Toward Accurate Anomaly Detection in Industrial Internet of Things Using Hierarchical Federated Learning

Xiaoding Wang, Sahil Garg, *Member, IEEE*, Hui Lin, Jia Hu, Georges Kaddoum, *Senior Member, IEEE*, Md. Jalil Piran, *Senior Member, IEEE*, and M. Shamim Hossain, *Senior Member, IEEE*

***Abstract*—The Industrial Internet of Things (IIoT) is an emerging technology that can promote the development of industrial intelligence, improve production efficiency, and reduce manufacturing costs. However, anomalies of IIoT devices might expose sensitive data about users of high authenticity and validity, resulting in security and privacy threats to the IIoT applications. That suggests the significance of anomaly detection executed by proper authorities. To address these problems, in this paper, we propose a reliable anomaly detection strategy for IIoT using federated learning. Specifically, we apply the federated learning technique to build a universal anomaly detection model with each local model trained by the deep reinforcement learning (DRL) algorithm. Since local data sets are not required during the federated learning, the chance of privacy leakage is reduced. In addition, by introducing privacy leakage degree and action relation to anomaly detection design, we can greatly improve the detection accuracy. The validation experiments indicate that the proposed strategy achieves high throughput, low latency, and high anomaly detection accuracy for privacy preservation in various IIoT scenarios.**

***Index Terms*—Anomaly detection, deep reinforcement learning (DRL), federated learning (FL), Industrial Internet of Things (IIoT), privacy preservation.**

I. Introduction

IN industrial domain, as an evolutionary trend of automation and machine-typed data exchange, the Industrial Internet of Things (IIoT) [1] enables remotely managed machines or devices to achieve valuable objectives, i.e., improving the competitiveness of industrial automation [2]. By integrating environment cognition, data analytic, machine-to-machine communication, and robotic automation, IIoT has a wide range of applications in many fields, such as intelligent manufacturing, energy harvest, and intelligent transportation [3]. However, the proliferation of smart IIoT devices and the subsequent overwhelming amount of sensitive business and personal data collected for analytic solutions and industrial operation optimizations have posed serious privacy and security issues [4], [5].

Widely distributed IIoT devices are vulnerable to a variety of attacks due to the lack of efficient security protection. Thereby, establishing a trust network to ensure the reliability of data source and to guarantee the authenticity and validity of data against security and privacy threats has become an urgent issue in IIoT, which is the prerequisite to ensure IIoT implementation in various production fields, and it is also an important foundation and guarantee for industrial security and national security [6]. Anomaly detection, as an effective solution, have been used in IIoT to realize the attack anticipation so as to reduce the attack possibility effectively [7]. However, how to ensure the efficiency, accuracy, and effectiveness in anomaly detection with massive data generated by IIoT devices has become a challenging issue.

To address such critical issues the deep reinforcement learning (DRL)-based solution is developed using the deep deterministic policy gradient (DDPG) algorithm, as an instance [8]. However, there exists a potential risk of privacy leakage about such methods, due to the requirement of private data aggregation in universal model training for model generalization. As a new distributed learning paradigm, the federated learning (FL) enables the decentralized training of a prediction model in a collaborative way [9]–[11]. More importantly, rather than sharing and disclosing the training data set with the server, the model parameters are optimized collaboratively by a great number of local interconnected devices to ensure privacy preservation. That indicate the advantages of the federated DRL technology [12], which represents the deep amalgamation between FL and DRL.

According to above analysis, in this paper, we present an anomaly detection architecture. As shown in Fig. 1, anomaly detections are applied to a variety of industrial applications, i.e., smart manufacture, the Internet of Vehicles (IoV), smart

Manuscript received December 21, 2020; revised March 19, 2021; accepted April 2, 2021. Date of publication April 20, 2021; date of current version May 9, 2022. This work was supported by the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia, through the Vice Deanship of Scientific Research Chairs: Research Chair of Pervasive and Mobile Computing. *(Corresponding authors: Hui Lin; Md. Jalil Piran.)*

Xiaoding Wang and Hui Lin are with the College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China (e-mail: wangdin1982@fjnu.edu.cn; linhui@fjnu.edu.cn).

Sahil Garg and Georges Kaddoum are with the Department of Electrical Engineering, École de Technologie Supérieure, Montreal, QC H3C 1K3, Canada (e-mail: sahil.garg@ieee.org; georges.kaddoum@etsmtl.ca).

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

Md. Jalil Piran is with the Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea (e-mail: piran@sejong.ac.kr).

M. Shamim Hossain is with the Research Chair of Pervasive and Mobile Computing and with the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia (e-mail: mshossain@ksu.edu.sa).

Digital Object Identifier 10.1109/JIOT.2021.3074382

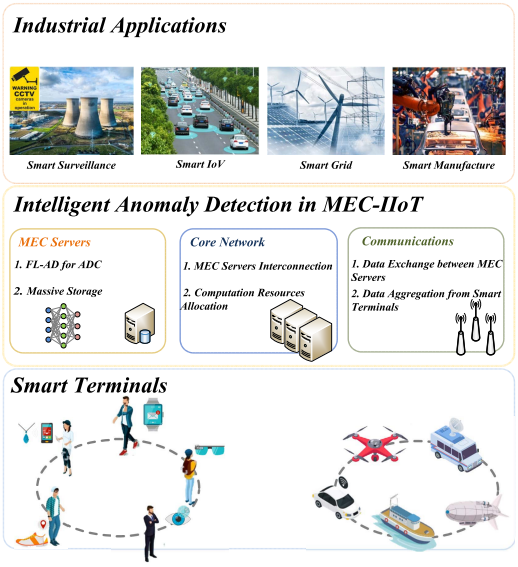


Fig. 1. Architecture of the proposed FLAD.

surveillance, smart grid, etc., due to security and privacy concerns. Because of the computational resources is in great demand for mobile edge computing (MEC) [13] enabled IIoT (MEC-IIoT), the anomaly detection centers (ADC) employ MEC servers for artificial intelligence (AI) empowered accurate anomaly detection using massive relevant data collected by smart terminals. Both interconnection and computation resource allocation between MEC servers are administrated by the core network. However, ADCs might also be abnormal, i.e., misreporting the anomalies of users.

In general, there exists two problems in the efficient and accurate anomaly detection. The first one is how to realize the privacy-preserving anomaly detection. The second one is how to improve the anomaly detection accuracy with abnormal ADCs. In this article, based on the proposed architecture, a federated DRL-empowered anomaly detection method, named the federated deep reinforcement learning empowered anomaly detection (FLAD), is proposed for IIoT. We summarize the contributions of this article as follows.

1) To achieve anomaly detection on users, while considering the alleviation on discrepancies between ADCs, a federated DRL algorithm is design based on privacy leakage degree to construct a universal user anomaly detection model without accessing user private information for the purpose of user privacy preservation and anomaly detection generalization.
2) To improve user anomaly detection accuracy, given the fact that ADCs might be abnormal, the anomaly detection on ADCs is performed first. By conducting internal/external action comparisons, abnormal ADCs are detected. Moreover, an appeal mechanism is introduced to reclaim the nonanomaly of misjudged users for detection accuracy improvement.
3) Through extensive experiments, we demonstrate that the proposed anomaly detection method FLAD can accurately detect abnormal users with high system throughput and low latency in IIoT.

The rest of the paper is organized as follows. Section II presents the related work. Section III introduces the system model. Section IV gives the implementation details of the proposed FLAD. Section V provides the performance evaluation. Section VI draws the conclusions.

II. Related Work

FL for IIoT has attracted a great interest from academia and industry in recent. Wang *et al.* [14] proposed an anomaly detection method based on a composite auto encoder model, in which anomalies are detected according to error distribution to find abnormal devices. Genge *et al.* [15] called the gradual decay of IIoT's physical dimension as the aging process, with which an anomaly detection system is developed to detect aging IIoT. The FL is integrated with deep anomaly detection by Liu *et al.* [16], in which a convolutional neural network (CNN) model with the long short term memory is developed to improve the detection accuracy and meanwhile the gradient compression is utilized in FL for communication cost reduction and communication quality improvement. Yi *et al.* [17] proposed an FL-based deep anomaly detection scheme in IIoT utilizing an attention mechanism empowered convolutional neural network model of long short term memory for detection accuracy improvement. And the gradient compression is introduced for communication efficiency improvement in FL. Both the Paillier encryption mechanism and FL are employed by Li *et al.* [18] to design an intrusion detection model based on secure communication protocols. Due to the FL, the intrusion detection model is jointly trained with privacy preservation. Taghavinejad *et al.* [19] investigated the intelligent intrusion detection problem and propose the decision tree-based solution. Wang [20] applied cooperative games to abnormal action detection in perception environments for IIoT. Wu *et al.* [21] developed a Gaussian Naive Bayes-based anomaly detection model with long short term memory, in which Gaussian Naive Bayes model is employed to detect outliers.

Although the above discussed previous studies could achieve efficient anomaly detections, there still exist the following problems, which are required to be investigated. First, how to accomplish the privacy-preserving anomaly detection is an open problem? Second, how to improve the anomaly detection accuracy, given the fact that the ADCs might be abnormal, poses a great challenge? In this article, the FLAD method is developed to solve above mentioned problems in IIoT.

III. System Model

A. System Architecture

To solve the privacy leakage problem caused by abnormal users, in this article, an FLAD method is proposed for IIoT. In this framework, we consider the following entities: the global anomaly detection center (GADC), the local anomaly detection center (LADC), the regional anomaly detection center (RADC), and the users.

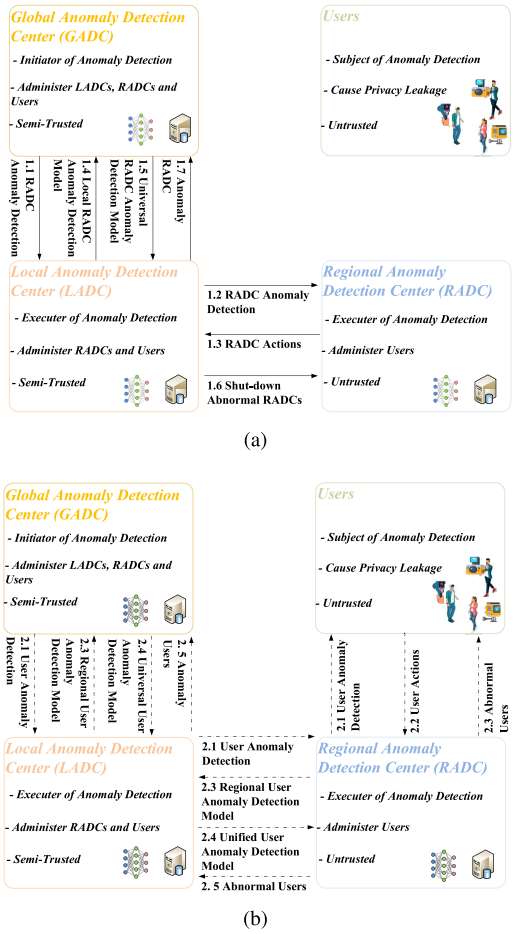


Fig. 2. System model of FLAD. (a) Phase-one RADC anomaly detection. (b) Phase-two user anomaly detection.

The proposed FLAD aims to detect abnormal users, whose actions might give away victims' positions, identities, relations, privacies, or other sensitive information either accidentally or deliberately. Because each user belongs to a specific region, the anomaly detection on internal users is conducted by the corresponding RADC. It is difficult to develop an accurate anomaly detection model by each RADC using machine learning technologies. For some RADCs incapable of the training anomaly detection model, the universal model trained by the GADC is applicable. We assume both GADC and LADC are semi-trusted, i.e., GADC or LADC might be honest but curious about users' sensitive information. That suggest the FL, the nature of which is to build the universal model without gathering the local data set from each participant, can be utilized to build the universal anomaly detection model for user privacy preservation. In addition, we assume RADCa are untrusted, i.e., some RADCs might be abnormal.

In fact, the anomalies of RADCs can be detected by discovering the differences between RADC actions. Specifically, the action difference between the suspicious action and the regular one of an RADC, which is called the internal difference, is dominant in RADC anomaly detection; while the relation between the suspicious action of a specific RADC and the regular ones of other RADCs, which is called the external difference, is another dominant factor for RADC anomaly detection. Thereby, the anomaly detection on RADC should be made based on both internal difference and external difference.

The above analysis suggests the anomaly detection should be implemented in multiphases. In this article, the proposed FLAD is a two-phase anomaly detection framework, which is summarized in Fig. 2.

In phase one, as shown in Fig. 2(a), the abnormal RADCs are detected by evaluating corresponding actions and then they are shutdown to ensure high accurate anomaly detection on users. Specifically, the GADC initializes the anomaly detection on RADCs. Then, each LADC that are capable of model training, trains the local RADC anomaly detection model and sends it to the GADC. Next, the GADC combine each local model to build the universal one and dispatch it to each LADC. For the LADCs unable to train their own models, this universal model is applied to detect abnormal RADCs directly.

In phase two, as shown in Fig. 2(b), the abnormal users are detected by normal RADCs using the privacy leakage-based anomaly detection, which is different from the RADC anomaly detection. Since the aim of user anomaly detection is to prevent privacy leakage, the privacy leakage degree should be integrated with the anomaly detection design. Specifically, the normal RADCs train their own user anomaly detection models and then send them to the GADC to build the universal user anomaly detection model. Then, the universal model is dispatched to all RADCs.

B. Attack Model

Privacy Leakage Attack: Some users might expose victims' sensitive information, i.e., names, ages, sexes, occupations, etc., through abnormal actions either intentionally or accidentally. However, such a severe privacy violation is not tolerable. That suggests the abnormal actions of users should be properly detected. Once the users are detected exposing other users' privacy, these users will be isolated to prevent further damage on victims.

The privacy leakage attack is indirectly launched by anomaly RADCs. Specifically, abnormal RADCs tend to report abnormal users randomly and frequently. That suggests if the abnormal actions are detected as the normal one, then these actions might cause the privacy leakage; furthermore, once normal actions are assessed as abnormal ones, it will hinder users from behaving normally. Thereby, the abnormal RADCs should be detected first by LADCs before conducting user anomaly detection.

Moreover, the LADCs might expose RADCs and users sensitive information due to all LADCs are assumed to be semi-trusted. That suggests during the RADC anomaly detection all sensitive information about RADCs and users of an LADC should be kept from the other LADCs. Owing to the FL, RADC anomaly detection model can be trained without directly accessing private information about RADCs and users.

IV. Implementation of the Proposed FLAD

A. Phase-One RADC Anomaly Detection

In phase one, each LADC that are capable of model training applies the DDPG algorithm for local anomaly detection

model training. Note that we choose the DRL algorithm over both the RL and deep learning (DL) algorithms due to the following reason. As the deep integration of DL and RL, the DRL is designed with the principle of RL and the structure of DL. In addition, the neural networks of the DRL provide extraordinary capabilities to solve both continuous state-action space searching problem and Q-table fitting problem.

1) DDPG-Based Local RADC Anomaly Detection: By evaluating the actions of each RADC, the GADC is able to determine whether this RADC is an abnormal one or a normal one, if the rating of the ith RADC R_RADC_i falls within the range of (0, 0.5) or [0.5, 0.1]. In addition, we denote the action of the ith RADC A_RADC_i by a tuple, i.e., $A\_RADC_i = (FA_i, FN_i, DA_i, DN_i)$, where FA_i and FN_i denote the frequencies of judging a user as the abnormal one and a normal one, respectively; DA_i and DN_i represent the anomaly degree and the normality degree of the user's actions, respectively.

According to the analysis given in the previous section, we know that the deviations between current actions A_RADC_is and regular ones RA_RADC_is, i.e., $RA\_RADC_i = (RFA_i, RFN_i, RDA_i, RDN_i)$, of the ith RADC and regular ones of other RADCs RA_RADC_js, $i \neq j$, can be used to determine whether the ith RADC is abnormal or not. Specifically, we calculate the rating R_RADC_i of the ith RADC by integrating both internal difference In_D and external difference Ex_D as

$$R\_RADC_i = 1 - (In\_D + Ex\_D), \tag{1}$$

where

$$\begin{aligned} In\_D = {} & \alpha * (FA_i - RFA_i) + (1-\alpha) * (FN_i - RFN_i) \\ & + \beta * (DA_i - RDA_i) + (1-\beta) * (DN_i - RDN_i) \end{aligned} \tag{2}$$

$$\begin{aligned} Ex\_D = \sum_j \big[& \alpha * \big(FA_i - RFA_j\big) + (1-\alpha) * \big(FN_i - RFN_j\big) \\ & + \beta * \big(DA_i - RDA_j\big) + (1-\beta) \\ & * \big(DN_i - RDN_j\big)\big] \end{aligned} \tag{3}$$

α and β denote anomaly scale factors. (1) suggests if an RADC keeps reporting abnormal actions of users or the degrees of which are always high or low, then the RADC might be abnormal with a pair of anomaly scale factors (α, β).

In this article, the optimal pair of anomaly scale factors are obtained utilizing the DRL algorithm DDPG that includes three factors (i.e., state, action, and reward). Note that if an RADC is judged as an abnormal one, then it should be shutdown for a period of time. Thereby, we let RADCs' on/off-line conditions consist of the state s, i.e., $s = (C\_RADC_1, C\_RADC_2, \ldots, C\_RADC_n)$, where $C\_RADC_i = 1$ indicates the ith RADC is online without being shutdown, otherwise $C\_RADC_i = 0$. Any pair of anomaly scale factors (α, β) consist of the action a. In general, the action a_t is chosen for the state s_t, according to the reward r_t, to detect abnormal RADCs if their ratings fall within the range of (0, 0.5). Then, abnormal RADCs will be shut down and all internal users will join normal RADCs. And the resulted on/off-line conditions of RADCs compose of the next state s_{t+1}. We denote the number of users of abnormal RADCs by NU_ARADC. To guarantee the anomaly detection accuracy, considering a limited number of users NU_RADC_i that the ith RADC can accommodate, we introduce the Softmax function to calculate the extra number of users ENU_RADC_i reassigned to the ith normal RADC without exceeding the NU_RADC_i, i.e.,

$$ENU\_RADC_i = NU\_ARADC * \frac{e^{R\_RADC_i}}{\sum_i e^{R\_RADC_i}}, \tag{4}$$

otherwise

$$ENU\_RADC_i = NU\_RADC_i - CNU\_RADC_i, \tag{5}$$

where CNU_RADC_i denotes the current number of users of the ith RADC. We repeat the reassigning process until each normal RADC has reached user number constrain. We are aware that the increasing number of user anomaly detections might cause the system failure of normal RADCs. Besides, the anomalies of RADCs might occur from time to time. Thereby, we consider to put some abnormal RADCs back on line only if these RADCs will not cause more serious anomalies by adding a time-frame parameter κ to the action a_t. Specifically, if an RADC is shut down, then after κ timeslots this RADC is back on line again. Thus, we rewrite the action a_t as $a_t = (\alpha, \beta, \kappa)$. According to the ratings of RADCs, we give the reward r_t by

$$r_t = -\sum_i^{n'} R\_RADC_i. \tag{6}$$

As observed from (6), we know that a better reward indicates a less anomaly for each RADC. Then, experience (s_t, a_t, r_t, s_{t+1}) is stored in experience poor $\mathcal{P}$. As the training of the DDPG, N experience are sampled from $\mathcal{P}$ for critic network update through the loss function

$$\mathcal{L}\left(\vartheta^Q\right) = \frac{1}{N}\sum_i^N \left[Q\left(s_i, a_i|\vartheta^Q\right) - \mathcal{Y}_i\right]^2, \tag{7}$$

where

$$\mathcal{Y}_i = r_i + \gamma\left(Q\left(s_{i+1}, \pi\left(s_{i+1}|\vartheta^{\pi'}\right)|\vartheta^{Q'}\right)\right). \tag{8}$$

And the actor network is updated by gradient ascent as

$$\begin{aligned} \nabla_{\vartheta^\pi} J = \frac{1}{N}\sum_i^N \Big[& \nabla_a Q\left(s, a|\vartheta^Q\right)|s = s_i, a = \pi\left(s_i|\vartheta^\pi\right) \\ & \nabla_{\vartheta^\pi}\pi\left(s|\vartheta^\pi\right)|s = s_i\Big]. \end{aligned} \tag{9}$$

Furthermore, target networks are updated based on the parameters of the updated actor network and the critic network with a certain learning rate τ.

2) Universal RADC Anomaly Detection Model Training With Federated Learning: The GADC can use the FL [22] to build the universal RADC anomaly detection model based on local RADC anomaly detection models trained by LADCs. Note that the reason for choosing the FL over the transfer learning is as follows. Although both FL and transfer learning can achieve privacy-preserving model training due to both learning methods do not require each participant to directly access other participants' sensitive information (i.e., the data sets), the transfer learning is mainly applied for reducing the model training time while the FL is employed for constructing

the universal model to solve the model generalization problem. Basically, the FL constructs a universal model by integrating local models with a proper set of weights for model generalization. Although taking the average of all local models provides the simplest solution to the universal model construction, the average model might cause severe performance degradation. For example, the average model will result in unacceptable false alarm rate (FAR) and missing detection rate (MDR) on abnormal RADC detection. That suggests the significance of choosing integration weights properly.

In this article, we introduce the feedback mechanism, which exploits the characteristic of each local anomaly detection model, to improve the detection accuracy of the universal model. Specifically, the rating feedback is considered in the optimal integration weight discovery utilizing the DRL algorithm DDPG. Similarly, we let the LADC participation condition consists of the state s. It is evident that the set of integration weights consist of the action a, i.e., $a = (\omega_1, \omega_2, \ldots, \omega_n)$. According to above analysis, we know that the reward function should be designed based on rating deviations. To be specific, the rating deviation is evaluated by the ratio between the universal model-based rating and the local model-based rating. That is, we calculate the reward r by

$$r = \sum_{i}^{n} \frac{R\_RADC_i}{UR\_RADC_i}, \tag{10}$$

where R_RADC_i and UR_RADC_i denote the ratings on the ith RADC utilizing the local model and the universal model, respectively. Once the learning process is converged, the optimal set of integration weights are obtained such that the universal RADC anomaly detection model is constructed.

Usually, when and how to determine the learning process is converged or not is an open problem. Although reaching a preset time frame or a certain accuracy threshold, the learning process of the FL might converge. However, in our scenario, there exists no public data sets to verify whether the universal RADC anomaly detection model can satisfy the accuracy threshold. Besides, setting a time frame in advance cannot guarantee the maximization of the reward. Thereby, to improve the RADC anomaly detection accuracy, the stabilization of the reward (10) is considered as the end of the FL learning process.

B. Phase-Two User Anomaly Detection

Since abnormal RADCs are detected, normal ones can help the GADC to identify abnormal users.

1) DDPG-Based User Anomaly Detection: The aim of user anomaly detection is to prevent privacy leakage caused by abnormal users. Different from RADC anomaly detection implemented based on action deviations, the privacy leakage is considered in user anomaly detection. Specifically, we evaluate the action UA_i of the ith user with a tuple that consist of the exposure of both sensitive information and nonsensitive information, i.e., $\mathrm{UA}_i = (\mathrm{SI}_{ij}, \mathrm{NSI}_{ij})$, where SI_{ij} represents the sensitive information (e.g., names, ages, sexes, occupations, etc.) of the jth user exposed by the ith user, while NSI_{ij} represents the nonsensitive ones exposed by the ith user. Thus, the privacy leakage degree PLD_i of the ith user can be quantified by

$$\mathrm{PLD}_i = \sum_{j} \gamma * \mathrm{SI}_{ij} + (1-\gamma) * \mathrm{NSI}_{ij}, \tag{11}$$

where γ represents the privacy leakage scale factor.

By evaluating the actions of each user, the RADC is able to determine whether this user is an abnormal one or a normal one, if the privacy leakage degree falls within the range of (0, 0.0.5) or [0.5, 1]. Similar to RADC anomaly detection, we define the state, the action, and the reward, respectively, for the DDPG-based user anomaly detection. Let the state s consist of the anomaly conditions of K users of each normal RADC, i.e., $s = (\mathrm{AC\_User}_1, \mathrm{AC\_User}_2, \ldots, \mathrm{AC\_User}_K)$, where $\mathrm{AC\_User}_i = 1$ denotes the ith user is abnormal, otherwise $\mathrm{AC\_User}_i = 0$. Accordingly, the privacy leakage scale factor γ serves as the action a. The action a_t is chosen based on the current state s_t according to the reward r_t. Once the privacy leakage scale factor is obtained, abnormal users will be detected if their privacy leakage degrees fall within the range of (0, 0.5). Then, abnormal users will be isolated such that they cannot expose others' sensitive information anymore. Thereby, the reward r_t should be designed referring to the overall privacy leakage as

$$r_t = -\sum_{i}^{K} \mathrm{PLD}_i. \tag{12}$$

And the collection of resulted anomaly conditions compose of the next state s_{t+1}. We are aware that normal users might be misjudged as abnormal ones. Therefore, we introduce the appeal mechanism for each misjudged normal user to reclaim their nonanomalies. Specifically, for each pair of users User_i and User_j, we evaluate the action difference using the Mahalanobis distance $M(UA_i, UA_j)$, i.e.,

$$M\left(UA_i, UA_j\right) = \sqrt{\left(\vec{UA_i} - \vec{UA_j}\right)^T \Sigma^{-1} \left(\vec{UA_i} - \vec{UA_j}\right)}. \tag{13}$$

Accordingly, the overall action difference between the ith user to the others, denoted by OAD_i, is calculated as $OAD_i = \sum_{j \neq i} M(UA_i, UA_j)$. Then, the all OAD_i are sorted in the ascending order. For the ith user, who claims the nonanomaly, the position of the OAD_i should be higher than the median of the OAD sequence. Obviously, such an appeal mechanism helps to improve user anomaly detection accuracy. The training process is executed similarly as that of DDPG-based RADC anomaly detection. When the learning process converges, each user is rated reasonably.

2) Universal User Anomaly Detection With Federated Learning: Similar to universal RADC anomaly detection, the GADC employs user anomaly detection models trained by RADCs to build the universal one with the FL. In general, the GADC generates a set of integration weights, based on which all user anomaly detection models are integrated into a universal one using the DRL algorithm DDPG within the FL framework. To minimize both FAR and MDR in user anomaly detection, we introduce the feedback mechanism for model generalization.

TABLE I
EXPERIMENT SETUP

Parameter	*Description*	*Range*
Num_GADC	Number of GADCs	1
Num_LADC	Number of LADCs	10
Num_RADC	Number of RADCs	[30,100]
Num_User	Number of Users	[400,750]
AR_U	Anomaly Rate of Users	[0.05, 0.15]
AR_R	Anomaly Rate of RADCs	[0.05, 0.15]
γ	Discount Factor	0.9
τ	Learning Rate	0.1

To be specific, the participation condition of RADCs consists of the state s, while the set of integration weights compose of the action a. Note that the universal user anomaly detection model design should mitigate the negative effect on detection accuracy due to the differences among regional models. To address this problem, we introduce the model-based feedback mechanism, in which the ratio between the universal model-based privacy leakage degree and the regional model-based privacy leakage degree is considered, to design the reward function. We then give the reward r as

$$r = \sum_{i}^{n'} \frac{\text{PLD}_i}{\text{UPLD}_i}, \tag{14}$$

where n' denotes the number of normal RADCs and PLD_i and UPLD_i denote the privacy leakage degree of the ith user utilizing the regional model and the universal model, respectively. Once the learning process converges, the universal user anomaly detection model is built.

V. PERFORMANCE EVALUATION

A. Simulation Setup

We conduct the experiment to evaluate the performance of the proposed scheme FLAD in Python on computers, each of which is equipped with Intel Core i7 processor, 64 G running memory, CPU frequency 6.4 GHz 64-bit win7 system. In addition, we use a computer to simulate the GADC. And up to ten computers are deployed, each of which is used to simulate an LADC and all regional ADC within. Table I gives the parameters of this simulation.

B. Performance Metrics

We validate the performance of FLAD in terms of system throughput, average latency, and anomaly detection accuracy while considering different *Num_RADC*, *Num_User*, *AR_U* and *AR_R*, respectively.

1) *System Throughput:* In general, a higher anomaly rate of users or RADCs will result in a lower system throughput.
2) *Latency:* Similar to the system throughput, the average latency will increase if the anomaly rate grows.
3) *Anomaly Detection Accuracy:* Both FAR and MDR are used to measure the anomaly detection accuracy.

C. Experiment Results

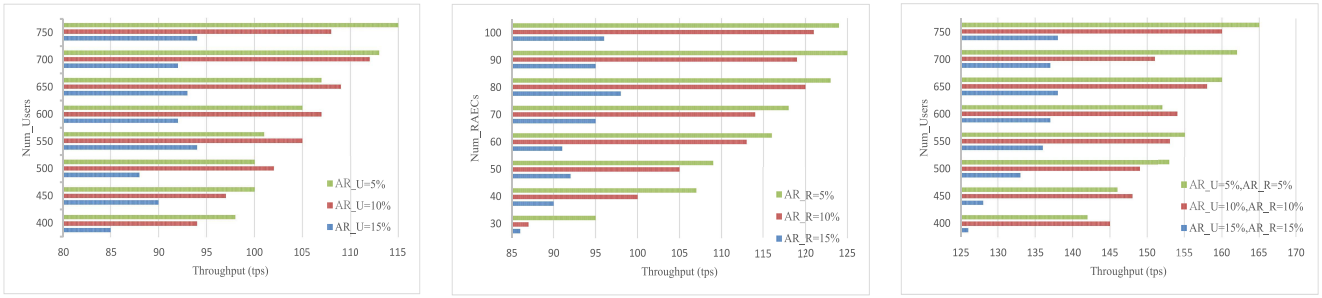
1) System Throughput: Fig. 3 gives the throughput of RADC, LADC, and GADC with the varying number of users *Num_User* and number of RADCs *Num_RADC*. As shown in Fig. 3(a), it is evident that the throughput rises as the *Num_User*. And the average throughput of RADC is about 106 tps, 102 tps, and 93 tps for $AR_U = 5\%$, $AR_U = 10\%$, and $AR_U = 15\%$, respectively. That indicates the fact that a lower AR_U contributes to a higher RADC throughput. The reason for that is once an user is detected as an abnormal one, the isolation of this user will result in the throughput drop. Once abnormal users are detected, the RADC throughput rises with the increasing *Num_User*. This is why the highest RADC throughput of $AR_U = 5\%$ is almost twice as much as that of $AR_U = 15\%$.

In Fig. 3(b), as the number of RADCs *Num_RADC* increases the LADC throughput grows due to the fact that more RADCs will contribute to more RADC anomaly detections of the LADC. In addition, we observe that with less RADC anomaly rate *AR_R* the LADC tends to achieve higher throughput. That is, for $AR_R = 15\%$, $AR_R = 10\%$, and $AR_R = 5\%$, the average throughput of the LADC is about 93 tps, 107 tps, and 113 tps, respectively. Note that although abnormal RADCs might misjudge users' actions, the proposed FLAD can shut down abnormal RADCs and let the users join other normal RADCs. However, normal RADCs might not be able to take in users from abnormal RADCs such that the LADC throughput still drops. there is a chance that normal RADCs might not be able to accept users from abnormal RADCs. Due to the effectiveness and efficiency of the proposed FLAD the highest LADC throughput reaches 125 tps.

Observed from Fig. 3(c), we know that as *Num_User* increases the GADC throughput increases. We define different combination of user anomaly rate and RADC anomaly rate as different cases for clarity, i.e., case 1 for $AR\_U = AR\_R = 5\%$, case 2 for $AR\_U = AR\_R = 10\%$, and case 3 for $AR\_U = AR\_R = 15\%$ throughout this article. The highest GADC throughput is 138 tps, 160 tps, and 165 tps for case 1, case 2, and case 3, respectively. This is because if there are less abnormal users and RADCs, then the GADC will isolate less abnormal users and shut down less abnormal RADCs such that the throughput still increases. Note that the proposed FLAD can detect abnormal users and RADCs efficiently such that the average GADC throughput is about 150 tps, 147 tps, and 135 tps for case 1, case 2, and case 3, respectively.

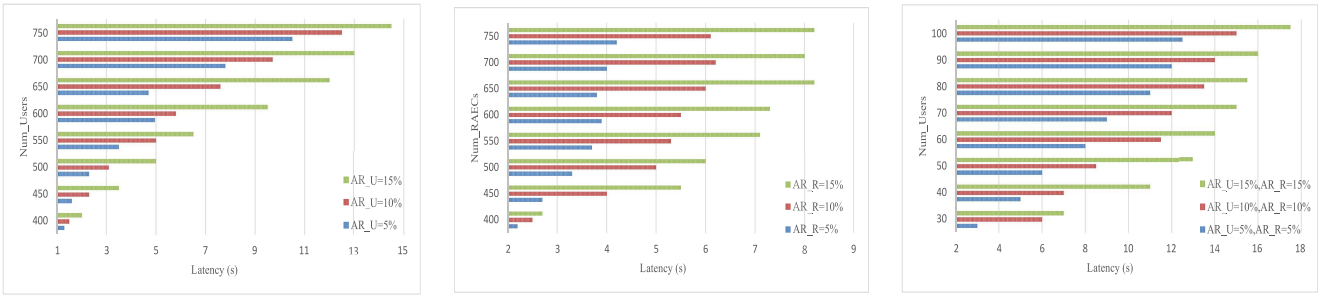
2) Latency: In Fig. 4, the latency comparison, considering both user anomaly rate *AR_U* and RADC anomaly rate *AR_R*, is presented with the varying number of users *Num_User* and number of RADCs *Num_RADC*. As shown in Fig. 4(a), we find that the latency increases as the *Num_User*. And a higher latency is resulted from a higher *AR_U*, i.e., the RADC latency is about 10.5 s, 12.5 s, and 14.5 s, when *AR_U* reaches 5%, 10%, and 15%, respectively. This is because when more users are involved, more users' actions should be compared by the RADC to distinguish abnormal ones. Therefore, the latency increases. Besides, a higher *AR_U* suggests more abnormal users should be detected such that the average latency rises. Moreover, the average latency of RADC is only 7 s when the *AR_U* reaches 15%.

Observed from Fig. 4(b), we find that the latency rises as *Num_RADC* and the average latency is about 3.5 s, 5 s, and



(a) (b) (c)

Fig. 3. System throughput while varying the number of users *Num_User* and the number of RADCs *Num_RADC*. (a) Throughput of RADC. (b) Throughput of LADC. (c) Throughput of GADC.



(a) (b) (c)

Fig. 4. Average latency while varying while varying the number of users *Num_User* and the number of RADCs *Num_RADC*. (a) Latency of RADC. (b) Latency of LADC. (c) Latency of GADC.

6.5 s for $AR_R = 5\%$, $AR_R = 10\%$, and $AR_R = 15\%$, respectively. That indicates a higher AR_R results in a higher LADC average latency. The reason for that is when more RADCs become abnormal, more RADC action comparisons should be executed by the LADCs for the anomaly detection such that the latency rises. With more abnormal RADCs shut down, the average latency grows slowly. In addition, we observed that the highest average latencies are only 4.2 s, 6.3 s, and 8.2 s for $AR_R = 5\%$, $AR_R = 10\%$, and $AR_R = 15\%$, respectively.

As shown in Fig. 4(c), for all cases, the increment of *Num_User* leads to the growth of GADC latency. Case 3 achieves the 17.8 s latency, which is highest one, compared with 15 s of case 2 and 12.3 s of case 1. The reason for that is as *Num_User* increases more abnormal users and RADCs occur. That suggest more comparisons should be done between users' actions and between RADCs' actions by the GADC. Therefore, the GADC latency rises. Furthermore, even the abnormal users are isolated, shutting down abnormal RADCs definitely imposes the burdens on normal RADCs such that the latency increases for both RADCs and the GADC. However, even the latency rises as *Num_User*, the average latency for each case is only 13.5 s, 11 s, and 9 s. That suggests the proposed FLAD can efficiently detect abnormal users and RADCs to preserve privacy.

3) Anomaly Detection Accuracy: Fig. 5 compares the anomaly detection accuracy in FAR and MDR, considering both user anomaly rate *AR_U* and RADC anomaly rate *AR_R*, while varying the number of users *Num_User* and the number of RACDs *Num_RADCs* with/without the RADC anomaly detection. As shown in Fig. 5(a)–(c), we find that as *Num_User* increases both FAR and MDR increase at first and level off eventually. Without the RADC anomaly detection (RADC_AD), the average FAR is around 22% of case 1, 26% of case 2, and 29% of case 3, respectively, compared with 3% of case 1, 5% of case 2, and 6% of case 3 of the average FAR with the RADC anomaly detection, respectively. In addition, without the RADC anomaly detection, the average MDR reaches 13.5% of case 1, 14% of case 2, and 16% of case 3, compared with 2% of case 1, 3% of case 2, and 6% of case 3 of the average MDR with the RADC anomaly detection, respectively. That is, as *Num_User* increases, either FAR or MDR is much lower with the RADC anomaly detection. The reason for that is abnormal RADCs tend to misjudge users' actions. And that also explains why a higher RADC anomaly rate and a higher user anomaly rate will result in higher FAR and MDR. Observed from Fig. 5(d)–(f), we find that as *Num_RADC* increases both FAR and MDR fluctuate. And without the RADC anomaly detection, the average FAR is about 23% of case 1, 25% of case 2, and 26% of case 3, respectively, compared with 5% of case 3, 3.5% of case 2, and 5% of case 3 of the average FAR with the RADC anomaly detection, respectively. In addition, without the RADC anomaly detection, the average MDR reaches 12% of case 1, 14% of case 2, and 14.5% of case 3, compared with 2.5% of case 1, 3% of case 2, and 4% of case 3 of the average MDR with the RADC anomaly detection, respectively. That is, as *Num_RADC* increases, either FAR or MDR is much lower with the RADC anomaly detection. Observed from Fig. 5(c) and (d), we find that a higher RADC anomaly rate and a higher user anomaly rate do not always result in

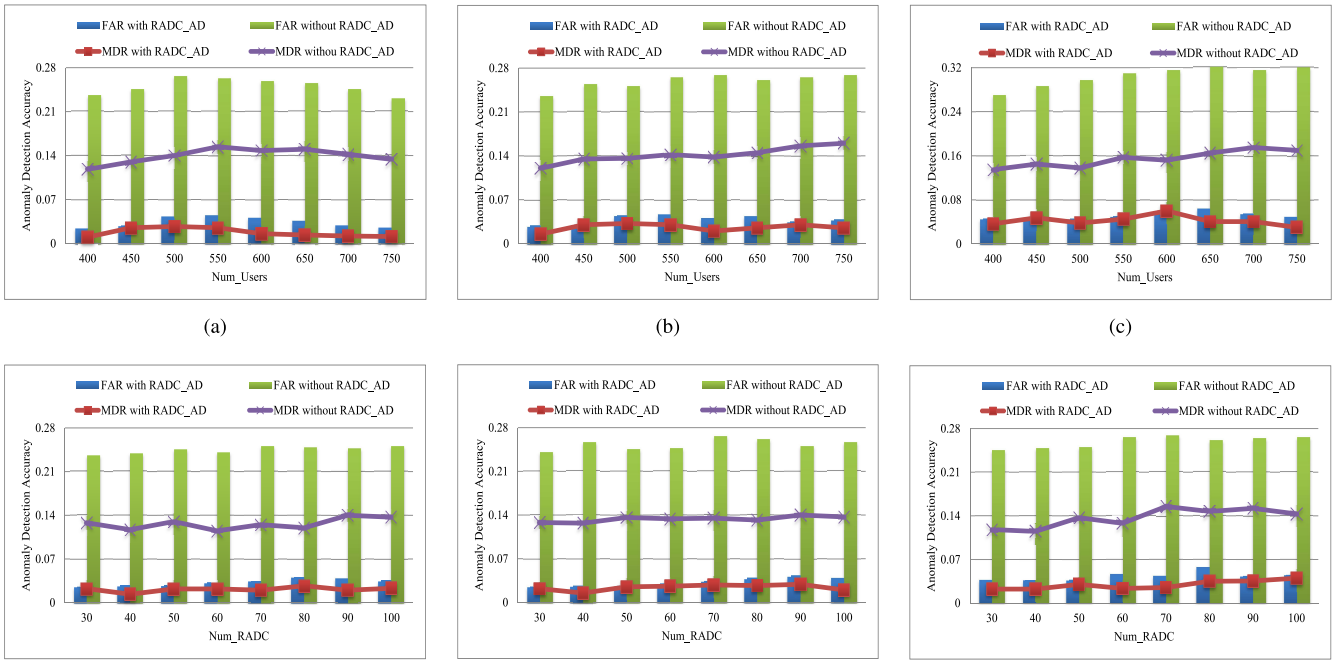


Fig. 5. FAR and MDR while varying the number of users *Num_Users* and the number of RADCs *Num_RADCs*, considering both user anomaly rate *AR_U* and RADC anomaly rate *AR_R*, with/without RADC anomaly detection. (a) AR_U = AR_R = 5%, 400 ≤ Num_User ≤ 750. (b) AR_U = AR_R = 10%, 400 ≤ Num_User ≤ 750. (c) AR_U = AR_R = 15%, 400 ≤ Num_User ≤ 750. (d) AR_U = AR_R = 5%, 30 ≤ Num_RADC ≤ 100. (e) AR_U = AR_R = 10%, 30 ≤ Num_RADC ≤ 100. (f) AR_U = AR_R = 15%, 30 ≤ Num_RADC ≤ 100.

higher FAR and MDR with the RADC anomaly detection. This is because the proposed FLAD can accurate detect anomalies of RADCs and users such that both FAR and MDR are greatly reduced.

VI. Conclusion

To enhance privacy and security for various IIoT applications, we propose a federated DRL empowered anomaly detection scheme for IIoT. Specifically, the anomaly detection is implemented utilizing the FL without aggregating local data sets for user privacy preservation. To increase the detection accuracy, the degree of privacy leakage is quantified and the relation between actions is discovered. Then, abnormal users are accurately and efficiently detected using the federated DRL algorithm. The experiment results indicate that the proposed scheme achieves the accurate anomaly detection for user privacy preservation with high throughput and low latency in IIoT.

References

[1] U. Sendler, *The Internet of Things: Industrie 4.0 Unleashed*. Berlin, Germany: Springer, 2018, doi: 10.1007/978-3-662-54904-9_9.

[2] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial Internet of Things: Challenges, opportunities, and directions," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4724–4734, Nov. 2018.

[3] S. Mumtaz, A. Alsohaily, Z. Pang, A. Rayes, K. F. Tsang, and J. Rodriguez, "Massive Internet of Things for industrial applications: Addressing wireless IIoT connectivity challenges and ecosystem fragmentation," *IEEE Ind. Electron. Mag.*, vol. 11, no. 1, pp. 28–33, Mar. 2017.

[4] M. Aazam, K. A. Harras, and S. Zeadally, "Fog computing for 5G tactile Industrial Internet of Things: QoE-aware resource allocation model," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3085–3092, May 2019.

[5] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in Internet of Things: Towards secure and privacy-preserving fusion," *Inf. Fusion*, vol. 51, pp. 129–144, Nov. 2019.

[6] J. H. Park, "Advances in future Internet and the Industrial Internet of Things," *Symmetry*, vol. 11, no. 2, p. 244, 2019.

[7] K. Zhang, Y. Zhu, S. Maharjan, and Y. Zhang, "Edge intelligence and blockchain empowered 5G beyond for the Industrial Internet of Things," *IEEE Netw.*, vol. 33, no. 5, pp. 12–19, Sep./Oct. 2019.

[8] H. Lin, S. Garg, J. Hu, G. Kaddoum, M. Peng, and M. S. Hossain, "Blockchain and deep reinforcement learning empowered spatial crowdsourcing in software-defined Internet of Vehicles," *IEEE Trans. Intell. Transp. Syst.*, early access, Oct. 15, 2020, doi: 10.1109/TITS.2020.3025247.

[9] S. Savazzi, M. Nicoli, and V. Rampa, "Federated learning with cooperating devices: A consensus approach for massive IoT networks," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4641–4654, May 2020.

[10] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, "Federated learning for ultra-reliable low-latency V2V communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–7.

[11] S. Wang *et al.*, "Adaptive federated learning in resource constrained edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1205–1221, Jun. 2019.

[12] J. Pang, Y. Huang, Z. Xie, Q. Han, and Z. Cai, "Realizing the heterogeneity: A self-organized federated learning framework for IoT," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3088–3098, Mar. 2021. doi: 10.1109/JIOT.2020.3007662.

[13] J. Wang, J. Hu, G. Min, A. Zomaya, and N. Georgalas, "Fast adaptive task offloading in edge computing based on meta reinforcement learning," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 1, pp. 242–253, Jan. 2021.

[14] C. Wang, B. Wang, H. Liu, and H. Qu, "Anomaly detection for industrial control system based on autoencoder neural network," *Wireless Commun. Mobile Comput.*, vol. 2020, pp. 1–10, Aug. 2020, doi: 10.1155/2020/8897926.

[15] B. Genge, P. Haller, and C. Enăchescu, "Anomaly detection in aging Industrial Internet of Things," *IEEE Access*, vol. 7, pp. 74217–74230, 2019.

[16] Y. Liu, N. Kumar, Z. Xiong, W. Y. B. Lim, J. Kang, and D. Niyato, "Communication-efficient federated learning for anomaly detection in industrial Internet of Things," in *Proc. GLOBECOM*, 2020, pp. 1–6.

[17] L. Yi *et al.*, "Deep anomaly detection for time-series data in industrial IoT: A Communication-efficient on-device federated learning approach," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6348–6358, Apr. 2021, doi: 10.1109/JIOT.2020.3011726.
[18] B. Li, Y. Wu, J. Song, R. Lu, T. Li, and L. Zhao, "DeepFed: Federated deep learning for intrusion detection in industrial cyber-physical systems," *IEEE Trans. Ind. Informat.*, early access, Sep. 11, 2020, doi: 10.1109/TII.2020.3023430.
[19] S. M. Taghavinejad, M. Taghavinejad, L. Shahmiri, M. Zavvar, and M. H. Zavvar, "Intrusion detection in IoT-based smart grid using hybrid decision tree," in *Proc. 6th Int. Conf. Web Res. (ICWR)*, 2020, pp. 152–156.
[20] C. Wang, "IoT anomaly detection method in intelligent manufacturing industry based on trusted evaluation," *Int. J. Adv. Manuf. Technol.*, vol. 107, no. 1, pp. 993–1005, 2020.
[21] D. Wu, Z. Jiang, X. Xie, X. Wei, W. Yu, and R. Li, "LSTM learning with Bayesian and Gaussian processing for anomaly detection in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5244–5253, Aug. 2020.
[22] J. Mills, J. Hu, and G. Min, "Communication-efficient federated learning for wireless edge intelligence in IoT," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5986–5994, Jul. 2020.

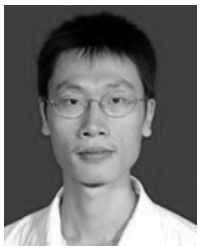
Xiaoding Wang received the Ph.D. degree from the College of Mathematics and Informatics, Fujian Normal University, Fuzhou, China, in 2016.

He is an Associate Professor with Fujian Normal University.

His main research interests include network optimization and fault tolerance.

Sahil Garg (Member, IEEE) is currently working as a Postdoctoral Research Fellow with the Department of Electrical Engineering, École de technologie supérieure, Université du Québec, Montreal, QC, Canada. Some of his research findings are published in top-tier journals, such as IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING, IEEE INTERNET OF THINGS JOURNAL, *IEEE Network Magazine*, *IEEE Communications Magazine*, *IEEE Wireless Communications Magazine*, *IEEE Consumer Electronics Magazine*, *Future Generation Computer Systems* (Elsevier), *Information Sciences* (Elsevier), and various International conferences of repute, such as IEEE Globecom, IEEE ICC, IEEE WCNC, IEEE VTC, IEEE Infocom Workshops, ACM MobiCom Workshops, and ACM MobiHoc Workshops. His research interests include machine learning, big data analytics, knowledge discovery, cloud computing, Internet of Things, software defined networking, and vehicular *ad hoc* networks.

Mr. Garg was a recipient of the prestigious Visvesvaraya Ph.D. Fellowship from the Ministry of Electronics and Information Technology under Government of India from 2016 to 2018. For his research, he also received the IEEE ICC Best Paper Award in 2018 at Kansas City, USA. He serves as the Managing Editor for *Human-Centric Computing and Information Sciences* (Springer) and an Associate Editor of *IEEE Network Magazine*, *Applied Soft Computing* (Elsevier), and *Wiley's International Journal of Communication Systems* (Elsevier). He also serves as the Workshops and Symposia Officer for the IEEE ComSoc Emerging Technology Initiative on Aerial Communications.



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

Jia Hu received the B.Eng. and M.Eng. degrees in electronic engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2004 and 2006, 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, University of Exeter, Exeter, U.K. He has published over 80 research papers within these areas in prestigious international journals and reputable international conferences. His research interests include edge-cloud computing, resource optimization, applied machine learning, and network security.

Dr. Hu has received the Best Paper Awards at IEEE SOSE'16 and IUCC14. He serves on the editorial board of *Computers and Electrical Engineering* (Elsevier) and he has guest-edited many special issues on major international journals (e.g., IEEE Internet of Things Journal, *Computer Networks*, and *Ad Hoc Networks*). He has served as the General Co-Chair of IEEE CIT'15, IUCC'15, and the Program Co-Chair of IEEE ISPA'20, ScalCom'19, SmartCity'18, CYBCONF'17, and EAI SmartGIFT'2016.

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 has published over 200 journal and conference papers and two pending patents.

Dr. Kaddoum was a recipient of the "Research Excellence Award of the Université du Quebec, 2018" and the "Research Excellence Award-emerging researcher" from ÉTS, 2019. Additionally, he is a co-recipient of the Best Papers Awards of the IEEE PIMRC 2017 and the IEEE WiMob 2014. Moreover, he received the "Exemplary Reviewer Award" from IEEE TRANSACTIONS ON COMMUNICATIONS twice in 2015 and 2017. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and the IEEE COMMUNICATIONS LETTERS. He held the ÉTS Research Chair in physical-layer security for wireless networks.

Md. Jalil Piran (Senior Member, IEEE) received the Ph.D. degree in electronics and information engineering from Kyung Hee University, Seoul, South Korea, in 2016.

He is currently an Assistant Professor with the Department of Computer Science and Engineering, Sejong University, Seoul, South Korea. He has published a substantial number of technical papers in well-known international journals and conferences in the area of intelligent information and communication technology, specifically in the fields of: wireless communications and networking; 5G/6G, Internet of Things, data science, machine learning, and security. He continued his research carrier as a Post-Doctoral Fellow of Information and Communication Engineering with the Networking Laboratory, KyungHee University.

Dr. Piran received the IAAM Scientist Medal of the year 2017 for notable and outstanding research in new age technology and innovation, Stockholm, Sweden. He has been recognized as the Outstanding Emerging Researcher by the Iranian Ministry of Science, Technology, and Research in 2017. His Ph.D. dissertation has been selected as the "Dissertation of the Year 2016" by the Iranian Academic Center for Education, Culture, and Research in the Engineering Group. In the worldwide communities, he has been, an Active Delegate from South Korea in the Moving Picture Experts Group since 2013, and an Active Member of the International Association of Advanced Materials since 2017.

M. Shamim Hossain (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Ottawa, Ottawa, ON, Canada, in 2019.

He is a Professor with the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He is also an Adjunct Professor with the School of Electrical Engineering and Computer Science, University of Ottawa. He has authored and coauthored more than 300 publications, including refereed journals conference papers, books, and book chapters. Recently, he Co-Edited a book on *Connected Health in Smart Cities* (Springer). His research interests include cloud networking, smart environment (smart city, smart health), AI, deep learning, edge computing, Internet of Things, multimedia for health care, and multimedia big data.

Prof. Hossain is the Chair of IEEE Special Interest Group on AI for Health with IEEE ComSoc eHealth Technical Committee. He is also the Co-Chair of the 1st IEEE GLOBECOM 2021 Workshop on Edge-AI and IoT for Connected Health. He is on the editorial board of the IEEE TRANSACTIONS ON MULTIMEDIA, IEEE MULTIMEDIA, *IEEE Network Magazine*, IEEE WIRELESS COMMUNICATIONS, IEEE ACCESS, *Journal of Network and Computer Applications* (Elsevier), and *International Journal of Multimedia Tools and Applications* (Springer). He also currently serves as a Lead Guest Editor for *IEEE Network Magazine*, *ACM Transactions on Internet Technology*, *ACM Transactions on Multimedia Computing, Communications, and Applications*, and *Multimedia Systems*. He is an IEEE ComSoc Distinguished Lecturer. He is a Senior Member of the ACM.