Enabling Privacy-Assured Fog-Based Data Aggregation in E-Healthcare Systems

Cheng Guo, *Member, IEEE*, Pengxu Tian, and Kim-Kwang Raymond Choo, *Senior Member, IEEE*

***Abstract*—Wearable body area network is a key component of the modern-day e-healthcare system (e.g., telemedicine), particularly as the number and types of wearable medical monitoring systems increase. The importance of such systems is reinforced in the current COVID-19 pandemic. In addition to the need for a secure collection of medical data, there is also a need to process data in real-time. In this article, we design an improved symmetric homomorphic cryptosystem and a fog-based communication architecture to support delay- or time-sensitive monitoring and other-related applications. Specifically, medical data can be analyzed at the fog servers in a secure manner. This will facilitate decision making, for example, allowing relevant stakeholders to detect and respond to emergency situations, based on real-time data analysis. We present two attack games to demonstrate that our approach is secure (i.e., chosen-plaintext attack resilience under the computational Diffie–Hellman assumption), and evaluate the complexity of its computations. A comparative summary of its performance and three other related approaches suggests that our approach enables privacy-assured medical data aggregation, and the simulation experiments using Microsoft Azure further demonstrate the utility of our scheme.**

***Index Terms*—COVID-19, data aggregation, e-healthcare, fog-based healthcare, privacy-preserving, wireless body area network (WBAN).**

Manuscript received February 16, 2020; revised April 10, 2020 and May 13, 2020; accepted May 13, 2020. Date of publication May 19, 2020; date of current version November 20, 2020. This work was supported in part by the National Science Foundation of China under Grant 61871064, Grant 61501080, Grant 61771090, and Grant 61601214, in part by the Fundamental Research Funds for the Central Universities under No. DUT19JC08, and in part by the Guangxi Key Laboratory of Trusted Software under Grant No. kx201903. The work of K.–K. Raymond Choo was supported by the Cloud Technology Endowed Professorship. Paper no. TII-20-0800. *(Corresponding author: Kim-Kwang Raymond Choo.)*

Cheng Guo is with the Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province and School of Software Technology, Dalian University of Technology, Dalian 116620, China, and also with the Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China (e-mail: guocheng@dlut.edu.cn).

Pengxu Tian is with the Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province and School of Software Technology, Dalian University of Technology, Dalian 116620, China (e-mail: pengxutian@mail.dlut.edu.cn).

Kim-Kwang Raymond Choo is with the Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249-0631 USA (e-mail: raymond.choo@fulbrightmail.org).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2020.2995228

I. Introduction

MEDICAL monitoring systems, one of the key components in the e-healthcare system, facilitate the collection of information vital to enhancing the quality of healthcare service delivery. In such a system, communications are generally performed over a wireless network, such as wearable body area network (WBAN) that comprises a set of medical sensors. The sensors can be embedded within a patient (e.g., embedded medical devices), worn by a patient (e.g., wearable devices), and/or installed in the healthcare premises (e.g., Internet of Things or Internet of Medical Things devices).

Information such as patients' health-related data can then be continuously and periodically collected and sent to a remote medical server; thus, allowing the analysis of the data and further medical diagnosis by healthcare professionals. However, such medical monitoring systems generally operate in an untrusted environment. This necessitates the protection of sensitive information (e.g., medical data) during transit.

Sensors generally communicate with other entities by using ZigBee or Bluetooth, both of which have relatively short communication ranges. Also, if all of the data collected by different sensors in different regions were to be transmitted to a medical cloud server (MCS), the server would take a long time to analyze the data due to the massive content and the delays in the transmission. It is generally impractical to directly transmit data wirelessly to a remote MCS. Thus, fog computing [1], a middleware interface to cloud computing, has been proposed to provide real-time computing and storage resources. In such a decentralized framework, loads are distributed over location-based servers. Generally, fog servers (FSs) are installed in the untrusted environment and the data (e.g., databases of the FSs) may be at risk. Therefore, both data-in-transit and data-at-rest at the FSs should remain encrypted.

Unlike most existing systems, we aim to deploy a time- or delay-sensitive medical monitoring system. In other words, such systems must be capable of responding in real-time in critical situations. Hence, we can use FSs to implement preliminary analysis on the medical data to facilitate triage and decision making. In such an infrastructure, WBANs collect the medical data and encrypt the data using homomorphic encryption, which ensures that the data are confidential and computable. This allows the FSs to process the encrypted medical data without the need for decryption. However, all known public-key homomorphic cryptographic constructions are time-consuming and thus, impractical.

In this article, we design a privacy-assured, fog-based, data-aggregation approach and apply it in a remote medical monitoring system. Different from traditional aggregation schemes [2], [3], our scheme supports preliminary analysis in FSs. To efficiently and securely implement analysis and aggregation operations in the ciphertext domain, we leverage existing efficient, symmetric homomorphic cryptosystem as introduced in [4]. In such a setting, the relevant devices collect and encrypt the medical data, prior to sending the encrypted medical data to the FSs. The FSs can execute the preliminary analysis and data aggregation operations while preserving the security of the medical data. Finally, the fog servers send the aggregated encrypted data to a MCS for decryption, and the server will be tasked with long-term storage and data analysis to inform detailed medical diagnosis.

There are two key contributions in this article, as explained below.

1) We design a new symmetric homomorphic encryption (SHE)-based data aggregation scheme for e-healthcare systems.
2) We use our proposed scheme in a fog-based architecture to support a time- or delay-sensitive medical monitoring system, as well as considering authentication of the data and data integrity. Specifically, the medical data can be analyzed at the FSs in a secure manner.

The rest of this article is organized as follows. Section II discusses the related data aggregation approaches. Sections III and IV present the problem statement and the relevant preliminaries, respectively. In Section V, we present our proposed scheme. In Section VI, we explain how it can be used for COVID-19 monitoring, and in Section VII, we evaluate the security of our approach using two attack games. In Section VIII we then evaluate the performance of our scheme using Microsoft Azure simulation and evaluate its computational complexity. Finally, Section IX concludes this article.

II. Related Work

Data aggregation can potentially inform decision making and enhance efficiency in a resource-constrained sensor network, although security and privacy are two key concerns due to the insecure communication in open channels (e.g., vehicular networks or e-healthcare system) and the inherent vulnerability of sensors (e.g., always deployed in hostile environments) [5]–[8].

Many approaches have been presented to execute data aggregation in privacy-preserving manners. Lu *et al*. [9], for example, designed a data aggregation framework for multidimensional data by leveraging Paillier encryption technique. To meet the fine-grained demands in smart grid, Li *et al*. [10] and Alsharif *et al*. [11] proposed two aggregation schemes that support multisubset data. In such schemes, the smart grid can collect users' electricity consumption data of different ranges and the number of users in each range. Song *et al*. [12] designed a dynamic data aggregation framework for smart grid architecture. Ke *et al*. [13] designed an architecture by leveraging the existing MapReduce framework. Zhou *et al*. [14] provided several aggregated statistics together with an efficient method for updating data in their three-party architecture. However, all of these schemes are based on cloud computing and hence, have the inherent limitation of network latency.

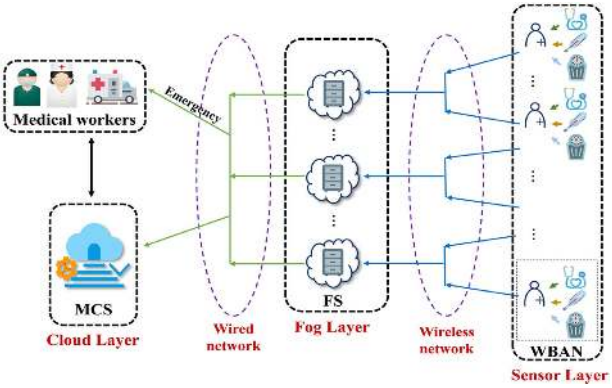


Fig. 1. System model.

To address the problem of network latency, fog computing has been utilized in some approaches. In [15], for example, the authors designed a two-layer encryption scheme in a privacy-preserving, fog-based data aggregation architecture. Their approach is designed to achieve minimal utility loss by distributing the noise generation with a Gaussian mechanism.

Although preserving privacy is of paramount importance for sensor-based systems, preserving the integrity of the data is also a key security requirement. By developing the ElGamal cryptosystem on data authentication, Ara *et al*. [2] designed a data-aggregation approach that also guarantees security. Sun *et al*. [16] designed a privacy-assured emergency-response approach in an E-healthcare framework. Specifically, the data integrity was ensured by leveraging the bilinear pairing technique. In [17], the scheme called P2DA was proposed by using Boneh-Goh-Nissim (BGN) encryption system, which is demonstrated to be secure against internal attacks and can preserve the integrity of the data. Shen *et al*. [18] designed a protocol that can aggregate multidimensional data, as well as preserving the privacy of the data. Also, they proposed a batch verification scheme to make the authentication of the data more secure and efficient.

However, none of the schemes mentioned above can satisfy all of the security requirements, and several schemes are too time-consuming to be practical. Thus, it is necessary to present a novel data aggregation approach to guarantee both the privacy and the efficiency requirements.

III. Problem Statement

A. System Model

The fog-based, health-monitoring system we are proposing is designed to collect medical data and monitor health conditions of patients at a remote location. Our model consists of four types of entities, which are shown in Fig. 1, i.e., medical workers, MCS, FS, and WBAN. For simplicity, we assume only one centralized (medical) cloud server in our system, which connects to m FSs. Each FS connects to n WBANs and each WBAN connects to l various medical sensors that collect l-dimensional medical data in real-time, denoted by $\{md_1, md_2, \ldots, md_l\}$. The communication between the FS and WBAN uses relatively inexpensive, open-wireless technology, and the communication between the MCS and the FSs can be via wired communication technologies, such as the Ethernet or optical fibers.

B. Threat Model and Security Requirements

In our framework, the MCS is regarded as fully trustworthy since it is established by a TTP (e.g., hospital). To design a time- or delay-sensitive medical monitoring system, the hospital needs to setup FSs to facilitate preliminary analysis. Therefore, the FSs can also be regarded as fully trustworthy. The procedure of data authentication can ensure the validity of the FSs. In other words, attackers can only attempt to compromise the database of a FS to obtain the encrypted medical data. In addition, wireless communication technologies are used to connect WBANs and FSs. Attackers can obtain the medical data generated from the medical sensors by eavesdropping the communication networks or tampering the medical data received by the FSs. The key security requirements are summarized below.

Privacy preserving: Attackers cannot obtain the content of the medical data in any stages of our scheme, even when they eavesdrop on the communication channel or compromise the database of the FSs.

Authentication and data integrity: Our proposed scheme can determine whether the received medical data are valid and integrated. In other words, malicious operations, such as forged or modified medical data, can be detected.

C. Design Goals

We aim to design a practical, privacy-assured, medical data aggregation approach and use it in a medical monitoring system. Specifically, our proposed scheme can achieve the following two design goals.

Security: During data transmission and comparing operations at the FSs, data confidentiality must be achieved.

Efficiency: Our proposed scheme is extremely efficient at each stage of the system, i.e., encryption, authentication, comparison, aggregation, decryption of the data.

IV. Preliminaries

A. Bilinear Pairing

$\mathbb{G}_1$ is an additive cyclic group and $\mathbb{G}_2$ is a multiplicative cyclic group with order, q, where q is a large prime number. The bilinear pairing is a map $e:\ \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ that has the characteristics as follows.

1) Computability: The map, e, can be computed effectively.
2) Bilinear: $e(aP, bQ) = e(P, Q)^{ab}\quad \forall P,\ Q \in\ \mathbb{G}_1$ and $\forall a,\ \ b \in \mathbb{Z}_q^*$.
3) Non-degeneracy: $e(P,\ \ P) \neq 1$ for $\exists P\ \in \mathbb{G}_1$.

B. Improved Symmetric Homomorphic Cryptosystem

The symmetric homomorphic cryptosystem was presented in [4]. However, Wang *et al.* [19] demonstrated that the security of the symmetric homomorphic cryptosystem should be looked at further. In our proposed scheme, we improve the original SHE of [19] and prove that it achieves a higher level of security (see Section VI). The improved SHE is described as follows.

$\mathrm{KeyGen}(\lambda)$: The input of the probabilistic key generation algorithm, $\mathrm{KeyGen}(\,)$, is a security parameter λ. The outputs of $\mathrm{KeyGen}(\lambda)$ are u and v, which are two prime numbers and satisfy $u \gg v$. The $\mathrm{KeyGen}(\,)$ randomly and uniformly chooses an integer, $s \in \mathbb{Z}_u^*$, and a parameter, d, called ciphertext degree. The symmetric key of the cryptosystem is an integer pair $K = (s,\ \ v,\ \ u,\ \ d)$.

$\mathrm{Enc}(K, m, r)$: The inputs of the encryption algorithm, $\mathrm{Enc}(\,)$, are the key, $K = (s,\ \ v,\ \ u,\ \ d)$, a plain-message, $m \in \mathbb{Z}_u$, a random integer, r, which satisfy $|r|_2 + |v|_2 < |u|_2$. $\mathrm{Enc}(K, m, r)$ encrypts the plain-message via the following equation:

$$c = \mathrm{Enc}\left(K,\ \ m,\ \ r\right) = s^d\left(rv + m\right) \mod u. \tag{1}$$

$\mathrm{Dec}(K, c)$: The inputs of decryption algorithm, $\mathrm{Dec}(\,)$, is the cipher-message, c, and the key, $K = (s,\ \ v,\ \ u,\ \ d)$. $\mathrm{Dec}(K, c)$ recovers the plain-message via the following equation:

$$m = \mathrm{Dec}\left(K,\ \ c\right) = \left(cs^{-d} \bmod u\right) \mod v. \tag{2}$$

Homomorphic Addition: The result of encrypting $(m_1 + m_2) \bmod v$ can be computed by the addition of c_1 and c_2 if $d_1 = d_2$

$$\begin{aligned} & c_1 + c_2 \\ & \quad = s^d\left(\left(r_1 + r_2\right)v + \left(m_1 + m_2\right)\right) \text{ if } d_1 = d_2 = d. \end{aligned} \tag{3}$$

If $|r_1 + r_2|_2 + |v|_2 + 1 < |u|_2$, the following equation holds:

$$\left(\left(c_1 + c_2\right)s^{-d} \bmod u\right) \mod v = m_1 + m_2. \tag{4}$$

Homomorphic Scalar Multiplication: The homomorphic scalar multiplication can be computed by $c_1 \times m_2$

$$c_1 \times m_2 = s^{d_1}\left(\left(r_1 m_2\right)v + m_1 m_2\right). \tag{5}$$

If $|r_1 m_2|_2 + |v|_2 + 1 < |u|_2$, the following equation holds:

$$\left(\left(c_1 \times m_2\right)s^{-d_1} \bmod u\right) \mod v = m_1 \times m_2. \tag{6}$$

V. Our Proposed Scheme

Our proposed scheme includes four phases, i.e., initialization of the system, generation of the medical report, aggregation of the medical report, and parsing of the data for simplicity, we assumed that the number of WBANs in each FS is a constant number, n. The main procedures of our scheme are presented in Fig. 2.

A. Initialization of the System

The following system parameters need to be generated in the system initialization stage.

Step 1: TTP chooses security parameter, k, generates $(q,\ P,\ \mathbb{G}_1,\ \mathbb{G}_2,\ e\ )$ for a bilinear map by calling $\mathrm{Gen}(k)$ and selects a secure cryptographic hash function, $H(\cdot):\{0,\ 1\}^* \rightarrow \mathbb{G}_1$.

Step 2: TTP releases the system's public parameters, $\{q,\ P,\ \mathbb{G}_1,\ \mathbb{G}_2,\ e,\ H(\cdot)\}$, to every legal entity in the system.

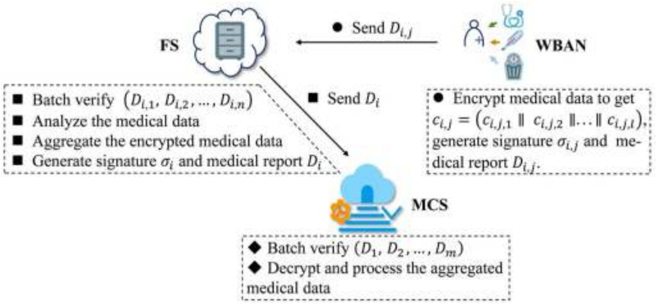

FS
● Send $D_{i,j}$
WBAN
■ Batch verify $(D_{i,1}, D_{i,2}, \dots, D_{i,n})$
■ Analyze the medical data
■ Aggregate the encrypted medical data
■ Generate signature σ_i and medical report D_i
■ Send D_i
● Encrypt medical data to get $c_{i,j} = (c_{i,j,1} \parallel c_{i,j,2} \parallel \dots \parallel c_{i,j,l})$, generate signature $\sigma_{i,j}$ and medical report $D_{i,j}$.
MCS
◆ Batch verify $(D_1, D_2, \dots, D_m)$
◆ Decrypt and process the aggregated medical data

Fig. 2. Our proposed scheme.

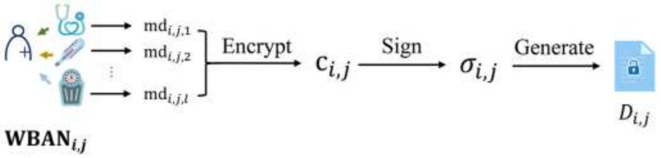


Fig. 3. Medical report generation.

Step 3: Fog server, FS$_i$, chooses $x_i \in \mathbb{Z}_q^*$ to be private key and $Y_i = x_i P$ to be the public key.

Step 4: TTP assigns a secret key $K_{i,j} = (s_{i,j},\ v_{i,j},\ u_{i,j},\ d_{i,j})$ to the wearable body area network, WBAN$_{i,j}$, which belongs to User$_{i,j}$ by calling GenKey(λ), where j is the ID of the WBAN in the area of the ith FS. Since FSs and the medical cloud server are established by TTP, each FS$_i$ possesses the user's secret keys $K_i = \{K_{i,1}, K_{i,2}, \dots, K_{i,n}\}$ where $1 \leq i \leq m$, and MCS possesses all of the secret keys $K = \{K_1, K_2, \dots, K_m\}$.

Step 5: WBAN$_{i,j}$ chooses $x_{i,j} \in \mathbb{Z}_q^*$ to be private key and $Y_{i,j} = x_{i,j} P$ to be public key.

Step 6: TTP deploys some preset ciphertext thresholds to the FSs for each kind of medical data. The thresholds in FS$_i$ are encrypted by $K_i = \{K_{i,1}, K_{i,2}, \dots, K_{i,n}\}$, respectively.

B. Generation of the Medical Report

Each WBAN collects and encrypts the l-dimensional medical data. To estimate the validity of the medical data, it also creates a digital signature to generate the transmitted medical report. The main procedures of medical report generation are presented in Fig. 3

Step 1: WBAN$_{i,j}$, in which i is in the range 1 to m and j is in the range 1 to n, collects medical data $(md_{i,j,1}, md_{i,j,2}, \dots, md_{i,j,l})$ continuously and encrypts medical data by utilizing the aforementioned improved symmetric homomorphic cryptosystem

$$c_{i,j,k} = \text{Enc}\left(K_{i,j},\ md_{i,j,k}\right) \tag{7}$$

where $1 \leq k \leq l$. Following this, WBAN$_{i,j}$ uses $x_{i,j}$ to generate signature $\sigma_{i,j}$

$$\sigma_{i,j} = x_{i,j} H\left(c_{i,j} \parallel ID_{FS_i} \parallel ID_{\text{WBAN}_{i,j}} \parallel Ts\right) \tag{8}$$

where $\parallel$ means the concatenation of each portion, $c_{i,j} = (c_{i,j,1} \parallel c_{i,j,2} \parallel \dots \parallel c_{i,j,l})$, Ts is the time stamp, ID_FS_i is the identity of the FS, and $ID\_\text{WBAN}_{i,j}$ is the identity of the WBAN. Finally, the WBAN$_{i,j}$ generates the medical report by computing

$$D_{i,j} = c_{i,j} \parallel ID\_\text{FS}_i \parallel ID\_\text{WBAN}_{i,j} \parallel Ts \parallel \sigma_{i,j}. \tag{9}$$

Step 2: WBAN$_{i,j}$ sends $D_{i,j}$ to FS$_i$.

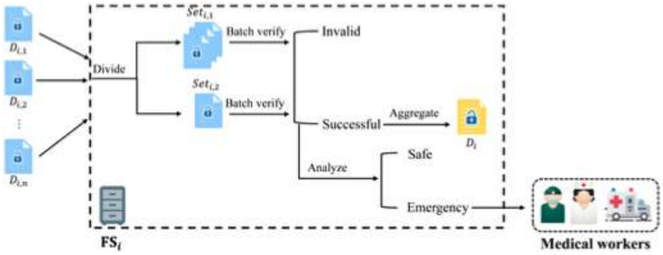


Fig. 4. Medical report aggregation.

C. Aggregation of the Medical Report

FS$_i (i = 1,\ 2, \dots, m)$ uses a batch authentication method to verify the received n medical reports $(D_{i,1},\ D_{i,2}, \dots, D_{i,n})$. Except for data aggregation, another important function in this fog-based medical monitoring system is that some alerts would be generated if a patient is in danger. Because of the huge computing resources of the fog servers, FS$_i$ determines whether the physiological parameters it receives are outside of normal ranges, which reflects the health conditions of patients. An emergency message will be sent to medical workers if a patient is in bad condition. However, all of the comparisons must be executed in the ciphertext domain to protect the security of the medical data. During each aggregation interval, FS$_i$ executes the aggregation operations on the same kind of medical data items. The main procedures of medical report aggregation are presented in Fig. 4.

Step 1: After receiving $(D_{i,1},\ D_{i,2}, \dots, D_{i,n})$, FS$_i$ verifies whether or not these reports are from legitimate WBANs. To make the verification more efficient, we use batch authentication. The FS$_i$ randomly divides Set$_i$ = $\{D_{i,1}, D_{i,2}, \dots, D_{i,n}\}$ into two subsets, i.e., Set$_{i,1}$ and Set$_{i,2}$, where $|\text{Set}_{i,1}| = n/2$, $|\text{Set}_{i,2}| = n/2$, and Set$_i$ = Set$_{i,1}$ + Set$_{i,2}$. FS$_i$ verifies the following equations:

$$e\left(P, \sum_{D_{i,j \in Set_{i,1}}} \sigma_{i,j}\right) = \prod_{D_{i,j \in Set_{i,1}}} e(Y_{i,j}, H(c_{i,j} \parallel \text{ID\_FS}_i \parallel \text{ID\_WBAN}_{i,j} \parallel Ts)). \tag{10}$$

$$e\left(P, \sum_{D_{i,j \in \text{Set}_{i,2}}} \sigma_{i,j}\right) = \prod_{D_{i,j \in \text{Set}_{i,2}}} e(Y_{i,j}\ , H(c_{i,j} \parallel \text{ID\_FS}_i \parallel \text{ID\_WBAN}_{i,j} \parallel Ts)). \tag{11}$$

Step 2: If all of the reports are determined to originate from legitimate WBAN, FS_i analyzes the medical data in a privacy-preserving manner. For example, we assume that $md_{i,j,k}$ represents the heart rate of $\text{User}_{i,j}$. We also assume that if the heart rate is less than a threshold, such as t, the user may need some medical assistance. We use the capital letter "T" to denote the ciphertext of "t", where $-T = \text{Enc}(K_{i,j},\ -t)$. FS_i selects a random positive integer, w, and computes:

$$Z = w \times (c_{i,j,k} + (-T))\,. \tag{12}$$

Then, FS_i decrypts the processed medical data as follows:

$$z = \text{Dec}\,(K_{i,j},\ Z)\,. \tag{13}$$

Based on the homomorphic properties, the decryption result is equal to $w \times (md_{i,j,k} - t)$. Since w is a positive integer, the values of z and $md_{i,j,k} - t$ have the same sign. We define each participating plaintext integer as less than $u/2$. Thus, z could be regarded as a positive number when $0 < z < u/2$, and z could be regarded as a negative number when $z > u/2$. If z holds negative for some time, an emergency message would be sent to medical workers.

Step 3: When receiving several integrated and legal medical reports, FS_i aggregates the same kind of encrypted medical data. The frequency of executing aggregation operations is defined as aggregation interval which is denoted as cn. The aggregation operation is formulated as follows:

$$\text{Agg}_{i,j,k} = c^1_{i,j,k} + c^2_{i,j,k} + \cdots + c^{cn}_{i,j,k}\,. \tag{14}$$

Step 4: When finishing the aggregating operations during an aggregation interval, FS_i uses the private key, x_i, to generate signature, σ_i, for Agg_i as

$$\sigma_i = x_i H\,(\text{Agg}_i \parallel \text{ID}_{\text{MCS}} \parallel \text{ID}_{FS_i} \parallel Ts)\,. \tag{15}$$

where $\text{Agg}_i = (\text{Agg}_{i,1,1} \parallel \text{Agg}_{i,1,2} \parallel \cdots \parallel \text{Agg}_{i,1,l} \parallel \text{Agg}_{i,2,1} \parallel \cdots \parallel \text{Agg}_{i,n,l})$ and ID_MCS is the identity of the MCS. Then, FS_i generates D_i by computing

$$D_i = \text{Agg}_i \parallel \text{ID}_{\text{MCS}} \parallel \text{ID}_{\text{FS}_i} \parallel Ts \parallel \sigma_i\,. \tag{16}$$

Step 4: FS_i sends D_i to the MCS.

D. Parsing of the Data

The MCS verifies the received m aggregated medical reports $(D_1,\ D_2, \ldots, D_m)$ in a batch authentication method. If all of the aggregated medical reports are integrated and received from legal FSs, the MCS decrypts the aggregated medical reports respectively. Finally, the MCS can store and analyze the medical data. The main procedures of parsing of the data are presented in Fig. 5.

Step 1: Like the batch authentication method described in Section V-C, the cloud server estimates whether all the aggregated medical reports are integrated and from legal FSs.

Step 2: Since the MCS holds all the secret keys, it can decrypt the aggregated medical reports for further storage and analysis. The aggregated medical reports are decrypted by the following equations:

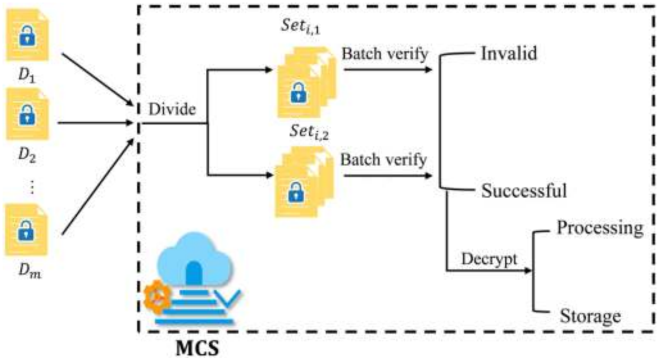


Fig. 5. Data parsing.

$$m_{i,j,k} = \text{Dec}\,(K_{i,j}, \text{Agg}_{i,j,k})\,. \tag{17}$$

The MCS can process the aggregated medical reports that have already been decrypted.

VI. Potential Use Case: COVID-19 Monitoring

In pandemics such as the current COVID-19 pandemic, data may exist in certain local clusters (e.g., a number of nursing homes in a county). Our proposed system allows the analysis of such data at the fog, in order to facilitate timely decision making (e.g., resource management), say by the county or state health departments, as well as COVID-19 monitoring.

In a hospital setting, data collected from devices within the WBANs (e.g., body temperature, cough frequency, and respiratory rate, and patient profile such as whether the patient has other medical conditions) can be used to facilitate preliminary diagnosis, for example, to determine whether further checks are required. However, such data is also sensitive. For example, there have been recent claims that the COVID-19 sufferers may be barred from serving in the military even after they have recovered [22]. Hence, data would also be securely sent to the medical cloud, for more in-depth analysis. This will provide a more comprehensive, global view of the pandemic, improve the quality of healthcare, and potentially minimize fatality rate.

VII. Security Analysis

A. Security of Individual Medical Data

According to the attack in [19], the inputs of their cryptanalytic algorithm were three plaintext/ciphertext pairs, $(m_1, c_1), (m_2,\ c_2), (m_3,\ c_3)$. Then, the attacker must take advantage of the public parameter u to compute $\text{c} = c_1^{-1} c_2 \bmod u$ and all convergent fractions of c/u. However, in our improved symmetric homomorphic cryptosystem, we have changed u as a part of secret key, K, rather than a public parameter. Therefore, the improved cryptosystem is secure against the attack proposed in [19].

In our proposed fog-based data aggregation scheme, we need to use the secret key, $K = (s,\ v,\ u,\ d)$, and a random integer, r, when encrypting plaintext. From a known $(m_i,\ c_i)$ pair, there are five unknown ingredients, i.e., s, v, u, d, and r_i, in the following equation:

$$c_i = s^d \left(r_i v + m_i \right) \bmod u. \tag{18}$$

If the attacker can get β pairs $(m_i,\ c_i)$, an underdetermined nonlinear system of β equations could be generated. However, there are still $\beta + 4$ unknown ingredients. The cryptosystem cannot be broken since it is converted to a nonlinear problem. Therefore, the symmetric homomorphic cryptosystem is semantically secure.

Attack Game 1 (CPA security). For the given (KeyGen, Enc, Dec), defined over $(\mathcal{K},\ \mathcal{M},\ \mathcal{C},\ \mathcal{R})$, we design two games, i.e., Game 0 and Game 1 between a challenger $\mathcal{C}$ and an adversary $\mathcal{A}$. For $b = 0, 1$, we have

Game b:

1) $\mathcal{C}$ randomly selects $k \leftarrow \mathcal{K}$.
2) $\mathcal{A}$ chooses a series of requests and sends them to $\mathcal{C}$. For $i = 1, 2, \ldots,$ the ith request consists of two plaintexts, $m_{i0}, m_{i1} \in \mathcal{M}$, that have the same length.
3) $\mathcal{C}$ randomly selects $r_{ib} \leftarrow \mathcal{R}$, computes $c_i \leftarrow \mathrm{Enc}(k,\ m_{ib},\ r_{ib})$, and responds to $\mathcal{A}$ with c_i.
4) $\mathcal{A}$ estimates whether $\hat{b} = 0$ or $\hat{b} = 1$.

W_b is defined that $\mathcal{A}$ outputs 1 in Game b where $b \in \{0, 1\}$. And the advantage of $\mathcal{A}$ can be expressed as

$$\mathrm{Adv}_{\mathcal{A}}^{\mathrm{CPA}} = \left| \Pr \left[W_0 \right] - \Pr \left[W_1 \right] \right|. \tag{19}$$

Definition 1: semantically secure against chosen plaintext attack (CPA security). We claim that a scheme is CPA secure if, for all polynomial-sized adversary, $\mathcal{A}$, the value, $\mathrm{Adv}_{\mathcal{A}}^{\mathrm{CPA}}$, is negligible.

Attack Game 2 (Bit Guessing). For the given (KeyGen, Enc, Dec), defined over $(\mathcal{K},\ \mathcal{M},\ \mathcal{C},\ \mathcal{R})$, the attack game is designed between a challenger $\mathcal{C}$ and an adversary $\mathcal{A}$

1) $\mathcal{C}$ randomly selects $k \leftarrow \mathcal{K}$.
2) $\mathcal{A}$ chooses a series of requests and sends them to $\mathcal{C}$. For $i = 1, 2, \ldots,$ the ith request consists of two plaintexts, $m_{i0}, m_{i1} \in \mathcal{M}$, that have the same length.
3) $\mathcal{C}$ randomly selects $r \leftarrow \mathcal{R}$ and a bit $b \leftarrow \{0,\ 1\}$, computes $c \leftarrow \mathrm{Enc}(k,\ m_b,\ r)$, and responds to $\mathcal{A}$ with c.
4) $\mathcal{A}$ estimates whether $\hat{b} = 0$ or $\hat{b} = 1$.

We claim that $\mathcal{A}$ wins the game if $\hat{b} = b$. The advantage that the adversary wins the bit guessing game is denoted by $\mathrm{Adv}_{\mathcal{A}}^{BG}$. In Attack Game 1, p_b is defined as the probability that $\mathcal{A}$ outputs 1 in Game b where $b \in \{0, 1\}$. Then, if we focus on the event that $b = 0$ in Attack Game 2, all of the corresponding values are the same as those in Game 0 of Attack Game 1, which is explained as follows:

$$\begin{aligned} \Pr \left[\hat{b} = 0 | b = 1 \right] &= p_0 \\ \Pr \left[\hat{b} = 1 | b = 1 \right] &= p_1. \end{aligned} \tag{20}$$

The probability of winning game is described as follows:

$$\begin{aligned} \Pr \left[\hat{b} = b \right] &= \Pr \left[\hat{b} = b | b = 0 \right] \Pr \left[b = 0 \right] \\ &\quad + \Pr \left[\hat{b} = b | b = 1 \right] \Pr \left[b = 1 \right] \\ &= \frac{1}{2} \left(\Pr \left[\hat{b} = 0 | b = 0 \right] + \Pr \left[\hat{b} = 1 | b = 1 \right] \right) \\ &= \frac{1}{2} \left(1 - p_0 + p_1 \right). \end{aligned} \tag{21}$$

Thus,

$$\mathrm{Adv}_{\mathcal{A}}^{BG} = \left| \Pr \left[\hat{b} = b \right] - \frac{1}{2} \right| = \frac{1}{2} \left| p_1 - p_0 \right| = \frac{1}{2} \mathrm{Adv}_{\mathcal{A}}^{\mathrm{CPA}} \tag{22}$$

which is negligible. Therefore, our proposed scheme is CPA security.

B. Security of Authentication and Data Integrity

Definition 2: Computational Diffie-Hellman assumption (CHD): Let g be a generator of a cyclic group $\mathbb{G}$ whose order is a prime number, p. The CDH problem is defined that it is difficult to compute g^{ab} from the tuple $(g,\ g^a,\ g^b)$ where $a,\ b \in \mathbb{Z}*_p$.

We make use of the BLS short signature [20] in our scheme. The medical data generated from the WBAN are signed by computing $\sigma_{i,j} = x_{i,j} H(c_{i,j} \parallel \mathrm{ID}_{\mathrm{FS}_i} \parallel \mathrm{ID}_{\mathrm{WBAN}_{i,j}} \parallel Ts)$ in which Ts is applied to guarantee security against replay attach. After receiving the signatures, the FS checks to determine whether or not they are legal. In the aggregation interval, if all of the batch verification is successful, another signature, $\sigma_i = x_i H(\mathrm{Agg}_i \parallel \mathrm{ID}_{\mathrm{MCS}} \parallel \mathrm{ID}_{\mathrm{FS}_i} \parallel Ts)$, would be generated for further verification. Since the BLS short signature [20] has been proven to be secure under CDH problem, the authentication and data integrity of our scheme are guaranteed by inheriting.

C. Security of Batch Verification

If there are k signatures that must be verified, $2k$ bilinear pairing operations are required using the ordinary verification approach, which is extremely time-consuming. Our scheme uses a signature to enable data integrity and uses the batch verification method to make the scheme more efficient. The batch-wise verification manner can improve the efficiency as follows:

$$e \left(P, \sum_{r=1}^{k} \sigma_r \right) = \prod_{r=1}^{k} e \left(Y_r, H \left(c_r \parallel ID \parallel T \right) \right). \tag{23}$$

However, forgery attacks would be executed to this kind of batch-wise verification manner, e.g., attacker, $\mathcal{A}$, can choose $a' \in \mathbb{Z}_q^*$ and generate two signatures, $\sigma_1^{'} - a'$ and $\sigma_2^{'} + a'$ which satisfy $\sigma_1 + \sigma_2 + \sigma_3 + \cdots + \sigma_k = \sigma_1' + \sigma_2' + \sigma_3 + \cdots + \sigma_k$.

Our scheme uses the method proposed in [18], where the k signatures are randomly distributed into two subsets, to resist forgery attacks. And the authors in [18] have also proven that the probability that $\mathcal{A}$ successfully forges two signatures (resp. the whole k signatures) is equal to $1/k(k-1)$ (resp. $\frac{((k/2)!)^4}{((k/4)!)^4 \times k!}$) where k is the total number of signatures.

VIII. Performance Evaluation

Our approach can be used to securely aggregate multidimensional medical data by taking advantage of the SHE instead of

TABLE I
TOTAL COMPUTATIONAL COST

Layer	Computational cost
WBAN	$mnlC_e + mnC_m$
Fog server	$m(n+2)C_p + mC_m$
Medical cloud server	$(m+2)C_p + mnlC_e$

TABLE II
COMPARISON OF AVERAGE COMPUTATION COST

Layer	SPPDA[2]	Scheme in [3]	Our scheme
SN/TD/WBAN	$3C_e$	$4C_e$	$lC_e + C_m$
LPU/ES/FS	$(n+1)C_p$ $+2(n-1)C_m$	$(n+1)C_p + C_e$	$(n+2)C_p$+C_m $/sn$
MS/PCC/MCS	$(2m+1)C_p$ $+m(n-1)C_m$	$(m+1)C_p$ $+O(\sqrt{\omega})C_e$	$\big((m+2)C_p$ $+mnlC_e\big)/sn$

TABLE III
COMPARISON OF OUR SCHEME WITH COMPETING DATA AGGREGATION SCHEMES

Layer	[2]	[3]	[20]	Our scheme
Confidentiality	☑	☑	☑	☑
Privacy-preserving	☑	☑	☑	☑
Data authentication	☑	☑	☑	☑
Data integrity	☑	☑	☑	☑
Batch verification	☑	☑	☒	☑
Flexibility	☒	☒	☒	☑
Multi-dimensional data aggregation	☒	☑	☑	☑
Homomorphic addition	☒	☑	☑	☑
Homomorphic multiplication	☑	☒	☒	☑
Preliminary analysis	☒	☒	☒	☑

using common time-consuming public key homomorphic encryption technologies such as ElGamal or other cryptosystems. We evaluated the computational cost of each procedures in our approach, and we also compared it with SPPDA [2] and the scheme in [3].

We assume that the time required for aggregation operations is negligible since the computation cost of addition in $\mathbb{Z}$ is extraordinarily small. And we did not consider the computation cost of comparison operations in FSs because of the varying requirements of the users. The computation cost of an exponentiation in $\mathbb{Z}$, a multiplication operation in $\mathbb{G}$, and a pairing operation are denoted as $C_e,\ C_m$, and C_p, respectively. We assume that there are m FS in our scheme, that each FS serves n WBANs, and that each WBAN contains l medical sensors. The total computational cost, which excludes comparing the data in FSs, is shown in Table I. However, the medical data are encrypted simultaneously by all of the mn WBANs. Thus, the encryption operations are performed by all WBANs concurrently. In addition, the signatures are generated concurrently by all mn WBANs. The average computational cost for a WBAN is $lC_e + C_m$. Similarly, the verification operations are performed by m FSs concurrently, and the FSs aggregate the medical data and generate the signature every sn times. Eventually the average computational cost for the FS is $(n+2)C_p + C_m/sn,$ and the average computational cost for the MCS is $((m+2)C_p + mnlC_e)/sn$.

Table II shows the comparison of the average computation cost for SPPDA [2], the scheme in [3], and our scheme where ω is the size of plaintext domain [21]. Both SPPDA and the scheme in [3] only support one-dimensional (1-D) data aggregation. However, our scheme supports multidimensional data aggregation. Also, SPPDA and the scheme in [3] only support homomorphic multiplication operation and homomorphic addition operation, respectively. Our scheme, on the other hand, supports both homomorphic addition and homomorphic multiplication operations. It can also achieve better performance in homomorphic multiplication aggregation. As observed in Table III, the computational cost of our scheme can be decreased by increasing the aggregation interval, cn. In other words, our scheme is flexible for data aggregation by varying the number of cn. In addition, we compare our scheme with three other data aggregation schemes, as shown in Table III.

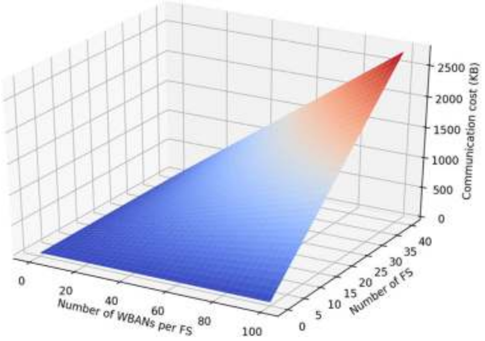


Fig. 6. Communication cost of our scheme.

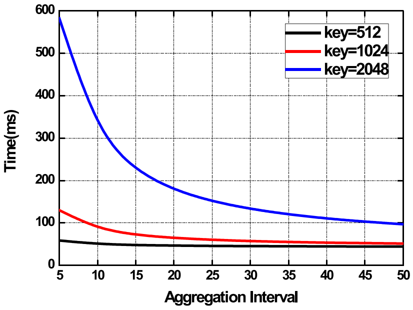
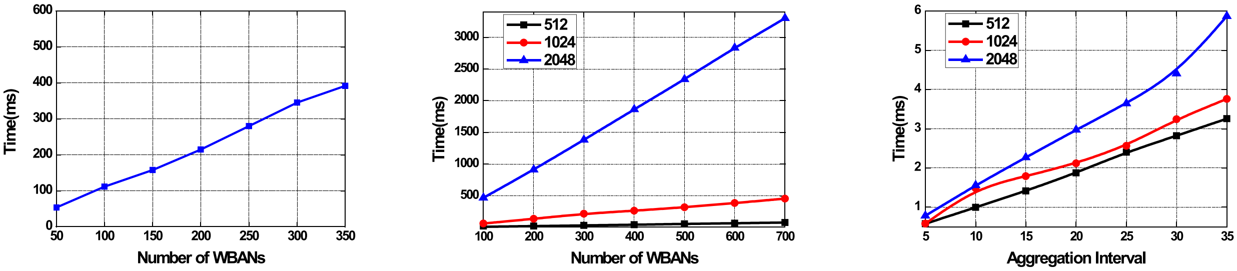


Fig. 7. Average computation cost for various aggregation intervals.

The WBANs is simulated using a PC, and we use Microsoft Azure to simulate the fog servers and cloud server in this article. The FSs and the cloud server are located in U.S., in the same VNET. The CPU configurations are summarized as follows.

1) PC: Intel(R) Core(TM) CPU i7-7700 @ 3.60 GHz.
2) FS: Intel(R) Xeon(R) CPU E5-2673 v3 @ 2.40 GHz 4 cores.
3) Cloud server: Intel(R) Xeon(R) CPU E5-2673 v4 @ 2.30 GHz 16 cores.

The physical memory capacity is 3.8, 8, and 32G, respectively, the operating system (OS) is Ubuntu, and the proposed scheme



(a) (b) (c)

Fig. 8. Partial computational cost in a FS.

is implemented in C. The runtime of the cryptosystems is from the benchmark results of PBC library and GMP library. For the input security parameter λ, the outputs of two primes u and v are with the size λ-bit and $\lambda/2$, respectively. And the bit lengths of s, d, r are $\lambda/2$, $\lambda/4$, and $\lambda/8$, respectively. If there is no explicit statement, the security parameter λ, which is regarded as the length of the key in our scheme, is set as 1024 by taking into account the tradeoff between the security and efficiency.

Fig. 6 is a 3-D diagram that shows the relationship between the communication cost, the number of FSs, (m), and the number of WBANs, (n), for each FS. The communication cost is caused mainly by the medical reports generated by WBANs and the aggregated reports by FSs. The reports are generated by the ciphertext of the medical data, the identities of each entities, the time stamps, and the signatures, whose sizes are 1024, 64, 32, and 1024-bit, respectively. They are used for transmitting the data and for verification of the integrity of the data. By fixing the number of aggregation intervals, the figure demonstrates that the communication cost increases linearly as the number of FSs and the number of WBANS for each FS increase.

Fig. 7 illustrates the average computational time (except the time for data comparing as mentioned above) when the aggregation interval is increased from 5 to 50 by fixing $m = 5$, $n = 50$, $l = 5$. The total computational cost consists of the time for data encryption, the time for signature generation, the time for signature verification, the time for data aggregation, and the time for data decryption. It demonstrates that the average computation cost will decrease as the aggregation interval increases, eventually converging to a certain value. Also, the results show that our approach is extremely efficient. However, the larger aggregation would result in lower accuracy of the medical data. Fig. 7 shows the relationship between the performance of executing the entire pipeline and security parameters, when the number of aggregation interval is fixed. This helps us to determine the tradeoff between efficiency and security.

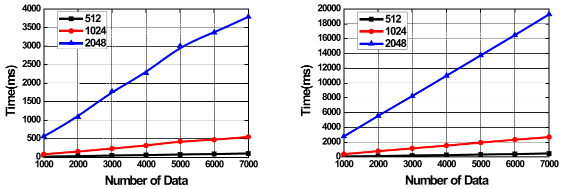
As discussed in Section V-C, the FSs would verify signatures, analyze the medical data, and aggregate the medical data when receiving medical reports. We evaluate the time required for each stage. Fig. 8 shows partial computational cost in a FS. To be specific, number of WBANs means that the number that users served by a FS and aggregation interval means the frequency of executing aggregation operations. By fixing the number of sensors in a WBAN, Fig. 8(a) shows the computation cost for signature verification and Fig. 8(b) shows the computation cost for data analysis, which contains generating Z and decrypting Z. Fig. 8(c) shows the computation cost for data aggregation. All the subfigures in Fig. 8 illustrate that the computation cost increases linearly with corresponding variable. Moreover, the cost of processing the data is rational for a FS.

TABLE IV
Communication Cost From WBAN to FS

l	2	4	6	8	10
$D_{i,j}$ (KB)	0.93	1.53	2.13	2.74	3.34
T (s)	0.398	0.341	0.371	0.343	0.366

TABLE V
Communication Cost From FS to MCS

$n \times l$	200	400	600	800	1000
D_i (KB)	60.49	120.64	180.80	240.95	301.11
T (s)	0.362	0.364	0.371	0.355	0.360

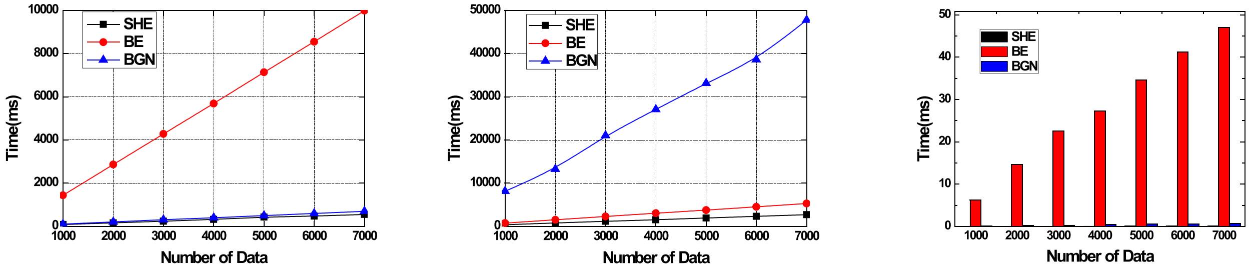


(a) (b)

Fig. 9. Time for data encryption and decryption.

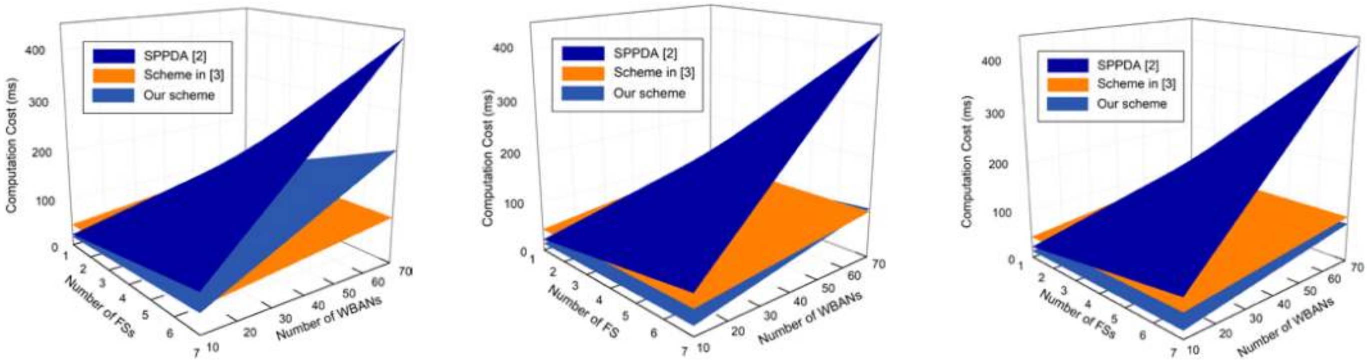
Tables IV and V present the communication cost from WBAN to FS and from FS to MCS, respectively. $D_{i,j}$ represents the size of a medical report to be transmitted and T represents the communication time in corresponding channel. Since the size of a report is relatively small, the propagation delay can be regarded as negligible against the transmission delay. Hence, the communication cost has no relationship with the size of medical report and is almost the same for each medical report. Fig. 9(a) and (b) illustrates the time required to encrypt the data on a WBAN and decrypt the data on the cloud. It is obvious that the time increases linearly as the number of data records. The computation cost is also within an acceptable range.

Fig. 10 shows the comparison between leveraging the improved SHE, bilinear ElGamal (BE) and BGN cryptosystem. The three subfigures in Fig. 10 show the time for encryption,



(a) (b) (c)

Fig. 10. Computation cost for comparison.



sn =5
(a)

sn =20
(b)

sn =30
(c)

Fig. 11. Computation cost for comparison.

decryption, and aggregation by varying different number of data records. To be specific, we evaluate the computation cost for the number of data that participating in computing. The encryption and aggregation operations for BE need to compute in an elliptic curve and the decryption for BGN cryptosystem need to solve the discrete logarithm by Pollard's lambda method. It is clear that the cryptography in our scheme is far more efficient than the others.

Fig. 11 shows the relationship between the computation cost, the number of FSs, (m), and the number of WBANs, (n), for three schemes. However, there is only one aggregated report generated in the scheme in [3] and the report only need to be decrypted once to get an average data for all users, which is different from the two other schemes. Although the decryption for BGN cryptosystem is extremely time-consuming, the scheme in [3] holds the minimum computation cost when $m = 7$ and $n = 70$, which can be observed in Fig. 11(a). Fig. 7 has proven that the computation cost of our scheme can decrease as the aggregation interval, sn, increases. The number of the aggregation interval in Fig. 11(a)–(c) are 5, 20, and 30, respectively. Fig. 11(b) illustrates that the computation cost is almost the same between the scheme in [3] and our scheme when $sn = 20$. Fig. 11(c) illustrates that the computation cost of our scheme is at its lowest when $sn > 20$.

IX. Conclusion

In this article, we focused on achieving privacy-preserving data aggregation in an e-healthcare system. Specifically, we proposed an SHE-based medical data aggregation scheme to achieve efficient homomorphic operations. To support time- or delay-sensitive e-healthcare applications, we used the fog-based architecture in our system model. We also demonstrated that our scheme satisfies key security properties, as well as evaluating its performance using Microsoft Azure.

References

[1] L. Vaquero and L. Merino, "Finding your way in the fog: Towards a comprehensive definition of fog computing," in *Proc. ACM SIGCOMM*, 2014, pp. 27–32.

[2] A. Ara, M. Al-Rodhaan, Y. Tian, and A. Al-Dhelaan, "A secure privacy-preserving data aggregation scheme based on bilinear elgamal cryptosystem for remote health monitor system," *IEEE Access*, vol. 99, pp. 12601–12617, Jun. 2017.

[3] X. Li, S. Liu, F. Wu, S. Kumari, and J. P. C. Rodrigues, "Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4755–4763, Jun. 2019.

[4] L. Li, R. Lu, K. K. R. Choo, A. Datta, and J. Shao, "Privacy- preserving-outsourced association rule mining on vertically partitioned databases," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1847–1861, Aug. 2016.

[5] Y. Wu *et al.*, "Secrecy-driven resource management for vehicular computation offloading networks," *IEEE Netw.*, vol. 32, no. 3, pp. 84–91, May/Jun. 2018.

[6] Z. Xu, F. Chen, Y. Wu, and Y. Gong, "A secure transmission scheme based on artificial fading for wireless crowdsensing networks," *Sensors*, vol. 18, no. 10, pp. 3500–3513, 2018.

[7] Y. Yang, X. Zheng, W. Guo, X. Liu, and V. Chang, "Privacy-preserving smart IoT-based healthcare big data storage and self-adaptive access control system," *Inf. Sci.*, vol. 479, pp. 567–592, 2019.

[8] A. Sohal, R. Sandhu, S. Sood, and V. Chang, "A cybersecurity framework to identify malicious edge device in fog computing and cloud-of-things environments," *Comput. Security*, vol. 74, pp. 340–354, 2018.

[9] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen, “EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1621–1631, Sep. 2012.

[10] S. Li, K. Xue, Q. Yang, and P. Hong, “PPMA: Privacy-preserving multisubset data aggregation in smart grid,” *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 462–471, Feb. 2018.

[11] A. Alsharif, M. Nabil, A. T. Sherif, M. E. A. Mahmoud, and M. Song, “MDMS: Efficient and privacy-preserving multidimension and multisubset data collection for AMI networks,” *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10363–10374, Dec. 2019.

[12] J. Song, Y. Liu, J. Shao, and C. Tang, “A dynamic membership data aggregation (DMDA) protocol for smart grid,” *IEEE Syst. J.*, vol. 14, no. 1, pp. 900–908, Mar. 2020.

[13] H. Ke, P. Li, S. Guo, and I. Stojmenovic, “Aggregations on the fly: reducing traffic for big data in the cloud,” *IEEE Netw.*, vol. 29, no. 5, pp. 17–23, Sep. 2015.

[14] G. Zhou, Q. Jia, L. Guo, M. Li, and P. Li, “Privacy-preserving verifiable data aggregation and analysis for cloud-assisted mobile crowdsourcing,” in *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun.*, 2015, pp. 1–9.

[15] L. Lyu, K. nandakumar, B. Rubinstrin, J. Jin, J. Bedo, and M. Palaniswami, “PPFA: Privacy preserving fog-enable aggregation in smart grid,” *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3733–3744, Feb. 2018.

[16] J. Sun, X. Zhu, and Y. Fang, “Privacy preserving in emergency response based on wireless body sensor networks,” in *Proc. IEEE GLOBECOM*, 2012, pp. 1–6.

[17] D. He, N. Kumar, S. Zeadally, A. Vinel, and L. Yang, “Efficient and privacy-preserving data aggregation scheme for smart grid against internal adversaries,” *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2411–2419, Sep. 2017.

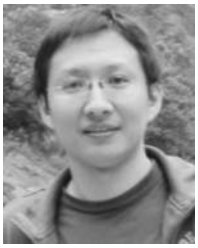
[18] H. Shen, M. Zhang, and J. Shen, “Efficient privacy-preserving cube-data aggregation scheme for smart grids,” *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 6, pp. 1369–1381, Jan. 2017.

[19] B. Wang, Y. Zhan, and Z. Zhang, “Cryptanalysis of a symmetric fully homomorphic encryption scheme,” *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1460–1467, Jan. 2018.

[20] M. Bellare and P. Rogaway, “Random oracles are practical: A paradigm for design efficient protocols,” in *Proc. 1st ACM CCS*, 1993, pp. 62–73.

[21] J. M. Pollard, “Kangaroos, monopoly and discrete logarithms,” *J. Cryptol.*, vol. 13, no. 4, pp. 437–447, 2000.

[22] [Online]. Available: https://www.npr.org/2020/05/07/852319458/the-military-ban-on-covid-19-patients-enlisting-is-yet-to-become-a-policy. Accessed on: May 13, 2020.



Cheng Guo (Member, IEEE) received the B.S. degree in computer science from Xi’an University of Architecture and Technology, Xi’an, China, in 2002, the M.S. and Ph.D. degrees in computer application and technology from the Dalian University of Technology, Dalian, China, in 2006 and 2009, respectively.

From July 2010 to July 2012, he was a Postdoc with the Department of Computer Science, the National Tsing Hua University, Hsinchu, Taiwan. Since 2020, he has been a Professor with the School of Software Technology, the Dalian University of Technology, Dalian, China. His current research interests include information security, cryptology, and cloud security.

Pengxu Tian received the B.S. degree in network engineering from Dalian University of Technology, Dalian, China, in 2016, where he is currently working toward the Ph.D. degree at the School of Software Technology.

His current research interests include searchable encryption, secure database, and fog computing.

Kim-Kwang Raymond Choo (Senior Member, IEEE) received the Ph.D. degree in information security from Queensland University of Technology, Brisbane, QLD, Australia, in 2006.

He currently holds the Cloud Technology Endowed Professorship at The University of Texas at San Antonio (UTSA), San Antonio, TX, USA.

He and his team won the Digital Forensics Research Challenge organized by Germany’s University of Erlangen-Nuremberg, in 2015. He was the recipient of the 2019 IEEE Technical Committee on Scalable Computing (TCSC) Award for Excellence in Scalable Computing (Middle Career Researcher), 2018 UTSA College of Business Col. Jean Piccione and Lt. Col. Philip Piccione Endowed Research Award for Tenured Faculty, British Computer Society’s 2019 Wilkes Award Runner-up, 2019 EURASIP Journal on Wireless Communications and Networking Best Paper Award, Korea Information Processing Society’s Journal of Information Processing Systems Survey Paper Award (Gold) 2019, IEEE Blockchain 2019 Outstanding Paper Award, Inscrypt 2019 Best Student Paper Award, IEEE TrustCom 2018 Best Paper Award, ESORICS 2015 Best Research Paper Award, 2014 Highly Commended Award by the Australia New Zealand Policing Advisory Agency, Fulbright Scholarship in 2009, 2008 Australia Day Achievement Medallion, and British Computer Society’s Wilkes Award in 2008. He is also a Fellow of the Australian Computer Society, and Co-Chair of IEEE Multimedia Communications Technical Committee’s Digital Rights Management for Multimedia Interest Group.