Privacy-Aware Access Control in IoT-Enabled Healthcare: A Federated Deep Learning Approach

Hui Lin, Kuljeet Kaur, *Member, IEEE*, Xiaoding Wang, Georges Kaddoum, *Senior Member, IEEE*, Jia Hu, and Mohammad Mehedi Hassan, *Senior Member, IEEE*

***Abstract*—The traditional healthcare is overwhelmed by the processing and storage of massive medical data. The emergence and gradual maturation of Internet-of-Things (IoT) technologies bring the traditional healthcare an excellent opportunity to evolve into the IoT-enabled healthcare of massive data storage and extraordinary data processing capability. However, in IoT-enabled healthcare, sensitive medical data are subject to both privacy leakage and data tampering caused by unauthorized users. In this article, an attribute-based secure access control mechanism, coined (SACM), is proposed for IoT-Health utilizing the federated deep learning (FDL). Specifically, we manage to discover the relationship between users' social attributes and their trusts, which is the trustworthiness of users rely on their social influences. By applying graph convolutional networks to the social graph with the susceptible–infected–recovered model-based loss function, users' influences are obtained and then are transformed to their trusts. For each occupation, users' trusts allow them to access specific medical data only if their trusts are higher than the corresponding threshold. Then, the FDL is applied to obtain the optimal threshold and relevant access control parameters for the improvement of access control accuracy and the enhancement of privacy preservation. The experimental results show that the proposed SACM achieves accurate access control in IoT-enabled healthcare with high data integrity and low privacy leakage.**

***Index Terms*—Access control, federated learning, graph convolutional networks (GCNs), Internet-of-Things (IoT)-enabled healthcare, social attributes.**

Manuscript received 23 March 2021; revised 22 June 2021 and 15 July 2021; accepted 27 August 2021. Date of publication 15 September 2021; date of current version 6 February 2023. This work was supported by the King Saud University, Riyadh, Saudi Arabia, through the Researchers Supporting Project under Grant RSP2023R18. *(Corresponding authors: Xiaoding Wang; Jia Hu.)*

Hui Lin and Xiaoding Wang are with the College of Computer and Cyber Security, Fujian Normal University, Fuzhou, China, and also with the Engineering Research Center of Cyber Security and Education Informatization, Fujian Province University, Fuzhou 350117, China (e-mail: linhui@fjnu.edu.cn; wangdin1982@fjnu.edu.cn).

Kuljeet Kaur and Georges Kaddoum are with the Electrical Engineering Department, École de Technologie Supérieure, Montreal, QC H3C 1K3, Canada (e-mail: kuljeet.kaur@ieee.org; georges.kaddoum@etsmtl.ca).

Jia Hu is with the Department of Computer Science, University of Exeter, Exeter EX4 4PY, U.K. (e-mail: j.hu@exeter.ac.uk).

Mohammad Mehedi Hassan is with the Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia (e-mail: mmhassan@ksu.edu.sa).

Digital Object Identifier 10.1109/JIOT.2021.3112686

I. Introduction

With the outbreak of the new coronavirus, the existing healthcare system is facing huge challenges in data processing and storage [1], [2]. As the solution, Internet-of-Things (IoT) technology is the core of future smart healthcare. IoT realizes intelligent identification, access control, data processing by comprehensively applying sensor technology, network technology, artificial intelligence technology [3], etc., to the entire healthcare management for information exchange and communication, so as to establish a real-time, efficient, and secure reinforced healthcare (IoT-Healthcare) [4], [5]. However, in IoT-Health, there exist serious privacy leakage problem and data tampering problem [6], both of which are caused by medical data access from unauthorized personnel. That suggests the significance of secure access control for important medical data.

The identity-based access control mechanisms used in traditional centralized computing environments, such as role-based access control (RBAC) and access control lists (ACLs), can only solve the security problems, i.e., privacy leakage and data tampering, of specific systems to a certain extent. Distributed access control of medical data in an open environment poses severe challenges to traditional access control models and mechanisms. For example, how to authenticate and authorize resource requesters based on their identities and how to solve the interoperability problem between different security systems based on centralized access control models. Compared with RBAC and ACL, attribute-based access control (ABAC) relies on the authorization of the subject's attributes and is an effective way to establish trust relationships between unfamiliar parties. In ABAC, attributes of related entities (such as subjects, resources, and environments) are used as the basis for authorization, rather than identities only. Thereby, ABAC is particularly suitable for authorization and access control in open and distributed medical systems. Plenties of ABAC mechanisms have been investigated for decades. Apart from the access control mechanism that utilizes the attribute-based encryption [7], the trust-based one [8] applies to medical data access control, i.e., the doctor of a higher trust should be granted the authority to access patients' medical data. However, the proper trust evaluation is an open problem.

Note that users' trusts are closely related to their social influences. This is because social activities usually take place between users of high social similarities [9]. Besides, influential users are more trustworthy due to the fact that the

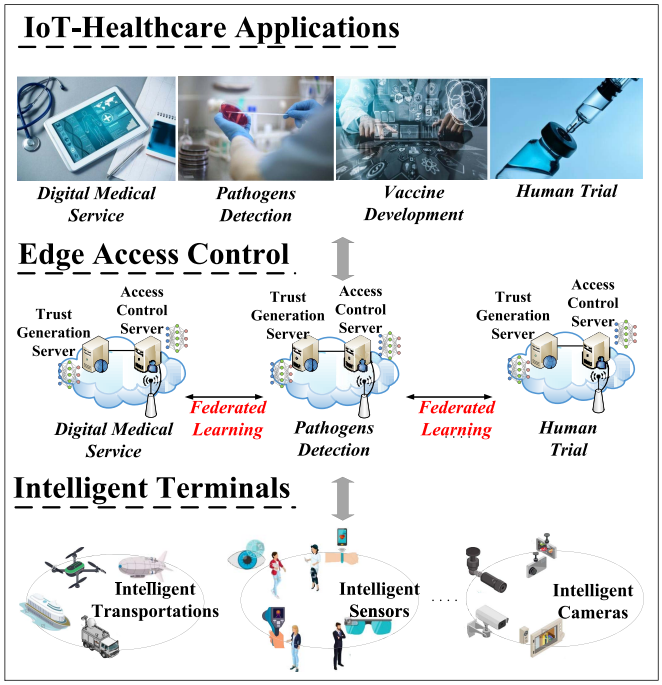


Fig. 1. GCN and federated learning-based access control architecture for IoT-health.

participation in any malicious activity might contribute to serious influence degradation. In [10], a user's influence identification model InfGCN that integrates social data, graph convolutional networks (GCNs), and susceptible–infected–recovered (SIR) model is proposed. However, InfGCN disregards the construction of the social network about users. Moreover, the trust-based access control demands of the access control threshold are elaborately designed for privacy protection and data integrity preservation.

Based on the above analysis, we give a secure access control architecture (see Fig. 1) based on machine learning technologies [11] for IoT-Health. This architecture can be divided into three layers: 1) the IoT-Health application layer; 2) the edge access control layer; and 3) the intelligent terminal layer. Specifically, the edge access control layer is composed of trust generation servers and access control servers responsible for granting users specific authorities and dealing with data access requests using machine learning technologies (i.e., GCN and federated learning) to support a variety of IoT-healthcare applications with data provided by intelligent terminals [12].

Based on this architecture, in this article, we propose an attribute-based secure access control mechanism, named (SACM), for IoT-Health using federated deep learning (FDL). The main contribution of this article is summarized as follows.

1) To achieve the secure access control, we grant each user a unique authority to specific medical data. Specifically, we introduce the social network about users, in which each edge weight represents the connection probability of a specific pair of users according to social similarities. Then, the adjacent matrix of the neighbor graph of each user in the social network and the corresponding features of each node (i.e., the degree centrality, the betweenness centrality, the closeness centrality, and the eigenvector centrality) are used as the input of the GCN that employs the SIR-based loss function to obtain the user's influence and the trust. Based on the trust and occupation of the user, a specific authority is given for access control.
2) To improve the access control accuracy, we adopt the FDL to learn relevant access control parameters. Specifically, by integrating the federated learning framework and the deep reinforcement learning method (i.e., Twin Delayed Deep Deterministic policy gradient algorithm TD3), the access control threshold is learned, considering the privacy preservation of patients and the integrity maintenance of medical data, and the accuracy of access control is significantly improved.
3) The validation experiment is conducted on the real data set. The experimental results indicate that the proposed SACM can achieve secure access control on users in IoT-Health with high data integrity and low privacy leakage.

The remainder of this article is organized as follows. The related work is presented in Section II. Both system model and attack model are introduced in Section III. The implementation details of the proposed SACM are given in Section IV. The performance of the SACM is evaluated in Section V. We conclude this article in Section VI.

II. Related Work

There is an increasing interest in the access control problem for IoT-Healthcare and many excellent works have been proposed. Yang *et al.* [13] solved the problem of accessing encrypted medical data by proposing both ABAC policy and break-glass access control policy. The ABAC only requires workers to satisfy the specific attribute set for medical data access, while the timely access is supported by the break-glass mechanism. Roy *et al.* [14] proposed a fine-grained access control mechanism for cloud computing-enabled healthcare. They also provide a provable authentication mechanism for user access control. In [15], the collusion-resistant access control in ehealth is achieved by Edemacu *et al.* for secure medical data sharing with the consideration of the revocation of attributes and users. Liu *et al.* [16] developed a multiauthority-based access control mechanism for medical services in healthcare. This mechanism is lightweight against collision attack for privacy preservation. Zhang *et al.* [17] achieved the fine-grained access control for the e-healthcare system. They design an encryption scheme of two layers to guarantee the attribute-based medical data access control and the privacy preservation for role attributes and access policies with both cloud-based computation and blind data retrieving protocol. In [18], a secure SDN-based framework is proposed by Meng *et al.* for sharing data in healthcare. This framework provides authorized services to patients by effectively authenticating user devices' MAC addresses against identity theft. Jiang *et al.* [19] developed an access control and medical data sharing mechanism for personal healthcare utilizing the symptom matching technology. This mechanism can achieve granular symptom matching based on a blind signature for privacy preservation. Xu *et al.* [20] used the

blockchain technology to realize the fine-grained access control for large-scale medical data. In this mechanism, authorized doctors are added/revoked based on user transactions and medical data cannot be tamper against medical disputes. In [21], both access control policies and user attributes are transformed into vectors of proper lengths, respectively, by Sun *et al.* to reduce the overhead of the access process for encrypted medical data. In [22], access control and data sharing is achieved by Fan *et al.* using the blockchain technology for nonrepudiation and user self-certification.

All these works are devoted to the access control problem in IoT-Health, however, there remain two problems: 1) how to obtain users' influences and trusts based on users' social data and 2) how to achieve accurate and secure access control according to users' trusts and occupations without exposing users' privacy. In this article, an attribute-based SACM is proposed for IoT-Health using FDL to address these problems.

III. System Model

In IoT-Health, there are serious privacy leakage and data tampering issues, both of which are caused by unauthorized personnel accessing medical data. This indicates that the importance of secure access control to important medical data. To preserve user privacy and data integrity in IoT-Health, the trust-based access control is considered, in which there exist three important entities, i.e., users, trust generation servers, and access control servers. In general, each access control server executes the access control on a number of users based on their trusts obtained by the trust generation servers. We assume both access control servers and trust generation servers are semitrusted while the users are untrusted. Thereby, in this article, two type of attacks, namely, the data tampering attack and the privacy leakage attack, are considered.

1) *Data Tampering Attack:* Such attack is launched by unauthorized users who aim to interfere with the medical diagnosis by altering sensitive medical data either randomly or maliciously. Since data tampering might endanger patients' lives, only authorized users are allowed to access patients' medical data. To this end, the trust-based access control is considered. To be specific, to access important medical data, users should be adequately trustworthy to prevent data tampering attack.
2) *Privacy Leakage Attack:* Unauthorized accesses to the sensitive medical data in IoT-Healthcare will cause severe privacy leakage about patients. For example, anyone rather than the doctors, who reads the medical record of the patient, will seriously violate the patient's privacy. That suggests the significance of the access control. However, to build the unified access control model, users' social data should be provided such that users' privacy is exposed. Thereby, instead of actually access patients' medical data, local access control servers provide local access control models during the FDL to construct the universal access control model against the privacy leakage attack on patients.

IV. Implementation of the Proposed SACM

As an intelligent access control mechanism [23], the proposed strategy SACM composes of two important modules, namely, the social graph-based influence and trust evaluation module that utilizes GCN and the trust-based access control module that employs the FDL technology.

A. Social Graph-Based Influence and Trust Evaluation Utilizing Graph Convolutional Networks

1) Social Graph Construction: Recall that the social activities usually take place between users who have social similarities. That suggests we can construct the social graph about users, in which each edge is associated with a connection probability determined by the social similarity of the end user of this edge. To be specific, we calculate the connection probability CP_{ij} of a pair of user nodes UN_i and UN_j using the cosine similarity by

$$\mathrm{CP}_{ij} = \frac{\mathrm{UN}_i \cdot \mathrm{UN}_j}{\|\mathrm{UN}_i\| \|\mathrm{UN}_j\|} \tag{1}$$

where each user node UN_i is an N-dimensional vector that consists of the user's social data, i.e., the education background, the occupation, the social service condition, the religion, the partisanship, etc. However, how to decide if there exists an edge between a pair of user nodes is an open problem. In this article, we introduce the threshold such that the edge $\mathrm{UN}_i\mathrm{UN}_j$ exists in the social graph only if $\mathrm{CP}_{ij} \geq 0.5$.

Note that the kth representation of a node's neighbors is related to the $(k+1)$th representation of this node in GCN. That suggests the k-step network that is the k-neighbor graph of a node is related to the kth representation of this node. Therefore, we find the k-neighbor graph of each user node by performing the breadth-first search from it to get its neighbors, and then introduce the neighbor network of those neighbors.

2) Trust Evaluation Using GCN and SIR Model: According to the previous analysis, we adopt the GCN to measure the influence of each user node. Unlike the InfGCN model [10], the features of each node, i.e., the degree centrality, the betweenness centrality, the closeness centrality, and the eigenvector centrality, are considered.

1) *Degree Centrality:* The degree of a node can be used to measure the centrality, i.e., a node that has more social connections suggests its high influence.
2) *Betweenness Centrality:* If a node is located on multiple shortest paths between other nodes, then that this node is of high influence.
3) *Closeness Centrality:* The closeness centrality uses the characteristics of the entire network, i.e., the node position in the entire structure. Compared with the betweenness centrality, the closeness centrality is closer to the geometric center position.
4) *Eigenvector Centrality:* The basic idea of the eigenvector centrality is that the centrality of a node is a function of the centrality of adjacent nodes. In other words, a node is more influential if this node connects to other influential nodes.

We use the GCN can learn the representation of nodes using the graph structures and features, i.e.,

$$R^{i+1} = \phi\left(LR^i W^i + B^i\right) \tag{2}$$

where R^i denotes the nodes' representations at the ith GCN layer; L denotes the Laplacian of the neighbor graph which is normalized symmetrically; W_i and B_i are weights and bias, respectively, and ϕ represents the nonlinear activation function, i.e., ELU. In addition, the SIR-based loss function is employed in GCN. The reason for that is as follows. In SIR, each node is of three states, namely, susceptible, infectious, and recovered. Infectious nodes can infect susceptible neighbor nodes and get recovered with infection rate β and recovery rate λ. Susceptible neighbor nodes can get infected by infectious nodes with infection rate β, while any infectious node can get recovered with recovery rate λ. A recovered node cannot infect other neighbor nodes and get infected [24]. Let one node be the first infected node, while the rest nodes are set to be susceptible. The infection scale is used to measure the influence of the first infected node. Thereby, we can add the LogSoftMax module to classify the GCN outputs, the result of which is compared with the ground truth obtained through the SIR experiment as the loss function. In this article, we only consider two types of users who are the most influential users and the much less influential users. Thereby, the LogSoftMax module only has to provide a two-category classification. Once the user's influence is obtained, the trust of who is calculated by

$$UT_i \propto UI_i \tag{3}$$

where UT_i and UI_i are the trust and the influence of the ith user, respectively.

B. Trust-Based Access Control Using Federated Deep Learning

Be aware that the GCN can generate each user's influence with a set of fixed parameters (e.g., the neighbor-graph size, the infection rate, and the recovery rate). Then, by setting a proper trust threshold, each user will be granted a specific authority for access control under the constraint of the user's occupation. For example, if the user is a doctor, whose authority is high enough, then this user is allowed to access patients' medical data; otherwise, the access is denied, i.e., a social work with an extraordinary high trust is forbidden from accessing medical data.

Note that the aim of implementing access control in IoT-Healthcare is to prevent both privacy leakage attack and data tampering attack. Therefore, we introduce the connection fading factor $\rho = (1 - PL_i + DI_i)/2$, where $PL_i \in [0, 1]$ and $DI_i \in [0, 1]$ represent the privacy leakage and the data integrity of the ith user. By introducing a fading factor ρ, the connection probability of the edge between the malicious user and another user in the social graph is significantly reduced as

$$\mathrm{CP}_{ij} \leftarrow \rho \mathrm{CP}_{ij} \tag{4}$$

thus resulting in a lower degree of the malicious user in the social graph.

Thereby, the access control threshold should be dynamically adjusted with fixed k, β, and λ, to minimize the privacy leakage and meanwhile maximize the data integrity. To this end, we employ the Twin Delayed Deep Deterministic policy gradient algorithm TD3 to learn the threshold for the construction of each local access control model. In addition, given the difficulty of model training for some access control servers, we apply the federated learning framework to the TD3 algorithm to build the universal access control model for user privacy preservation.

1) Local Access Control Model Construction Using TD3: The optimal access control threshold θ is discovered using the DRL algorithm TD3. To be specific, for each local access control server, the TD3 requires an actor network π, a target actor network π', two critic networks Q_1 and Q_2, and their target networks Q_1' and Q_1'. Basically, the actor network makes a choice about which action a should be taken for the state s, while the critic networks assess this choice and prevent the overestimation.

In the access control, each state s is presented by an N-dimensional vector of users' authorities UA_i, i.e., $s = (UA_1, UA_2, \ldots, UA_N)$, where $UA_i = 0$ denotes the ith user does not process the authority to access the medical data; otherwise, $UA_i = 1$. Then, the action a can be presented by $a = \theta$. For the current state s, we choose the action a according to the reward r. Since the access control is designed to prevent both privacy leakage attack and data tampering attack, the reward r is given to valuate access control outcome for medical data on N users by

$$r = \sum_{i}^{N} DI_i - PL_i. \tag{5}$$

In the training process of TD3, we randomly sample N experience to update the critic network with the loss function

$$\mathcal{L}\left(\vartheta^{Q_i}\right) = \frac{1}{N}\sum_{j}^{N}\left[Q_i\left(s_j, a_j|\vartheta^{Q_i}\right) - \mathcal{Y}_j\right]^2 \tag{6}$$

where

$$\mathcal{Y}_j = r_j + \gamma\left[Q_i'\left(s_{i+1}, \pi\left(s_{i+1}|\vartheta^{\pi'}\right)|\vartheta^{Q_i'}\right)\right]_{i=1,2}. \tag{7}$$

Thereby, we have

$$\vartheta^{Q_i} \leftarrow \vartheta^{Q_i} - \eta\frac{\partial\mathcal{L}\left(\vartheta^{Q_i}\right)}{\partial\vartheta^{Q_i}}. \tag{8}$$

Then, we update the actor network π by optimizing the objective function

$$J\left(\vartheta^{\pi}\right) = \sum_{j}^{N}\left[Q_1\left(s, a|\vartheta^{Q_1}\right)\pi\left(s_j|\vartheta^{\pi}\right)|s = s_j, a = \pi\left(s_j|\vartheta^{\pi}\right)\right] \tag{9}$$

with

$$\vartheta^{\pi} \leftarrow \vartheta^{\pi} + \iota\frac{\partial J(\vartheta^{\pi})}{\partial\vartheta^{\pi}}. \tag{10}$$

Next, the parameters of target networks $\vartheta^{Q'}$ and $\vartheta^{\pi'}$ are updated with a learning rate κ. When the TD3 learning process

is converged, the local access control model is constructed. Thereby, for each user, only if his trust is higher than the access control threshold and his occupation is a doctor, then he can access the patients' medical data.

2) Universal Access Control Model Construction Using Federated Learning: Considering the difficulties in access control model training, the federated learning technology is employed. Basically, the federated learning is a unique machine learning technology that only requires the trained model from each federated learning participant, which is the access control server in this article, instead of the private data set for participants' privacy preservation, so as to build the universal model. To be specific, each local access control server trains its own access control model using the DRL algorithm TD3 as we described in the previous section and sends its own model to a fusion server. The fusion server generates a set of weights, each of which is assigned to a local model, to construct a synthetic model and distribute this model to each access control server for further training. The process is performed iteratively until the universal model is converged. In this article, we apply the DRL algorithm TD3 in the federated learning framework to develop a TD3-based universal model learning algorithm rather than trial and error. That is, for each participation condition during the federated learning, the TD3 finds the optimal set of weights to aggregate local models.

Specifically, we let the state s consist of the participation condition SP of each local access control server, i.e., $s = (SP_1, SP_2, \ldots, SP_{N'})$, where $SP_i = 0$ denotes the ith access control server does not join the federated learning; otherwise, $SP_i = 1$. No doubt that the action a is the set of aggregation weights, i.e., $a = (\omega_1, \omega_2, \ldots, \omega_{N'})$. For each state s, the action a is chosen based on the reward r. Since both privacy leakage and data integrity are considered in local access control model training, we then give the reward r of the federated learning by

$$r = \sum_{i}^{N'} \sum_{j}^{N} DI_{i,j} - PL_{i,j} \tag{11}$$

where $PL_{i,j} \in [0, 1]$ and $DI_{i,j} \in [0, 1]$ represent the privacy leakage and the data integrity caused by the ith user of the kth access control server to the medical data. The parameter update of the neural networks of the TD3-based FDL is similar to that given in the previous section.

C. Case Study

The Office for Civil Rights (OCR) of the United States Department of Health and Human Services (HHS) is a department that implements the Health Insurance Portability and Accountability Act (HIPAA). It is mainly responsible for protecting some basic rights of people, mainly including the right against discrimination, the right of religious freedom, the right of the privacy of patients' medical information, etc., and is responsible for investigating cases of HIPAA violations. In 2018, Fresenius Medical Care in the United States was investigated for the leakage of patients' privacy. As a large-scale medical group in the United States, the company is mainly engaged in kidney disease medical products and medical services. It has 60 000 employees and serves 170 000 patients, including kidney dialysis centers, cardiovascular centers, emergency centers, and nursing centers. The company has many operating points across the country, but the communication and data exchange between these operating points cannot guarantee the safety of patients' medical information, and unauthorized personnel is very easy to access.

The access control algorithm SACM proposed in this article is designed to prevent the occurrence of the above situation. The algorithm is deployed on multiple access control servers and trust generation servers, where each access control server performs access control for specific users based on their trusts. Specifically, each user provides social data, such as educational background, occupation, social service conditions, religious beliefs, etc., to the trust generation server. Based on these data, the trust generation server first constructs a social graph, where nodes represent users, edges represent social connections between users, and the connection probability is determined by the users' social similarity. Then, the trust generation server uses the social graph as the input of the GCN and SIR model to obtain the user's influence and trust. Next, the access control server grants the user corresponding authorities based on the user's trust and occupation, so that the user can access the corresponding medical data. The outcome of the access control on users can be measured by both privacy leakage and data integrity on medical data, the trust-based access control results are observed after users access the medical data according to their authorities. The above process is performed iteratively until the privacy leakage is minimized and the data integrity is maximized. For some access control servers with difficulties in model training, FDL is used to construct a unified access control model to prevent patients' privacy leakage and medical data tampering. Fig. 2 gives the access control framework for the "Fresenius Medical Care" case.

V. Performance Evaluation

A. Experimental Setup

We evaluate the performance of the proposed SACM in Python on computers equipped with i7 processor, 16-GB memory, 3.2-GHZ CPU, and 64-bit win7 system. The data set we choose is the Facebook-like Social Network, which is available at "https://toreopsahl.com/datasets/". The Facebook-like Social Network originate from an online community for students at the University of California, Irvine. The data set includes the users that sent or received at least one message. This network has also been described in Patterns and Dynamics of Users' Behavior and Interaction. In addition, this data set contains many nodal attributes (e.g., gender, age, and course attended). We randomly generate a number of malicious users based on this data set with the data about 20% deviated from the original ones. Similar to [10], in this experiment, there are four layers in the GCN, which is the hidden layer of 8 units and three fully connected layers of 16, 8, and 2 units, respectively. We train all parameters using the Adam optimizer with $1e^{-4}$ weight decay, 0.0001 learning rate, and 32 mini-batch size. Table I gives the parameters of this experiment.

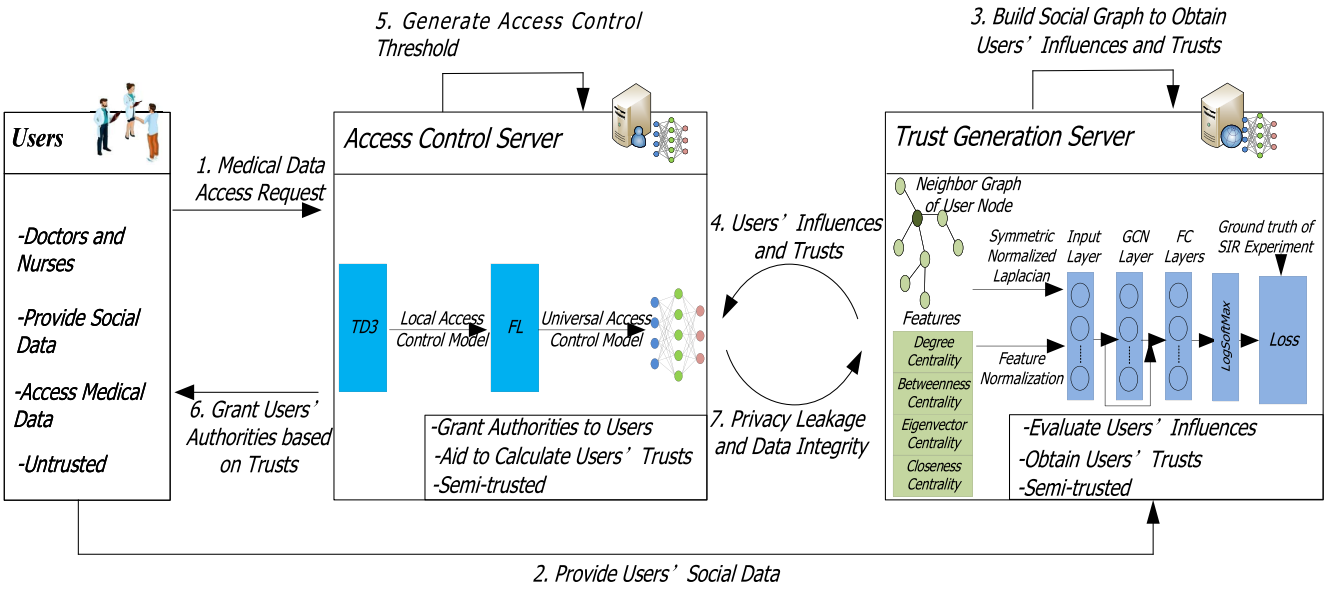
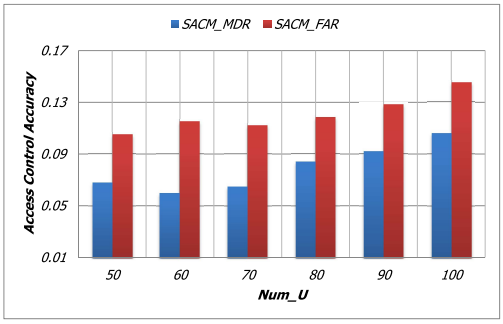


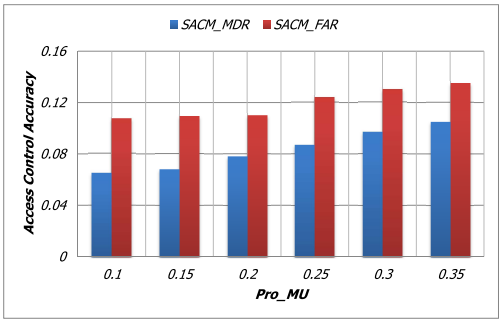
Fig. 2. Access control framework for the "Fresenius Medical Care" case.

TABLE I
PARAMETER SETUP

Parameter	*Description*	*Range*
Num_U	number of users	[50,100]
Pro_MU	probability of malicious users	[0.5,1]
Pro_PE	probability of privacy exposed	[0.5,1]
Pro_IC	probability of integrity compromised	[0.5,1]
Num_Ser	number of access control servers	10
γ	reward factor	0.3
β	learning rate	0.04
λ	recovery rate	1
k	hops of neighbor graph	5



(a)



(b)

Fig. 3. Access control accuracy of the SACM with different (a) numbers of users and (b) probabilities of malicious users.

B. Performance Metrics

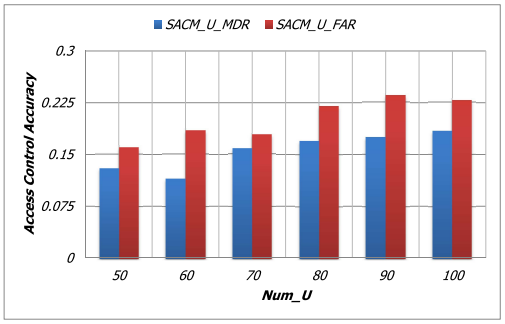
The performance of the SACM is evaluated by access control accuracy, privacy leakage degree, and data integrity with different numbers of users, numbers of medical data, and percentages of malicious users. To be specific, we first evaluate the access control accuracy in false alarm rate (FAR) and miss detection rate of SACM and SACM_U, where SACM_U denotes the universal access control model built by the SACM. Then, we compare both privacy leakage degree and data integrity between SACM, SACM_U, and k-BGP [25].

1) *Access Control Accuracy:* Both FAR and miss detection rate consist of the access control accuracy.
2) *Privacy Leakage:* The privacy leakage degree is measured by the percentage of the private data exposed to the overall data.
3) *Data Integrity:* The data integrity represents the percentage of data that remains unaltered.

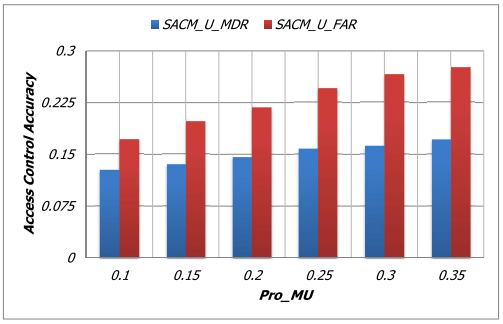
C. Experimental Results

1) Access Control Accuracy: The access control accuracy is measured in Fig. 3 for SACM and in Fig. 4 for SACM_U, while considering different numbers of users and different percentages of malicious users.

Observed from Fig. 3(a), we find that as the number of users increases both MDR and FAR grow. The maximum MDR and FAR are about 11% and 14.5%, compared with the minimum MDR of 6% and 10%. Note that both MDR and FAR of the SACM are less than 15% at any number of users. This is because the trust-based access control is achieved utilizing the DRL algorithm with users' trusts obtained from their social data through the GCN with users' social data. Then,



(a)



(b)

Fig. 4. Access control accuracy of the SACM_U (a) with different numbers of users and (b) probabilities of malicious users.

with the percentage of malicious users fixed, the SACM is able to authorize honest users to access medical data.

As shown in Fig. 3(b), we know that both MDR and FAR increase as the percentage of malicious users. When there are 35% malicious users, the SACM obtains roughly 13.5% of FAR and 10% of MDR. No doubt that the highest percentage of malicious users results in the highest FAR and MDR with a fixed number of users. The reason for that is as follows. Even one-third of users are malicious, the social data of users can be used to efficiently determine who are trustworthy through GCN. Furthermore, both privacy leakage and data integrity are considered in the construction of the access control model, therefore, the trusts of malicious users can hardly be higher than the access control threshold obtained by the DRL algorithm.

In Fig. 4(a), it is clear that both FAR and MDR gradually grow with the number of users. The maximum FAR and MDR are around 22.5% and 18.5%, respectively, when there are 100 users. Note that both FAR and MDR are obtained using the universal access control model of the SACM. Therefore, compared with 14.5% of FAR and 11% of MDR of the SACM that utilizes the local access control model, both FAR and MDR obtained by the SACM_U that uses the universal access control model are at least 7% higher. This is because the universal access control model is built utilizing the federated learning, the accuracy of which depends on the differences between local data sets. That explains the accuracy degradation on the universal model with a specific local dataset. However, the maximum FAR and MDR are less than 23%.

Observed from Fig. 4(b), we find that, with the percentage of malicious users increase, the FAR and the MDR increase. In addition, the SACM_U obtains 22% of FAR and 14.5% of MDR on average with the highest ones equal to 26% and 18%, respectively. Although the SACM_U applies the universal access control model, trustworthy users can be granted authorities to access medical data with both FAR and MDR nearly 8% higher than that of the SACM. Again, the federated learning process explains the accuracy degradation.

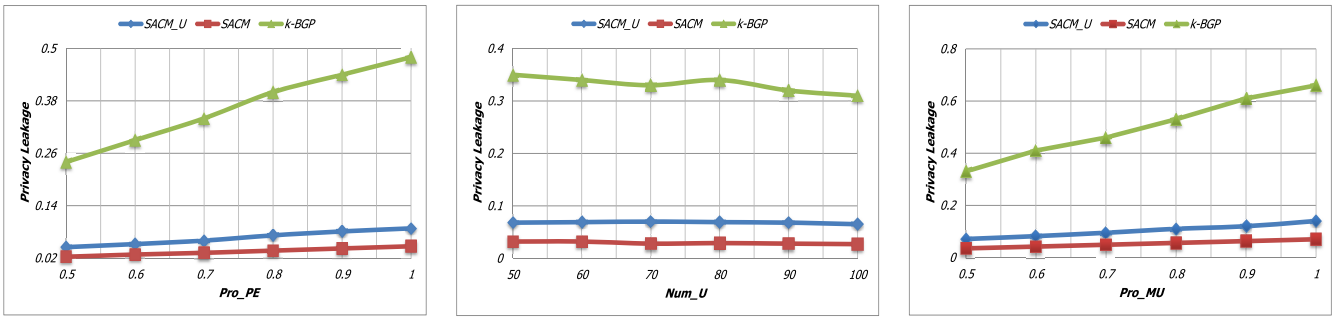
Figs. 3 and 4 suggest the SACM can improve the access control accuracy for IoT-Healthcare.

2) Privacy Leakage: Fig. 5 gives the privacy leakage comparison between SACM, SACM_U, and k-BGP considering different probabilities of privacy exposed, numbers of users, and probabilities of malicious users.

As shown in Fig. 5(a), it is evident that privacy leakage degree increases as the probability of privacy exposed for all approaches with the number of users equals to 50 and the probability of malicious users equals to 0.5. In addition, the privacy leakage degree increases by almost 25% for the k-BGP, compared with 2% of the SACM_U and 1% of the SACM, respectively. This is because although the k-BGP can achieve access control, the privacy leakage is not considered. Besides, both SACM and SACM_U adopt the DRL algorithm to determine the proper threshold for the access control to minimize the privacy leakage. The results shown in Fig. 5(b) are as we expected with the probability of privacy exposed equals to 0.7 and the probability of malicious users equals to 0.5. For example, only 7% and 3% of maximum privacy leakage for SACM_U and SACM, respectively, compared with 35% of k-BGP, due to k-BGP disregards the privacy leakage degree rather than both SACM_U and SACM. Observed from 5(c), we find that with the percentage of malicious users grows the privacy leakage degree increases with the probability of privacy exposed equals to 0.7 and the number of users equals to 50. For example, nearly 65% of privacy will be exposed on average by k-BGP, compared with 9% of SACM_U and 7% of SACM. Fig. 5 indicates that the SACM can prevent privacy leakage for IoT-Healthcare.

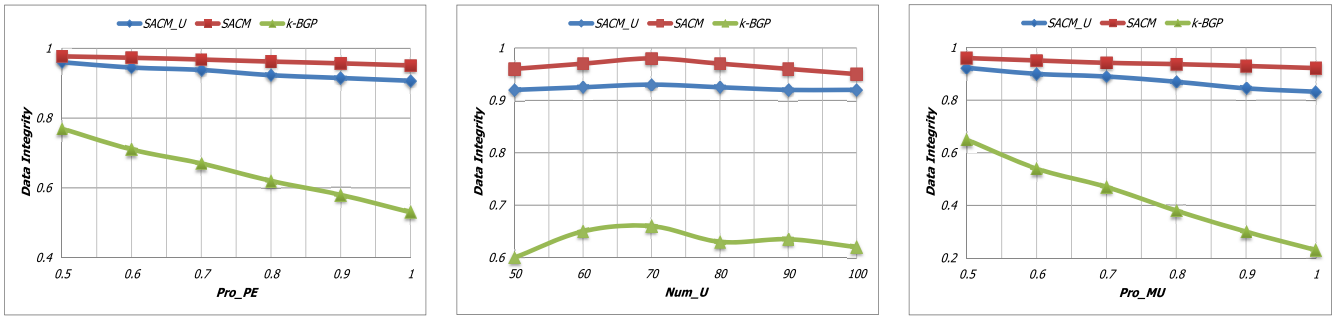
3) Data Integrity: Fig. 6 presents the data integrity comparison between SACM, SACM_U, and k-BGP considering different probabilities of privacy exposed, numbers of users, and probabilities of malicious users.

As shown in Fig. 6(a), it is clear that the data integrity decreases as the number of medical data increases for all approaches with the number of users equals to 50 and the probability of malicious users equals to 0.5. There is a 23% drop in data integrity for k-BGP, compared with 6% of SACM_U and 2% of SACM, respectively. Note that only SACM and SACM_U employ the DRL-based access control mechanism with the consideration of reducing both privacy leakage (see Fig. 5) and data integrity. Therefore, k-BGP obtains a much less data integrity. Fig. 6(b) shows the negative effect of the number of users on the data integrity for all approaches with the probability of integrity compromised equals to 0.8 and the probability of malicious users equals to 0.5. Obviously, SACM and SACM_U manage to maintain the 96% and 92% data integrity, respectively, on average, while compared with 63% of k-BGP. Observed from Fig. 6(c), we find that the data integrity drops with the increasing percentage of malicious



(a) (b) (c)

Fig. 5. Privacy leakage comparison between SACM, SACM_U, and k-BGP with different (a) probabilities of privacy exposed, (b) numbers of users, and (c) probabilities of malicious users.



(a) (b) (c)

Fig. 6. Data integrity comparison between SACM, SACM_U, and k-BGP with different (a) probabilities of privacy exposed, (b) numbers of users, and (c) probabilities of malicious users.

users as we expected with the number of users equals to 50 and the probability of integrity compromised equals to 0.8. Once again, SACM obtains the highest data integrity, that is, 95%, compared with 92% of SACM_U and 65% of k-BGP. Fig. 6 suggests the SACM can improve the data integrity for IoT-Healthcare.

VI. Conclusion

To prevent patients' privacy leakage and maintain medical data integrity in IoT-Healthcare, in this article, we propose an attribute-based SACM using FDL. Specifically, given the fact that an influential user is considerably trustworthy, we introduce the social graph about users, in which each edge weight stands for the connection probability of a specific pair of users according to their social similarities. Then, we feed the GCN with both neighbor graph and features of each user to the user's influence and trust utilizing an SIR-based loss function. Next, the secure access control is accomplished by giving each user a specific authority according to their trusts and occupations. Furthermore, the FDL technology is applied to learn the access control threshold for the privacy preservation of patients and the integrity maintenance of medical data. The experimental results indicate that: 1) the proposed SACM can achieve secure access control on users in IoT-Healthcare and 2) the SACM performs excellently with high data integrity and low privacy leakage.

References

[1] F. Al-Turjman and B. Deebak, "Privacy-aware energy-efficient framework using the Internet of medical things for COVID-19," *IEEE Internet Things Mag.*, vol. 3, no. 3, pp. 64–68, Sep. 2020.

[2] S. Misra, V. Tiwari, and M. S. Obaidat, "Lacas: Learning automata-based congestion avoidance scheme for healthcare wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 4, pp. 466–479, May 2009.

[3] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 912–921, May/Jun. 2021, doi: 10.1109/TCBB.2020.2994780.

[4] M. Raza, M. Awais, N. Singh, M. Imran, and S. Hussain, "Intelligent IoT framework for indoor healthcare monitoring of Parkinson's disease patient," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 2, pp. 593–602, Feb. 2021.

[5] S. Misra, A. Roy, C. Roy, and A. Mukherjee, "DROPS: Dynamic radio protocol selection for energy-constrained wearable IoT healthcare," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 2, pp. 338–345, Feb. 2021.

[6] H. Liu, X. Yao, T. Yang, and H. Ning, "Cooperative privacy preservation for wearable devices in hybrid computing-based smart health," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1352–1362, Apr. 2019.

[7] Z. Guan, J. Li, L. Zhu, Z. Zhang, X. Du, and M. Guizani, "Toward delay-tolerant flexible data access control for smart grid with renewable energy resources," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3216–3225, Dec. 2017.

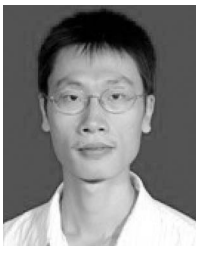
[8] J. Xia, G. Cheng, S. Gu, and D. Guo, "Secure and trust-oriented edge storage for Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4049–4060, May 2020.

[9] S. Peng *et al.*, "An immunization framework for social networks through big data based influence modeling," *IEEE Trans. Dependable Secure Comput.*, vol. 16, no. 6, pp. 984–995, Nov./Dec. 2019.

[10] G. Zhao, P. Jia, A. Zhou, and B. Zhang, "InfGCN: Identifying influential nodes in complex networks with graph convolutional networks," *Neurocomputing*, vol. 414, pp. 18–26, Nov. 2020.

[11] N. Pathak, S. Misra, A. Mukherjee, and N. Kumar, "HeDI: Healthcare device interoperability for IoT-based e-Health platforms," *IEEE Internet Things J.*, early access, Jan. 18, 2021, doi: 10.1109/JIOT.2021.3052066.

[12] S. Misra, S. Moulik, and H.-C. Chao, "A cooperative bargaining solution for priority-based data-rate tuning in a wireless body area network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2769–2777, May 2015.
[13] Y. Yang, X. Liu, and R. H. Deng, "Lightweight break-glass access control system for healthcare Internet-of-Things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3610–3617, Aug. 2018.
[14] S. Roy, A. K. Das, S. Chatterjee, N. Kumar, S. Chattopadhyay, and J. J. P. C. Rodrigues, "Provably secure fine-grained data access control over multiple cloud servers in mobile cloud computing based healthcare applications," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 457–468, Jan. 2019.
[15] K. Edemacu, B. Jang, and J. W. Kim, "Collaborative Ehealth privacy and security: An access control with attribute revocation based on OBDD access structure," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 10, pp. 2960–2972, Oct. 2020.
[16] J. Liu, H. Tang, R. Sun, X. Du, and M. Guizani, "Lightweight and privacy-preserving medical services access for healthcare cloud," *IEEE Access*, vol. 7, pp. 106951–106961, 2019.
[17] W. Zhang, Y. Lin, J. Wu, and T. Zhou, "Inference attack-resistant E-healthcare cloud system with fine-grained access control," *IEEE Trans. Services Comput.*, vol. 14, no. 1, pp. 167–178, Jan./Feb. 2021.
[18] Y. Meng, Z. Huang, G. Shen, and C. Ke, "SDN-based security enforcement framework for data sharing systems of smart healthcare," *IEEE Trans. Netw. Service Manag.*, vol. 17, no. 1, pp. 308–318, Mar. 2020.
[19] S. Jiang, M. Duan, and L. Wang, "Toward privacy-preserving symptoms matching in SDN-based mobile healthcare social networks," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1379–1388, Jun. 2018.
[20] J. Xu *et al.*, "Healthchain: A blockchain-based privacy preserving scheme for large-scale health data," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8770–8781, Oct. 2019.
[21] J. Sun, H. Xiong, X. Liu, Y. Zhang, X. Nie, and R. H. Deng, "Lightweight and privacy-aware fine-grained access control for IoT-oriented smart health," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6566–6575, Jul. 2020.
[22] K. Fan *et al.*, "A secure and verifiable data sharing scheme based on blockchain in vehicular social networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 5826–5835, Jun. 2020.
[23] X.-L. Huang, Y.-X. Li, Y. Gao, and X.-W. Tang, "Q-learning-based spectrum access for multimedia transmission over cognitive radio networks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 110–119, Mar. 2021.
[24] I. M. Foppa, W. O. Kermack, and A. G. McKendrick, "A seminal contribution to the mathematical theory of epidemics (1927)," in *A Historical Introduction to Mathematical Modeling of Infectious Diseases*, vol. 115. Boston, MA, USA: Academic, 2017, pp. 59–87.
[25] M. U. Arshad, M. Felemban, Z. Pervaiz, A. Ghafoor, and W. G. Aref, "A privacy mechanism for access controlled graph data," *IEEE Trans. Dependable Secure Comput.*, vol. 16, no. 5, pp. 819–832, Sep./Oct. 2019.



Hui Lin received the Ph.D. degree in computing system architecture from the College of Computer Science, Xidian University, Xi'an, China, in 2013.

He is a Professor with the College of Computer and Cyber Security, Fujian Normal University, Fuzhou, China, where he is currently an M.E. Supervisor. He has published more than 50 papers in international journals and conferences. His research interests include mobile cloud computing systems, blockchain, and network security.

Kuljeet Kaur (Member, IEEE) received the B.Tech. degree in computer science and engineering from Punjab Technical University, Kapurthala, India, in 2011, and the M.E. degree in information security and the Ph.D. degree in computer science and engineering from Thapar Institute of Engineering and Technology (Deemed to be University), Patiala, India, in 2015 and 2018, respectively.

She worked as an NSERC Postdoctoral Research Fellow with the École de technologie supérieure (ETS), Université du Québec, Montreal, QC, Canada, from 2018 to 2020. She is currently working as an Assistant Professor with the Electrical Engineering Department, ETS, and a Visiting Researcher with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. She has secured several research articles in top-tier journals, such as IEEE WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON CLOUD COMPUTING, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON SMART GRID, IEEE SYSTEMS JOURNAL, IEEE INTERNET OF THINGS JOURNAL, *IEEE Communications Magazine*, IEEE NETWORK, IEEE TRANSACTIONS ON POWER SYSTEMS, *Future Generation Computer Systems*, *Journal of Parallel and Distributed Computing*, and *Peer-to-Peer Networking and Applications* (Springer), and various international conferences, including IEEE Globecom, IEEE ICC, IEEE PES GM, IEEE WCNC, IEEE Infocom Workshops, ACM MobiCom Workshops, and ACM MobiHoc workshops. During her Ph.D., she received two prestigious fellowships, i.e., INSPIRE Fellowship from the Department of Science and Technology, India, in 2015, and a Research Scholarship from Tata Consultancy Services from 2016 to 2018. Her main research interests include cloud computing, energy efficiency, smart grid, frequency support, and vehicle to grid.

Dr. Kaur received the IEEE ICC Best Paper Award in 2018 from Kansas City, USA, the 2019 Best Research Paper Award from Thapar Institute of Engineering and Technology, India, and the 2020 IEEE SYSTEMS JOURNAL Best Paper Award. She serves as an Associate Editor for *Security and Privacy* (Wiley), *Journal of Information Processing Systems*, and *Human-Centric Computing and Information Sciences* (Springer) and a Guest Editor for special issues in IEEE TRANSACTION ON INDUSTRIAL INFORMATICS and IEEE OPEN JOURNAL OF THE COMPUTER SOCIETY. She is a Website Co-Chair of the N2Women Community. She also serves as the Vice-Chair of the IEEE Montreal Young Professionals Affinity Group. She has also been the TPC Co-Chair for IEEE Infocom in 2020 and ACM MobiCom in 2020 workshops on DroneCom. She is a member of the IEEE Communications Society, IEEE Computer, IEEE Women in Engineering, IEEE Software Defined Networks Community, IEEE Smart Grid Community, ACM, and IAENG.

Xiaoding Wang received the Ph.D. degree from the College of Mathematics and Informatics, Fujian Normal University, Fuzhou, China, in 2016.

He is an Associate Professor with the College of Computer and Cyber Security, Fujian Normal University, FuZhou, China. His main research interests include network optimization and fault tolerance.

Georges Kaddoum (Senior Member, IEEE) received the Ph.D. degree (Hons.) in signal processing and telecommunications from the National Institute of Applied Sciences, Toulouse, France, 2008.

He held the ETS Research Chair in physical-layer security for wireless networks. He published over 200 journal and conference papers and two pending patents.

Prof. Kaddoum is the recipient of the Research Excellence Award of the Université du Quebec in 2018 and the Research Excellence Award-Emerging Researcher from ETS in 2019. He is also a co-recipient of the Best Papers Awards of the IEEE PIMRC in 2017 and the IEEE WiMob in 2014. He received the Exemplary Reviewer Award from IEEE TRANSACTIONS ON COMMUNICATION twice in 2015 and 2017. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and IEEE Communications Letters.

Jia Hu received the M.Eng. and B.Eng. degrees in electronic engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2004, respectively, and the Ph.D. degree in computer science from the University of Bradford, Bradford, U.K, in 2010.

He is a Senior Lecturer of Computer Science with the University of Exeter, Exeter, U.k. His research interests include edge–cloud computing, resource optimization, applied machine learning, and network security. He has published over 80 research papers within these areas in prestigious international journals and reputable international conferences.

Dr. Hu has received the Best Paper Awards at IEEE SOSE'16 and IUCC14. He serves on the Editorial Board of *Computers & Electrical Engineering* (Elsevier) and has guest-edited many special issues on major international journals, such as IEEE INTERNET OF THINGS JOURNAL, *Computer Networks*, and *Ad Hoc Networks*. He has served as the General Co-Chair of IEEE CIT'15 and IUCC'15 and the Program Co-Chair of IEEE ISPA'20, ScalCom'19, SmartCity'18, CYBCONF'17, and EAI SmartGIFT'2016.

Mohammad Mehedi Hassan (Senior Member, IEEE) received the Ph.D. degree in computer engineering from Kyung Hee University, Seoul, South Korea, in 2011.

He is currently a Full Professor with the Information Systems Department, College of Computer and Information Sciences, King Saud University (KSU), Riyadh, Saudi Arabia. He has authored and coauthored more than 260 publications, including refereed journals (over 218 SCI/ISI-Indexed journal papers, four ESI highly cited papers, and one hot paper), conference papers, books, and book chapters. His research interests include cloud/edge computing, Internet of Things, artificial intelligence, body sensor network, big data, mobile computing, cyber security, smart computing, 5G/6G network, and social network.

Dr. Hassan is a recipient of a number of awards, including the Distinguished Research Award from College of Computer and Information Sciences, KSU, in 2020, the Best Conference Paper Award from IEEE International Conference on Sustainable Technologies for Industry 4.0 in 2020, the Best Journal Paper Award from IEEE SYSTEMS JOURNAL in 2018, the Best Conference Paper Award from CloudComp in 2014 conference, and the Excellence in Research Award from College of Computer and Information Sciences, KSU, in 2015 and 2016. He has served as the Chair and the Technical Program Committee Member in numerous reputed international conferences/workshops, such as IEEE CCNC, ACM BodyNets, and IEEE HPCC. He is listed as one of the top 2% Scientists of the world in Networking and Telecommunication field. He is one of the top computer scientists in Saudi Arabia as well. He is on the Editorial Board of several SCI/ISI-indexed journals. He has also played role of the guest editor of several international ISI-indexed journals.