, and how we can derive the overall trust to evaluate the reputation of a node. Section V discusses the environmental configuration and analyzes the experimental results. Specifically, we investigate the proposed approach under both simulated and real-world CIDN environments (in collaboration with an IT organization). Finally, we conclude the work in Section VI.

II. Related Work

In practice, how to improve the detection performance of an IDS against various threats is a long-term problem. A single detector would fail to identify some complex attacks without a global view of the network [42]. To address this issue, many distributed and collaborative detection approaches have been proposed like [2, 11, 37, 43]. A distributed or collaborative system can enable a single detector connecting with other detectors, and hence constructed a collaborative intrusion detection network (CIDN).

However, for the sake of the distributed architecture, such collaborative networks may be susceptible to insider attacks. For example, Li *et al.* [12] found that the scalability is a common issue for a majority of distributed IDSs. To mitigate this issue, they then introduced a framework of distributed IDS by considering the location and the routing structure. However, as the framework trusts each node, it cannot defeat insider threats.

To establish a suitable trust management scheme in CIDSs is an important and effective solution against insider attacks. An early P2P-based overlay IDS was proposed by Duma *et al.* [4], which could aggregate alarms via a trust engine and compute trustworthiness in an adaptive manner. A challenge-based CIDS was introduced by Fung *et al.* [7], which used a kind of message called *challenge* to explore the reputation of other nodes. They then introduced a Dirichlet-based trust management scheme to evaluate the node's trustworthiness by considering the mutual experience [8]. Li *et al.* [13, 14] introduced a trust management scheme based on intrusion sensitivity, which could derive the reputation by considering the capability of a node in identifying a particular attack. A Bayesian inferenc-based trust management was proposed by Meng *et al.* [29, 30] for medical smartphone networks, which could detect malicious internal nodes according to the traffic status, e.g., normal or abnormal. Sharma *et al.* [38] presented a trust management scheme for Mobile Internet of Things, by judging the energy consumption during the data transfer. Such scheme could help improve response time and reduce delays in authentication.

With the fast evolution of adversarial tools, many attackers focus on how to compromise existing trust mechanisms. For challenge-based CIDNs, *passive message fingerprint attack (PMFA)* [15, 20] is an advanced collusion attack, enabling malicious nodes to identify normal requests and maintain their trustworthiness by behaving faithfully to challenges. Another example is Special On-Off Attack (*SOOA*) [18, 19], in which a adversarial node can behave faithfully to some nodes, while acting untruthfully to other nodes. This would affect the trust computation process when a third node (target node) aggregates decisions. Meng *et al.* [28] described a kind of random poisoning attack on challenge-based CIDNs, in which a malicious node could send a challenge in a random possibility. Then, another kind of collusion attack - Bayesian Poisoning Attack [35] could figure out whether the received messages have a better possibility of being a normal request through using the aggregated appearance probability. Then, malicious nodes can make a response strategy to bypass the examination of challenges.

To further improve the robustness of trust management,

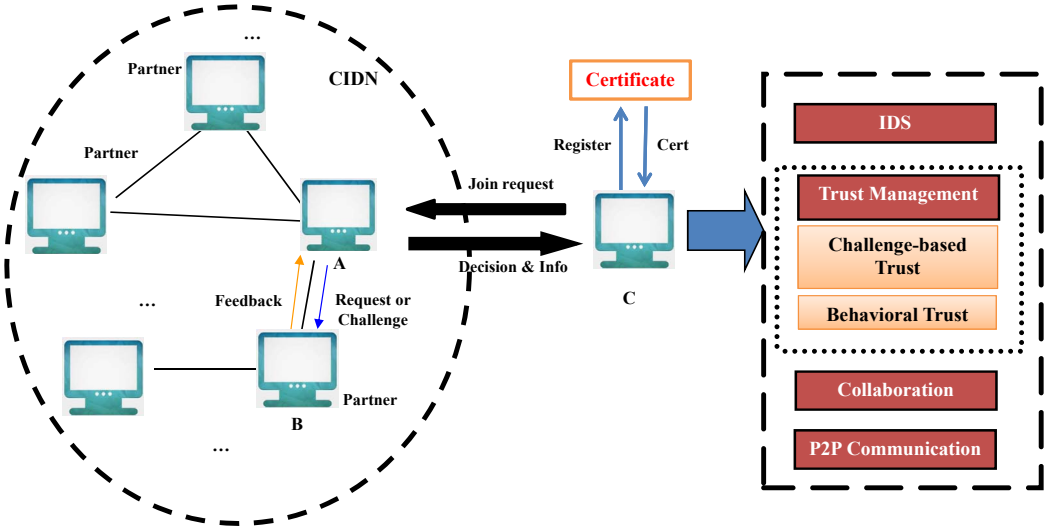

Figure 1. The high-level architecture of a challenge-based CIDN with our updated trust management component.

there is a need to consider additional level (types) of trust. For instance, a blockchain-based challenge-based CIDN [21] was proposed through leveraging the benefits of blockchain in checking the integrity of shared data, which could improve the detection of malicious internal nodes. The blockchain can also be used to build a blockchain-based trust [32, 36] to better evaluate the reputation of a CIDN node, by examining the deviation in the feedback pair between the challenge and the request.

Followings are some related studies on collaborative detection systems [6, 10], trust management scheme [1, 5, 41], trust-based filtration [30, 31], and trust enhancement such as machine learning [17, 27], and traffic sampling [33, 34].

III. Challenge-based CIDN Architecture with Updated Component

The main purpose of challenge-based trust mechanisms is to safeguard such collaborative detection systems against common insider attacks by evaluating and verifying a node's reputation via challenges in a cyclical manner. Based on the previous work [16], Fig. 1 depicts the high-level architecture of a challenge-based CIDN. The major components include an *IDS module*, *trust management component*, *collaboration component* and *P2P communication*.

- *IDS component* provides the basic function of identifying malicious events based on either signatures (rules) or pre-defined algorithms.
- *Trust management component* attempts to calculate the trust value for CIDN nodes. The challenge-based trust model can measure the trustworthiness by judging how the received responses suffice the prospective feedback. Challenge is a special kind of message that includes a set of IDS alarms for severity labeling. In this work, we further consider *an additional trust level of behavioral trust*, and thus update this component through adding a sub-component to help measure *behavioral trust*. The derivation of behavioral trust will be discussed later.
- *Collaboration component* is responsible for collecting and sharing necessary information (e.g., normal request, challenge), with the purpose of conjecturing the trustworthiness of a target node, by means of the received feedback. This component can be also used to send self responses to a received message. An example is given in Fig. 1, where node *B* will reply to node *A*'s messages, including both challenges and normal requests.
- *P2P communication.* This component aims to establish the physical connection with other CIDN nodes, and to provide management according to the defined policies and guidelines.

Network Interactions: To establish a partner list, each node can select and determine partner nodes based on their own criteria, during for a period of time. Then before joining a CIDN, a node needs to first get its credit like private and public keys, which can be obtained from a trusted certificate authority (CA) (see node *C* in Fig. 1). For a newly joined node, an initial list of partners can be provided. As described earlier, a CIDN node can use two types of messages within the network: namely normal request and challenge.

- *Normal request* is a kind of message utilized by a node to exercise the alarm aggregation, which is an essential characteristic of collaborative detection systems. When a node receives such message, it has to share the information (e.g., severity) on the queried alarms. Normally, only highly trusted nodes would be participated in the process of aggregating alarms.
- *Challenge* is a special kind of message used by a node to examine the reputation of another node in a CIDN.

It can consist of several alarms and be sent for severity labeling. The trustworthiness or satisfaction level can be derived by measuring how the received answers satisfy the expected feedback.

IV. Trust Computation

This section firstly describes how to calculate challenge-based trust and behavioral trust, and then introduce a single metric to integrate them.

A. Challenge-based Trust

In practice, a CIDN node with the challenge-based trust management model can send a challenge in an average rate of ε, according to the requirements. Usually, the rate should be low for those nodes who have a high trust value. For other nodes, the rate should be high in order to provide superior confidence in how the target node is trusted or not. To ensure the effectiveness, a pseudo random generation process can be used to send such challenges.

Node expertise. Different detectors may have their own detection superiority. This work thus accepts three levels of expertise: low (0.1), medium (0.5) and high (0.95), which can be formed by using the following function [8, 14].

$$\begin{aligned} f(p|\alpha,\beta) &= \frac{1}{B(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1} \\ B(\alpha,\beta) &= \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt \end{aligned} \tag{1}$$

where p ($\in [0,1]$) indicates how much it is possible to figure out an intrusion, $f(p|\alpha,\beta)$ describes how to compute the possibility by considering expertise level of l and difficulty level of d ($\in [0,1]$). Intuitively, a bigger l can result in a higher possibility, whereas a bigger d can decrease the possibility. The calculation of α and β is shown as below.

$$\begin{aligned} \alpha &= 1 + \frac{l(1-d)}{d(1-l)} r \\ \beta &= 1 + \frac{l(1-d)}{d(1-l)} (1-r) \end{aligned} \tag{2}$$

where $r \in \{0,1\}$ represents the prospective result. Given a fixed d, the detection performance of a node should depend on its particular level of proficiency. That is, an expert node should perform better than a non-expert node.

Node Trust Evaluation. As explained above, the ultimate trust value should rely on how the received answers satisfy the expected feedback. Based on the previous work [7, 17], the trust value of a node i according to node j can be derived as below.

$$T_{ct}^{i,j} = (w_s \frac{\sum_{k=0}^{n} F_k^{j,i} \lambda^{tk}}{\sum_{k=0}^{n} \lambda^{tk}} - T_s)(1-x)^d + T_s \tag{3}$$

where n indicates how many responses (or answers) are received, λ is a factor to allocate more weight on the recent response, $F_k^{j,i} \in [0,1]$ represents the satisfaction level on the received answer of k (number), and w_s is an adaptive weight: a) $w_s = \frac{\sum_{k=0}^{n} \lambda^{tk}}{m}$ if the received answers are less than a threshold of m; and b) $w_s = 1$ if the received answers are larger than m. x counts how many "don't know" answers are received within a given time slot, and d is a parameter to control how much punishment should be given.

B. Behavioral Trust

Challenge-based trust mechanism has proven to be sound to detect common insider attacks, but some advanced attacks would still be probable, such as PFMA [16]. In this work, we advocate that the integration of an additional trust level can help enhance the detection of advanced insider threats, since malicious insider nodes may often perform differently from a normal node, i.e., dropping many packets or delaying messages exchanged in a CIDN to prepare attacks.

In a real-world scenario, monitoring a node's behavior can benefit the trust evaluation. For instance, a node could have a high probability to be malicious if a great change occurs in its behavior as compared with normal nodes. This makes it feasible to identify insider attacks by analyzing the spatial correlation among partner nodes. As a study, we identify and revise a localized algorithm in prior work [22]. The selection is due to the following reasons.

- This algorithm does not require any former knowledge about either normal or malicious nodes. This is a very desirable and vital feature, as nodes' behavior could be very dynamic in practice (not easy to model).
- This algorithm can consider a node's behavior in many views such as packet sending rate, packet dropping rate, and forwarding time delay, etc.
- It could provide a balance between detection accuracy and false alarms.

Suppose a network deployed in a $a \times a$ square located in the two dimensional Euclidean plane R^2, with a number of NS sensors that are uniformly distributed. Let $NS_1(x)$ be a bounded closed set of R^2, which could be monitored by node x directly, as x's one-hop partner. Let $NS(x) \supseteq NS_1(x)$ denote another closed set of R^2, which includes node x and additional $n-1$ nearest nodes.

If node x monitors and assembles a dataset $D(X)$ about the partner node $NS(x)$, then we cannot ensure that $D(X)$ describes the truthful activities, especially when $NS_1(x) \subset NS(x)$ and $D(X)$ has indirect observations. Based on the direct observations, node x allocates a trust value to each partner node $x_i \in NS_1(x)$. Typically, let $T_{bt}(x_i) \in [0,1]$ denote the trust value of node x_i. Since we believe that nodes should behave similarly in the close proximity, $T_{bt}(x_i)$ can be derived based on the degree of how x_i's behavior differs from the partner nodes.

For each $x_i \in NS_1(x)$, node x calculates the maximum attribute component $D'_M(x_i) = \max\{D'_j(x_i)|1 \leq j \leq q)\}$, which indicates the biggest deviation of node x_i' behavior from the partner nodes. Then, the trust value $T_{bt}(x_i)$ can be computed as below.

$$T_{bt}(x_i) = \frac{D^m_M}{D'_M(x_i)} \tag{4}$$

where $D^m_M = \min\{D'_M(x_i)|x_i \in NS_1(x)\}$. More details like how to compute $D'_j(x_i)$ can refer to previous work [22].

C. Hybrid Trust

To facilitate the trust evaluation of a node, we introduce a single metric (T_{ovl}) to integrate both challenge-based trust and behavioral trust, which can be reckoned as below.

$$T_{ovl} = W_1 \times T_{ct} + W_2 \times T_{bt} \tag{5}$$

where W_1 and W_2 ($W_1 + W_2 = 1$) are weight utilized to control the emphasis on each trust. Given a threshold of t, we can determine the status of a node (malicious or not).

- If $T_{ovl} \geq t$, then the node is considered as trusted.
- If $T_{ovl} < t$, then the node is considered as malicious.

V. Evaluation

In this section, we launch two experiments to examine the capability of our approach under both a simulated and a real network environment.

- *Experiment-1.* This experiment establishes a simulated CIDN environment with the aim to study our approach under common insider attacks.
- *Experiment-2.* This experiment investigates the performance of our approach in a practical network environment, in collaboration with an IT organization.

In this work, we selected three typical factors aiming to compute the behavioral trust, such as message dropping rate, message sending rate and message time delay.

- *Message dropping rate.* After sending the message like packets and normal requests, testing node should check whether the partner nodes forward the message or not, i.e., via a buffer and watchdog [23].
- *Message sending rate.* The testing node should monitor and compute the amount of messages that tested nodes have sent out within a time period.
- *Message time delay.* The testing node can measure this metric by identifying the difference between the time of partner nodes receiving the message and the time of partner nodes forwarding it out.

A. Experiment-1

1) CIDN Settings: In the simulated environment, a total of 40 nodes were randomly distributed in a 10×10 grid area. The open-source IDS, namely Snort [39], was accepted in every CIDN node. The partner nodes could be discovered after a period of time. To facilitate the comparison, according to prior work [8, 15, 17], Table I summarizes the parameters for the simulated environment.

Table I
Parameters for the simulated CIDN environment.

Parameters	Value	Description
ε_l	10/day	Low request frequency
ε_h	20/day	High request frequency
λ	0.9	Forgetting factor
t	0.8	Trust threshold
T_s	0.5	Trust value for newcomers
m	10	Lower limit of received feedback
d	0.3	Severity of punishment

2) Satisfaction Measurement: Let $e \in [0, 1]$ denote the prospective response and $r \in [0, 1]$ denote the received response. Then we can use the following function F ($\in [0, 1]$) to measure the satisfaction level.

$$F = 1 - \left(\frac{e - r}{max(c_1 e, 1 - e)}\right)^{c_2} \quad e > r \tag{6}$$

$$F = 1 - \left(\frac{c_1(r - e)}{max(c_1 e, 1 - e)}\right)^{c_2} \quad e \leq r \tag{7}$$

where c_1 manages the severity of punishment for incorrect estimates, and c_2 manages the satisfaction sensitivity. A large c_2 means the satisfaction is more sensitive to the feedback. Similarly, we adopted $c_1 = 1.5$ and $c_2 = 1$.

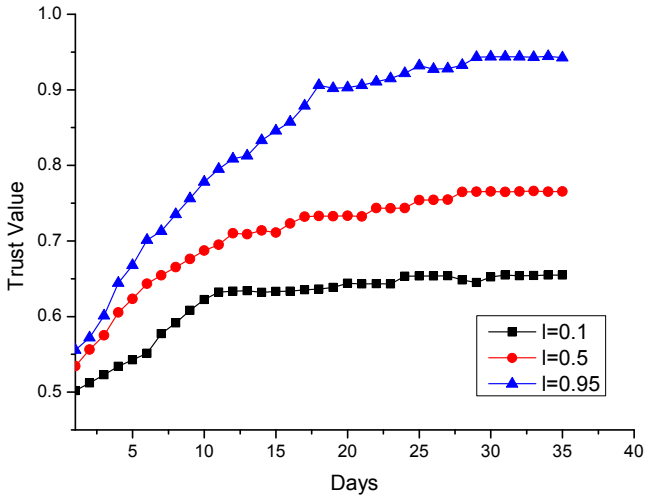

Figure 2. Convergence of trust values regarding three expertise levels.

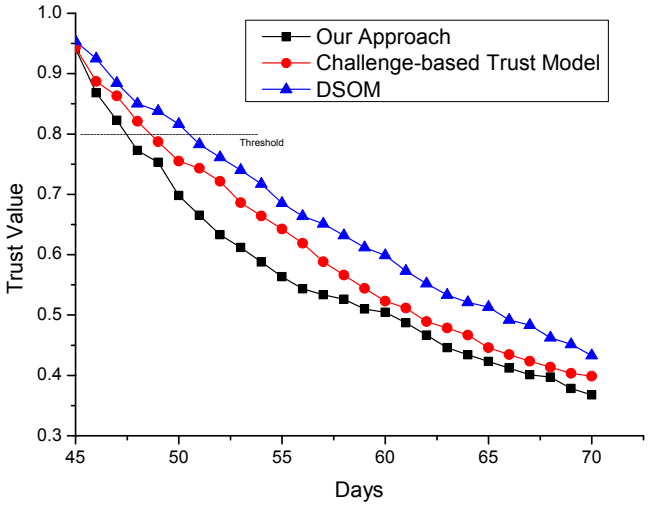

Figure 3. Trust values of malicious nodes under betrayal attack for different trust models.

3) Results and Analysis: In the simulated environment, we first explore the convergence of trust values. Fig. 2 shows the trust convergence according to particular expertise level of nodes. Intuitively, the nodes with higher expertise could reach better trust values, i.e., the high expertise nodes ($I = 0.95$) outperformed the other nodes. This result validates the observations in former research [8, 17].

Newcomer attack. Differently, under our configuration, the network required around 25 days to become stable (i.e., for converging trust values), which means up to 5 more days than the previous studies like [8, 15, 17]. This is because our hybrid trust combines two trust types, which evaluates the trustworthiness of a node in more aspects. That is, a node can be regarded as normal only if it can both respond correctly to the challenges and act truthfully with other partner nodes. This indicates that our approach is more robust to *newcomer attacks* (or called new entry attack), where an attacker tries to register a new entity to discard its bad history.

Betrayal attack. Newcomer attack is the first and basic step for performing a betrayal attack, where a trusted entity unexpectedly becomes malicious. Based on this, we randomly selected three nodes with $I = 0.95$ to conduct a betrayal attack. Fig. 3 indicates the trust values of malicious nodes on average under different trust models. In the comparison, we selected two similar approaches: DSOM trust model [4] and original challenge-based trust model [8].

It is visible that the reputation of betrayal nodes was declining under all three trust models. In particular, challenge-based trust model could outperform DSOM trust model by using a forgetting factor, which can emphasize the impact of neoteric behavior. While our approach could reach better performance than the other trust approaches by decreasing the reputation of malicious nodes quickly, i.e., the decrease of trust could be one day faster as compared with the original challenge-based trust model.

The results in the simulated environment demonstrate that our approach is realistic and encouraging to ameliorate the challenge-based trust mechanism against common insider attacks, i.e., our approach can improve the detection efficiently of identifying malicious nodes.

B. Experiment-2

In this experiment, we co-worked with an IT organization (with over 50 personnel) to study our approach in a practical wired CIDN environment including 43 nodes. The high-level network architecture is depicted in Fig 4. There is a DMZ and all CIDN nodes can connect with the Internet. Similarly, we employed the same parameter settings in the simulated experiment, and ran the whole environment to become stable (i.e., waiting for nodes' trust values to become converged). The experiment was supported and examined under the help from IT administrators in the participating organization due to privacy and security rules.

Newcomer attack. We mainly explore the nodes' reputation under newcomer attack among different trust models. Fig. 5 depicts the trust values among the DSOM trust model, the original challenge-based trust model, and our approach. It is identified that the convergence speed of DSOM is faster than others, which could reach the threshold more quickly. In comparison, our approach could delay this procedure by 5 days and 7 days than the challenge-based trust model and DSOM, respectively. This validates that our approach could defeat such kind of attack in practice.

Betrayal attack. In collaboration with the IT administrators, we selected a total of 6 nodes with high expertise to launch a betrayal attack by expressing untruthfully from the 51st day, i.e., by sending malicious packets, dropping traffic, delaying message transmit, etc. Fig. 6 describes the average trust value of malicious nodes for different trust models. The main observations are discussed below.

- It is visible that the trend of trust values could drop below the threshold under all trust models. The challenge-based trust model could outperform DSOM trust model, whereas our approach could provide better performance than the others, through decreasing the trustworthiness of malicious nodes below the threshold more quickly, i.e., 2 days and 3 days faster than the challenge-based trust model and DSOM model, respectively.
- The reputation gap between our approach and the original challenge-based trust model was not significant than that in the simulated experiment as observed in Fig. 3. This is mainly because the traffic is more dynamic in a real network environment.

Our evaluation demonstrates that our approach could be more robust to newcomer and betrayal attack than similar approaches in practice, by integrating nodes' behavioral trust based on spatial correlation. It is found that behavioral trust is more sensitive to a node's behavior than challenge-based

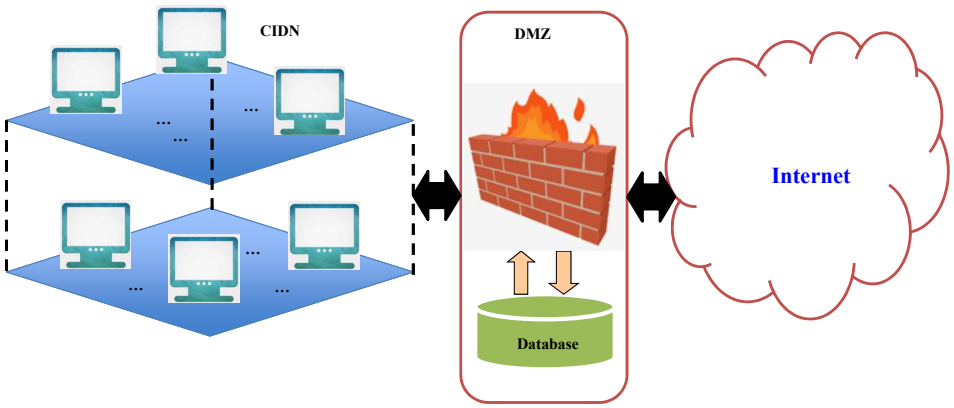

Figure 4. The high-level network architecture of the practical participating organization.

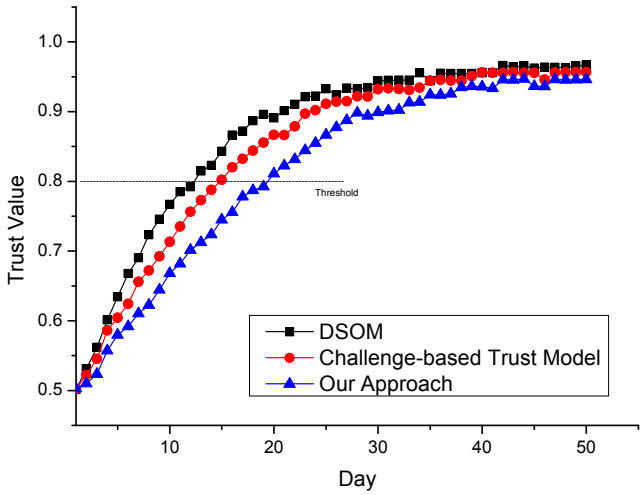

Figure 5. Trust values of malicious nodes under the newcomer attack for different trust models in the simulated CIDN.

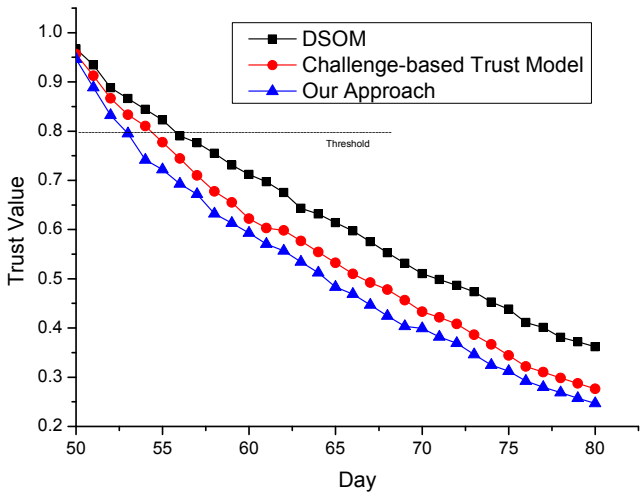

Figure 6. Trust values of malicious nodes under betrayal attack for different trust models in the real CIDN.

trust. If there is any abnormal action, the hybrid trust would be affected timely. The observations were also confirmed by the participating organization.

VI. Conclusion

To defend computer networks against threats, challenge-based CIDNs are a promising solution. However, such kind of trust mechanism may still be vulnerable to some advanced attacks. In this work, we focus on challenge-based CIDNs and advocate that the integration of an additional trust level can help strengthen its robustness. In particular, we propose an enhanced trust management model by considering behavioral trust that considers spatial correlation amongst different nodes, such as time delay, message sending rate and message dropping rate. In the evaluation, we explore the performance of our approach in both simulated and real network environments against newcomer and betrayal attacks. Experimental results demonstrate that our approach can outperform similar trust models in identifying malicious nodes, i.e., diminishing the reputation of treacherous nodes 2-3 days faster. Future work could include investigating the integration of other trust types in enhancing the challenge-based trust mechanism, and evaluating some advanced insider attacks like PMFA.

Acknowledgment

We would like to thank IT administrators from the participating organization for their assistance and support in deploying our mechanism. This work was partially supported by the National Natural Science Foundation of China (No. 61802077). K.-K. R. Choo was supported only by the Cloud Technology Endowed Professorship.

References

[1] F. Bao, I.R. Chen, M. Chang, and J.H. Cho, "Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection," *IEEE Transactions on Network and Service Management* 9(2), pp. 169-183, 2012.

[2] B. Chun, J. Lee, H. Weatherspoon, and B.N. Chun, "Netbait: A distributed worm detection service," Technical Report IRB-TR-03-033, Intel Research Berkeley, 2003.

[3] J. Douceur, "The sybil attack," *In: Proceedings of IPTPS*, vol. 2429. Springer, Heidelberg, 2002.

[4] C. Duma, M. Karresand, N. Shahmehri, and G. Caronni, "A Trust-Aware, P2P-Based Overlay for Intrusion Detection," *In: Proceedings of DEXA Workshop*, pp. 692-697, 2006.

[5] Z.M. Fadlullah, T. Taleb, A.V. Vasilakos, M. Guizani, and N. Kato, "DTRAB: combating against attacks on encrypted protocols through traffic-feature analysis," *IEEE/ACM Trans. Netw.* 18(4), pp. 1234-1247, 2010.

[6] I. Friedberg, F. Skopik, G. Settanni, and R. Fiedler, "Combating advanced persistent threats: From network event correlation to incident detection," *Computers & Security* 48, pp. 35-57, 2015.

[7] C.J. Fung, O. Baysal, J. Zhang, I. Aib, and R. Boutaba, "Trust Management for Host-Based Collaborative Intrusion Detection," *In: Proceedings of DSOM*, pp. 109-122, 2008.

[8] C.J. Fung, O. Baysal, J. Zhang, I. Aib, and R. Boutaba, "Robust and scalable trust management for collaborative intrusion detection," *In: Proceedings of the 2009 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, pp. 33–40, 2009.

[9] F. Gong, Next Generation Intrusion Detection Systems (IDS). McAfee Network Security Technologies Group, 2003.

[10] J. Hong and C.-C. Liu, "Intelligent Electronic Devices With Collaborative Intrusion Detection Systems," *IEEE Trans. Smart Grid* 10(1), pp. 271-281, 2019.

[11] A.P. Lauf, R.A. Peters, and W.H. Robinson, "A distributed intrusion detection system for resource-constrained devices in ad-hoc networks," *Ad Hoc Networks* 8(3), pp. 253-266, 2010.

[12] Z. Li, Y. Chen, and A. Beach, "Towards Scalable and Robust Distributed Intrusion Alert Fusion with Good Load Balancing," *In: Proceedings of the 2006 SIGCOMM workshop on Largescale attack defense (LISA)*, pp. 115-122, 2006.

[13] W. Li, Y. Meng, and L.F. Kwok, "Enhancing Trust Evaluation Using Intrusion Sensitivity in Collaborative Intrusion Detection Networks: Feasibility and Challenges," *In: Proceedings of the 9th International Conference on Computational Intelligence and Security (CIS)*, pp. 518-522, 2013.

[14] W. Li, Y. Meng, and L.F. Kwok, "Design of Intrusion Sensitivity-Based Trust Management Model for Collaborative Intrusion Detection Networks," *In: Proceedings of IFIP International Conference on Trust Management (IFIPTM)*, pp. 61-76, 2014.

[15] W. Li and Y. Meng, "Enhancing Collaborative Intrusion Detection Networks Using Intrusion Sensitivity in Detecting Pollution Attacks," *Information and Computer Security* 24(3), pp. 265-276, 2016.

[16] W. Li, W. Meng, L.F. Kwok, and H.H.S. Ip, "PMFA: Toward Passive Message Fingerprint Attacks on Challenge-based Collaborative Intrusion Detection Networks," *In: Proceedings of the 10th International Conference on Network and System Security*, pp. 433-449, 2016.

[17] W. Li, W. Meng, L.F. Kwok, H.H.S. Ip, "Enhancing Collaborative Intrusion Detection Networks Against Insider Attacks Using Supervised Intrusion Sensitivity-Based Trust Management Model," *Journal of Network and Computer Applications*, vol. 77, pp. 135-145, 2017.

[18] W. Li, W. Meng, and L.F. Kwok,, "SOOA: Exploring Special On-Off Attacks on Challenge-Based Collaborative Intrusion Detection Networks," *In: Proceddings of GPC*, pp. 402-415, 2017.

[19] W. Li, W. Meng, and L.F. Kwok, "Investigating the Influence of Special On-Off Attacks on Challenge-based Collaborative Intrusion Detection Networks," *Future Internet*, vol. 10, no. 1, pp. 1-16, 2018.

[20] W. Li, W. Meng, L.-F. Kwok, and H.H.S. Ip, "Developing Advanced Fingerprint Attacks on Challenge-based Collaborative Intrusion Detection Networks," *Cluster Computing*, vol. 21, no. 1, pp. 299-310, 2018.

[21] W. Li, Y. Wang, J. Li, and M.H. Au, "Towards A Blockchain-based Framework for Challenge-based Collaborative Intrusion Detection," *International Journal of Information Security*, pp. 1-13, 2020. https://doi.org/10.1007/s10207-020-00488-6

[22] F. Liu, X. Cheng, and D. Chen, "Insider Attacker Detection in Wireless Sensor Networks," *In: Proceddings of INFOCOM*, pp. 1937-1945, 2007.

[23] S. Marti, T.J. Giuli, K. Lai, and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks," *In: Proceddings of ACM MOBICOM*, pp. 255-265, 2000.

[24] McAfee, Threats Predictions Report, 2019. https://securingtomorrow.mcafee.com/other-blogs/mcafee-labs/mcafee-labs-2019-threats-predictions/

[25] Y. Meng and L.F. Kwok, "Enhancing False Alarm Reduction Using Voted Ensemble Selection in Intrusion Detection," *International Journal of Computational Intelligence Systems*, vol. 6, no. 4, pp. 626-638, 2013.

[26] W. Meng, W. Li, and L.F. Kwok, "EFM: Enhancing the Performance of Signature-based Network Intrusion Detection Systems Using Enhanced Filter Mechanism," *Computers & Security*, vol. 43, pp. 189-204, 2014.

[27] W. Meng, W. Li, and L.F. Kwok, "Design of Intelligent KNN-based Alarm Filter Using Knowledge-based Alert Verification in Intrusion Detection," *Security and Communication Networks* 8(18), pp. 3883-3895, 2015.

[28] W. Meng, X. Luo, W. Li, and Y. Li, "Design and Evaluation of Advanced Collusion Attacks on Collaborative Intrusion Detection Networks in Practice," *In: Proceedings of the 15th TrustCom*, pp. 1061-1068, 2016.

[29] W. Meng, W. Li, Y. Xiang, and K.K.R. Choo, "Bayesian Inference-based Detection Mechanism to Defend Medical Smartphone Networks Against Insider Attacks," *Journal of Network and Computer Applications* 78, pp. 162-169, 2017.

[30] W. Meng, W. Li, and L.F. Kwok, "Towards Effective Trust-based Packet Filtering in Collaborative Network Environments," *IEEE Transactions on Network and Service Management*, vol. 14, no. 1, pp. 233-245, 2017.

[31] W. Meng, W. Li, and L.F. Kwok, "Towards Effective and Robust List-based Packet Filter for Signature-based Network Intrusion Detection: An Engineering Approach," *HKIE Transactions*, vol. 24, no. 4, pp. 204-215, 2017.

[32] W. Meng, E.W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When Intrusion Detection Meets Blockchain Technology: A Review," *IEEE Access*, vol. 6, no. 1, pp. 10179-10188, 2018.

[33] W. Meng, W. Li, C. Su, J. Zhou, and R. Lu, "Enhancing Trust Management for Wireless Intrusion Detection via Traffic Sampling in the Era of Big Data," *IEEE Access*, vol. 6, no. 1, pp. 7234-7243, 2018.

[34] W. Meng, "Intrusion Detection in the Era of IoT: Building Trust via Traffic Filtering and Sampling," *IEEE Computer* 51(7), pp. 36-43, 2018.

[35] W. Meng, W. Li, L. Jiang, K.K.R. Choo, and C. Su, "Practical Bayesian Poisoning Attacks on Challenge-Based Collaborative Intrusion Detection Networks," *In: Proceedings of the 24th European Symposium on Research in Computer Security (ESORICS)*, pp. 493-511, 2019.

[36] W. Meng, W. Li, L.T. Yang, and P. Li, "Enhancing Challenge-based Collaborative Intrusion Detection Networks Against Insider Attacks using Blockchain," *International Journal of Information Security*, vol. 19, no. 3, pp. 279C290, 2020.

[37] T. Peng, C. Leckie, and K. Ramamohanarao, "Information sharing for distributed intrusion detection systems," *Journal of Network and Computer Applications* 30(3), pp. 877-899, 2007.

[38] A. Sharma, E.S. Pilli, A.P. Mazumdar, and P. Gera, "Towards trustworthy Internet of Things: A survey on Trust Management applications and schemes," *Comput. Commun.* 160, pp. 475-493, 2020.

[39] Snort: an open source network intrusion prevention and detection system (IDS/IPS). Homepage: http://www.snort.org/

[40] Symantec, Internet Security Threat Report, 2019. https://www.symantec.com/content/dam/symantec/docs/reports/istr-24-executive-summary-en.pdf

[41] T.A. Tuan, "A Game-Theoretic Analysis of Trust Management in P2P Systems," *In: Proceedings of ICCE*, pp. 130-134, 2006.

[42] Y.-S. Wu, B. Foo, Y. Mei, and S. Bagchi, "Collaborative Intrusion Detection System (CIDS): A Framework for Accurate and Efficient IDS," *In: Proceedings of ACSAC*, pp. 234-244, 2003.

[43] Y. Zhang, L. Wang, W. Sun, R.C. Green II, and M. Alam, "Distributed Intrusion Detection System in a Multi-Layer Network Architecture of Smart Grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 796-808, 2011.