Computers and Electrical Engineering 111 (2023) 108884

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A blockchain-assisted security management framework for collaborative intrusion detection in smart cities ☆

Wenjuan Li [a,b,*], Christian Stidsen [c], Tobias Adam [c]

[a] Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, China
[b] Institute of Artificial Intelligence and Blockchain, Guangzhou University, China
[c] Faculty of IT and Design, Aalborg University, Denmark

ARTICLE INFO

Keywords:
Intrusion detection
Trust management
Blockchain technology
Internet of Things
Smart city

ABSTRACT

Aiming to safeguard a decentralized setup such as smart cities, collaborative intrusion detection system (CIDS) has become a mainstream security mechanism to protect different types of computer networks, especially decentralized computing platforms such as Internet of Things (IoT). The main benefit of CIDS relies on the information sharing process among devices, nodes, software and hardware entities. However, traditional CIDS often requires a trusted third partner, e.g., a centralized computing server, to help build up a trusted communication channel among various entities. Such requirement is not practical in real-world implementation, making the integrity of shared information compromised easily. With the wide adoption, blockchain technology has given a solution to protect the distributed/collaborative detection system. In the current market, blockchain technology has been extensively researched across many detection scenarios, but there is a need to explore how such technology can overall contribute to CIDS and a general distributed detection system. In this work, we introduce a blockchain-assisted security management framework for CIDS, which summarizes and provides an integrated protection given by blockchain. In the case study, we evaluate our proposed framework in both a simulated and a real CIDS setup with challenge-based mechanism. The results demonstrate the promising benefits provided by blockchain in CIDS.

1. Introduction

Internet of Things (IoT) is currently driving the trend of society digitization in many novel ways, such as autonomous machines, self-driving vehicles, remote medical diagnosis, and more. According to Statista report,[1] the IoT devices worldwide were estimated to be almost increased from 9.7 billion in 2020 to over 29 billion by the end of 2030. Also, the report predicted that China will have the largest IoT market with around 5 billion consumer devices at that time. Similarly, the report from IoT-Analytics expected the global IoT market to reach 14.4 billion active connections with a growth rate of 18%. By the end of 2025, due to the impact from supply constrains and COVID-19, up to 27 billion IoT devices could be Internet-enabled for easy control and communication.[2]

Though providing various convenience, IoT devices and networks are often suffered from many security threats [1], including ransomware, distributed-denial-of-service (DDoS) attacks, botnet, and the expanded attack surface of threats. For protection,

☆ This paper is for special section VSI-spci. Reviews were processed by Guest Editor Dr. Weizhi Meng and recommended for publication.

* Corresponding author at: Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, China.
E-mail address: wenjuan.li@polyu.edu.hk (W. Li).

[1] https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/.
[2] https://iot-analytics.com/number-connected-iot-devices/.

https://doi.org/10.1016/j.compeleceng.2023.108884
Received 9 May 2023; Received in revised form 18 July 2023; Accepted 20 July 2023
Available online 17 August 2023
0045-7906/

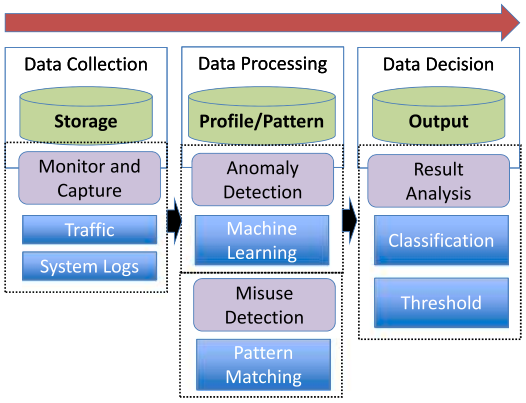

Fig. 1. The workflow of an IDS with rule-based scheme and anomaly-based scheme.

intrusion detection system (IDS) is one typical and essential solution, which can monitor the whole network and identify malicious events. Basically, an IDS can be generally classified into two types: rule-based scheme and anomaly-based scheme (see Fig. 1). The former can identify potential attacks via signature matching process such as exclusive matching [2] and regular expression matching [3]. On the other hand, the latter can detect unusual events via building normal profiles, which can use machine learning [4] or statistic method [5]. A threshold will be then used to determine whether the current event is malicious or not. These two detection approaches can be integrated as well.

With the network structure to be distributed, a traditional IDS cannot identify many complex attacks such as DDoS attacks, hence several IDS nodes should cooperate to exchange required information—namely distributed IDS (DIDS) or collaborative IDS (CIDS). For example, Wu et al. [6] introduced an early CIDS framework with multiple IDS nodes by layers including application, network and kernel level. The cooperation between IDS nodes can help detect complicated threats. A kind of CIDS based on intrusion sensitivity was proposed by Li et al. [7], which can measure a node's detection capability in the aspect of identifying particular attacks. This aims to actually give more weight on the expert nodes.

IoT is the basis for building smart city environments, hence CIDS can help protect such distributed environment, by allowing different sensors to sense and exchange predefined data. For example, Meng et al. [2] presented that a more efficient signature-matching process–exclusive signature matching can degrade the computing requirement and maintain the detection accuracy for a CIDS in the smart city environment.

Motivation. The key feature of a CIDS is to allow the data and information exchanged with each other in the network, but an inside intruder may have the chance to manipulate the exchanged information in-between. For example, an attacker can perform a multiple-mix-attack to modify the exchanged data [8]. Thus, it is a big problem when sharing the data via a channel, especially how to ensure the data integrity. For this issue, trust management is believed to be an effective method of detecting insider attacks, but it also needs to monitor the shared information via the channel. With the development of blockchain-based solutions, it provides a new and emerging opportunity to enhance the overall robustness of CIDS. The main benefits provided by blockchain are: (1) it does not need a trusted third partner and allows the communication to be enabled among various nodes, and (2) the chain-like structure can identify any unauthorized changes to the on-chain data.

Contributions. In the literature, the main blockchain applications have been explored in collaborative intrusion detection, such as building a trusted pattern database [9] and learning model [10]. In practice, we observe that blockchain has a potential to provide a multi-level benefit to the life-circle of collaborative intrusion detection. Motivated by this trend, in this work, we summarize the existing relevant studies and introduce a blockchain-assisted security management framework for a CIDS. Our contributions can be listed as follows.

- We introduce a blockchain-assisted security management framework for a CIDS that can handle and improve traffic filtration, rule database construction, learning model sharing, and false alarm filtration. With the integration of InterPlanetary File System (IPFS)–kind of distributed file storage protocol, our framework can be flexible and scalable in real-world implementation.
- Then we perform a study, e.g., considering challenge-based CIDS, and implement the framework in collaboration with an IT company. The experimental results indicate that our blockchain-assisted framework can enhance the CIDS performance in the aspects of detection effectiveness and robustness.

It is worth noting that our framework can enhance the performance of CIDS in various aspects, such as traffic filtration, rule database setup, learning model sharing, and false alarm filtration. These make our framework different from existing research studies.

The reminder of this work is structured as follows. Section 2 reviews state-of-the-art on trust management and blockchain in collaborative intrusion detection. Section 3 presents the proposed blockchain-assisted security framework in detail, including typical structure, framework layers and a case implementation. In Section 4, we show an evaluation of our framework in a simulated and a real CIDS network. We describe and discuss open challenges and future work in Section 5. Section 6 provides a conclusion.

2. Related work

In this section, we first introduce some challenges for CIDS (e.g., trust management and insider attacks), and then discuss the blockchain-enabled solutions. Generally, CIDS can be classified into three types:

- Centralized system: This kind of CIDS will feature a central server to help collect and analyze the data within the network.
- Decentralized system: This kind of CIDS will have several main servers to help collect and analyze the data at multiple level within the network.
- Distributed system: This kind of CIDS does not contain any central server, and each CIDS node can collect and analyze the monitored data.

2.1. CIDS: Trust management and challenges

The capability of sharing information among IDS nodes is the main benefit for a CIDS, but this also makes it suffered from insider attacks (e.g., betrayal attack, newcomer attack), similar to many other distributed infrastructures [6].

Trust management. For this challenge, trust management is believed to be a promising solution. *Trust* can be regarded as a belief level that one entity can treat on another entity for some particular actions via either direct or indirect observations. The design of a trust-based CIDSs has been widely studied by the society.

Fung et al. [11] introduced a kind of challenge-based CIDN, which can compute the trustworthiness of a node by measuring the received feedback according to the expected feedback. Liu et al. [8] introduced an approach—Perceptron Detection (PD), which could utilize both perceptron and K-means method to measure the reputation of an IoT node and detect a multiple-mix-attack where insider attackers can perform actions collaboratively. They further proposed a method called Distributed Consensus based Trust Model (DCONST), which could decide whether an IoT node is trusted through sharing a special type of information called cognition [12]. Meng et al. [13] put an emphasis on the detection of insider attacks in healthcare SDNs, and devised a trust-based approach based on Bayesian inference. Their results indicated that their approach could reduce the trustworthiness of malicious nodes faster than similar approaches. Marche and Nitti [14] introduced a trust model in a decentralized way via a learning algorithm, by considering the scores and measuring usefulness, perseverance and goodness.

For most traditional schemes, all nodes have the same weight on the final decision, which may degrade the effectiveness of trust management. To improve this issue, Li et al. [7] proposed a notion of *intrusion sensitivity*, aiming to measure the capability of a detection method by given a special attack type. It can intuitively give more weight on the contributions from expert nodes. With such method, it could provide better detection performance and sensitivity, as compared with traditional approaches.

Advanced insider threat. As mentioned in previous work et al. [6], insider attackers can perform many malicious actions, i.e., kicking out a normal node from the network. Also, most existing trust management schemes can defend against common insider attacks such as newcomer attack, but may be vulnerable to advanced insider tricks. Meng et al. [15] presented an advanced insider attack called *random poisoning attack*, in which a malicious action can be performed with a possibility, which can greatly impact the effectiveness of trust measurement.

Li et al. [16] showed a more complicated collusion attack - *Passive Message Fingerprint Attack (PMFA)*, where the malicious nodes can collect the messages and identify the alarm ranking request. Then these malicious nodes can provide a false response to degrade the alarm aggregation. They also introduced another type of collusion attack called *Special On–Off Attack* (*SOOA*), which can provide truthful feedback to selected node(s), but will give false information to other nodes [17]. This will cause a conflict between nodes within the same network. Meng et al. [18] introduced *Bayesian Poisoning Attack* by using the Bayesian model to describe the appearance possibility of normal messages, which could achieve a high success rate of bypassing the detection.

With the development of Social Internet of Things (SIoT) paradigm [19], the above advanced threat may cause a large impact. This requires to develop corresponding security mechanisms. For example, Salimitari et al. [20] presented a composite framework to defend against On–Off attack by considering two types of fusion rule at IoT hub: namely optimistic and conservative. Li et al. [21] introduced a lightweight message verification approach through adding a verification alarm into each normal request. As it is hard for attackers to guess the correct location of the verification alarm, any untruthful feedback can be identified. Liu et al. [22] formulated the insider threat as a multivariate multiple linear regression problem and applied the K-means method to detect and classify malicious nodes.

More related work regarding trust-based collaborative intrusion detection can refer to a recent survey [23].

2.2. Blockchain in CIDS

Blockchain technology has been studied in CIDS, by providing benefits in the aspects of rule database construction and learning model sharing [24].

Meng et al. [25] introduced a blockchain-enabled method for challenge-based CIDNs against advanced insider attacks. It develops a special kind of blockchain-based trust. They also examined their approach in both simulated and real network environments. Cao et al. [26] design a token-based access control mechanism for smart contracts with an IDS for identifying attacks against smart contracts. Khan et al. [10] focused on Unmanned Aerial Vehicle (UAV) and introduced a decentralized machine learning framework based on blockchain, which could improve the data integrity and storage for intelligent decision making between several UAVs. The blockchain is used to achieve decentralized predictive analytic.

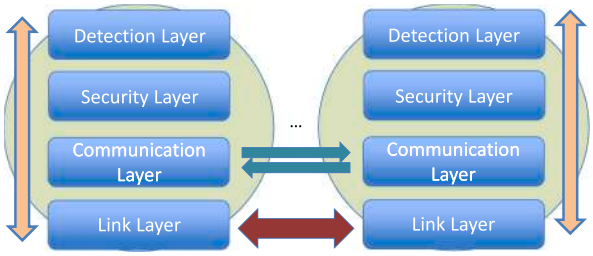

Fig. 2. The typical architecture of a CIDS with major components.

Sarhan et al. [27] introduced a hierarchical blockchain-based federated learning framework for a CIDS in IoT environment, with secure and privacy-preserved collaboration. It allowed sharing cyber threat intelligence among different IoT networks and providing better detection performance. They considered model updates as transactions and the smart contract will verify the conformance of performed tasks. A hybrid blockchain-enabled detection model was also proposed [28], which could exchange signatures from one node to the other for a CIDS. It also used a cryptosystem to encrypt the data stored in blocks to improve the security aspect.

3. Our proposed approach

In this section, we describe the typical architecture of a CIDS, and then present our proposed blockchain-assisted security management framework. We later provide an implementation case based on challenge-based CIDS.

3.1. Typical architecture of CIDS

Fig. 2 illustrates the typical architecture and major components of a CIDS, such as detection layer, security layer, communication layer, and link layer.

- *Detection Layer.* This is the top layer in a CIDS node, which is responsible for examining incoming events (e.g., traffic) with particular detection methods (e.g., rule-based or anomaly-based detection). Alarm aggregation is an important decision making process in this layer. This layer can also include various additional schemes to enhance the detection performance such as traffic filter and alarm filter.
- *Security Layer.* This layer is responsible for protecting the security aspect of a CIDS. For example, we can implement a trust model in this layer to identify insider attacks and malicious events or input. It can also implement particular schemes to identify sensitive data between IDS nodes, i.e., a query component can be used to compute the intrusion sensitivity of a node. For instance, given three sensitivity levels, e.g., such as high, medium and low, the testing node can derive the sensitivity value based on the responses received from other nodes.
- *Communication Layer.* This layer aims to help exchange pre-defined information between IDS nodes in a CIDS. For example, it can help send a request for alarm aggregation and then can get back the relevant responses. Also, it is able to collect the required data for performing trust evaluation in the network. For a challenge-based CIDS, this layer can deliver a challenge to the tested node and then wait for the returned response.
- *Link Layer.* This is the bottom layer, which is responsible for boosting the physical connection among nodes (i.e., building a peer-to-peer link), and keep the link available and stable.

Node Registration. To establish a list of peers and nodes, each node in a CIDS can select their peers based on defined rules. To enable a node to join such network, there is a need to get its identity and contact the surrounding nodes. Based on the predefined policies, a decision can be made whether to allow the joining application.

Exchanged Messages. In a CIDS, several types of messages can be exchanged among nodes, including consultation messages, information messages (or construction messages) and response message [23].

- *Consultation message* is normally sent by a node to request an alarm ranking or alarm severity during alarm aggregation process. This is a distinguished feature of a CIDS, where the inputs from trusted nodes will be considered. After receiving this type of message, the recipient has to get back a list of alarms or severity rank based on the requirement.
- *Construction message* is often constructed by a node to exchange required information or data for either external detection or insider detection. For example, when designing a trust management scheme, a node can build up a construction message and collect the information to evaluate a node's reputation. For challenge-based detection, challenges—a type of message can be used to explore the trustworthiness of other nodes and to derive a trust value by analyzing the returned feedback.
- *Response message* is actually a corresponding response to the received messages, i.e., offering a response to either a consultation message or a construction message.

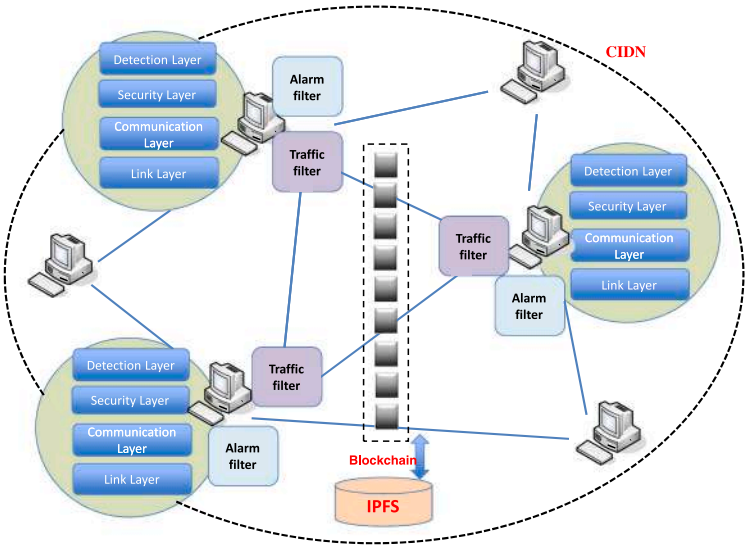

Fig. 3. Our proposed blockchain-assisted security management framework for a CIDS or collaborative intrusion detection network (CIDN).

3.2. Blockchain-assisted security management framework

In the literature, many studies tried to apply blockchain technology into some aspects of CIDS such as rule sharing and learning model sharing, but we figure out that blockchain can also be beneficial to the overall performance of detection including traffic filtration and alarm reduction. Motivated by our observation, we extend the state-of-the-art and propose a blockchain-assisted security management framework for a CIDS, as shown in Fig. 3.

As a security management framework, our proposed structure can interact with blockchain and provide benefits to not only threat detection, but also traffic filtration and false alarm reduction.

- *Threat detection.* For external attacks, our framework allows to establish a robust rule database and learning model via the blockchain, which can enhance the detection accuracy. For insider attacks, our framework allows to build a robust trust management scheme by interacting with blockchain. With the help from IPFS, our framework can host and share some small pieces of information among nodes.
- *Malicious traffic filtration.* With the increasing size of computer networks, Internet traffic has become large, especially under the era of big data and future Internet. Our framework can enable traffic filtration by using blockchain and IPFS. For example, a more robust blacklist can be established by sharing information among nodes, and the IPFS can host a small part of blacklisted data or sources [29].
- *False alarm reduction.* In practice, false alarms (false positives) can greatly degrade the detection performance and effectiveness. Hence reducing false positives can provide many benefits, e.g., lowering the workload of analysts and saving budget of hiring additional domain experts. Our framework can help establish false alarm filter via blockchain and IPFS, which can refine the output from a detector.

Overall, our proposed framework offers many benefits of using blockchain in collaborative intrusion detection, especially ensuring the integrity of sharing data and information with each other.

3.3. An implementation case

To investigate our framework's performance, we consider a user study with challenge-based CIDS and realize our blockchain-assisted security management framework as depicted in Fig. 4.

Trust evaluation in challenge-based CIDS. To compute the trust value, a testing node should periodically deliver a challenge message to the tested node via a random generation process. Based on the previous similar research [11,16,17,25], the trust evaluation under challenge-based mechanism, e.g., node i according to node j, can be derived as below:

$$T_{value}^{i,j} = w_s \frac{\sum_{k=0}^{n} F_k^j \lambda^{tk}}{\sum_{k=0}^{n} \lambda^{tk}} \tag{1}$$

In the above equation, n describes the totally received responses, w_s indicates a *significant value* depending on the totally received responses. There is a condition: if the received responses are lower than a predefined value m, then $w_s = \frac{\sum_{k=0}^{n} \lambda^{tk}}{m}$; otherwise w_s is 1.

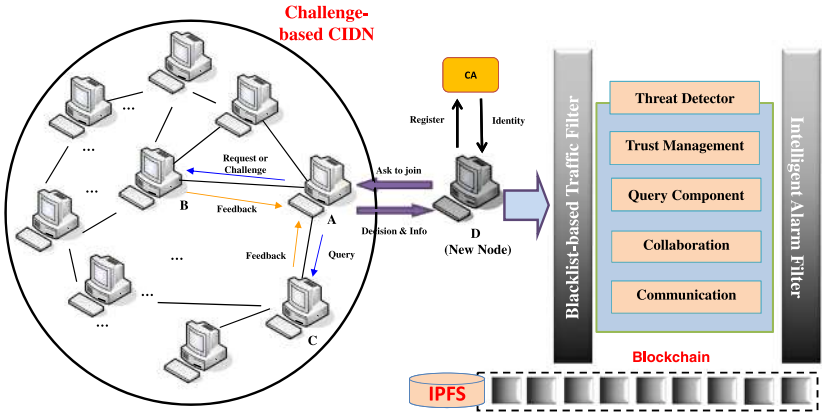

Fig. 4. An implementation case of our framework.

λ indicates a *forgetting factor* by giving more weight to recent actions and degrade the impact of historical data. Finally, $F_k^j \in [0,1]$ indicates the satisfaction level by considering k responses.

In terms of weighted majority approach, we can calculate a node's reputation, e.g., node j, as below.

$$T_j = \frac{\sum_{T \geq r} T_{value}^{i,j} D_i^j I_s^i}{\sum_{T \geq r} T_{value}^{i,j} D_i^j} \tag{2}$$

In the above equation, r means a trust threshold, $I_s^i (\in [0,1])$ describes the intrusion sensitivity of node i, $D_i^j (\in [0,1])$ describes the hops between two nodes, and $T_{value}^{i,j} (\in [0,1])$ computes the trustworthiness of node i according to node j.

Satisfaction measurement. To derive the satisfaction level, we have to consider two items: the received response ($r \in [0,1]$) and the expected response ($e \in [0,1]$). It is similar to an exam scenario, the teacher, who knows the right answers, will expect the correct answer from the students. Here, we can use a function F ($\in [0,1]$) to calculate the satisfaction:

$$F = 1 - \left(\frac{e-r}{max(c_1 e, 1-e)}\right)^{c_2} \quad e > r \tag{3}$$

$$F = 1 - \left(\frac{c_1(r-e)}{max(c_1 e, 1-e)}\right)^{c_2} \quad e \leq r \tag{4}$$

Here if we want to modify the penalty levels for wrong estimates, we can adjust c_1. If we want to modify the sensitivity, we can adjust c_2. For better comparison with former studies such as [11], we selected $c_1 = 1.5$ and $c_2 = 1$.

Query component: Intrusion sensitivity allocation. As mentioned earlier, intrusion sensitivity is used to judge whether a CIDS node is sensitive to some special cyber-attacks. To derive the sensitivity value, we can use a kind of messages—queries and request other nodes for help. After getting a set of responses, we can use supervised machine learning tools to assign the value for each CIDS node. The whole procedure is similar to do a classification including training and classification [7]:

- *Training step.* In this step, we need to obtain a set of labeled data and then decide suitable classifiers. Each classifier can be used to establish a learning model.
- *Classification step.* In this step, we use the pre-established model to assign sensitivity value for each CIDS ndoe.

Blacklist-based traffic filtration. As a study, we deploy a blacklist packet filter [30] to identify unwanted traffic and reduce the workload for a CIDS. For the incoming packets, we have to check their IP address according to the blacklist: if there is a match, then the packet will be considered as malicious.

To avoid a high false rates, we compute the IP reputation by adopting a weighted ratio-based method, as shown in Eq. (5). Here W indicates a weight value, k indicates how many malicious packets and i indicates the total amount of benign packets.

$$IP\ reputation = \frac{i}{\sum_1^k W \times k} \quad (i, k \in \mathbf{N}) \tag{5}$$

Intelligent false alarm filter. As a study, we adopted an intelligent false alarm filter [31] that can refine alarm positives by choosing the most suitable classifiers. For the steps, the filter firstly define the necessary features from the alarms, and then pre-processes all the data according to the format. Then algorithm training and testing will be performed and the most suitable algorithm will be decided. The filter output will treat as true alarms.

Table 1
Parameter settings in the simulated environment.

Parameters	Value	Description
λ	0.9	Forgetting factor
$T_{dir,initial}$	0.5	Trust value for newcomers
r	0.8	Trust threshold
μ_1	15/day	Arrival rate for challenges
μ_2	5/day	Arrival rate for queries
$k1$	5	Satisfaction levels
m	10	Lower limit of received feedback
$k2$	10	Intrusion sensitivity levels

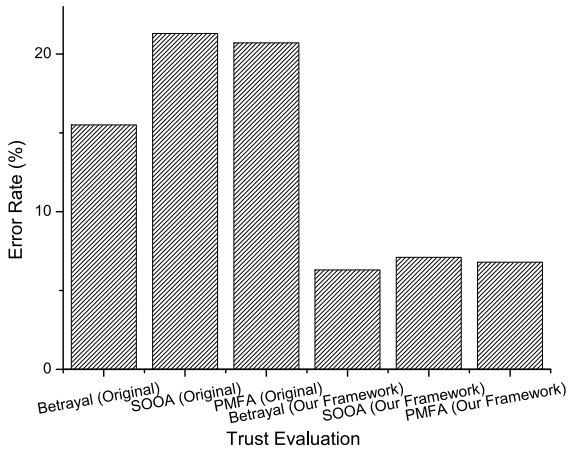

Fig. 5. The error rate on trust evaluation.

4. Evaluation and discussion

This section investigates the performance of our approach under both a simulated and a real CIDS environment. There are two experiments:

- *Experiment-1:* Under this test, we setup a simulated CIDS network and explore the framework performance (e.g., error rate) under different adversarial scenarios, such as betrayal attack, Special On–Off Attack (SOOA), and Passive Message Fingerprint Attack (PMFA).
- *Experiment-2:* In this test, we investigated the framework performance (e.g., error rate) in a real network environment under the above adversarial conditions.

For the CIDS, we employed Snort (http://www.snort.org/) as the typical detector. We also adopted three priority levels for the produced alarms: high, medium and low.

4.1. Experiment-1

The purpose of this experiment is to study the framework performance in the aspect of error rate under attacks, relating to trust computation, query allocation, traffic filtration, and alarm reduction. Similar to previous work [7], we adopted *intrusion sensitivity* (I_s^i) to be 10 levels including bad (0.1), very low (0.2), low (0.3), not good (0.4), neural (0.5), good (0.6), high (0.7), very high (0.8), excellent (0.9), and expert (1.0). The CIDS environment consisted of 15 nodes, and Table 1 summarizes the settings and parameters.

During the experiment, we launched three types of attacks: (1) betrayal attack where a trusted node becomes malicious, (2) SOOA: a malicious node behaves truthfully from time to time, and (3) PMFA: several malicious nodes work together to share false information. The results are shown in Figs. 5 6 7 8.

- *Trust evaluation.* Fig. 5 depicts the error rate on trust evaluation under different attacks. It is found that the overall errors were ranged from 15% to 24%. For instance, the error rate was around 17% under betrayal attack but increased to over 20% under SOOA and PMFA. This is because challenge-based CIDS could identify betrayal attack by measuring the received responses to challenges. By contrast, the errors under our framework could reach a value below 8%, thanks to the blockchain technology, i.e., detecting untruthful feedback via consensus process among nodes.

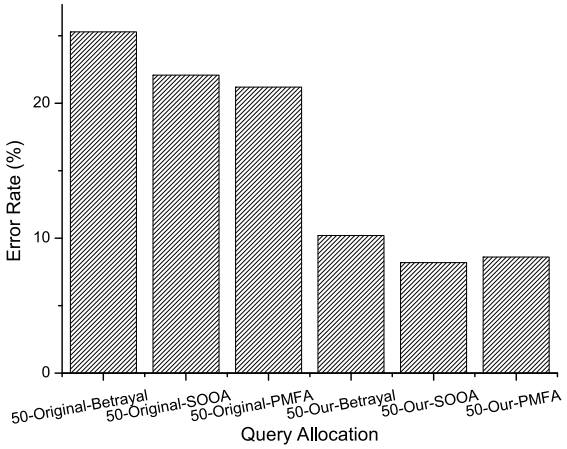

Fig. 6. The error rate on query allocation.

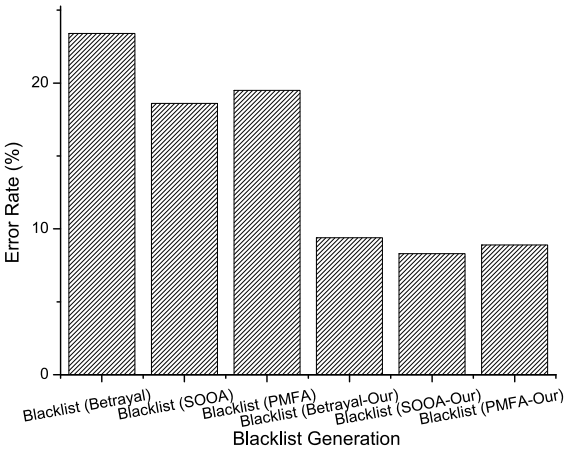

Fig. 7. The error rate on blacklist generation.

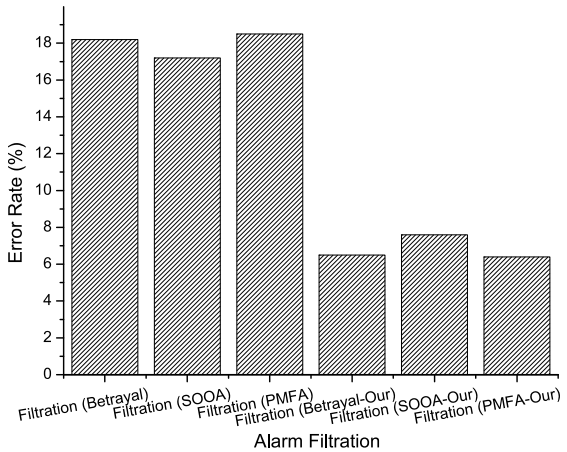

Fig. 8. The error rate on alarm filtration.

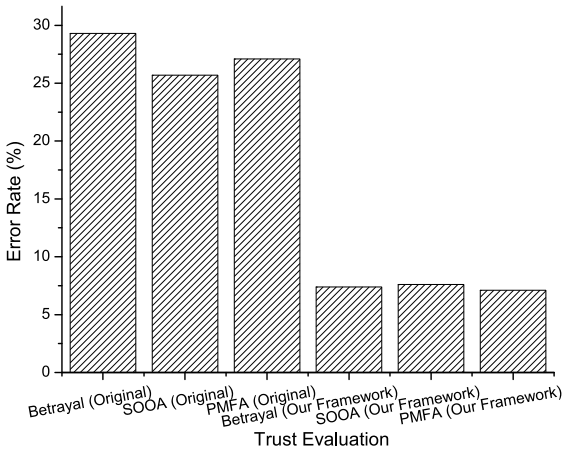

Fig. 9. The error rate on trust evaluation.

- *Query allocation.* Fig. 6 describes the error rate on query allocation (with 50 alarms) under different attacks. It is seen that insider attacks could significantly affect the accuracy of query assignment. For example, the error rate was over 25% under betrayal attack, and ranged around 23%–24% under SOOA and PMFA. With the assistance of blockchain, our framework could greatly decrease the error rate below 10%. This is because malicious query information could be identified among different nodes.
- *Blacklist generation.* Fig. 7 shows the error rate on blacklist generation. The blacklist accuracy will determine the effectiveness of traffic filtration. It is seen that the error rate was over 23% and around 20% under betrayal attack, SOOA and PMFA, respectively. Betrayal attack achieved a higher error rate due to the maximum harm model where the attackers always behave maliciously. By contrast, our blockchain-assisted framework could lower the error rate below 8% in total. With the blockchain, the malicious inputs could be identified via a consensus process among nodes.
- *Alarm filtration.* Fig. 8 presents the error rate on alarm filtration. The intelligent alarm filter can select a suitable machine learning classifier relying on the shared information. Under attacks, the error rate was reached around 18% for betrayal attack, SOOA and PMFA. Under our framework, the error rate could be decreased below 8%. The malicious input could be figured out via a consensus process among nodes.

Our obtained results indicate that a big and negative impact could be caused by internal attacks on the overall performance of collaborative intrusion detection, while our proposed blockchain-assisted framework can greatly reduce the error rates and enhance the system performance in the aspects of trust evaluation, query allocation, traffic filtration and alarm reduction.

4.2. Experiment-2

The purpose of this experiment is to explore the practical performance; thus, we worked with an IT company to launch the tests in an industry CIDN used by around 120 professional staff. The wired CIDS environment consisted of 75 nodes and due to privacy regulations, the IT company deployed our framework for testing.

During the experiment, we considered the same insider attacks as Experiment-1 and explored the error rates on trust evaluation, query allocation, traffic filtration and alarm reduction. Figs. 9 10 11 12 present the error rate respectively.

- *Trust evaluation.* Fig. 5 describes that the trust estimation errors could reach an average rate over 25% under different attacks, i.e., as betrayal attack adopted the maximum harm model, it could cause the error rate to be around 30%. The error rate was around 25% and 27% under SOOA and PMFA, respectively. As a comparison, our framework could decrease the error rate below 7.5%.
- *Query allocation.* Fig. 6 presents the error rate on query allocation (with 50 alarms) under different attacks. It is found that the error rate could reach around 33% under betrayal attack, and around 26% for SOOA and PMFA. With our blockchain-assisted framework, the error rate could be controlled under 10%.
- *Blacklist generation.* Fig. 7 shows the error rate on blacklist generation, i.e., the error rate was over 22% under betrayal attack, SOOA and PMFA. The rate was higher than that in Experiment-1, as the blacklist generation requires a period of time, any error could be propagated. By contrast, our framework could decrease the error rate below 8%, due to the consensus process among nodes.
- *Alarm filtration.* Fig. 8 shows that the error rate on alarm filtration was around 20%, i.e., it was 22%, 21% and 20% for betrayal attack, SOOA, and PMFA, respectively. The errors made during alarm reduction may miss the important alarms raised by attacks, and make the whole system under risk. By contrast, our framework could lower the error rate below 8%.

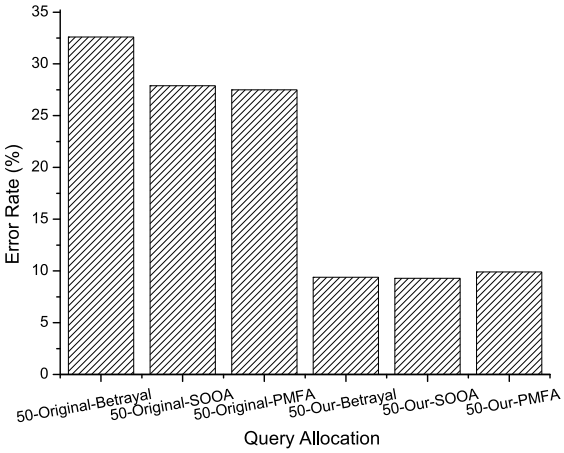

Fig. 10. The error rate on query allocation.

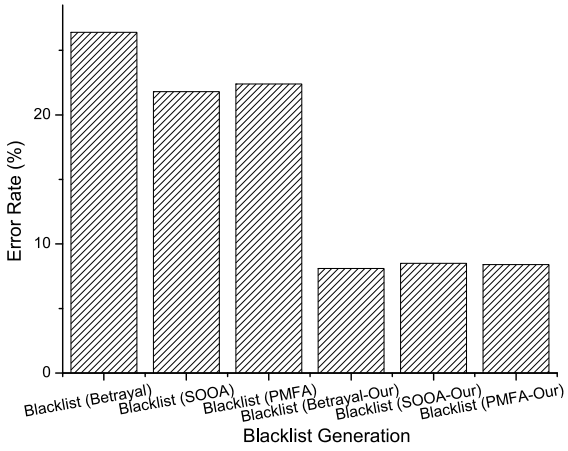

Fig. 11. The error rate on blacklist generation.

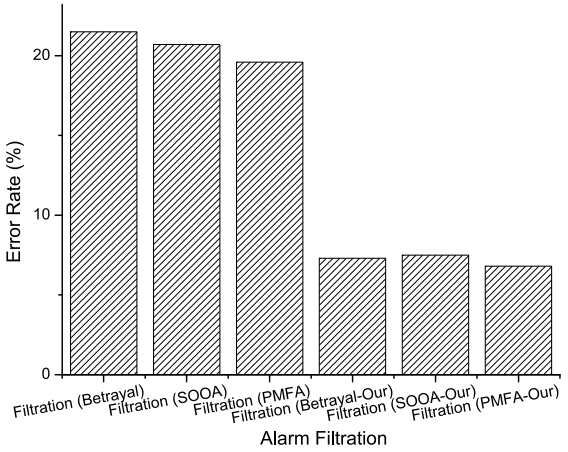

Fig. 12. The error rate on alarm filtration.

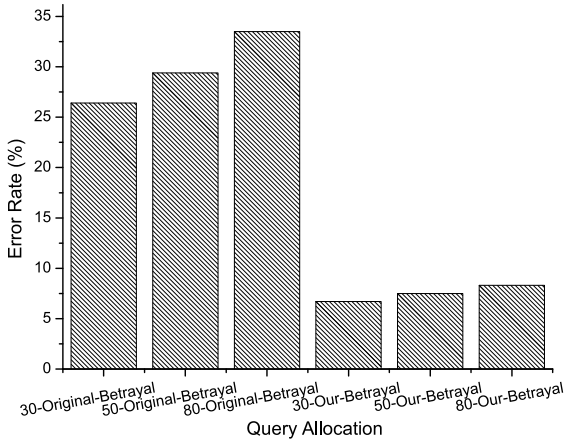

Fig. 13. The error rate on query allocation varied by alarm numbers.

In the practical test, it is found that the error rate could be higher than that in the simulated environment. This is probably due to the complicated network events and traffic in practice. On the whole, these practical results validate and demonstrate that our proposed framework can indeed provide a performance enhancement to the traditional CIDS.

4.3. Discussion

To explore further, based on the simulated CIDS environment, we aim to explore the impact of alarm numbers and node numbers on query allocation and trust evaluation, respectively. We have the following observations:

- Fig. 13 shows the error rate on query allocation varied by alarm numbers, e.g., 30, 50 and 80 alarms. It is found that the error rate was higher for more alarms. For example, the error rate was around 26%, 29% and 33% for 30, 50 and 80 alarms. With the help of blockchain, our framework can reduce the error rate to below 7%, and the error rate was still a bit higher for bigger number of alarms.
- Fig. 14 presents the error rate on trust evaluation varied by node numbers, e.g., 30, 50 and 70 nodes. It is visible that the error rate was higher with few nodes, as several false messages can quickly degrade the trust computation. For example, the error rate was reached roughly 35% for 30 nodes. Under our framework, the error rate was reduced to below 10%, and similarly, the error rate was still higher for 30 nodes than that with 50 and 70 nodes.

Our findings demonstrate that error rate in CIDS would be affected by many factors, especially the alarm amount used in the training as well as the node amount. Our blockchain-assisted framework can generally lower the error rate and enhance the robustness of collaborative intrusion detection.

5. Challenges and future work

This section presents some research challenges and issues in our current work.

Threat model. In our evaluation, we mainly considered three types of insider attacks: betrayal attack, SOOA and PMFA. Among these, betrayal attack adopted a native threat model—maximum harm model where a malicious node always tries to behave untruthfully to cause a larger damage to the network. Instead, SOOA and PMFA adopted an advanced strategy to compromise the CIDS and hide the malicious nodes. In practice, betrayal attack can cause a larger impact on the network performance, but can be identified more quickly than SOOA and PMFA.

In a real network environment, attackers can always choose an advanced/dynamic strategy to intrude a network, hence we aim to investigate the proposed framework under different threat models.

Network scale. In the evaluation, we explored the impact caused by the number of nodes on the trust evaluation. It is found that the error rate would be affected with the network scale. For our blockchain-assisted framework, scalability is one of the major challenges. This could be an interesting and important topic in our further work.

Query size. In the evaluation, we explored different query sizes such as 30, 50 and 80 alarms. It is found that error rate would be varied with different sizes, i.e., the rate is smaller with few alarms. However, different classifiers may have their own optimal query size, this is an important open challenge.

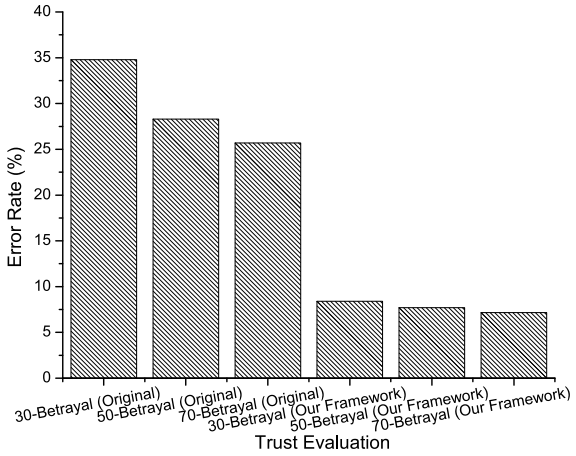

Fig. 14. The error rate on trust evaluation varied by node numbers.

Trust model and platform. For our current work, we consider the challenge-based CIDS into a study, but many different trust models (e.g., analyzing the performance of the target node) are available. As there is no standard platform to compare these different models, it is a big problem to evaluate the framework performance under different tests. This is one of our future work to investigate this challenge.

Intelligent alarm reduction. The main idea is to evaluate and select the best suitable classifier given the incoming data, which can mitigate the weakness of a single classifier. In this work, we mainly considered the traditional supervised learning, but there are more advanced classifiers available like deep learning. One promising topic is to explore the suitability of our blockchain-assisted framework under these newly developed learning schemes.

Blockchain limitations. Basically, blockchain can be treated as a kind of distributed database technique, which may also bring many issues into a blockchain-based system, such as energy consumption, latency, privacy issues, etc. To explore such issues can be one of our future work.

Machine leaning. Our framework can adopt machine learning techniques (e.g., deep learning) into several components such as alarm filtration, query allocation and traffic filtration. While deploying an algorithm, there is a need to explore the workload and hardware requirement over an IoT device. This can be an interesting and important topic in future.

6. Conclusion

This work proposed a blockchain-assisted security management framework for collaborative intrusion detection, which can provide an overall enhancement to the CIDS performance, in the aspects of trust evaluation, query allocation, traffic filtration and alarm reduction. We also provided a case study with challenge-based mechanism. More specifically, we evaluated our proposed framework under several attacks: betrayal attack, SOOA and PMFA. Our experimental results indicate that these attacks would cause a high error rate on trust evaluation, query allocation, traffic filtration and alarm reduction. As a comparison, our framework could greatly lower the error rate below 10% and increase the robustness of collaborative intrusion detection. Our work shows that blockchain can be beneficial to not only detection accuracy, but also detection efficiency.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

Thanks for the great support from the participating organizations and managers. The work is funded by National Natural Science Foundation of China with No. 62102106.

References

[1] Rondon LP, Babun L, Aris A, Akkaya K, Uluagac AS. Survey on enterprise internet-of-things systems (e-IoT) a security perspective. Ad Hoc Netw 2022;125:102728.
[2] Meng W, Li W, Tug S, Tan J. Towards blockchain-enabled single character frequency-based exclusive signature matching in IoT-assisted smart cities. J Parallel Distrib Comput 2020;144:268–77.
[3] Meiners CR, Patel J, Norige E, Torng E, Liu AX. Fast regular expression matching using small TCAMs for network intrusion detection and prevention systems. In: USENIX security symposium. 2010, 2010, p. 111–26.
[4] Jamalipour A, Murali S. A taxonomy of machine-learning-based intrusion detection systems for the internet of things: A survey. IEEE Internet Things J 2022;9(12):9444–66.
[5] Meng W, Li W, Wang Y, Au MH. Detecting insider attacks in medical cyber-physical networks based on behavioral profiling. Future Gener Comput Syst 2020;108:1258–66.
[6] Wu YS, Foo B, Mei Y, Bagchi S. Collaborative intrusion detection system (CIDS) a framework for accurate and efficient IDS. In: Proceedings of ACSAC. 2003, p. 234–44.
[7] Li W, Meng W, Kwok LF, Ip HHS. Enhancing collaborative intrusion detection networks against insider attacks using supervised intrusion sensitivity-based trust management model. J Netw Comput Appl 2017;77:135–45.
[8] Liu L, Ma Z, Meng W. Detection of multiple-mix-attack malicious nodes using perceptron-based trust in IoT networks. Future Gener Comput Syst 2019;101:865–79.
[9] Li W, Tug S, Meng W, Wang Y. Designing collaborative blockchained signature-based intrusion detection in IoT environments. Future Gener Comput Syst 2019;96:481–9.
[10] Khan AA, Khan MM, Khan KM, Arshad J, F. Ahmad: A blockchain-based decentralized machine learning framework for collaborative intrusion detection within UAVs. Comput Netw 2021;196:108217.
[11] Fung CJ, Zhang J, Aib I, Boutaba R. Robust and scalable trust management for collaborative intrusion detection. In: Proceedings of IM. 2009, p. 33–40.
[12] Ma Z, Liu L, Meng W. Towards multiple-mix-attack detection via consensus-based trust management in IoT networks. Comput Secur 2020;96:101898.
[13] Meng W, Choo K-KR, Furnell S, Vasilakos AV, Probst CW. Towards Bayesian-based trust management for insider attacks in healthcare software-defined networks. IEEE Trans Netw Serv Manag 2018;15(2):761–73.
[14] Marche C, Nitti M. Trust-related attacks and their detection: A trust management model for the social IoT. IEEE Trans Netw Serv Manage 2021;18(3):3297–308.
[15] Meng W, Luo X, Li W, Li Y. Design and evaluation of advanced collusion attacks on collaborative intrusion detection networks in practice. In: Proceedings of the 15th IEEE international conference on trust, security and privacy in computing and communications (TrustCom). 2016, p. 1061–8.
[16] Li W, Meng W, Kwok LF, Ip HHS. Developing advanced fingerprint attacks on challenge-based collaborative intrusion detection networks. Cluster Comput 2018;21(1):299–310.
[17] Li W, Meng W, Kwok LF. SOOA: Exploring special on-off attacks on challenge-based collaborative intrusion detection networks. In: Proceedings of GPC. 2017, p. 402–15.
[18] Meng W, Li W, Jiang L, Choo K-KR, Su C. Practical Bayesian poisoning attacks on challenge-based collaborative intrusion detection networks. In: Proceedings of the 24th European symposium on research in computer security. 2019, p. 493–511.
[19] Magdich R, Jemal H, Ayed MB. A resilient trust management framework towards trust related attacks in the Social Internet of Things. Comput Commun 2022;191:92–107.
[20] Salimitari M, Bhattacharjee S, Chatterjee M, Fallah YP. A prospect theoretic approach for trust management in IoT networks under manipulation attacks. ACM Trans Sens Netw. 2020;16(3):26:1–26.
[21] Li W, Meng W, Wang Y, Han J, Li J. Towards securing challenge-based collaborative intrusion detection networks via message verification. In: The 14th international conference on information security practice and experience. 2018, p. 313–28.
[22] Liu L, Yang J, Meng W. Detecting malicious nodes via gradient descent and support vector machine in internet of things. Comput Electr Eng 2019;77:339–53.
[23] Li W, Meng W, Kwok LF. Surveying trust-based collaborative intrusion detection: State-of-the-art, challenges and future directions. IEEE Commun Surv Tutor 2022;24(1):280–305.
[24] Meng W, Tischhauser E, Wang Q, Wang Y, J. Han: When intrusion detection meets blockchain technology: A review. IEEE Access 2018;6:10179–88.
[25] Meng W, Li W, Yang LT, P. Li: Enhancing challenge-based collaborative intrusion detection networks against insider attacks using blockchain. Int J Inf Sec 2020;19(3):279–90.
[26] Cao S, Dang S, Zhang Y, Wang W, N. Cheng: A blockchain-based access control and intrusion detection framework for satellite communication systems. Comput Commun 2021;172:216–25.
[27] Sarhan M, Lo WW, Layeghy S, M. Portmann: HBFL: A hierarchical blockchain-based federated learning framework for collaborative IoT intrusion detection. Comput Electr Eng 2022;103:108379.
[28] Khonde SR, V. Ulagamuthalvi: Hybrid intrusion detection system using blockchain framework. EURASIP J Wireless Commun Networking 2022;2022(1):58.
[29] Li W, Wang Y, J. Li: Enhancing blockchain-based filtration mechanism via IPFS for collaborative intrusion detection in IoT networks. J Syst Archit 2022;127:102510.
[30] Meng W, Li W, Kwok LF. EFM: Enhancing the performance of signature-based network intrusion detection systems using enhanced filter mechanism. Comput Secur 2014;43:189–204.
[31] Meng Y, Kwok LF. Adaptive false alarm filter using machine learning in intrusion detection. In: The 6th international conference on intelligent systems and knowledge engineering. 2011, p. 573–84.

Wenjuan Li obtained the Ph.D degree from City University of Hong Kong (CityU) in 2019. She is currently a Research Assistant Professor in the Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, China. She is a senior member of IEEE. Her research interests include network management and security, intrusion detection, trust management, and blockchain security.

Christian Stidsen obtained his master degree from Aalborg University, Denmark, and is working as a research assistant. His research focus is on security and blockchain application.

Tobias Adam obtained his master degree from Technical University of Denmark, and is working as a researcher at Aalborg University, Denmark. His research focus is on system security and optimization, malware detection and blockchain.