Privacy-Enhanced Data Fusion for COVID-19 Applications in Intelligent Internet of Medical Things

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

***Abstract*—With the worldwide large-scale outbreak of COVID-19, the Internet of Medical Things (IoMT), as a new type of Internet of Things (IoT)-based intelligent medical system, is being used for COVID-19 prevention and detection. However, since the widespread use of IoMT will generate a large amount of sensitive information related to patients, it is becoming more and more important yet challenging to ensure data security and privacy of COVID-19 applications in IoMT. The leakage of private information during IoMT data fusion process will cause serious problems and affect people's willingness to contribute data in IoMT. To address these challenges, this article proposes a new privacy-enhanced data fusion strategy (PDFS). The proposed PDFS consists of four important components, i.e., sensitive task classification, task completion assessment, incentive mechanism-based task contract design, and homomorphic encryption-based data fusion. The extensive simulation experiments demonstrate that PDFS can achieve high task classification accuracy, task completion rate, task data reliability and task participation rate, and low average error rate, while improving the privacy protection for data fusion under COVID-19 application environments based on IoMT.**

***Index Terms*—COVID-19, data fusion, deep reinforcement learning, Internet of Medical Things (IoMT), privacy protection.**

I. Introduction

With the application and development of the Internet of Things (IoT) technology in the medical field, the Internet of Medical Things (IoMT) that collects, processes, and analyzes the medical data generated by various IoT devices, has also seen rapid progress [1]. IoMT can effectively improve the accessibility and efficiency of disease treatment, reduce errors, improve patient experience, and provide lower costs [2]. Recently, the worldwide large-scale outbreak of COVID-19 has brought tremendous pressure and challenges to existing medical detection systems, and has put forward new requirements for the timeliness, accuracy, and reliability of medical detection data. In this situation, IoMT is expected to be used to collect and analyze the main symptoms of COVID-19 patients by providing large-scale real-time detection data and tracking the source of the disease outbreak [3]. However, since the widespread use of IoMT especially during the data fusion process will generate a large amount of sensitive information related to patients, and the leakage of private information will cause serious problems and affect people's willingness to contribute data in IoMT, it is becoming more and more important yet challenging to design privacy-enhanced data fusion technologies for ensuring data privacy for COVID-19 applications in IoMT [4]–[6].

In the IoMT-based COVID-19 applications, various IoT-based medical detection equipment will generate a large amount of application data [7], [8]. According to the actual needs of the COVID-19 applications, these data need real-time processing for fast decision making. However, most existing IoMT systems store data in the cloud, completely relying on remote cloud servers for data processing and analysis [9]. As the number of IoMT devices and the generated data continues to increase, the network pressure grows and the delay increases, which may lead to failures or erroneous diagnosis and seriously affect COVID-19 data detection and service response. To overcome the above deficiencies and provide better services for COVID-19 applications, the traditional cloud IoMT architecture needs to be improved. Based on the above analysis, this article builds a new IoMT architecture MEC-IoMT combining IoMT and multiaccess edge computing (MEC) [10], [11]. As shown in Fig. 1, the new MEC-IoMT consists of three important components, i.e., intelligent medical data collection terminals, the multiaccess edge network, and the remote COVID-19 applications and services center. Specifically, the intelligent medical data collection terminals, such as cameras, electronic thermometer and wearable detection sensors, and so on, are responsible for intelligent collection of COVID-19 disease detection data. The multiaccess

Manuscript received July 28, 2020; revised September 17, 2020; accepted October 7, 2020. Date of publication October 22, 2020; date of current version October 22, 2021. 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: Chair of Pervasive and Mobile Computing. *(Corresponding authors: Jia Hu; Xiaoding Wang.)*

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

Sahil Garg is with the Department of Electrical Engineering, École de technologie supérieure, Montreal, QC H3C 1K3, Canada (e-mail: sahil.garg@ieee.org).

Jia Hu is with the School of Mathematics and 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 Chair of Pervasive and Mobile Computing and 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.2020.3033129

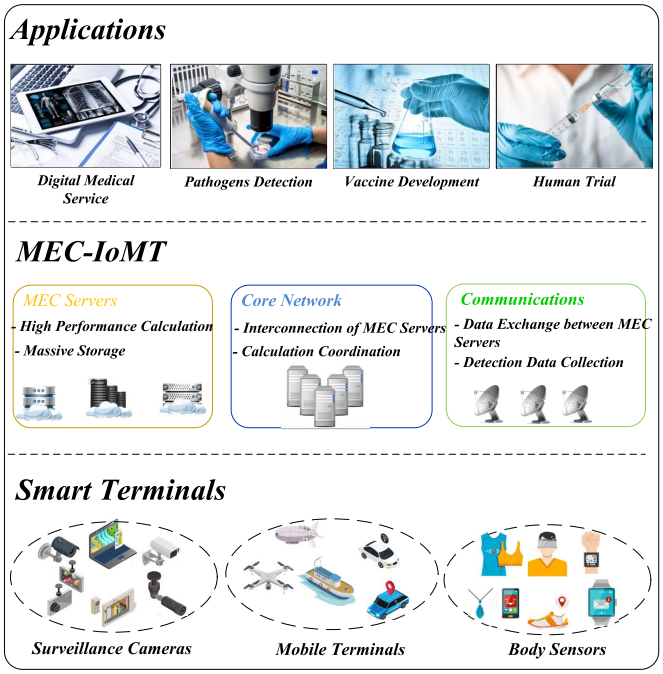


Fig. 1. Architecture of MEC-IoMT.

edge network composed of MEC servers and wireless access stations is employed to provide various COVID-19 applications, implement communication, and data exchange among MEC servers. Furthermore, the remote COVID-19 applications and services centers are responsible for the storage, processing, and analysis of large-scale COVID-19 disease detection data or detection results. In MEC-IoMT, first, pervasive smart data collection terminals are employed to collect various COVID-19 disease detection data by using the mobile-edge crowdsensing technology. Next, the collected data is transmitted over the wireless network to the COVID-19 applications deployed on the edge network servers, which implements the large scale and effective COVID-19 data fusion, analysis, and processing. Finally, the results of analysis and processing are sent to the individuals and COVID-19 control centers, which stores these data, further analyzes and processes the data, and forecasts the development trend of the COVID-19 disease. The MEC-IoMT architecture migrates services located in the remote cloud center to the distributed edge network, closer to the user, thereby reduces the service delay and communication overhead [12].

In the prevention and treatment of COVID-19, the data fusion process consists of data collection, analysis, and processing. It is one of the most important links that decides whether the spreading can be controlled in time. At the same time, it is also the premise and basis for patients to be treated in time. However, the heterogeneous data in the data fusion process contain much privacy information related to patients, while the collection terminal and processing center cannot be fully trusted [13], [14]. They may launch active attacks, or passive attacks after being captured by the attacker, which lead to the leakage of private information. To address this challenge, we integrate artificial intelligence technology, such as deep deterministic policy gradient (DDPG) [11] into privacy strategy and propose a new privacy-enhanced data fusion strategy (PDFS) for IoMT. The main contributions of our work are as follows.

1) To protect sensitive information from malicious test subjects, a novel task security level-based privacy-aware data fusion tasks classification mechanism is proposed to assure that a task can be accepted and the sensitive information in the task can be obtained only if the test subject's security level is higher than that of the task.
2) To design a reasonable task contract, a deep reinforcement learning algorithm, DDPG, is applied to reward/punish test subjects for their superior/poor performances in task completion and meanwhile to help the center for disease control (CDC) to set the appropriate payments with respect to test subjects' performances. In addition, the data reliability is validated utilizing DDPG as well for validation accuracy improvement.
3) To protect the test subject's security from both the fusion center and the detection center, a homomorphic encryption-based data fusion mechanism is proposed by introducing random numbers for real identities perturbation to hide the true identities of test subjects in the data fusion process.
4) The theoretical analysis and validation experiments demonstrate that: a) PDFS has advantages in task classification accuracy, average error rate, completion rate, data reliability, and participation rate compared with baseline strategies and b) PDFS is efficient against both task privacy attack and identity privacy attack.

II. Related Work

Data fusion, as one of the most important links in the prevention and treatment of the COVID-19 in Intelligent IoMT, has received extensive attention, and some relevant research results have emerged. In [15] and [16], researchers summarize the data security and privacy protection requirements and challenges in IoMT systems, and suggest the future directions for research on security and privacy. Tang *et al.* [5] proposed a privacy protection and incentives-based data fusion strategy to implement privacy security and fair incentives for contributing patients in the process of health data collection. Guan *et al.* [17] proposed a privacy preserving and authentication-based data fusion scheme to provide device-oriented anonymous privacy protection in fog-aided IoMT. Yang *et al.* [18] proposed a differential privacy and machine-learning-based multifunctional data fusion strategy to provide statistical fusion functions for IoMT applications. Li *et al.* [19] proposed a data fusion scheme supporting privacy preserving and publicly fusion result verification in IoT application systems such as IoMT. Wu *et al.* [20] proposed a novel data fusion mechanism to assist the fusion servers to realize the privacy-preserving data fusion by integrating fog computing and homomorphic encryption techniques. Yang *et al.* [21] proposed a blockchain-based privacy preservation framework, which uses the anonymous nature of blockchain to protect workers privacy during data fusion. Li *et al.* [22] proposed a

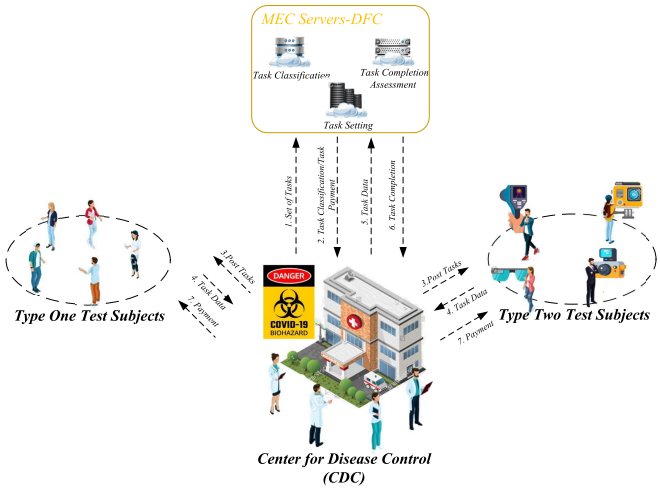


Fig. 2. System model of the proposed PDFS.

novel decentralized framework to realize privacy preservation through the use of blockchain, pseudonym, and encryption-based distributed storage technology.

Although these works contribute to privacy-preserving data fusion in IoMT, there still remain two challenges: 1) how to prevent the leakage of private information contained in sensitive tasks and 2) how to prevent the different task receivers who accept the task from colluding to share the fragmented privacy contained in the task. In this article, a new PDFS is proposed to address these two problems.

III. System Model

In this article, a new PDFS is proposed to defend the collusion attacks and prevent the leakage of private information contained in sensitive tasks. As shown in Fig. 2, the system model of PDFS mainly considers three entities, i.e., the CDC, the test subjects, and the data fusion center (DFC).

Specifically, the single CDC case is considered in this article. For COVID-19 surveillance, the CDC posts a series of health condition fusion task $\{\text{Task}_i\}$. Each task Task_i has one of K specific security levels and a corresponding payment. Once the task is completed, the CDC will determine whether to pay the test subjects based on the reliability reports provided by the DFC. In addition, PDFS includes two types of test subjects are considered, i.e., the type one test subjects, the ith of which is denoted by $\text{TS}_{1,i}$, are only responsible for providing the personal health condition data; and the type two test subjects, the ith of which is denoted by $\text{TS}_{2,i}$, provide health condition data of type one test subjects and their own as well. Due to the privacy concern, each test subject has a security level SL, i.e., $SL_{\text{TS}_{1,i}}$ for a type one test subject while $SL_{\text{TS}_{2,i}}$ for a type two test subject. A task Task_i is accepted only if the security level of the task SL_{Task_i} is inferior to that of a test subject, which is $SL_{\text{Task}_i} \leq SL_{\text{TS}_{1,i}}$ or $SL_{\text{Task}_i} \leq SL_{\text{TS}_{2,i}}$. The DFC, which is employed by the CDC and built on MEC servers with high-performance computing infrastructures, classifies the task set into groups according to the privacy requirements and performs the task completion assessment. Moreover, based on the assessment report, the DFC aids the task contract design by calculating the payment to the test subject who claims to have completed the task.

In general, to prevent the COVID-19, the CDC posts a series of data fusion tasks, i.e., gathering personal health data, for the COVID-19 vaccine development. Then, test subjects, who sign the contract of data fusion tasks, provide health data of their own or others' to the CDC via DFC. Due to the privacy concern, test subjects might provide inaccurate health data or even fake ones. In this case, DFC assesses the reliability of each data to determine whether the test subject should be given the payment, i.e., the test subject whoever provides the reliable data will receive the payment. Due to the computation resources required in data reliability validation and contract design for data fusion task, all computations are implemented on MEC servers connected by the core networks.

We consider the privacy disclosure problem among all three entities: 1) the CDC; 2) the test subjects; and 3) the DFC. In fact, the CDC suffers the task privacy attack from both malicious test subjects and the DFC, while the test subjects are vulnerable to the identity privacy attack.

1) *Task Privacy Attack:* The task privacy attack is defined as the unauthorized access launched by the malicious test subjects and the DFC to the sensitive information contained in tasks.
2) *Identity Privacy Attack:* The identity privacy attack is defined as the test subjects' true identities is disclosed and theft by malicious DFC and center for disease control.

IV. Implementation Details of the PDFS

The proposed PDFS consists of four important components, i.e., sensitive task classification, task completion assessment, task contract design, and homomorphic encryption-based data fusion.

A. Sensitive Task Classification

We develop a K-means-based privacy-preserving classification mechanism for sensitive tasks partition with the task privacy hidden from DFC. Note that the traditional K-means-based classification simply assigns each data to the nearest group center. For example, the Euclidean distance between a group G_j and a task Task_i is calculated as

$$D_{ij} = \left(\text{Task}_i - g_j\right)^T \left(\text{Task}_i - g_j\right) \tag{1}$$

where g_j denotes the center of group G_j. Then, we compute the deviation of the distance from Task_i to group G_j and the distance from Task_i to group G_l as

$$D_{ij} - D_{il} = g_j^T g_j - g_l^T g_l - \text{Task}_i^T \left(g_j - g_l\right) - \left(g_j - g_l\right)^T \text{Task}_i. \tag{2}$$

Obviously, Task_i should be assigned to group G_j if $D_{ij} - D_{il} < 0$; otherwise, Task_i belongs to group G_l. The closest group center to each task Task_i can be repeatedly identified $(k-1)$ times. However, the DFC can directly obtain sensitive information contained in task Task_i during the process that results in the privacy disclosure.

To solve this problem, a privacy-preserving k-means strategy is proposed to ensure the privacy protection of each task data Task_i. The details are given as follows. DFC generates a set of fake centers $\{g'_j\}_{j=1}^{k-1}$ under the constraint that

$$\begin{cases} \left(g'_j - g'_l\right)^T \left(g'_j - g'_l\right) = 0 \\ \left|g'_j - g'_l\right| \neq 0 \end{cases} \tag{3}$$

where $j \neq l$. Then, DFC sends the set of distances between fake centers, i.e., $\{(g'_j - g'_l)\}$, to the CDC. For each task Task_i, the CDC calculates the perturbed task Task_{ijl} as

$$\text{Task}_{ijl} = \text{Task}_i + t_{ijl}\left(g'_j - g'_l\right) \tag{4}$$

where t_{ijl} represents a random number used to perturb $g'_j - g'_l$ against the exposure of Task_i. Then, the CDC sends each perturbed task Task_{ijl} to the DFC. Accordingly, the distance deviation can be rewritten as

$$\begin{aligned} D_{ij} - D_{il} = &\left(\text{Task}_{ijl} - g_j\right)^T \left(\text{Task}_{ijl} - g_j\right) \\ &- \left(\text{Task}_{ijl} - g_l\right)^T \left(\text{Task}_{ijl} - g_l\right). \end{aligned} \tag{5}$$

Similarly, if $D_{ij} - D_{il} < 0$, then Task_i is assigned to group G_j otherwise G_l. This process is still repeated $(k-1)$ times to determine the nearest center to Task_i. Through the data perturbation, the DFC knows neither the original task Task_i nor the group it belongs to such that the privacy of the CDC is preserved. The correctness of the propose privacy-preserving k-means-based classification is proved in the following theorem.

Theorem 1: Both fake centers and data perturbation will not affect the result of the proposed privacy-preserving k-means strategy.

Proof: We prove this theorem, by verifying whether the result of $(D_{ij} - D_{il})$ calculated by our strategy is the same as that calculated by the original k-means algorithm.

For each perturbed task $Task_{ijl}$, the distance deviation between $Task_{ijl}$ to center g_j and g_l can be calculated in the exactly form of (2) as

$$\begin{aligned} D_{ij} - D_{il} &= g_j^T g_j - g_l^T g_l - \left(\text{Task}_i + t_{ijl}\left(g'_j - g'_l\right)\right)^T \left(g_j - g_l\right) \\ &\quad - \left(g_j - g_l\right)^T \left(\text{Task}_i + t_{ijl}\left(g'_j - g'_l\right)\right) \\ &= \left(\text{Task}_i - g_j\right)^T \left(\text{Task}_i - g_j\right) \\ &\quad - \left(\text{Task}_i - g_l\right)^T \left(\text{Task}_i - g_l\right). \end{aligned}$$

■

Once the classification is completed, each group G_j, $1 \leq j \leq K$, is given a specific security level such that test subjects are only allowed to accept the tasks of corresponding security levels.

B. Task Contract Design

How much money should be paid for task fulfillment is determined by the CDC. In fact, two dominant factors should be considered in the payment determination. For example, test subjects might be reluctant accepting the tasks even if their security levels are above that of the task. That suggests the CDC should give an "appeal" offer to attract test subjects. On the other hand, even if test subjects are paid enough, their works could fail to reach the requirement, i.e., data provided by test subjects are unreliable. These two cases suggest the payment to the test subjects should be set dynamically with respect to the assessment of task completion.

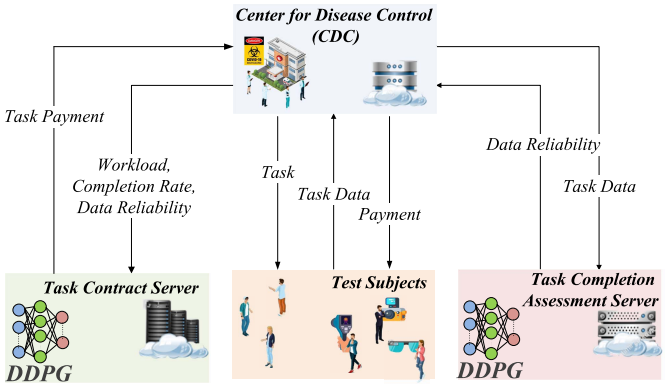


Fig. 3. DDPG-based task contract design.

To solve this problem, we develop a DRL-based task contract design utilizing the DDPG to meet the satisfaction of both CDC and test subjects. In fact, reinforcement learning algorithms, i.e., DQN, could be useful in discovering optimal strategies. However, the optimal threshold searching is implemented in a continuous space to ensure the assessment accuracy. That suggests traditional reinforcement learning-based methods cannot meet our requirement. For example, DQN can only work well in a discretized action space. We therefore apply the DDPG to learn the optimal threshold. In general, a DDPG decision system consists of a critic network Q, a target critic network Q', an actor networks π, and the corresponding target actor networks π, the parameters of which are denoted by ϑ^Q, $\vartheta^{Q'}$, ϑ^π, and $\vartheta^{\pi'}$, respectively. In addition, the experience is stored in an experience pool $\mathcal{P}$ in the form of a transition of a quadruple (s_t, a_t, r_t, s_{t+1}), the structure of the DDPG-based task contract design is given in Fig. 3.

As a DRL, DDPG requires three basic components, i.e., state, action, and reward. The action is given at a state to obtain the reward and then the next state is observed from the environment. In the incentive mechanism design, we consider the triple (WL, CP, DR) as a state s, where WL, CR, and DR represent the workload, completion rate, and data reliability of the task, respectively. The payment Pay from the CDC to the test subject is used as the action a, while the reward is denoted by r. This is because a test subject TS_i who accepts a task Task_j might not complete Task_j as he/she claims. Besides, the data aggregated in the Task_j could be unreliable. That indicates the ratio between the utility of the CDC Uti_j and the payment Pay_j to the test subject on this task Task_j is eligible to determine whether the task completed by the test subject is worth the payment. Thus, we calculate the Uti_j with the consideration of the workload of a specific task WL_j, the ratio of task completion CR_j, and the reliability of data collected in the task DR_j, i.e., $Uti_j = WL_j * CR_j * DR_j$. Accordingly, in timeslot t, the reward r_t is calculated based on the state–action pair

Algorithm 1 Task Contract Design With DDPG

for $E = 1, T_{max}$ **do**
  for $t = 1, T$ **do**
    Select action via (7)
    Execute action a_t, calculate reward r_t via (6) and observe next state s_{t+1}
    Store (s_t, a_t, r_t, s_{t+1}) in experience poor $\mathcal{P}$
    Randomly sample N experiences from experience poor $\mathcal{P}$
    Update critic network via (8) and (9)
    Update actor network via (10)
    Update target networks via (11) and (12)
  end for
end for

(s_t, a_t) by

$$r_t = \sum_j \frac{Uti_{j,t}}{\text{Pay}_{j,t}}. \tag{6}$$

As a DRL, the goal of DDPG is to find the optimal action a_t for each state s_t in order to maximize the reward r_t. Therefore, we choose the one of the maximal Q value, which is

$$a_t = \arg\max_{a_t \in A} Q(s_t, a_t). \tag{7}$$

Then, experience (s_t, a_t, r_t, s_{t+1}) is stored in experience poor $\mathcal{P}$.

In the training process, we sample N experience from $\mathcal{P}$ to update the critic network utilizing the following loss function:

$$\mathcal{L}\left(\vartheta^Q\right) = \frac{1}{N}\sum_i^N \left[Q\left(s_i, a_i|\vartheta^Q\right) - \mathcal{Y}_i\right]^2 \tag{8}$$

where

$$\mathcal{Y}_i = r_i + \gamma\left(Q\left(s_{i+1}, \pi\left(s_{i+1}|\vartheta^{\pi'}\right)|\vartheta^{Q'}\right)\right). \tag{9}$$

Accordingly, we update π utilizing policy gradient as

$$\nabla_{\vartheta^\pi} J = \frac{1}{N}\sum_i^N \Big[\nabla_a Q\left(s, a|\vartheta^Q\right)|s = s_i, a = \pi\left(s_i|\vartheta^\pi\right) \nabla_{\vartheta^\pi}\pi\left(s|\vartheta^\pi\right)|s = s_i\Big]. \tag{10}$$

Target networks are copies of the actor π and critic Q networks of different update rules. Once networks πs and Q are updated, we then update the parameters of target networks $\vartheta^{Q'}$ and $\vartheta^{\pi'}$ with a learning rate κ

$$\vartheta^{Q'} = \kappa\vartheta^Q + (1-\kappa)\vartheta^{Q'} \tag{11}$$

$$\vartheta^{\pi'} = \kappa\vartheta^\pi + (1-\kappa)\vartheta^{\pi'}. \tag{12}$$

It is worth to mention that the budget of the CDC is limited. That suggests if the test subject who accepts task $Task_j$ fails to provide reliable personal health data, then the payment Pay_j will be shared by other test subjects as a *reward–punishment mechanism*. We summarize task contract design in Algorithm 1.

C. Task Completion Assessment

Once a test subject TS_i claims the data fusion task $Task_i$ is complete, the task completion should be assessed to determine whether the TS_i should be paid, i.e., the test subject who provides reliable data is paid according to the task contract designed in the previous section. Since more than one test subject of type two might be responsible for data fusion on the same group of test subjects of type two, the personal health data collected by each test subject should not deviate much from each other. That suggests the hypothesis test can be applied to validate the data reliability. We use H_0 to represent the hypothesis of the data being reliable, while the data being unreliable is represented by the hypothesis H_1. We define the false alarm rate (FAR) as the probability of judging a reliable data as an unreliable one. And the missing detection rate (MDR) is defined as the probability of judging an unreliable data as a reliable one. Then, we construct the test static as

$$L = \left\|x_i^k - \hat{x}_i^k\right\|^2. \tag{13}$$

Note that no matter the test statistic is constructed as the deviation between the test data and the reference or the ratio of the deviation, the DRL algorithm DDPG employed by the proposed strategy PDFS can discover the optimal threshold considering the test cost, the FAR, and the MDR. Besides, all test data are normalized within the range of 0 to 1. According to (13), we know that the test statistic falls into the range of (0, 1) such that the searching space is significantly narrowed down. The hypothesis test is then followed as

$$L \underset{H_1}{\overset{H_0}{\lessgtr}} \phi. \tag{14}$$

We update the reference $\hat{x}_i^k$ by $\hat{x}_i^k \leftarrow x_i^k$ only if the data x_i^k are reliable; otherwise, we let $\hat{x}_i^k$ equal to x_i^{k-1}. The reason for that is as follows. It is difficult to determine the true value of the reference in each reliability validation process. Besides, there exists a certain resemblance between the data provided in two sequential timeslots, i.e., the kth timeslot and the $k-1$th timeslot. Specifically, the data provided by two test subjects on a specific test subject should be resemble. That suggests the reliable data in the previous timeslot can be used as the reference in the current timeslot for data reliability validation at the beginning until another reliable data is found. Furthermore, if all data provided in the current timeslot fail to pass the validation process, then the reliable data in the previous timeslot could be used as the reference for the reliability validation in the next timeslot.

The utility of the CDC, denoted by $u(\phi, p)$, is calculated as

$$\begin{aligned} u(\phi, p) = &\sum_{i=1}^N p_i[G_0(1 - MDR(\phi)) - G_1(1 - FAR(\phi))] \\ &+ C\sum_{i=1}^N p_i[FAR(\phi) - MDR(\phi)] \\ &+ (G_1 + C)FAR(\phi) + G_1 \end{aligned} \tag{15}$$

where C represents the validation cost; G_1 and G_0 denote the gain of receiving a reliable data and an unreliable data, respectively; and $p = \{p_i\}$ represents the probability set of data being unreliable.

The hypothesis test in (14) determines the reliability of each data based on the test threshold ϕ. Again, we apply DDPG for the optimal threshold estimation due to its advantage in continuous space searching. Accordingly, we first introduce the state space and action space. Let the state in the tth time slot consist of the FAR and the MDR of validation in the previous time slot be denoted by $s_t = [FAR_{t-1}, MDR_{t-1}]$. For each state s_t, each potential action ϕ_t is chosen by

$$\phi_t = G\left(s_t|\theta^{\pi}\right) \tag{16}$$

where $0 \leq \xi \leq 1$. Based on ϕ_t, we obtain the utility u_t as

$$u_t = \sum u(\phi, p). \tag{17}$$

Let $\phi = \{\phi_t\}$ represent the action set. Thus, we choose the action of the maximal Q value by

$$\phi_t = \arg\max_{\phi_t \in \phi} Q(s_t, \phi_t). \tag{18}$$

Once the action ϕ_t is taken, we can calculate utility u_t and observe the next state s_{t+1} from the environment, meanwhile, $(s_t, \phi_t, r_t, s_{t+1})$ is stored in experience poor $\mathcal{P}$.

In the training process, we sample N experience from $\mathcal{P}$ to update the critic network through the following loss function with N randomly sampled experience from $\mathcal{P}$ as:

$$\mathcal{L}\left(\theta^{\mathcal{Q}}\right) = \frac{1}{N}\sum_{i}^{N}\left[\mathcal{Q}\left(s_i, a_i|\theta^{\mathcal{Q}}\right) - \mathcal{Y}_i\right]^2 \tag{19}$$

where

$$\mathcal{Y}_i = r_i + \delta\left(\mathcal{Q}\left(s_{i+1}, \eta\left(s_{i+1}|\theta^{\eta_i'}\right)|\theta^{\mathcal{Q}'}\right)\right). \tag{20}$$

Then, the actor network is updated utilizing the policy gradient

$$\nabla_{\theta^{\eta}} J = \frac{1}{N}\sum_{i}^{N}\Big[\nabla_a \mathcal{Q}\left(s, a|\theta^{\mathcal{Q}}\right)|s = s_i, a = \eta\left(s_i|\theta^{\eta}\right) \nabla_{\theta^{\eta}} \eta\left(s|\theta^{\eta}\right)|s = s_i\Big]. \tag{21}$$

Then, the parameters of target networks $\theta^{\mathcal{Q}'}$ and $\theta^{\eta'}$ are updated with a learning rate τ as

$$\theta^{\mathcal{Q}'} = \tau\theta^{\mathcal{Q}} + (1-\tau)\theta^{\mathcal{Q}'} \tag{22}$$

$$\theta^{\eta'} = \tau\theta^{\eta} + (1-\tau)\theta^{\eta'}. \tag{23}$$

We summarize the DDPG-based data reliability validation in Algorithm 2.

Algorithm 2 Data Reliability Validation Based on DDPG

for $E = 1, T_{max}$ **do**
  for $t = 1, T$ **do**
    Select action via (16) and (18)
    for $i = 1, N$ **do**
      Calculate L via (14)
      if $L \leq \theta_n$ **then**
        $\hat{x_i} \leftarrow x_i$
        Output the ith data is reliable
      else
        Output the ith data is unreliable
      end if
    end for
    Calculate the utility according to (17) and observe the s_{t+1}
    Store (s_t, a_t, r_t, s_{t+1}) in experience poor $\mathcal{P}$
    Randomly sample N experiences from $\mathcal{P}$
    Update critic network via (19) and (20)
    Update actor network via (21)
    Update target networks via (22) and (23)
  end for
end for

D. Homomorphic Encryption-Based Data Fusion

Once the reliability of personal health data is validated by the task completion assessment, all reliable data should be sent to the CDC. Due to the privacy concern, each test subject is unwilling to use his/her real identity. To achieve identity privacy, homomorphic encryption is employed in data fusion. And, the details are given as follows.

1) *Encryption:* CA generates a pair of public key $p_k(n, g)$ and private key $p_r = \lambda$ to encrypt the data $x \in N$ with a random number $r \in N$ as

$$E(x) = g^m r^n \bmod n^2$$

2) *Decryption:* The data x can be obtained by decrypting ciphertext $E(x)$ with private key λ as

$$D(E(x)) = \frac{L\left(E(x)^{\lambda} \bmod n^2\right)}{L\left(g^{\lambda} \bmod n^2\right)} \bmod n$$

where $L(x) = (x - 1/n)$.

In fact, a Paillier cryptosystem has the following property:

$$E(x_1 + x_2) = E(x_1) * E(x_2). \tag{24}$$

where $E(.)$ is the encryption function. Equation (24) allows us to design a three-steps data fusion mechanism without identity exposure. Let Id_i denote the identity of the ith test subject. Then, we introduce the data fusion mechanism as follows.

1) DFC chooses a random number RN_i and encrypt it with the public key of the CDC, then sends $E(RN_i)$ to the test subject.
2) The test subject calculates a perturbed identity $E(\mathrm{Id}_i + RN_i)$ by $E(\mathrm{Id}_i + RN_i) = E(\mathrm{Id}_i) * E(RN_i)$ with CDC's public key, then sends the data–identity pair $(E(\mathrm{Id}_i + RN_i), E(\mathrm{data}))$ back to the DFC.
3) DFC signs the encrypted data–identity pair and relay it to the CDC.

According to (24), the DFC cannot decrypt the $E(\mathrm{Id}_i + RN_i)$ due to the lack of the private key of the CDC, meanwhile, the CDC can only obtain the perturbed identity $\mathrm{Id}_i + RN_i$ instead of the genuine one of the ith test subject Id_i. It is worth to mention that the perturbed identity generated by adding a random number to a specific real identity might accidentally be

TABLE I
EXPERIMENT PARAMETER SETUP

Parameter	*Description*	*Range*
DR	Data radius of normalized data	[0.05,0.5]
Pay	Normalized payment to test subjects	[0.1, 1]
Num_TR	Number of task releaser	1
Num_TS	Number of test subjects	500
Num_MTS	Number of malicious test subjects	[55, 100]
Num_T	Number of tasks	[50, 500]

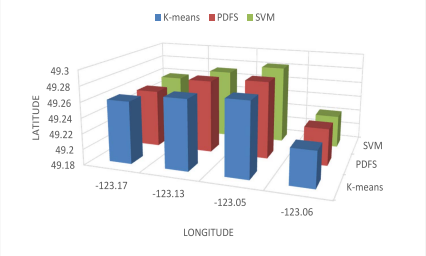


Fig. 4. Clustering accuracy comparison.

identical to the real identity of a certain test subject. In this case, according to such perturbation rule, the CDC might mistake that real identity as one of the perturbed ones. Thereby, the real identity is still kept from the CDC. However, in order to prevent this case and further improve the privacy preservation of test subjects, the random number is generated each time while new data are sent to the CDC.

V. PERFORMANCE EVALUATION

A. Simulation Setup

The simulation is implemented to validate the performance of the proposed strategy PDFS in Python on a computer equipped with Intel Core i7 processor, 64-GB running memory, CPU frequency 6.4GHZ 64-bit win7 system. Table I gives the parameters of this simulation. The data set utilized is about the survival of 306 patients who experienced breast cancer surgery [25].

1) Performance Index: We validate the performance of PDFS in clustering accuracy, average error rate, completion rate, data reliability, and participation rate while considering different values of payment, the number of tasks, and data radius, respectively.

1) *Classification Accuracy:* The deviation between the classification result of different classification algorithms.
2) *Average Error Rate:* Both FAR and MDR compose the average error rate.
3) *Completion Rate:* The ratio of the quantity of completed work to the entire work load.
4) *Data Reliability:* The deviation between a data and the reference should be less than a proper threshold.
5) *Participation Rate:* The percentage of test subjects who participate the data fusion tasks.

B. Experiment Results

1) Classification Accuracy: Fig. 4 shows the comparison result of the classification accuracy among PDFS, *K*-means clustering, and SVM through calculating the positions of group centers established by these four methods. From the comparison result, it can be found that each pair of group centers are close to each other. Compared with *K*-means clustering and SVM, although PDFS adds perturbation into task information for privacy-preserving classification, the established task groups only differ slightly from that of either *K*-means clustering or SVM. That suggests PDFS is more suitable for privacy-enhanced data fusion in MEC-IoMT.

2) Average Error Rate: Fig. 5 shows how *DR* and *Num_T* affect FAR and MDR while either DDPG, DQN, or *Q*-learning is employed by PDFS. The comparison results of FAR of DDPG, DQN, and *Q*-learning are shown in Fig. 5(a), it can be found that because less data provided will results in a larger FAR, so the FAR of all three methods increase as the data radius at first and eventually gets stabilized for each approaches. In addition, it is obvious that the FAR of the proposed PDFS is lower than that of either DQN or Q-learning. The reason lies in that enough data will make the reliability validation more accurate, and then the FAR will drop. In PDFS, by using the DDPG, it can discover the optimal threshold for accurate data reliability assessment. So the FAR of the proposed PDFS is the lowest. The comparison results of MDR of the three methods are shown in Fig. 5(b). As observed from Fig. 5(b), we know that the data radius affects the DDPG much less than that to either DQN or Q-learning due to the similar reason of DDPG having a much less FAR compared with DQN and Q-learning. In Fig. 5(c) and (d), it is clear that the FAR of DDPG is only 6% on average compared with DQN 28% and *Q*-Learning 34%, while the MDR of DDPG is less than 5% compared with DQN 10% and Q-Learning 16% on average as the growth of *Num_T*. This is because once the unreliable data is detected by either DDPG, DQN, or Q-Learning the reward–punishment mechanism will be executed to ensure the test subjects to provide more reliable data. The results in Fig. 5 indicate that DDPG is more effective in data reliability validation compared with DQN and *Q*-Learning.

3) Completion Rate: Fig. 6 shows the completion rate of PDFS, PPCC [23], and REAP [24] under different value of payment and number of tasks. First, we compared the completion rates of the three methods with different values of *Pay*. As shown in Fig. 6(a), since the more rewards can get, the more people are willing to participate in test, so the completion rate of all the three methods increases with the increase of payment. Especially, the completion rate of PDFS is higher than the other two methods due to the reward–punishment mechanism employed by PDFS. Then, we compare the completion rate with different *Num_T*, the results are shown in Fig. 6(b), it is clear that PDFs obtain 78% completion rate on average compared with 55% of REAP and 58% of PPCC. Next, the completion rate is compared among three methods with different *Num_TS* as shown in Fig. 6(c). It is obvious that PDFS outperforms baseline approaches with the highest completion rate 85% compared with 57% of PPCC and 53% of REAP. The reason is the proposed PDFS employs the DRL-based data reliability assessment for unreliable data detection, which makes the malicious test subjects receiving no payment as a punishment once they are detected providing unreliable data.

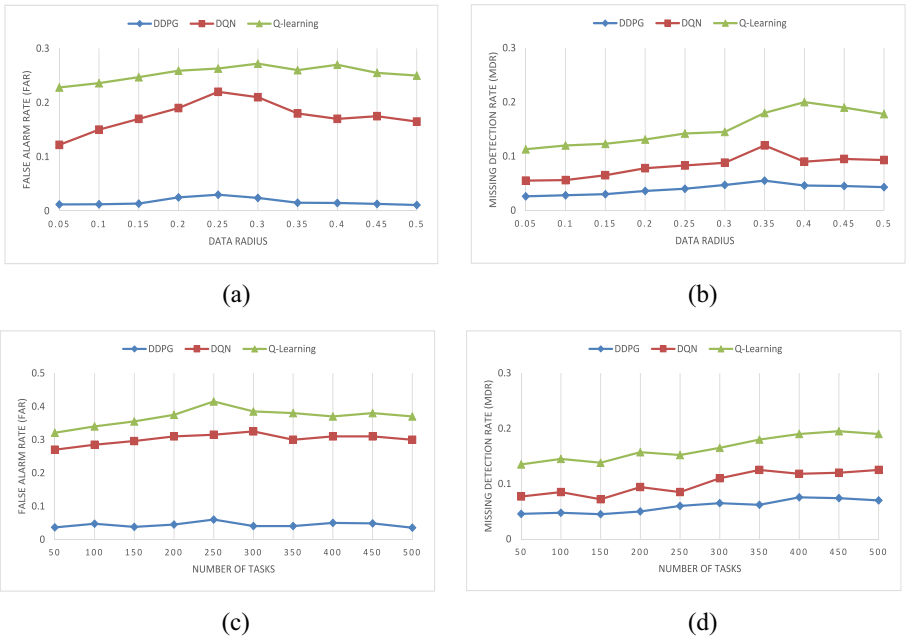


Fig. 5. FAR and MDR while varying (a) and (b) data radius DR, (c) and (d) number of tasks Num_T.

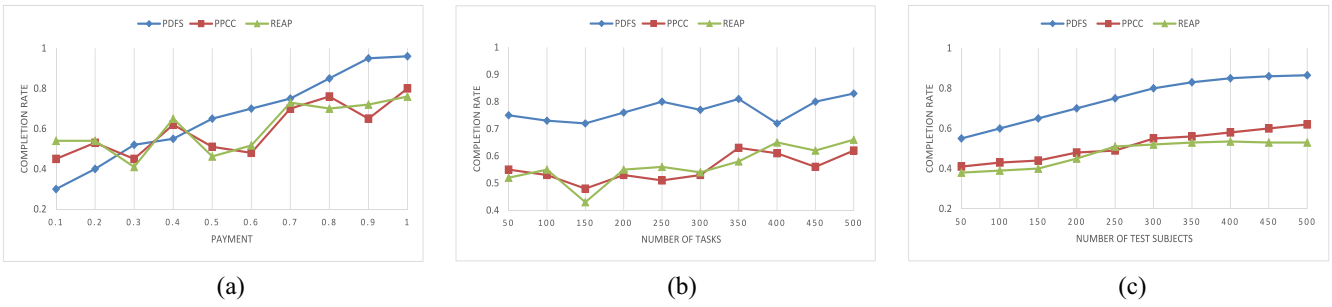


Fig. 6. Completion rate while varying (a) payment *Pay*, (b) number of tasks *Num_T*, and (c) number of test subjects *Num_TS*.

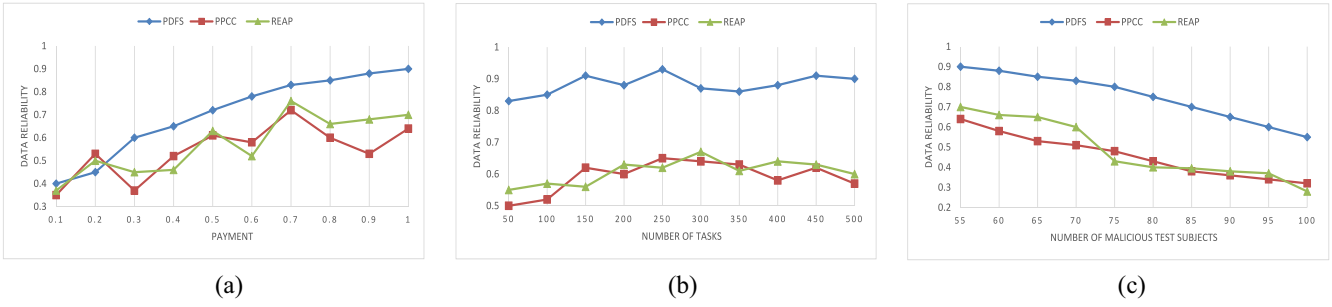
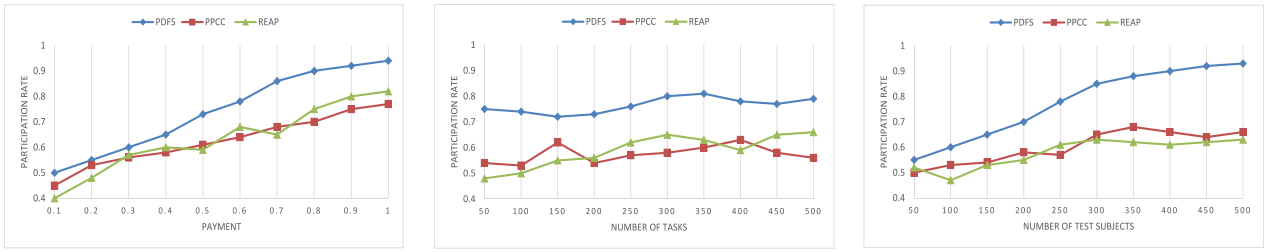


Fig. 7. Data reliability while varying (a) payment *Pay*, (b) number of tasks *Num_T*, and (c) number of malicious test subjects *Num_MTS*.

So the PDFS is capable of ensuring the highest completion rate.

4) Data Reliability: The reliability of the data will have a great impact on the accuracy and validity of the COVID-19 disease test results. Therefore, we compare the reliability of the data of the PPCC and REAP, PDFS with different values of payment and number of tasks. Fig. 7(a) shows the data reliability comparison results of the three methods under different values of payment, and it is obvious that data reliability increases with the payment for each approaches. That is because each test subject is willing to provide reliable if the CDC pays them enough. In addition, we can find that compared with PPCC and REAP, the proposed PDFS obtains the highest data reliability because it adopts the reward–punishment mechanism while the other methods do not use any measures to encourage test subjects participation. Moreover, from the results in Fig. 7(b), we can observed that PDFS achieves the highest data reliability 88% under different number of tasks, while PPCC and REAP only accomplish the data reliability of 58% and 60%, respectively. The impact of *Num_MTS* on data reliability is shown in Fig. 7(c). It is obvious that although data reliability decrease for all



(a) (b) (c)

Fig. 8. Participation rate while varying (a) payment *Pay*, (b) number of Tasks *Num_T*, and (c) number of Test Subjects *Num_TS*.

approaches as the *Num_MTS* increases, the proposed PDFS still has the highest data reliability. The reason behind that is as follows. Even more malicious test subjects join in the data fusion tasks providing unreliable data, the proposed PDFS can efficiently detect unreliable data and refuses to pay malicious test subjects. No doubt that it is effective to impede malicious test subjects to provide unreliable data.

5) Participation Rate: Fig. 8 shows the participation rate of PDFS, PPCC, and REAP under different values of *Pay*, *Num_T*, and *Num_TS*. First, we compared the participation rates of the three methods with different values of *Pay*. As shown in Fig. 8(a), more rewards will encourage more test subjects to participate in data fusion task, so the participation rate of all the three methods increases with the growth of *pay*. Obviously, the participation rate of PDFS is higher than the other two methods due to the reward–punishment mechanism employed by PDFS. Then, we compare the participation rate with different *Num_T*, the results are shown in Fig. 8(b). Finally, the participation rate is compared between all three methods with different *Num_TS* as shown in Fig. 8(c). It is obvious that the participation rate increases as the *Num_TS* and gets stable eventually. The proposed PDFS obtains the highest participation rate 93% compared with 63% of PPCC and 60% of REAP due to PDFS adopts the dynamic payment instead of that depending on availability and privacy degree only as PPCC and REAP.

VI. Conclusion

In this article, aiming at the lack of internal attacks defense and privacy protection faced by the data fusion process, and the need to ensure the real time and accuracy of massive data collection, analysis, and processing required by the COVID-19 detection based on IoMT, a new PDFS is proposed. In PDFS, *K*-means-based privacy-preserving classification mechanism, DRL-based incentive mechanism, DDPG-based task completion assessment method, and homomorphic encryption-based data fusion are deeply integrated into the data fusion process of the COVID-19 detection application to achieve the internal attacks defense and privacy protection of the CDC and test subjects. The simulation experimental results show that PDFS has advantages in task classification accuracy, average error rate, task completion rate, task data reliability, and task participation rate compared with contemporary strategies. The PDFS is efficient against internal collusion attack to enhance privacy security during the data fusion process while improve system performance according to actual needs.

References

[1] S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K. S. Kwak, "The Internet of Things for health care: A comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[2] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of Things for smart healthcare: Technologies, challenges, and opportunities," *IEEE Access*, vol. 5, pp. 26521–26544, 2017.

[3] T. Yang, M. Gentile, C. F. Shen, and C. M. Cheng, "Combining point-of-care diagnostics and Internet of Medical Things (IoMT) to combat the COVID-19 pandemic," *Diagnostics*, vol. 10, no. 4, p. 224, 2020.

[4] R. Wang, H. Liu, H. Wang, Q. Yang, and D. Wu, "Distributed security architecture based on blockchain for connected health: Architecture, challenges, and approaches," *IEEE Wireless Commun.*, vol. 26, no. 6, pp. 30–36, Dec. 2019.

[5] W. Tang, J. Ren, K. Deng, and Y. Zhang, "Secure data aggregation of lightweight E-healthcare IoT devices with fair incentives," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8714–8726, Oct. 2019.

[6] Y. Sun, F. P. W. Lo, and B. Lo, "Security and privacy for the Internet of Medical Things enabled healthcare systems: A survey," *IEEE Access*, vol. 7, pp. 183339–183355, 2019.

[7] M. S. Hossain, G. Muhammad, and N. Guizani, "Explainable AI and mass surveillance system-based healthcare framework to combat COVID-I9 like pandemics," *IEEE Netw.*, vol. 34, no. 4, pp. 126–132, Jul./Aug. 2020.

[8] M. A. Rahman, M. S. Hossain, and N. A. Alrajeh, "Adversarial examples—Security threats to COVID-19 deep learning systems in medical IoT devices," *IEEE Internet Things J.*, early access, Aug. 3, 2021, doi: 10.1109/JIOT.2020.3013710.

[9] M. S. Hossain, "Cloud-supported cyber–physical localization framework for patients monitoring," *IEEE Syst. J.*, vol. 11, no. 1, pp. 118–127, Mar. 2017.

[10] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 450–465, Feb. 2018.

[11] X. Guo, H. Lin, Z. Li, and M. Peng, "Deep-reinforcement-learning-based QoS-aware secure routing for SDN-IoT," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6242–6251, Jul. 2020.

[12] B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, and K. Mankodiya, "Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare," *Future Gener. Comput. Syst.*, vol. 78, pp. 659–676, Jan. 2018.

[13] D. He, R. Ye, S. Chan, M. Guizani, and Y. Xu, "Privacy in the Internet of Things for smart healthcare," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 38–44, Apr. 2018.

[14] J. Hu, H. Lin, X. Guo, and J. Yang, "DTCS: An integrated strategy for enhancing data trustworthiness in mobile crowdsourcing," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4663–4671, Dec. 2018.

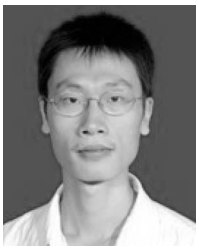
[15] J. J. Hathaliya and S. Tanwar, "An exhaustive survey on security and privacy issues in healthcare 4.0," *Comput. Commun.*, vol. 153, pp. 311–335, Mar. 2020.

[16] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, and G. Wang, "Security and privacy in the medical Internet of Things: A review," *Security Commun. Netw.*, vol. 2018, Mar. 2018, Art. no. 5978636.

[17] Z. Guan *et al.*, "APPA: An anonymous and privacy preserving data aggregation scheme for fog-enhanced IoT," *J. Netw. Comput. Appl.*, vol. 125, pp. 82–92, Jan. 2019.

[18] M. Yang, T. Zhu, B. Liu, Y. Xiang, and W. Zhou, "Machine learning differential privacy with multifunctional aggregation in a fog computing architecture," *IEEE Access*, vol. 6, pp. 17119–17129, 2018.

[19] T. Li, C. Gao, L. Jiang, W. Pedrycz, and J. Shen, "Publicly verifiable privacy-preserving aggregation and its application in IoT," *J. Netw. Comput. Appl.*, vol. 126, pp. 39–44, Jan. 2019.
[20] H. Wu, L. Wang, G. Xue, "Privacy-aware task allocation and data aggregation in fog-assisted spatial crowdsourcing," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 589–602, Jan.–Mar. 2020.
[21] M. Yang, T. Zhu, K. Liang, W. Zhou, and R. H. Deng, "A blockchain-based location privacy-preserving crowdsensing system," *Future Gener. Comput. Syst.*, vol. 94, pp. 408–418, May 2019.
[22] M. Li *et al.*, "CrowdBC: A blockchain-based decentralized framework for crowdsourcing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 30, no. 6, pp. 1251–1266, Jun. 2019.
[23] X. Wang, J. He, P. Cheng, and J. Chen, "Privacy preserving collaborative computing: Heterogeneous privacy guarantee and efficient incentive mechanism," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 221–233, Jan. 2019.
[24] Z. Zhang, S. He, J. Chen, and J. Zhang, "REAP: An efficient incentive mechanism for reconciling aggregation accuracy and individual privacy in crowdsensing," *IEEE Trans. Inf. Forensics Security*, vol. 13, pp. 2995–3007, 2018.
[25] T. S. Lim. (1999). *Haberman's Survival Data Set*. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Haberman



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 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.

Sahil Garg (Member, IEEE) received the Ph.D. degree from the Thapar Institute of Engineering and Technology, Patiala, India, in 2018, where he was the recipient of a prestigious Visvesvaraya Ph.D. Fellowship from the Ministry of Electronics and Information Technology, Government of India.

He is currently a Postdoctoral Research Fellow with École de Technologie Supérieure, Montreal, Canada, and a MITACS Intern with the Global AI Accelerator, Ericsson, Montreal, QC, Canada. He is also a Visiting Researcher with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests are mainly in the areas of machine learning, big data analytics, knowledge discovery, cloud computing, Internet of Things, software defined networking, and vehicular ad-hoc networks. He has over 60 publications in high ranked journals and conferences, including more than over 40 top-tier journal papers and over 20 reputed conference articles. He has been awarded the IEEE ICC Best Paper Award in 2018. He is currently a Managing Editor of *Human-Centric Computing and Information Sciences* journal (Springer). In addition, he also serves as the Workshops and Symposia Officer for the IEEE ComSoc Emerging Technology Initiative on Aerial Communications. He also serves as the special Sessions/Workshop Chair and Publication Chair for CCCI'20 and ICICC'20. He is also the Workshop Chair/Publicity Co-Chair for several IEEE/ACM conferences, including IEEE Infocom, IEEE Globecom, IEEE ICC, and ACM MobiCom. He Guest Edited/Editing a number of special issues in top-cited journals, including IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE INTERNET OF THINGS JOURNAL, *IEEE Network Magazine*, *Future Generation Computer Systems* (Elsevier), and *Neural Computing and Applications* (Springer). He is also an Associate Editor of *IEEE Network Magazine*, IEEE SYSTEMS JOURNAL, *Future Generation Computer Systems* (Elsevier), *Applied Soft Computing* (Elsevier), and *International Journal of Communication Systems* (Wiley). He is a Member of ACM and IAENG, and also actively involved in various technical societies, including IEEE Communications Society, IEEE Computer Society, IEEE Industrial Electronics Society, and IEEE Smart Grid Community.

Jia Hu received the B.Eng. and M.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 70 research papers within the above 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 and 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.

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 School of Fujian Normal University. His main research interests include network optimization and fault tolerance.

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

He is an Assistant Professor with the Department of Computer Science and Engineering, Sejong University, Seoul. Subsequently, he continued his work as a Postdoctoral Research Fellow in the field of Resource Management and Quality of Experience in 5G and Beyond and Internet of Things with the Networking Laboratory, Kyung Hee University. He published substantial number of technical papers in well-known international journals and conferences in research fields of wireless communications and networking, Internet of Things, multimedia communication, applied machine learning, security, and smart grid.

Dr. Piran received the IAAM Scientist Medal of the year 2017 for notable and outstanding research in the field of New Age Technology and Innovation, Stockholm, Sweden. Moreover, he has been recognized as the Outstanding Emerging Researcher by the Iranian Ministry of Science, Technology, and Research in 2017. In addition, 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 field of Electrical and Communications Engineering. In the worldwide communities, he has been an Active Member of the Institute of Electrical and Electronics Engineering since 2010, an Active Delegate from South Korea in 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 *Connected Health in Smart Cities* (Springer). His research interests include cloud networking, smart environment (smart city and smart health), AI, deep learning, edge computing, Internet of Things, multimedia for healthcare, and multimedia big data.

Prof. Hossain currently serves as a Lead Guest Editor of *IEEE Network*, the *ACM Transactions on Internet Technology*, the *ACM Transactions on Multimedia Computing, Communications, and Applications*, and *Multimedia Systems*. He is on the editorial board of several SCI/ISI-indexed journals/transactions, including the IEEE TRANSACTIONS ON MULTIMEDIA, IEEE MULTIMEDIA, IEEE NETWORK, *IEEE Wireless Communications*, IEEE ACCESS, the *Journal of Network and Computer Applications* (Elsevier), and the *International Journal of Multimedia Tools and Applications* (Springer). He is a Senior Member of ACM.